Electronic Journal for Research in
Science & Mathematics Education

*Flagship Journal of the International Consortium for
Research in Science & Mathematics Education (ICRSME)*

ICRSME

CONTENTS

RESEARCH / EMPIRICAL

ICRSME
International Consortium for Research in
Science & Mathematics Education

ICRSME Moments: International Science and Mathematics Collaborations Fostered by ICRSME Consultations

Mark A. Bloom
Dallas Baptist University

Sarah Quebec Fuentes
Texas Christian University

In this editorial, we share the voices of three long-time ICRSME friends. We interviewed each to get their personal experiences of collaborating with educators in ICRSME Consultation host countries. In what follows, Dr. Pamela Fraser-Abder recollects how ICRSME was conceived to foster international research collaborations. Then, Dr. Michael Kamen and Dr. Kathy Horak Smith share their personal collaborative experiences with host-country schools that emerged from ICRSME Consultations.

Pamela Fraser-Abder

Dr. Pamela Fraser-Abder, Professor Emeritus of Science Education at New York University, has been involved with ICRSME from its inception. She shared with us how the first ICRSME consultation developed from an identified need to bring various educational researchers together to collaborate and support each other's work. In 1984, both Dr. Fraser-Abder and Dr. Arthur White, Emeritus Professor of Science, Mathematics, and Technology Education at The Ohio State University, were independently researching education throughout Latin America and the Caribbean. This similar research interest was what initiated the idea of ICRSME:

> It started with [Dr. White], in the U.S., getting funding to come down to Trinidad to work with teachers … he was also working in other Latin America and Caribbean countries. I was in Trinidad at that time and had gotten some funding from UNESCO to do a project on research going on in the Caribbean and Latin America. And what I was finding was that people were literally sitting in their offices, next to each other, and not knowing what research [each other] were doing. And many times their research could be supportive of one another. So, based on that, I got UNESCO to fund a source book on science education research and the Caribbean.

Soon after, Dr. Fraser-Abder presented her work at a meeting of the National Association for Research in Science Teaching (NARST), which is where she and Dr. White met. They also interacted with educators from Panamá and Costa Rica and talks began about how they could bring together these various silos of educational research to share and collaborate with one another.

> [The group] went to Ohio and talked about our research and what we're doing and how we can work together and how we can form this group. [Someone asked] 'Why don't we have a meeting?' And I said, Okay, 'I'll do it in Trinidad.' So that's how the first ICRSME meeting was started.

As Dr. Fraser-Abder described it, the vision that developed was to:

> Pull people at the university and school level together so that we could learn from each other and share the research we had been doing at universities with teachers in classrooms. And the idea was to develop a sort of conference, which was different from the run of the mill conference where you went, you sat down, you listened to people, and then you left. We thought we needed to be more immersed in the culture of the country, so that we got a better understanding of what was happening in each country.

After moving to the U.S., Dr. Fraser-Abder's vision of what ICRSME could be began to develop further:

> I was teaching in New York City, and I would look around my class and I would have [education students] literally from around the world in the classroom … people from Latin America, the Caribbean, India, Europe … and it was their first interaction with somebody who didn't look like them. And they had no idea about the culture of the different peoples that they were going to be teaching. You know, where they came from, what did they eat? What was their religion? What were their views on certain things? So, that sort of made me focus even more on ICRSME as a place where we can learn about the cultures of people who we were meeting in our classrooms.

In what follows, two ICRSME friends describe their own collaborative experiences that developed after attending ICRSME Consultations.

Michael Kamen

Dr. Michael Kamen, Professor of Education at Southwestern University, has been involved with ICRSME for almost as long as Dr. Fraser-Abder. He first attended an ICRSME Consultation in 1991 (the third ICRSME Consultation), while still a doctoral student at the University of Texas (UT) at Austin. In our interview, he shared with us how he and another graduate student had heard that some professors from UT were going to go to a meeting in Merida, Mexico and decided to join them. He described how this decision serendipitously led to collaboration between graduate student and established science educator:

> I remember the first meeting, Pamela [Fraser-Abder] was presenting on something related to my dissertation. And so I ended up connecting with her and talking about it. I don't think I had done a proposal, but I ended up doing a piece of the presentation with her.

Dr. Kamen described how this represented the qualities of ICRSME that kept him involved ever since:

> What I loved about [ICRSME] from the start is it was small, and inclusive. I felt like I got to know science and math educators … at all levels, from some of the very well published names in the field, undergrad students, and everything in between.

He also emphasized how much he enjoyed visiting local schools in host countries, a common activity at ICRSME Consultations that sometimes lead to international collaborative opportunities. He describes one such experience that happened while attending the tenth ICRSME Consultation in Concepción, Chile in 2004:

> The Thomas Jefferson School was a place we visited … I had met Greg [Trezbiatowski], who was the headmaster of the school … And I ended up teaching a summer course for Southwestern [University] students in Concepción. I took [five or six] students down … it's an English immersion school for Chilean kids and they're eager to have native speakers of English at the school … and so my students were placed in classrooms … Greg had arranged for families for them to stay with, they were there for three weeks. … It's one thing neat about South America – our students are off for summer and you can do a summer course and their schools are in session.

The students were able to experience teaching in a foreign country and were able to contrast the learning environment with those with which they were familiar at home. Such rich experiences can only help developing teachers recognize the variation among learners in diverse environments.

Kathy Horak Smith

Dr. Kathy Horak Smith, Professor at Tarleton State University, shared one of her interests with us: "Wherever I travel, I like to be able to see the schools." Attending an ICRSME Consultation provided her this possibility. When we asked how ICRSME is different from other organizations, Dr. Smith shared:

> I see [ICRSME] as different in the fact that it is international, and it is not International and in the United States, it is [hosted in a country outside of the United States]. And when we go there, whatever country it is, teachers from the public schools and [faculty from] the universities [attend]. We get to meet people that are teaching in the public schools in those countries. … That for me is the biggest excitement about ICRSME.

Dr. Smith's experience attending her first ICRSME Consultation in Nassau, Bahamas in 2006 exemplifies the opportunities she described. She came to the ICRSME Consultation not knowing anyone and left with the foundation for a long-term collaboration.

While in the Bahamas, Dr. Smith and several other ICRSME attendees, including Dr. Michael Kamen, asked Bahamian teacher attendees if they could visit their local schools. The school visits initiated a partnership between Dr. Smith and one of the teachers. After the Consultation, Dr. Smith reached out to her Bahamian colleague and asked if she could return for a visit. For the next five years, Dr. Smith visited this teacher's school annually for a week at a time. On her visits, she taught model lessons (Figure 1) and conducted professional development, addressing needs identified by her colleague (e.g., problem solving). Dr. Smith also talked about what she learned from these experiences:

> I learned from the Consultations whether I go into the schools or just listen to the teachers, during sessions or [outside of the] sessions. We visit to learn about their schools and how they [teach] and then I can talk about that in my own classroom [as a teacher educator].

She further explains that mathematics and mathematics instruction are not universal. ICRSME provides attendees the opportunity to learn about different cultures beyond a typical tourist experience through its site-based nature with chances to engage with local teachers and schools.

Figure 1
Dr. Smith Teaching a Mathematics Lesson in the Bahamas

Conclusion

As we wrote about in the previous editorial, collaboration is an integral part of ICRSME, both in how ICRSME was formed and in the experiences of Consultation attendees. The stories shared in the present editorial describe such collaborations from an organizational and individual level. Dr. Fraser-Abder described how the origin of ICRSME stemmed from similar research being conducted internationally. Both Drs. Kamen and Smith talked about partnerships resulting from school visits while attending ICRSME Consultations. These collaborations flourished beyond the Consultations themselves.

In our conversation with Dr. Kamen he talked about *ICRSME moments*, a term he coined to mean "moments of interaction, culturally and unexpected." For Dr. Kamen, examples of such moments included meeting Dr. Fraser-Abder as a graduate student or spontaneously asking local teachers if he and some ICRSME colleagues could visit classrooms. We hope that you have the opportunity to attend an ICRSME Consultation, whether ICRSME XVI in Panamá City, Panamá or a future Consultation, to interact with ICRSME friends, potentially establish collaborative experiences, and make your own ICSRME moments.

ICRSME
International Consortium for Research in
Science & Mathematics Education

Houston-Louis Stokes Alliance for Minority Participation: Findings from 17 years of a Multi-institutional Consortium Focused on Building Minority Student Success in STEM

Dina Ghazzawi (iD)
University of Houston

Donna Lynn Pattison (iD)
University of Houston

Catherine Horn (iD)
University of Houston

Bobby Wilson (iD)
Texas Southern University

ABSTRACT

This rich, longitudinally descriptive study provides an examination of program outcomes, student characteristics, and workforce outcomes of the Houston-Louis Stokes Alliance for Minority Participation (H-LSAMP) program. Utilizing data from the University of Houston's Education Research Center, this study offers a detailed analysis of the long-term effects of participation in the H-LSAMP program, from high school to university and into the workforce. Findings from this study revealed that only 25% of the high school campuses attended by H-LSAMP students had a high proportion of economically disadvantaged students. In addition, nearly 75% of undergraduate students enrolled in the program graduated within the timeframe of the study, half of which were from Black and Latinx backgrounds.

Introduction

Wide disparities in science, technology, engineering, and mathematics (STEM) degree attainment across race and ethnicity remain a persistent concern across the nation (Dika & D'Amico, 2016; Hurtado et al., 2010; National Board of Sciences, 2010; Riegle-Crumb et al., 2019). Despite the nation's shifting demographics, student achievement in STEM fields fail to reflect the diverse backgrounds of the population (Allen-Ramdial & Campbell, 2014; James & Singer, 2016).

Under-represented minority (URM) students in the United States are formally defined as those students from Black, Mexican American, American Indian, Hawaiian Native, Alaskan Native, or Mainland Puerto Rican backgrounds (Association of American Medical Colleges, 2003; Page et al., 2013). Although they enter STEM majors with the same degree of interest as their peers, attrition rates among under-represented minority students are much higher (Chang et al.,

2014; Chen, 2013; National Academies of Sciences, Engineering and Medicine, 2016). For Black and Latinx Students, 6-year STEM undergraduate degree completion rates are 22% and 29%, respectively, compared to 43% for White students (National Academies of Sciences, Engineering and Medicine, 2016). According to a recent report by the U.S. Bureau of Labor Statistics (2021), employment in STEM occupations is projected to increase by approximately 8% by 2029, demanding an approximate increase of 80,000 jobs in the STEM workforce. National data project that by the year 2060, almost two-thirds of the U.S. population's youth will be of color (U.S. Census Bureau, 2019). Despite advances in efforts directed towards increasing diversity among the nation's STEM workforce, there still remains wide disparities in the diversity of the STEM workforce compared to that of the general population (National Center for Science and Engineering Statistics (NCSES), 2021. Although Latinx and Black students making up 18.5% and 13.4% of population demographics, respectively, only 15% and 9% of STEM undergraduate degrees were awarded to Latinx and Black students in the 2018/2019 academic year (NCSES, 2021). As the U.S. is rapidly transitioning into a non-White majority nation, these findings clearly portray the importance of URM students' educational outcomes and success in STEM to the nation's economic vitality and global prominence (National Academies of Sciences, Engineering and Medicine, 2016).

URM students face a multitude of challenges and barriers in their journey toward STEM degree completion (Chen et al., 2014). Lack of academic pre-college preparation, lack of support from faculty and peers, as well as alienation and racism have been identified as some of the barriers faced by minority students in STEM (Strayhorn, 2013). In addition, research studies suggest URM students, particularly first-generation, low-income students, are more likely to experience financial burdens as a result of the high cost of pursuing a STEM degree (Hurtado et al., 2010; Stephens et al., 2012). Consequently, economically disadvantaged URM students are more likely to take on debt and work in order to afford their living and tuition expenses compared to their peers (Stephens et al., 2012). Such pressures play a significant role in STEM attrition rates among URM students, as well as their adjustment and sense of belonging to their campus culture (Hurtado et al., 2010).

Another major barrier faced by URM student populations is the inadequate academic preparation they may receive in high school, particularly in the areas of mathematics, lowering their odds of persisting through their STEM degree (Chang et al., 2014; Lisberg & Woods, 2018; Riegle-Crumb et al., 2019). The wide disparities in academic preparation contribute not only to the significantly higher attrition rates among URM populations, but also create an unequal playing field among students even prior to their admission into college (Lisberg & Woods, 2018; Riegle-Crumb et al., 2019). A significant body of research highlights the vital role of student's precollege academic background on their success in STEM fields (Crisp et al., 2009; Saw et al., 2018; Shaw & Barbuti, 2010). Rigorous course-taking in math, particularly in the first year of study, can have a significant effect on a student's likelihood of persisting in a STEM degree (Chen, 2013; Whalen & Shelley, 2010).

The nation's economic prominence and position as a leader in STEM fields, along with innovation and technology, is inextricably linked to the quality, skill and diversity of its workforce. Improving the nation's STEM workforce has transcended merely increasing quantity and skill. In the current technological landscape, advances and discoveries in science thrive on teamwork and diverse skills and perspectives (National Board of Sciences, 2010). Approximately 99% of jobs in the STEM workforce require a post-secondary degree (Fayer et al., 2017). To meet this increasing demand, significant efforts must be made to recruit, retain, and assist students from URM backgrounds to successfully complete college and enter the STEM workforce (Holden et al., 2010).

Over the past decade, STEM intervention programs around the nation have embarked upon substantive efforts to increase the enrollment and collegiate success of under-represented minority students in STEM majors (Gilmer, 2007; LaCourse et al., 2017). The Houston-Louis Stokes Alliance for Minority Participation (H-LSAMP) represents a multi-institution alliance designed primarily to improve URM student persistence and graduation in STEM. Supporting research highlights the importance of STEM intervention programs at increasing the academic preparation, retention, and student success of URM students in STEM fields (Carpi et al., 2017; Jackson & Winfield, 2014). Furthermore, research emphasizes the value of certain intervention experiences in successful career outcomes of graduates (e.g., Griffith, 2010; Pascarella & Terenzini, 2005; Xu, 2013). For instance, extant literature on features of successful STEM intervention programs highlights their ability to foster minority students' sense of science identity and self-efficacy (Carlone & Johnson, 2007; Espinosa, 2011; Robinson et al., 2019). Xu (2013) found that science identity values among minority students were a significant predictor of STEM career retention. In addition, the academic support provided through STEM intervention programs, in the form of mentorship, advising, and academic preparation in math, play an integral role in STEM career pathway retention, particularly among under-represented student populations (Jelks & Crain, 2020; Stipanovic & Woo, 2017; Xu, 2013). With this evidence in mind, this study aims to demonstrate program achievements of the H-LSAMP in terms of bachelor, graduate degree attainment, and workforce outcomes across various socio-demographic and pre-college characteristics.

H-LSAMP

About the Alliance

The H-LSAMP program was established with the goal of increasing the number of qualified URM students who earn a baccalaureate degree in the STEM fields. As a partnership, H-LSAMP particularly strives to prepare minority students with the academic and career skill-sets to successfully pursue a graduate degree and/or career in a STEM field by working towards eliminating the two well-documented causes of attrition in STEM fields at the University level: financial need and academic support. To increase diversity among the STEM workforce, and bring its numbers in line with population demographics, the consortium relies on empirical research on student retention, academic integration, and collaborative research to guide program initiatives and components most beneficial to minority students' academic achievement.

At its core, H-LSAMP was founded with the goal of providing academic support to increase student retention. To achieve this enrichment core, program components reflect the principles of Treisman (1992), centered on collaborative learning communities (CLC), that promote academic integration and sense of belonging among URM students. Research highlights the tendency of URM students to self-isolate from peers, depending on themselves to succeed through their academic journey (Treisman, 1992). The CLC model emphasizes the added value of peer group learning techniques, faculty mentorship, and academic supplementary support on URM academic achievement and persistence in STEM. The success of the CLC model in promoting the academic integration, sense of belonging, and academic achievement of URM students is well-documented and has been widely implemented in various institutions across the nation (Bonsangue et al., 2018; Chinn et al., 2007; Drew, 2011; Duncan & Dick, 2000). Guided by the CLC model, partner institutions dedicate significant resources to provide H-LSAMP scholars with the right academic success skills that enable them to succeed in their respective STEM major, and withstand the various pressures associated with pursuing a degree in STEM. At every partner institution, a dedicated space is provided for H-LSAMP students to gather, study and attend workshops. In addition, partner institutions provide student skills training, which

assists students in building better study habits and time management skills, goal setting capabilities, and communication skills. Personal and professional development workshops are also offered to provide support in areas such as student commitment to their major and career, as well as work ethics, behaviors, and attitudes.

In terms of core academic development, collaborative learning experiences are promoted through both formal and informal group learning practices, peer-facilitated workshops, both peer and faculty mentors, and tutors. Peer and faculty mentorship is a crucial component of the H-LSAMP program, particularly in its benefit for URM student success, academic integration, and sense of belonging (Holland et al., 2012; Kendricks et al., 2013). Collective research findings emphasize the importance of faculty mentorship in predicting minority student academic performance in STEM (Kendricks et al., 2013, Wilson et al, 2010). In addition, informal peer mentors can also ease the transition of students into their STEM field of study, and increase the sense of camaraderie among students, which ultimately builds URM students' socio-academic integration and sense of belonging to their academic community (Carlone & Johnson, 2007; Holland et al., 2012).

Program Components

Components of the program are carefully designed to recruit, prepare and support high school students and incoming freshmen from disadvantaged, under-represented backgrounds.

1) **High School Recruitment:** the alliance focuses on active recruitment efforts directed towards high schools with a high percentage of minority enrollment. These include recruitment efforts targeted towards high school counselors and events such as College Night Out, that attract students from across the state. In addition, the alliance connects with schools on a yearly basis to participate in STEM competitions that motivate students' interests in STEM fields.
2) **Summer Camps:** each year, the alliance conducts engineering camps that introduce both middle and high school students to different engineering fields and careers through hands-on experience and conversations with industry leaders and professors.
3) **Pre-freshman Enrichment Program (PREP):** to combat inadequate pre-college academic preparation, particularly in the areas of mathematics, the alliance runs a yearly, academic intensive, seven-week program for high school students aiming to enter a STEM college major. The program includes SAT preparation and science workshops, collaborative learning environments, and STEM career awareness talks to introduce and prepare potential students for a successful career in their chosen STEM field.
4) **Summer Bridge Program:** focused on increasing the first-year success of at-risk students enrolled in STEM fields, the summer bridge program provides an intensive month-long calculus preparation, followed by four more weeks of course work that stimulates their readiness for university courses. Through the use of self-guided studying, collaborative learning environments, mentorship, and additional faculty support and guidance, the summer bridge program provides at-risk students with invaluable preparation and skills to persist through their STEM degree.
5) **Financial Support and Assistantships:** Students participating in the H-LSAMP are awarded stipends based on their academic credentials and financial need. Despite the strong academic support offered through the H-LSAMP program, research strongly suggests that socio-economic difficulties are the main force behind student attrition rates in STEM fields (Barton, 2003; Malcom & Dowd, 2008). Stipends are awarded annually, for up to five years, and are contingent upon students' active participation in the H-LSAMP program activities. In

addition, and in line with the collaborative learning model which encourages social integration among students, level one students funded through stipends are encouraged to offer three hours each week to the learning community. Such support could be in the form of a workshop facilitator, tutor, or computer lab assistant.

Collectively, the different elements within the H-LSAMP program work in tandem to equip high school graduates and freshmen students with the foundational tools, as well as the academic and financial support to prepare them for success in their respective STEM majors and a seamless transition into the STEM workforce.

Purpose of the Study

This study provides a rich descriptive longitudinal study of the graduation and workforce outcomes of students enrolled in the H-LSAMP program from the first graduating cohort in the 2004/2005 academic year until 2019/2020. As part of a broader set of on-going longitudinal studies that examine program participation effects on student success outcomes, this study provides essential foundational understanding of H-LSAMP program components, initiatives, and outcomes related to student achievement, retention, and graduation. This study will provide the foundational descriptive basis for which future, more detailed analyses of the impact of math and financial support, among other variables, could potentially impact student success. Despite the program being in effect for over 20 years, few studies have explored overall, multi-institutional program impacts in terms of degree completion, and matriculation to graduate school and workforce outcomes. This study uniquely adds to the existing body of literature on STEM intervention outcomes by tracking students' progress, graduation outcomes, and mobility across an extended period of time and across multiple participating institutions in a single urban area. Such expansive data availability lends the study the ability to follow H-LSAMP scholars from high school and into the workforce, offering a broad yet detailed description of program outcomes over time. In particular, our study addresses the following research questions:

1) Descriptively, to what extent do H-LSAMP graduates come from economically disadvantaged high-school backgrounds?
2) What are the H-LSAMP undergraduate and graduate degree outcomes, particularly across race, field of study, and gender?
3) To what extent do intervention efforts lead to careers in STEM fields as identified by the Standard Industrial Classification (SIC) codes?

Theoretical Framework

The development of the H-LSAMP was empirically supported by Treisman's (1992) Mathematics Workshop Model. The model centralizes around characterizing and promoting minority students' academic skills in both mathematics and science courses, through the establishment of a collaborative learning environment with a strong and supportive peer network. Treisman's model is founded upon concepts of sense of belonging and academic integration, enabling minority students to learn science and mathematics in a more effective manner through implementing various support resources for these students, including advising support, peer led team learning, and faculty mentorship. The model was initially implemented at the University of California, Berkeley, and led to successful academic outcomes among participants of the program. Subsequently the program, along with the collaborative learning techniques, have been successfully implemented in various other STEM intervention programs across the nation (Chin et al., 2006; Duncan & Dick, 2000). To that end, the

concepts of Treisman's model, along with the supporting evidence of the success of undergraduate STEM intervention programs on the retention and graduation of minority students, provide a conceptual lens from which to assess the success of the program at supporting minority student success in STEM.

Data Source & Methods

Data for this rich, longitudinally descriptive study were obtained through the University of Houston's Education Research Center, which contains key longitudinal student-level information obtained from the state education agencies, higher education agencies as well as the state workforce commission. The data sample used for this analysis contained graduates from the H-LSAMP program beginning from the first undergraduate graduating class in 2004/2005 until 2019/2020 (n=2,044). Participant data analyzed in this study was limited to level-one H-LSAMP students, who were funded by the H-LSAMP program across the consortium's partner institutions. Using data from the University of Houston's Education Research Center, this study allowed for the tracking of students over time using longitudinal data beginning from high school, to higher education, and into the workforce.

Data Analysis

Given that the scope of this study is to provide a rich descriptive analysis of H-LSAMP student characteristics and graduation outcomes, data analysis involved the process of descriptive statistics in the form of cross-tabulations and frequency tables to assess the proportion of H-LSAMP graduating students across major, race, and similarly across workforce occupation. To assess the percentage of H-LSAMP students from disadvantaged backgrounds, school accountability ratings were merged with graduation data in order to match H-LSAMP students with their respective high school campus and rating. Finally, H-LSAMP graduation files were merged with workforce data, where SIC codes were identified to describe the workforce occupation of students following their completion of the program.

Limitations

Although this study presents data indicative of students' postsecondary and workforce outcomes, results of this analysis only provide data for students who remained in the state after graduation. For this reason, students who pursued graduate degrees in other states would not be captured in this analysis. Therefore, graduate degree outcomes presented in this study may not be an accurate representation of program results outside of the state. In addition, though a key indicator of students' socio-economic status, pell eligibility of H-LSAMP scholars could not be captured in this analysis due to the abundance of missing data in this field. Several research studies posit that this is due to the reduced likelihood of students, particularly those from socio-economically disadvantaged backgrounds, to file the Free Application for Federal Student Aid (De La Rosa, 2006; Feeney & Heroff, 2013). For this reason, the lack of inclusion of pell eligibility status limits the results of this analysis. However, other indicators of socio-economic status were included such as high school accountability ratings, and the proportion of economically disadvantaged students at high school campuses. One of the main objectives of the program, and similar STEM intervention programs across the nation, is to equip students with the financial and academic support needed to access and successfully complete a STEM undergraduate degree (Lisberg & Woods, 2018). Recruitment initiatives for the H-LSAMP undertaken by program administrators focus on targeting students from economically disadvantaged backgrounds (Ghazzawi et al., 2021). Despite the unavailability of precise

data on student socio-economic status, this study uses the increased probability of consortium students being from socio-economically disadvantaged backgrounds as a proxy for financial need.

Results

Findings from this study found that the majority of high school campuses (67%) attended by H-LSAMP students had a favorable accountability rating. In addition, only 25% of high school campuses attended by H-LSAMP scholars had a high proportion of economically disadvantaged students. Under-graduate degree outcomes portrayed that approximately half of all H-LSAMP graduates were from Black and Latino backgrounds. Results also showed that nearly half of all under-graduate degrees obtained were in the field of Natural Science and Mathematics (NSM), and 34.5% of under-graduate degrees were obtained in Engineering fields. In terms of graduate degree outcomes, our findings indicate that 319 H-LSAMP undergraduates went on to pursue graduate degrees, 56% of which were in STEM fields of study. Finally, analysis of workforce outcomes demonstrated that 38% of H-LSAMP graduates were employed in STEM or STEM-related occupations, while 62% were not.

Research Question 1: High School Background, Accountability Ratings and Socio-economic Status

Given supporting evidence of the association between high-school academic experiences and collegiate success of minority students, an important component of this descriptive study involved the overview of high school accountability ratings as well as the percentage of economically disadvantaged students enrolled at those campuses attended by H-LSAMP scholars. Minority students are more likely to attend schools with a lower socio-economic status (Goldsmith, 2011), inevitably leading to a lack of resources and less instruction in advanced courses such as mathematics and science (Martinez & Guzman, 2013). As one of the major recruitment objectives of the alliance is to reach minority students from economically disadvantaged, low-performing high schools, descriptive data concerning the high school background of H-LSAMP scholars offers key indicators on the extent to which recruitment tactics are targeting the intended student population.

State campus accountability ratings were collected for the 2018/2019 academic year. These ratings evaluate school performance based on student achievement, school progress, and closing gaps across racial ethnic groups. Children at Risk (n.d.) ranks campuses according to student achievement, with a particular attention to student performance comparisons across campuses with similar low-SES levels, as well as student growth and college readiness. Campuses also receive an overall rating ranging from A to D, and an F rating is assigned to those campuses that did not meet performance criteria to earn at least a D rating (Texas Education Agency, 2021). High school campus ratings of H-LSAMP scholars are presented in Table 1. In addition to campus accountability ratings, descriptive data concerning the percentage of economically disadvantaged students at each campus were collected. Results are presented in Table 2. A total of 265 high school campuses were rated, including 18 campuses with ratings unavailable during 2019. Approximately 25% of high school campuses received a rating of C or below, while 67% of campuses received a rating of B- or higher.

Table 2 represents the proportion of campuses with above 75% of students from socio-economically disadvantaged backgrounds. The cut-off rate of 75% was chosen for the purpose of this study in accordance with the National Center for Education Statistics (2021) as a proxy to identify high-poverty schools. Results demonstrate that nearly 25% of high school campuses attended by H-LSAMP graduates had a high percentage of economically disadvantaged students, while the majority of high school campuses (75.5%), had a low percentage of economically disadvantaged students.

Table 1

H-LSAMP Scholars High School Campus Ratings

Accountability Rating	N	%
A	46	17.4
A-	21	7.9
B	64	24.2
B-	46	17.4
C	21	7.8
C-	30	11.3
D	13	4.9
D-	*	*
F	*	*
N/A	18	6.8
Total	265	

Note. *Data value was masked due to small sample size

Table 2

Proportion of Economically Disadvantaged Students – High School Campuses Attended by H-LSAMP

Percentage of Economically Disadvantaged Students at High School Campus	N	%
Below 75%	200	75.5
Above 75%	65	24.5
Total	265	

Research Question 2: Undergraduate and Graduate Degree Outcomes

As of the 2019/2020 academic year, there was a total of 2,839 students enrolled in the program. Approximately 74% of students graduated over the course of the timeframe of this study (2004-2020). The H-LSAMP has graduated 2,044 level one students with bachelor's degrees. Approximately 46% of graduates were female and 54% were male. Nearly 50% of bachelor's degree recipients were from Black and Latino backgrounds. Figure 1 presents the number and percentage of bachelor degrees conferred each year from 2004-2019. Table 3 presents the percentage of graduates from each racial/ethnic category. The percentage of graduates varied over time, likely due to a number of factors such as changes in level of financial support between entities such as the National Science Foundation, corporate sponsorship, donor and institutional financial support, which has a subsequent effect on cohort sizes each year.

Figure 1

Bachelor Degrees Conferred by the H-LSAMP*: 2004 -2019*
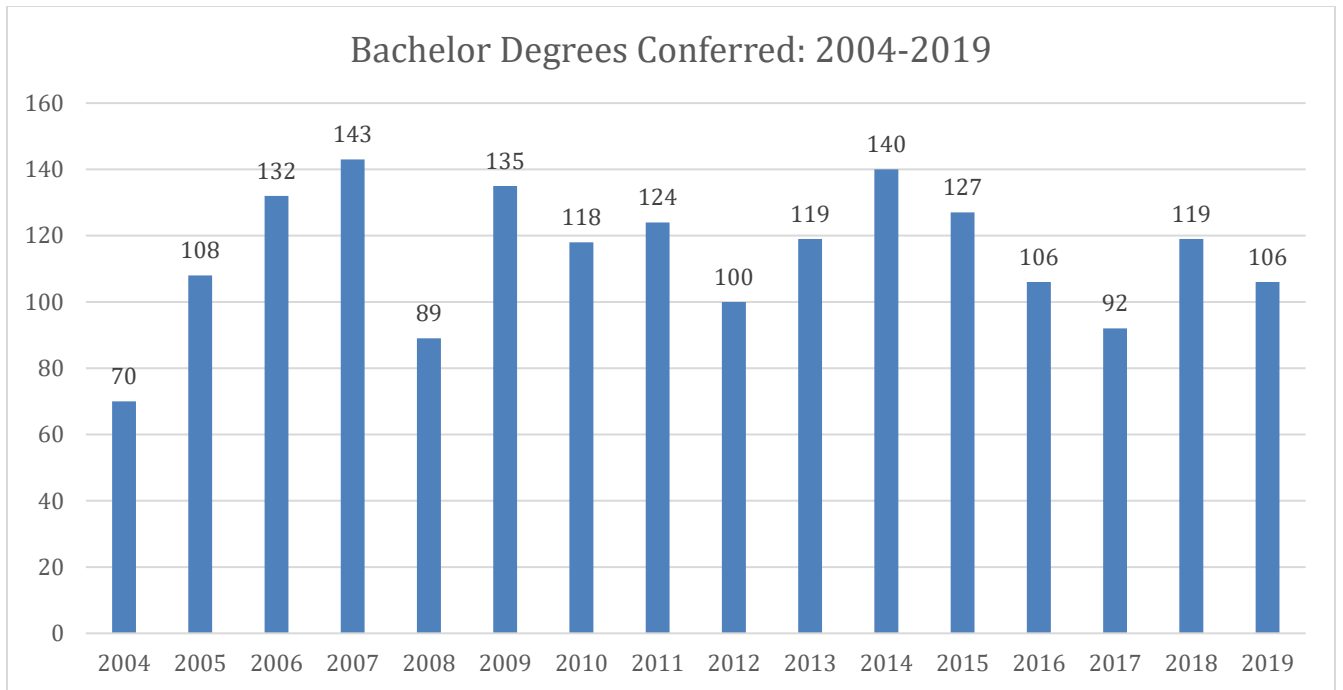
Note. Bachelor degrees conferred (all degrees) includes H-LSAMP students. Copyright 2020 by the Urban Education Research Center.

The largest proportion of H-LSAMP scholars were Black students, accounting for approximately 27% of total enrollment in the program, followed by Latino students (23%). Asian students made up 18% of total enrollment, followed by White students (9%). Degree completion mirrored program participation. See Table 3 for this information.

Table 3

Bachelor Degree Recipients by Race

Race	N	%
White	182	8.9
Black	551	26.9
Latinx	478	23.4
Asian	370	18.1
Other	463	22.7

Over 50% of bachelor degree recipients graduated in the field of NSM, and approximately 34.5% of recipients graduated with a major in Engineering. Only 8.5% of degrees awarded were in non-STEM fields of study. Bachelor's degree recipients by major/field of study are displayed in Table 4.

Table 4

Bachelor Degree Recipients by Major/Field of Study

	Frequency	%
NSM	1,042	51.0
Engineering	706	34.5
Computer/Technology	123	6.0
Other/non-STEM	173	8.5

Table 5a presents the proportion of students enrolled by major/field of study according to ethnicity. As shown in the table, approximately 55% of Black students and 40% of Latinx students were enrolled as NSM majors. Nearly 46% of Latinx students were enrolled in engineering degrees, while 24% of Black students chose engineering as a major. Table 5b displays bachelor's degree recipients by field of study and gender. In total, 54% of H-LSAMP undergraduate scholars were male and 46% were female. A larger proportion of bachelor degree recipients in NSM were female (56.8%), while 43% were male. A higher percentage of males received Engineering undergraduate degrees (69.5%), compared to females (30.5%).

Table 5a

Bachelor Degree Recipients by Major/Field of Study and Race

Race	NSM	Engineering	Computer Science	Other/Non-STEM	Total
	N	N	N	N	N
	%	%	%	%	%
White	84	80	9	9	182
	46.2	43.9	4.9	4.9	100
Black	303	131	53	64	551
	54.9	23.8	9.6	11.6	100
Latinx	187	219	22	50	478
	39.1	45.8	4.6	10.5	100
Asian	225	114	12	19	370
	60.8	30.8	3.2	5.1	100
Other	243	162	27	31	463
	52.5	34.9	5.8	6.7	100

Table 5b

Bachelor Degree Recipients by Major/Field of Study and Gender

Gender	NSM	Engineering	Computer Science	Other/non-STEM	Total
	N	**N**	**N**	**N**	**N**
	%	**%**	**%**	**%**	**%**
Female	592	215	37	96	940
	56.8	30.5	30.1	55.5	45.9
Male	450	491	86	77	1,104
	43.0	69.5	69.9	44.5	54.0

Graduate Degree Outcomes

As of the 2019/2020 academic year, 319 graduates from the H-LSAMP have successfully gone on to complete advanced degrees from state institutions, including 290 Master's degrees and 29 PhD degrees. In addition, 49 students received advanced graduate degrees in Pharmacy, Optometry, and Medicine. Approximately 56% of Masters degrees obtained were in STEM fields of study, including 32% of Masters degrees obtained in majors from NSM, 23% Engineering, and nearly 4% in Computer Science. Over half of PhD degrees awarded (55%) were in NSM fields. Among graduate degree recipients, nearly 35% of master's degree graduates and 48% of PhD graduates were Black, while 22% of master's degree recipients were Latinx students. Table 6 presents graduate degree outcomes by race.

Table 6

Graduate Degree Outcomes by Race

Degree Conferred	White	Black	Latinx	Asian	Other	Total
	N	**N**	**N**	**N**	**N**	
	%	**%**	**%**	**%**	**%**	
Masters	19	103	65	31	72	290
	6.6	35.3	22.4	10.7	24.8	
PhD	*	14	*	7	*	29
	*	48.3	3.5	24.1	*	
Pharmacy/Optometry/Medical Degrees	*	16	*	23	6	49
	*	32.7	*	47.0	12.2	

*Note. *Data value was masked due to small sample size*

Research Question 3: Workforce Outcomes

To examine the workforce outcomes of H-LSAMP STEM scholars after matriculation from college, graduation files were merged with data from the Workforce Commission enabling this study to track data on the industry that students were working in based on their corresponding SIC code.

Although this data provides the study with a wealth of information on career trajectories, a limitation to the results of this study is that the precise occupation of each student is not available. Consequently, our findings present workforce outcomes by industry level which may or may not accurately represent STEM-related careers within these industries. For instance, while education may not traditionally be defined as a STEM field of study, teachers of advanced biology, chemistry, or mathematics high school courses could potentially be categorized as working in a STEM field. With the inability to accurately account for these differences in mind, this section provides an overview of the industry outcomes of scholars after their entry into the state workforce. Table 7 presents the percentages of H-LSAMP graduates, who graduated in a STEM field of study, employed in both STEM and non-STEM related industries. As a whole, 25% of H-LSAMP STEM graduates in the state's workforce were employed in STEM occupations, while 75% of the graduates were not.

Table 7

H-LSAMP Graduates- Workforce Occupation Industries

Industry	N	%
Farming[1]	*	*
Oil/Natural Gas & Liquid Extraction[1]	23	1.44
Construction[1]	36	2.25
Transport and Product Manufacturing[1]	122	7.62
Wholesale, Retail & Entertainment[2]	703	43.88
Banking and Business Services[2]	168	10.49
Computer Related Services[1]	114	7.12
Medical Related Services[1]	112	6.99
Education & Public Service[2]	287	17.92
Engineering[1]	*	*
Unknown/Other[2]	35	2.18
Total	1,602	100

*Note. *Data value was masked due to small sample size
[1] STEM industry
[2] non-STEM industry

Discussion

Program Successes

Results from this descriptive study provide a broad overview of the goals of the H-LSAMP since the first graduating cohort of scholars in 2004. Given that the goals of the consortium are centered around increasing the academic preparedness and number of URM students in STEM fields, the results of this study demonstrate that out of the 2,044 undergraduate degrees obtained, approximately 27% were awarded to Black students and 24% to Latinx students. These findings provide clear evidence of program participation effects on the graduation, retention, and academic success of URM students in STEM fields and support previous studies that have demonstrated successful program outcomes (e.g., Duncan & Dick, 2000; Jiang et al., 2005; Lisberg & Woods, 2018). Findings also provide evidence of the success of the alliance at assisting students to successfully matriculate into Masters and PhD programs in STEM fields of study. From a national perspective, the

results of this study indicate the consortium supports the current imperative to produce a STEM workforce that truly mirrors the nations demographics (Stacy-Ann et al.; 2014; National Science Foundation, 2015). Given that part of the broader goals of the program, in line with national efforts, focus on creating a more level playing field for students from under-represented backgrounds, the results from this study, particularly as they pertain to the high proportion of students of color successfully completing their degree, support the notion of expanding access to and completion of undergraduate and graduate degrees in STEM. A substantive body of literature highlights the wide racial disparities that exist in STEM degree enrollment and completion (McGee, 2020; National Academics of Sciences, Engineering, and Medicine, 2019). In addition, the general atmosphere of STEM degrees and professions are widely known to have a biased, strict, and highly competitive culture that negatively impacts minority students in particular (Leath & Chavous, 2018). To that end, the results of this study indicate that institutional and program components, along with the academic support and financial aid provided to students, play a significant role in degree completion and successful matriculation of URM students into a STEM graduate degree and eventually into the workforce.

Recruitment Efforts

In terms of high school accountability ratings, results portray that a significant proportion of high school campuses from which H-LSAMP scholars graduated from were not classified as economically disadvantaged campuses. This does not necessarily mean individual students attending these schools were not themselves from low-income families, but it does raise the question of whether the participants being recruited are the ones most in need of the programming. Given that one of the major recruitment strategies of the alliance is to target students from primarily disadvantaged high schools, such findings highlight the need for further intentional recruitment efforts to reach students most in need of the financial assistance offered through the H-LSAMP program.

From an economic perspective, this study offers a preliminary understanding of intervention outcomes that could contribute to the upward economic mobility of disadvantaged and under-represented students (Xie & Killewald, 2012). Through learning mechanisms that focus on strong academic preparation in mathematics, financial and academic support in the form of peer and faculty mentorship, intervention programs such as the H-LSAMP are able to tackle the most significant elements that attribute to racial gaps in URM STEM degree attainment (Adelman, 2006; Chen, 2013; Lichtenberger & George-Jackson, 2013).

Implications

Findings from this study provide evidence of the effectiveness of STEM intervention efforts in diversifying and increasing the proportion of URM students successfully enrolling and graduating from STEM fields of study. However, these results do not show the success of such programs at increasing the percentage of STEM graduates employed in STEM fields of study. To that end, in support of national efforts to diversify the STEM workforce in a way that reflects today's population demographics, intervention programs are advised to further strengthen program content and components to bolster career preparedness and job training opportunities that could help create stronger networking skills and connections otherwise unavailable to this sub-population of students (Jelks & Crain, 2020; Stipanovic & Woo, 2017).

Directions for Future Research

Given the descriptive nature of this study, an ongoing set of studies are being conducted that utilize inferential inquiry to investigate program participation effects over time. Such analyses include the use of propensity score matching to match H-LSAMP scholars with similar non-participants of the program based on a series of baseline characteristics, and subsequently measure program participation effects after matching on covariates (Ghazzawi et al., 2021). The advantage to using techniques such as propensity score matching is the reduction of selection bias, increasing the robustness of our analyses (Winship & Morgan, 1999). Additionally, further studies being conducted include examining the impact of program participation on the wage differentials of H-LSAMP scholars over time. These studies aim to contribute to the existing literature by offering empirically proven results that support continuous calls for intervention efforts across the nation. In addition, through the use of longitudinal data that follows students through their educational trajectory and into the workforce, these studies will uniquely contribute to the extant literature by highlighting pre-college, academic, and workforce characteristics most significant to successful student and labor outcomes.

Conclusion

This study provides an essential foundational understanding of the outcomes of a multi-institutional consortium that strives to increase the academic achievement and completion rates of URM students enrolled in STEM fields. Through demonstrating both the outcomes and limitations of such programs, this study offers legitimacy to further increase efforts, both state and nationwide, in implementing STEM intervention programs that center around providing academic and faculty support, along with peer mentorship, to increase the representation of minority students in STEM.

Dina Ghazzawi (dalghazzawi@uh.edu) completed her PhD in 2019 from the University of Houston in Higher Education Leadership and Policy Analysis. Her research interests focus on STEM intervention programs and examining longitudinal patterns related to program efficacy, particularly on under-represented student populations.

Donna Lynn Pattison (dpattison@uh.edu) serves as Assistant Dean for Student Success in the College of Natural Sciences and Mathematics at the University of Houston. She is also an Instructional Professor in the Department of Biology and Biochemistry. Dr. Pattison earned a PhD in Biochemistry and Cell Biology at Rice University. Her current work focuses on curriculum development and the study of the impact of changes in teaching practices, curriculum, and programming designed to improve student outcomes, particularly for students who are among the underserved.

Catherine Horn (clhorn2@uh.edu) is the Moores Professor of Educational Leadership and Policy Studies and Executive Director of the Institute for Educational Policy Research and Evaluation at the University of Houston. She is also the Director for the Center for Research and Advancement of Teacher Education and the University of Houston Education Research Center. Dr. Horn, who received her PhD from Boston College, focuses on the systemic influences of secondary and postsecondary assessment and related policies on the learning trajectories of students, particularly those traditionally underserved by the education and social sectors.

Bobby Wilson (bobby.wilson@tsu.edu) earned a BS degree in Chemistry from Alabama State University in 1966 and the MS degree in Chemistry from Southern University in 1972. He culminated

his academic training by earning a Ph.D in Chemistry from Michigan State University in 1976. Dr. Wilson has distinguished himself as a professor, scientist and an administrator. Since 1976, Dr. Wilson has been employed at Texas Southern University in various positions from Assistant Professor to Acting President. Currently, Dr. Wilson is the L. Lloyd Woods Professor of Chemistry and the Shell Oil Endowed Chaired Professor of Environmental Toxicology.

References

Adelman, C. (2006). *The toolbox revisited: Paths to degree completion from high school through college.* US Department of Education.

Allen-Ramdial, S. A. A., & Campbell, A. G. (2014). Reimagining the pipeline: Advancing STEM diversity, persistence, and success. *BioScience, 64*(7), 612-618.

Barton, P. E. (2003). *Hispanics in science and engineering: A matter of assistance and persistence.* Educational Testing Service.

Bonsangue, M. V., Drew, D. E., & Gilmore, M. W. (2018). The impact of a supportive community experience on African-American students. *Chemistry and Chemical Engineering. Learning Assistance Review*, 23(2), 9-40.

Carlone, H. B., & Johnson, A. (2007). Understanding the science experiences of successful women of color: Science identity as an analytic lens. *Journal of Research in Science Teaching, 44*(8), 1187-1218.

Carpi, A., Ronan, D. M., Falconer, H. M., & Lents, N. H. (2017). Cultivating minority scientists: Undergraduate research increases self-efficacy and career ambitions for underrepresented students in STEM. *Journal of Research in Science Teaching, 54*(2), 169-194.

Chinn, D., Martin, K., & Spencer, C. (2007, March). Treisman workshops and student performance in CS. Association for Computing Machinery Special Interest Group Computer Science Education, *39*(1), 203-207. https://www.doi.org/10.1145/1227504.1227383

Chen, X. (2013). *STEM attrition: College students' paths into and out of STEM fields.* NCES 2014-001. National Center for Education Statistics.

Chang, M. J., Sharkness, J., Hurtado, S., & Newman, C. B. (2014). What matters in college for retaining aspiring scientists and engineers from underrepresented racial groups. *Journal of Research in Science Teaching, 51*(5), 555-580.

Children at Risk. (n.d.). *Speaking out + Driving change for children.* Children at Risk. https://childrenatrisk.org/

Crisp, G., Nora, A., & Taggart, A. (2009). Student characteristics, pre-college, college, and environmental factors as predictors of majoring in and earning a STEM degree: An analysis of students attending a Hispanic serving institution. *American Education Research Journal, 46*(4), 924-942.

De La Rosa, M. L. (2006). Is opportunity knocking? Low-income students' perceptions of college and financial aid. *The American Behavioral Scientist, 49*(12), 1670.

Dika, S. L., & D'Amico, M. M. (2016). Early experiences and integration in the persistence of first-generation college students in STEM and non-STEM majors. *Journal of Research in Science Teaching, 53*(3), 368-383.

Drew, D. E. (2011). *STEM the tide: Reforming science, technology, engineering, and math education in America.* JHU Press.

Duncan, H., & Dick, T. (2000). Collaborative workshops and student academic performance in introductory college mathematics courses: A study of a Treisman model math excel program. *School Science and Mathematics, 100*(7), 365-373.

Espinosa, L. (2011). Pipelines and pathways: Women of color in undergraduate STEM majors and the college experiences that contribute to persistence. *Harvard Educational Review, 81*(2), 209-241.

Fayer, S., Lacey, A., & Watson, A. (2017). STEM occupations: Past, present, and future. *Spotlight on Statistics, 1*, 1-35.

Feeney, M., & Heroff, J. (2013). Barriers to need-based financial aid: Predictors of rimely FAFSA completion among low-income students. *Journal of Student Financial Aid, 43*(2), 2.

Ghazzawi, D., Pattison, D. L., Horn, C., Hardy, J., and Brown, B. (2021). Impact of an intensive multi-disciplinary STEM enrichment program on underrepresented minority student success. *Journal of Applied Research in Higher Education. 14*(2), 660-678. https://www.doi.org/10.1108/JARHE-12-2020-0452

Gilmer, T. (2007). An understanding of the improved grades, retention and graduation rates of STEM majors at the Academic Investment in Math and Science (AIMS) program of Bowling Green State University. *Journal of STEM Education, 8*(1), 11-21.

Goldsmith, P. R. (2011). Coleman revisited: School segregation, peers, and frog ponds. *American Educational Research Journal, 48*(3), 508-535.

Griffith, A. L. (2010). Persistence of women and minorities in STEM field majors: Is it the school that matters? *Economics of Education Review, 29*(6), 911-922.

Holland, J. M., Major, D. A., & Orvis, K. A. (2012). Understanding how peer mentoring and capitalization link STEM students to their majors. *The Career Development Quarterly, 60(*4), 343-354.

Holden, J. P., Lander, E. S., & Varmus, H. (2010). *Prepare and inspire: K-12 education in science, technology, engineering and math (STEM) for America's future (Executive Report).* Presidents' Council of Advisors on Science and Technology. http://www.whitehouse.gov/administration/eop/ostp/pcast/docsreports

Hurtado, S., Newman, C. B., Tran, M. C., & Chang, M. J. (2010). Improving the rate of success for underrepresented racial minorities in STEM fields: Insights from a national project. *New Directions for Institutional Research, 2010(*148), 5-15.

Jackson, K. M., & Winfield, L. L. (2014). Realigning the crooked room: Spelman claims a space for African American women in STEM. *Peer review: emerging trends and key debates in undergraduate education, 16*(2), 9.

James, S. M., & Singer, S. R. (2016). From the NSF: The National Science Foundation's investments in broadening participation in science, technology, engineering, and mathematics education through research and capacity building. *CBE—Life Sciences Education, 15*(3), fe7

Jiang, X., Sarin, S., Williams, M., & Young, L. (2005). Assessment of the NC-LSAMP project: A longitudinal study. *Age, 10*, 1.

Jelks, S. M., & Crain, A. M. (2020). Sticking with STEM: Understanding STEM career persistence among STEM bachelor's degree holders. *The Journal of Higher Education, 91*(5), 805-831.

Kendricks, K., Nedunuri, K. V., & Arment, A. R. (2013). Minority student perceptions of the impact of mentoring to enhance academic performance in STEM disciplines. *Journal of STEM Education: Innovations and Research, 14*(2), 38-46.

LaCourse, W. R., Sutphin, K.L., Ott, L.E., Maton, K.I., McDermott, P., Bieberich, C. Farabaugh, P., & Rous, P. (2017). Think 500, not 50! A scalable approach to student success in STEM." *BMC Proceedings, 11(*12) 105-115.

Leath, S., & Chavous, T. (2018). Black women's experiences of campus racial climate and stigma at predominantly white institutions: Insights from a comparative and within-group approach for STEM and non-STEM majors. *The Journal of Negro Education, 87*(2), 125-139.

Lichtenberger, E., & George-Jackson, C. (2013). Predicting high school students' interest in majoring in a STEM Field: Insight into high school students' postsecondary plans. *Journal of Career and Technical Education, 28*(1), 19-38.

Lisberg, A., & Woods, B. (2018). Mentorship, mindset and learning strategies: an integrative approach to increasing underrepresented minority student retention in a STEM undergraduate program. *Journal of STEM Education, 19*(3), 14-20.

Malcom, L. E., & Dowd, A. C. (2012). The impact of undergraduate debt on the graduate school enrollment of STEM baccalaureates. *The Review of Higher Education, 35*(2), 265-305.

McGee, E. O. (2020). Interrogating structural racism in STEM higher education. *Educational Researcher, 49*(9), 633-644. https://www.doi.org/10.3102/0013189X20972718

Martinez, S., & Guzman, S. (2013). Gender and racial/ethnic differences in self-reported levels of engagement in high school math and science courses. *Hispanic Journal of Behavioral Sciences, 35*(3), 407-427.

National Academies of Sciences, Engineering, and Medicine. (2016). Barriers and opportunities for 2-year and 4-year STEM degrees: Systemic change to support students' diverse pathways. Author.
https://ebookcentral.proquest.com/lib/uh/detail.action?pqorigsite=primo&docID=455835

National Board of Sciences. (2010). *Preparing the next generation of STEM innovators: Identifying and supporting our nation's human capital.* National Science Foundation.

National Center for Science and Engineering Statistics. (2021). *Science and engineering degrees by race and ethnicity of recipients.* National Science Foundation.
https://www.nsf.gov/statistics/degreerecipients/

National Center for Education Statistics. (2021). *Concentration of public-school students eligible for free or reduced-price lunch.* Author. https://nces.ed.gov/programs/coe/indicator/clb

Page, K. R., Castillo-Page, L., Poll-Hunter, N., Garrison, G., & Wright, S. M. (2013). Assessing the evolving definition of underrepresented minority and its application in academic medicine. *Academic Medicine, 88*(1), 67-72.

Pascarella, E. T., & Terenzini, P. T. (2005*). How college affects students: A third decade of research. Volume 2.* Jossey-Bass.

Riegle-Crumb, C., King, B., & Irizarry, Y. (2019). Does STEM stand out? Examining racial/ethnic gaps in persistence across postsecondary fields. *Educational Researcher, 48*(3), 133-144.

Robinson, K. A., Perez, T., Carmel, J. H., & Linnenbrink-Garcia, L. (2019). Science identity development trajectories in a gateway college chemistry course: Predictors and relations to achievement and STEM pursuit. *Contemporary Educational Psychology, 56*, 180-192.

Saw, G., Chang, C. N., & Chan, H. Y. (2018). Cross-sectional and longitudinal disparities in STEM career aspirations at the intersection of gender, race/ethnicity, and socioeconomic status. *Educational Researcher, 47*(8), 525-531.

Shaw, E. J., & Barbuti, S. (2010). Patterns of persistence in intended college major with a focus on STEM majors. *NACADA Journal, 30*(2), 19-34.

Stephens, N. M., Townsend, S. S., Markus, H. R., & Phillips, L. T. (2012). A cultural mismatch: Independent cultural norms produce greater increases in cortisol and more negative emotions among first-generation college students. *Journal of Experimental Social Psychology, 48*(6), 1389-1393.

Stipanovic, N., & Woo, H. (2017). Understanding African American students' experiences in STEM education: An ecological systems approach. *The Career Development Quarterly, 65*(3), 192-206.

Strayhorn, T. L. (2013). Measuring race and gender differences in undergraduate students' perceptions of campus climate and intentions to leave college: An analysis in Black and White. *Journal of Student Affairs Research and Practice, 50*(2), 115-132.

Texas Education Agency (2021). *2021 accountability manual.* Author.
https://tea.texas.gov/sites/default/files/chapter-1-2021-accountability-overview.pdf

Treisman, U. (1992). Studying students studying calculus: A look at the lives of minority mathematics students in college. *The College Mathematics Journal, 5*(23), 362-372. https://doi.org/10.1080/07468342.1992.11973486

U.S. Bureau of Labor Statistics (2021). Employment in STEM occupations. Author. https://www.bls.gov/emp/tables/stem-employment.html

Whalen, D. F., & Shelley, M. C. (2010). Academic success for STEM and non-STEM majors. *Journal of STEM Education, 11*(1), 46-60.

Wilson, A., Sanner, S., & McAllister, L. (2010). An evaluation study of a mentoring program to increase the diversity of the nursing workforce. *Journal of Cultural Diversity, 17*, 144-150.

Winship, C., & Morgan, S. L. (1999). The estimation of causal effects from observational data. *Annual Review of Sociology, 25*(1), 659-706.

Xie, Y., & Killewald, A. A. (2012). *Is American science in decline?* Harvard University Press.

Xu, Y. J. (2013). Career outcomes of STEM and non-STEM college graduates: Persistence in majored-field and influential factors in career choices. *Research in Higher Education, 54*(3), 349-382.

The Impact of Internship on Undergraduate STEM Students' Interest in STEM Teaching

Kristin Cook iD
Bellarmine University

Akhtar Mahmood iD
Bellarmine University

Cody Nygard
Bellarmine University

Mira Gentry-Johnson iD
Kentucky Science Center

Mellisa Blankenship iD
Kentucky Science Center

ABSTRACT

STEM teachers are considered one of the most important factors in inspiring high school students in STEM to excite them about the dynamic nature of the STEM fields (PCAST, 2010). Increasing the number of qualified and diverse science and mathematics teachers, especially those working in high-need areas, is of paramount importance. The intention of this mixed methods study was to consider the ways in which a teaching-focused internship might impact an undergraduate STEM student's decision to consider entering the STEM teaching profession. Results indicated internships were an important recruitment tool for STEM majors to consider a career in science and/or mathematics teaching. Authors detail implications for supporting STEM interns at all points of their development as a teacher.

Keywords: internships, teacher recruitment, STEM

Introduction

Over the past two decades there has been much concern about the state of K-12 STEM education, instruction in the public school system, and its ripple effect on the future of the STEM workforce, US economic growth, and competitiveness in a global marketplace (Matthews, 2007). The quality of STEM teachers is increasingly recognized to be a critical factor, since teachers have a measurable impact on students' learning, success, and achievements (National Research Council (NRC), 2010; President's Council of Advisors on Science and Technology (PCAST), 2010). Indeed, STEM teachers are considered one of the most important factors in inspiring high school students in STEM to excite them about the dynamic nature of the STEM fields (PCAST, 2010). Highly qualified teachers lay the academic groundwork in STEM, thus inspiring and motivating K-12 students to major in a STEM field.

The demand for certified STEM teachers nationally has been rising, with schools across the country struggling to fill openings with qualified teachers. Fewer STEM students are entering the teaching profession, and school districts are struggling to attract and retain teachers (Barth et al., 2016). Compounding the problem, as public-school enrollments become increasingly diverse and the number of minority students enrolled is growing faster, with students of color comprising nearly half of the student population, the demographic composition of teachers has changed little (Almy & Theokas, 2010). In addition, teacher candidates do not reflect the demographic make-up of students in today's classrooms (American Association of College of Teacher Education (AACTE), 2013). Additionally, schools in high-poverty communities often do not have qualified STEM teachers or have teachers who teach outside of their certification fields. Classes in high-poverty schools are twice as likely to have an out-of-field teacher, compared to low-poverty schools. Moreover, out-of-field teachers are disproportionately teaching minority students and are assigned to teach students in high-poverty schools (Almy & Theokas, 2010). For these reasons, increasing the number of qualified and diverse science and mathematics teachers, especially those working in high-poverty areas, is of paramount importance.

Some states are experiencing these trends more than others and our study focused on the state of Kentucky, where the need is critical. The number of certified Kentucky high school STEM teachers in Math, Biology, Chemistry, and Physics has been declining at an alarming rate since 2010 (Department of Education, 2022). Due to the STEM teacher shortage in Kentucky, the vast majority of high school STEM teachers are teaching outside of their respective certified STEM discipline. Many districts across the state have an extensive shortage of teachers in the STEM fields. Several schools often have only one STEM teacher who is responsible for teaching all the STEM courses, and many of these teachers do not have the certification to teach in a specific STEM field. In order to meet the growing state and national STEM needs, barriers to interest in the STEM teaching profession, such as the perception of low pay and the difficulty of working with students, need to be considered.

In Kentucky, 34.8% of STEM teachers are currently certified to teach middle school Mathematics and 26% are certificated to teach high school Mathematics. Only 4.5% of high school STEM teachers have certification in Chemistry and only 3.3% in Physics. Since 2010, there has been a 44% decline in the total number of certified middle and high school STEM teachers in Kentucky. According to the latest Department of Education's report on nationwide teacher shortages, Kentucky reported teacher shortages in the STEM fields (Department of Education, 2022). There is an urgent need to increase the number of new certified teachers in Physics and Chemistry in particular, since high school Physics and Chemistry courses are often taught by teachers out of their area of certification who do not have a degree in either Chemistry or Physics. In recent years, there has been a significant decline in the number of certified Physics teachers in Kentucky, which mirrors national shortages (Heron & McNeil, 2016; White & Tyler, 2014).

As a part of a National Science Foundation (NSF) funded Noyce Capacity Building project aiming to inspire college-level STEM students to pursue a career in K-12 teaching, we formed a partnership with an urban science center to strengthen K-12 STEM teacher education pathways. We provided a paid summer internship opportunity at the Kentucky Science Center for STEM majors to design and implement hands-on, inquiry-based science activities with students from grade levels Pre-K through 8th grade to develop their teaching skills and interest in STEM education. Such interactions with younger students could have the effect of generating undergraduate interns' interest in pursuing a career in STEM teaching. The intention of this mixed methods study was to consider the ways in which a teaching-focused internship might impact an undergraduate STEM student's decision to consider entering the STEM teaching profession. Asking STEM interns to share their views of their internship experience, and the impact it had on their perception of teaching, gave researchers insight into additional variables which may influence undergraduate student interest in STEM teaching as a profession.

Literature Review

Interest in Teaching

Research has identified motivations people have when choosing a career in teaching that are broadly categorized as intrinsic, extrinsic, and altruistic (Richardson & Watt, 2014). The Organization for Economic Co-operation and Development (OECD, 2005) reported on studies independently conducted in France, Australia, Belgium, Canada (Québec), the Netherlands, the Slovak Republic, and the United Kingdom, which indicated that motivations for choosing teaching as a career were the desire to work with youth, the potential for intellectual fulfillment, and the wish to make a social contribution. The aspiration to work with children and adolescents has also been identified as important in many studies conducted over time in the United States (Moss, 2020), the United Kingdom, and Europe. However, we know that 'triggers' and 'drivers' to select teaching as a career can be different depending on regional context (Low et al., 2017). For example, in different sociocultural contexts, such as Africa and the Caribbean, "extrinsic motives" such as salary, job security, and career status have been found to be more prominent. Researchers seems to agree that interest can develop by having some teaching experience (Low et al., 2017).

Kyriacou and Coulthard's (2000) study on undergraduates' views of teaching as a career choice indicated that students who are seriously considering a career in teaching tend to have a much closer match of factors that are important to them and that they think are offered by teaching as a career, when compared with other students. These researchers surveyed previous studies on students' interest in teaching and found that the most motivating factors fall into three categories: altruistic reasons (e.g. desire to benefit society), intrinsic reasons (e.g. interest in subject matter and expertise), and extrinsic reasons (e.g. extended work breaks, level of pay, etc.) (Kyriacou & Coulthard, 2000). Teacher recruitment that focuses on factors influential to students who have already begun teacher training courses will inevitably miss reaching those students who are yet undecided on committing to teaching as a career. Finding an applicable means of reaching students who may have not yet even considered a teaching career was a focus of our project. We chose to examine the impact of a teaching internship on a student's consideration of the teaching profession as one possible influencing factor.

Taimalu et al. (2021) examined the influences on teaching as a career choice between students in Estonia and Finland. The researchers sought to minimize cultural differences as motivating factors and made clear from the start that both Finland and Estonia share similar cultures and courses for teacher education. However, each country exhibited very different evaluations of the teaching profession. For example, in Finland, 57% of teachers expressed that their profession is valued by society, compared to only 26% of teachers in Estonia (Taimalu et al., 2021). Although overall society between the two countries is very similar, the differences in the societies' evaluation of teaching may be attributed to the researchers' findings that in Estonia, until recently, teaching salaries were much lower than other professions. Finland has long offered adequate salaries for teachers. Estonian student perceptions of low pay may still be a contributing factor to the lack of interest in the profession. Additionally, the researchers proposed that career choices were more individually-motivated in Finland as opposed to Estonia, where a lower societal view of teaching may hinder students' interest in the career.

Australian public schools have suffered from a similar drought of qualified STEM teachers, similar to many places in the United States. As a result, the Office of the Chief Scientist launched the "Step Up Project" to stimulate interest in STEM teaching as a profession. McDonald (2017) explored the factors influencing secondary science teachers' decisions to pursue a career in teaching. Interestingly, societal factors, as discussed in the study above from Estonia and Finland, did not contribute significantly to participants' considerations. Instead, the research found that participants' experiences in high school and the lack of STEM-related jobs in other fields were the most influential

factors on their choice to pursue teaching as a career. Although the researcher acknowledged lower teaching salaries than many salaries in research-related STEM fields, the reality of a lack of jobs in those fields made the teaching profession more appealing to many of the participants. The availability of jobs, coupled with the generous time off and family-friendly work schedule, was highlighted by several participants as key factors in their choice of career.

In the United States, research has been conducted to determine promising students who may have not previously considered a teaching career. Moin et al. (2005) found that the best place to look for STEM students who may become interested in a science and/or mathematics teaching career was in their junior and senior years as an undergraduate STEM major. Surveying STEM students from two different universities to assess their interest in K-12 teaching, researchers found that students in their junior and senior year, as well as students who had mid-level academic performance, were the most promising in terms of becoming interested in a science and/or mathematics teaching career. Finally, with regard to type of major, mathematics majors were the most promising, followed by natural science majors, and engineering majors were the least promising.

Teaching Internships

Hutchinson's (2012) case study with new high school STEM teachers in the US revealed that the participants had a desire to become a teacher at very different times in their professional career preparation, and they selected different paths to enter the teaching profession. Despite information about the teaching profession not being readily available or promoted, each of the participants found ways to gain information about becoming a teacher. Beyond their recruitment, they needed support during their pre- and in-service teaching experiences. The participants referenced support and professional development activities from university mentors, mentoring from their instructional team members, support from campus leaders, and support through social networking as to what helped retain them as STEM teachers.

Regarding recruitment, Cerritos College recognized the potential impact of a STEM teaching internship on student interest in teaching. In 1999, the school developed their Teacher TRAC Program, which offers four different STEM-focused internship opportunities for STEM teachers-in-training. Parsons (2013) wrote that internship experiences are required by Cerritos College for graduation for STEM Education students, with the goal for students "to witness the application of contextualized curriculum in the classroom and create an environment of inter-disciplinary understanding" (p. 9). Our research emphasized the internship model with a critical eye to attempt to understand what factors of an internship had an impact on student perception and consideration of teaching as a career.

Borgerding's (2015) study highlighted the influence that summer internships for early STEM majors can have on student interest in teaching. Aspects of the internships such as the requirement for reflection, opportunities for teacher agency, and the balance between emphasis on classroom management and student learning were connected with interns' positive or negative teaching experiences. Furthermore, particular aspects of the interns themselves such as expectations for classroom management, a desire for authority, and interest/care for students also contributed to the impact of these internships for recruitment.

While teaching-focused internships are cited in many projects that aim to increase the interest of students who may not have previously considered teaching, the effectiveness of these internships, as far as recruitment into an education program, is debatable. From surveys of STEM students both prior to and following their summer experience, Morrell and Salomone (2017) found that very few of the interns actually did go into teaching. These researchers concluded that the internship program was not an effective use of Noyce funding to attract STEM undergraduates into the teaching profession. However, their data indicated that students did consider teaching as a viable career path sometime in

the future and ended the internship with positive views of teaching. This study suggests that the internship may have made a difference on career path choice in the long term for participants.

Still, other research points to the success of teaching internships for motivating STEM students to pursue a teaching career. For example, Mundy and Ratcliff (2021) incorporated teaching-focused internships in their Robert Noyce Summer Internship program for STEM students. Surveys of students indicated that the interns were satisfied with the Noyce Summer Internship Program, and some STEM majors or Noyce Interns decided to pursue a teaching certificate to become a high school STEM teacher. Researchers thereby contended that STEM students' interest in teaching in a STEM area increased because of the Noyce Summer Internship Program (Wong-Ratcliff & Mundy, 2019).

Methods

Internship Experience at the Kentucky Science Center

We formed a partnership with our regional Kentucky Science Center (KSC) to allow for internships that might inspire STEM students to pursue a career in STEM teaching. The KSC is a non-profit organization and hands-on science museum with an attendance of more than half a million people annually. In addition to the main gallery, the KSC has a number of hands-on exhibits such as sensory and magnet walls that offer ample experiences for experimentation. There is a MakerPlace exhibit that has a variety of skill-building activities with tools and is staffed with STEM professionals and interns who provide help and ask/answer questions. Additionally, there is a Discovery Gallery (a Natural History gallery) where students learn about anthropology, archaeology, and paleontology. There is also a curated collection of specimens, gems, and fossils. The KSC has an exhibit on NASA's second human space flight program and a Challenger Learning Center, where visitors can learn about the universe and space exploration. The Learning Center has a space mission simulation which shows how to solve real-world problems. Visitors can practice the skills utilized by the astronauts, scientists, and engineers who ensure each space mission's success.

The KSC's mission is to advance the statewide science literacy, encourage people of all ages to explore science in everyday life, and to reach all audiences through its mission to do hands-on science in ways that are engaging, educational, and entertaining. The KSC hosts numerous STEM summer camps and programs throughout the year for elementary, middle, and high school students. The KSC also holds a Youth Science Summit, which is an immersive, full-day workshop for middle and high-school students aiming to empower young adults into the STEM fields. The Youth Science Summits give high school students the chance to interact with and learn from industry professionals and faculty in academia. The KSC's educational methodology, practiced in all of its STEM-based summer camps and internship programs, uses problem-based learning methods focused on the science and engineering practices of the *Next Generation Science Standards* (NGSS, 2013*)*, thus helping students understand and implement best practices in STEM education. The KSC's summer camp internships are designed to immerse learners in inquiry-based exploration of STEM topics that are based on a relevant, real-world context that increase students' levels of interest and motivation in science. While not a replicate of a formal classroom, the standards-based focus and inquiry teaching practices emphasized in the summer camps provide a teaching and learning environment in which our interns could try out developing, implementing, and assessing lessons alongside experienced educators. Furthermore, our interns had the opportunity to learn how to manage large groups and build relationships with learners.

Funded by our Noyce Capacity Building grant, ten STEM students majoring in STEM fields took part as interns at a four-week summer camp experience (ranging from Pre-K to Grades 8) at the Kentucky Science Center so they could develop their skills in STEM education. We recruited the interns by sending out a position description to the Biology, Chemistry, Biochemistry & Molecular

Biology, Physics, and Mathematics departments at our university. Each student selected for the internship worked on average about 25 hours per week at the science center and received a stipend of $1000 for their internship, $500 as a housing stipend, and $368 towards food expenses. Interested STEM majors were asked to apply by providing a resume, statement of interest, and faculty recommendation. Of the 22 applicants, ten were selected based on GPA, experience working with youth, and faculty recommendation.

Once selected, interns were given a choice as to what grade level and content focus areas they were most interested in working with at the science center's summer camps. There were thirty summer camps offered to various grade levels ranging from PreK to Grade 8. These summer camps were called Planes, Trains and Automobiles, Cool Chemistry, Scales and Trails, Snowmen in Summer, Waterworks, Lego Design Challenge, Tall Tales, Wild By Design, Journey to Space, Speedometry, Once Upon a Design, Lego Robotics 2.0, Inventor's Lab, Matter Mysteries, Pokémon Science, Engineer Your World, Uniquely Human, Agents of STEAM, Magical Creatures, Code-Games-Create, Medical Mysteries, Lego NXT, Candy Chemistry, Animation Creation, Chemistry, Destruction Engineering, EV3 Robot Challenge, Virtual Reality, and Space Technologies.

Each summer camp was a week in duration, and each intern took part in at least three separate summer camps in order to diversify the content and grade level of their internship experience. Each applicant underwent a screening process and background check with the science center as part of their formal hiring process (all student applicants selected successfully passed this screening). Upon official hire, the ten interns completed training in preparation for their camp that was provided by the Science Center's experienced educators, which included inquiry-based teaching methods and classroom/materials management tutorials. Then, under the guidance and supervision by a science center staff member, interns facilitated all lesson plans provided by science center for up to 25 students in each summer camp class. Interns aided camp teachers by providing classroom management and guided instructional time. The interns used the engineering design processes, inquiry-based learning, and *Next Generation Science Standards* (NGSS, 2013) to conduct the summer camps and were given ongoing pedagogical support from certified educators who worked at the science center.

Demographic Information of Interns

The interns were selected from students within one of five STEM academic majors at a private Catholic university (i.e. Biology, Mathematics, Chemistry, Physics, and Biochemistry & Molecular Biology). Of the ten interns selected from a pool of interested applicants for this four-week summer internship at the science center, six interns (60%) were female and four (40%) were male. Regarding academic major, four were Biology students, two Chemistry students, two Physics students, and two Biochemistry & Molecular Biology students. Out of the ten interns, three were Freshman, four Sophomores, one Junior, and two graduating Seniors. Among the ten students, eight students were White (five were female and three were male), one student was Hispanic (male), and one student was Asian (female). Out of the ten interns, three indicated/expressed an interest in a teaching career as a high school teacher. The mean GPA of the interns was 3.35 and all had some experience working with youth in the area of tutoring specifically.

Data Collection and Analysis

The researchers utilized surveys, semi-structured interviews, and observational teaching assessments to collect data on participants' experiences before, during, and after the internship at the KSC. The interns took part in surveys and interviews conducted by researchers and by KSC staff educators that were collectively used to determine their interest in pursuing a career in STEM teaching. Interns were asked open-ended questions regarding their experience in the internship in the form of

a survey and in interviews (see Appendix A). Additionally, the science center summer camp supervisors conducted a quantitative and qualitative assessment of each student intern's performance at the summer camp (see Appendix B) using an evaluation tool common to their organization. Interviews were audio and video recorded and findings were triangulated with field notes written by the research assistant during the interviews. Science center staff also completed additional questionnaires related to their perceptions of our STEM majors' involvement in the internship. Following the completion of the study, the researchers analyzed all survey, interview, and observational data to determine interns' interest in pursuing a career in STEM teaching.

Quantitative data analysis was conducted on survey results collected by the research assistant. Qualitative analysis was conducted by examining frequencies of phrases and concepts amongst the group of interns both in surveys and interview transcripts. Researchers first coded the data independently and then came to consensus to categorize the codes thematically. Themes were then sorted into groups of positive and negative opinions about teaching. Member checking was conducted with interns to verify analysis. This thematic analysis revealed trends corresponding to interns' positive or negative perceptions about entering the STEM teaching profession and their positive or negative experiences as interns in the program. Qualitative analysis was used to triangulate variables that may impact interns' interest in teaching.

Findings

Results from Survey and Observational Data

Below we discuss the findings of key points of the survey data looking at the question of whether the internship motivated participants' interest in a teaching career and what factors influence that interest. We chose to focus our analysis on three areas: 1) Career interest of interns in K-12 teaching after the internship experience; 2) Willingness to pursue a Master of Arts in Teaching (MAT) certification if a NSF scholarship from the Noyce Program were provided for three years in the amount of $10,000 per year; and 3) Reasons for non-interest in a career in K-12 teaching. These three areas provided an understanding of interns' perceptions of entering the STEM teaching profession following their internship experience

Interns' interest in teaching at the start of the internship was 30%. See Table 1 for information on participants' interest in teaching.

Table 1

Interns' Pre and Post Interest in Teaching

Gender	Major	GPA	Year	Career Choice	Interest in Teaching (pre/post)
Male	PHYS	3.00	Senior	PhD in Appl. Physics	Y → Maybe
Male	BMB	3.81	Freshman	Pediatric Dentist	Y → Maybe
Male	CHEM	3.58	Junior	MS. In Chemistry	N
Male	PHYS	3.46	Sophomore	Astrophysicist	N → Maybe
Female	BMB	3.44	Sophomore	Public Health	N
Female	BIO	3.46	Senior	Genetics	N → Yes
Female	CHEM	3.50	Freshman	Ph.D. Chemistry	N
Female	BIO	3.60	Sophomore	Physician	N
Female	BIO	3.15	Sophomore	Pediatric Oncologist	N → Yes
Female	BIO	3.17	Freshman	Physician's Asst.	Y → Maybe

The final column shows the pre and post interest of the interns in K-12 teaching after their completion of the KSC internship. Some of the interns who began with an interest in teaching waned, while others who had no interest became interested in teaching after the internship.

We had a rather surprising response from the interns in the survey question – "If you were offered a scholarship in the amount of $10,000 in your Junior Year, $10,000 in your Senior Year, and $10,000 towards your teacher certification tuition expense, would you consider obtaining a high school teaching certification for a teaching career at a high school," six of the ten (60%) interns stated as "Yes" and 4 stated as "No" (40%). Figure 1 is a bar chart that shows the responses of the interns regarding their interest in pursuing the MAT Teaching certification degree if a scholarship were provided for the three years (Junior, Senior and the fifth MAT year).

Figure 1

Impact of Scholarship on Consideration of Pursuing MAT Teaching Degree
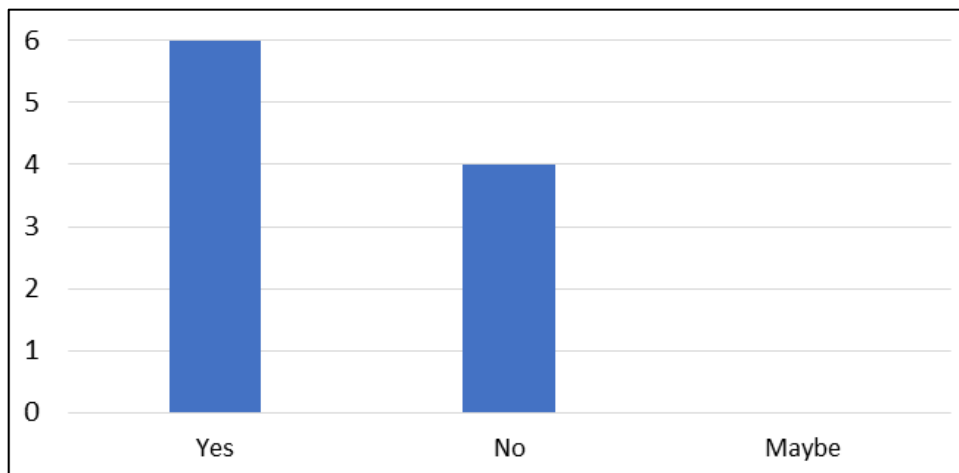
The interns' responses to this question are quite encouraging that scholarships can be an incentive for some STEM students to obtain a high school teaching certification for a K-12 teaching career at a high school.

Regarding intern reasoning for non-interest in teaching, we found a bias against the teaching profession based on perceived low wages and lack of opportunities for career development. Interns had exaggeratedly low perceptions of teacher salaries, and thus did not seriously consider teaching as a viable profession. Figure 2 shows the responses from all the interns. The top two responses were - "I wouldn't be good at it" (30%); "Not much opportunity for career development;" (30%) and "It does not initially pay as well as other occupations" (20%). Additionally, we asked "In what ways did your perceptions of teaching change during the internship?" We had varied responses; some were positive and some negative. We were particularly happy to get a rather encouraging response from one of the interns: "It went from something I never even considered as a career route to something I'm really looking into." The rationale behind this change in perceptions is explored further below with qualitative analysis of interns' responses to post-internships interview questions.

Figure 2

Reasons for Non-Interest in a K-12 Teaching Career
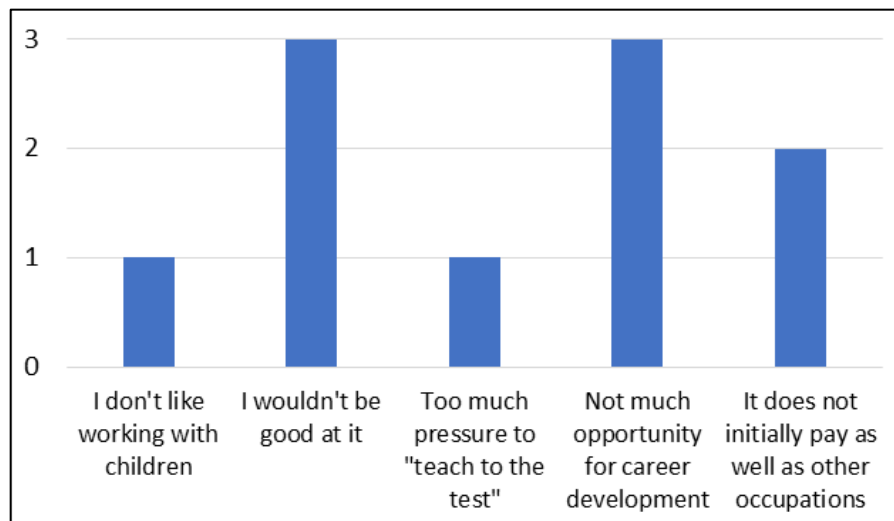
Observational Assessment Results

The staff supervisors at the KSC conducted evaluations of the interns and provided the summative results to interns at the end of their internship. The KSC staff of experienced educators are trained to use the evaluator instruments, and all interns are evaluated with this observational assessment . If interns have a low score (i.e. below 2.5) during these observations, formative supports would be instituted to help guide the interns. The results from the observational assessment of the interns focused on facilitation skills, including their interpersonal and communication/interaction skills. Student performance (Facilitation Skills) based on 12 separate parameters were rated on a scale of 1 - 5 (5-Outstanding, 4-Exeeds Expectations, 3-Meets Expectations, 2-Below Expectations, 1-Unsatisfactory). The results from the Facilitation Skills performance rating for the interns were used to gauge potential for how well students might do as K-12 teachers (to the extent these skills are transferable to a formal classroom environment) and interact with the younger students. This data shows that the vast majority of the ratings of each question were 4.0 or above and most ratings on the individual questions were at a 4.0. As shown in Table 2, the mean Facilitation Skills performance rating (based on the cumulative mean rating of 12 questions) of the interns ranged from 3.6 – 4.4 (out of a max. rating of 5.0), and the mean rating of each question also ranged from 3.7 – 4.3 (out of a max. rating of 5.0).

Figure 3 shows the distribution of each intern's facilitation skills performance ratings as a scatter plot. This figure shows that most of the interns scored above a 3.0 on each question of the Facilitations Skills performance rating. The interns scored the highest mean rating of 4.3 in question 5 (Is Courteous and Respectful to All Students and Staff), question 10 (Comfortable with Grade Level Content), and question 11 (Any Additional Content Provided from College Degree in STEM). All of the interns exceeded expectations in rating in questions 5 and 10, while 80% of the interns exceeded expectations in question 11. The interns scored the lowest mean rating of 3.7 and 3.8 in questions 9 ("Reads" Student Level of Engagement Effectively and Adjusts Facilitation Accordingly) and 7 (Confident in Engagement) respectively. It is understandable that interns scored somewhat lower in questions 7 and 9 compared to the other questions since this was their first experience participating in an internship in STEM education at various grade levels ranging from Pre-K to Grades 8. Even

though the mean rating of questions 7 and 9 was above 3.0, in question 7, 90% of the interns (9 out of ten) either met or exceeded expectations (out of these 9 interns, 7 exceeded expectation), while only 10% (1 intern) scored below expectation. In question 9, all (100%) of the interns either met or exceeded expectations with six of these ten interns exceeding expectations.

Despite being the first time any of these interns participated in an internship of this nature, the overall mean performance ratings of most of these interns were high for teaching in an informal classroom setting. Therefore, we concluded that these STEM students would be well suited to pursue a career as a K-12 STEM teacher and learn through a teacher preparation program to translate their teaching skills to a formal classroom setting.

Table 2

Facilitation Skills Performance Data of Summer Camp Interns

Facilitation Skills Performance Questions	S1	S2	S3	S4	S5	S6	S7	S8	S9	S10	Mean Rating for Each Question
Q.1: Follows Science Center Rules and Policies.	4.0	4.0	4.3	4.0	4.0	4.0	4.3	3.7	4.5	4.0	**4.1**
Q.2: Encourages Campers to Participate and Interact in Activities.	3.3	3.5	4.0	4.0	4.3	4.0	4.0	4.0	4.0	4.0	**3.9**
Q.3: Allows Students to Take Control of Learning But is There to Assist with Guiding Questions.	3.3	5.0	3.7	4.0	4.0	4.0	4.3	3.3	4.5	4.0	**4.0**
Q.4: Projects Positive Image and Attitude about STEM.	3.7	4.0	4.3	4.0	4.0	4.0	4.3	4.3	4.5	4.0	**4.1**
Q.5: Is Courteous and Respectful to All Students and Staff.	4.3	4.0	4.3	4.0	4.3	4.0	4.7	4.0	4.5	4.5	**4.3**
Q.6: Is Motivated and Takes Initiative.	3.0	3.5	3.7	4.0	4.3	4.0	4.7	3.0	4.5	4.0	**3.9**
Q.7: Confident in Engagement.	2.7	4.0	4.3	3.5	4.0	4.0	4.0	3.3	4.5	4.0	**3.8**
Q.8: Makes the Classroom a Fun Learning Environment.	3.3	4.0	4.5	4.0	4.0	4.0	4.7	3.7	4.0	4.0	**4.0**
Q.9: "Reads" Student Level of Engagement Effectively and Adjusts Facilitation Accordingly.	3.0	3.5	4.0	3.5	4.0	4.0	4.0	3.3	4.0	4.0	**3.7**
Q.10: Comfortable with Grade Level Content.	4.5	4.5	5.0	4.0	4.3	4.5	4.3	4.3	4.5	3.5	**4.3**
Q.11: Any Additional Content Provided from College Degree (STEM).	3.5	4.0	4.5	5.0	4.5	4.0	4.3	3.7	5.0	4.0	**4.3**
Q.12: Navigate Content Delivery Through Conversation Rather than Lecture.	4.5	3.5	4.0	4.0	4.0	4.0	5.0	4.0	4.0	4.0	**4.1**
Mean Cumulative Rating for Each Student (Scale: 0.0 - 5.0)	**3.6**	**4.0**	**4.2**	**4.0**	**4.1**	**4.0**	**4.4**	**3.6**	**4.0**	**4.0**	

Figure 3

Scatter Plot of Interns Ratings for All Questions
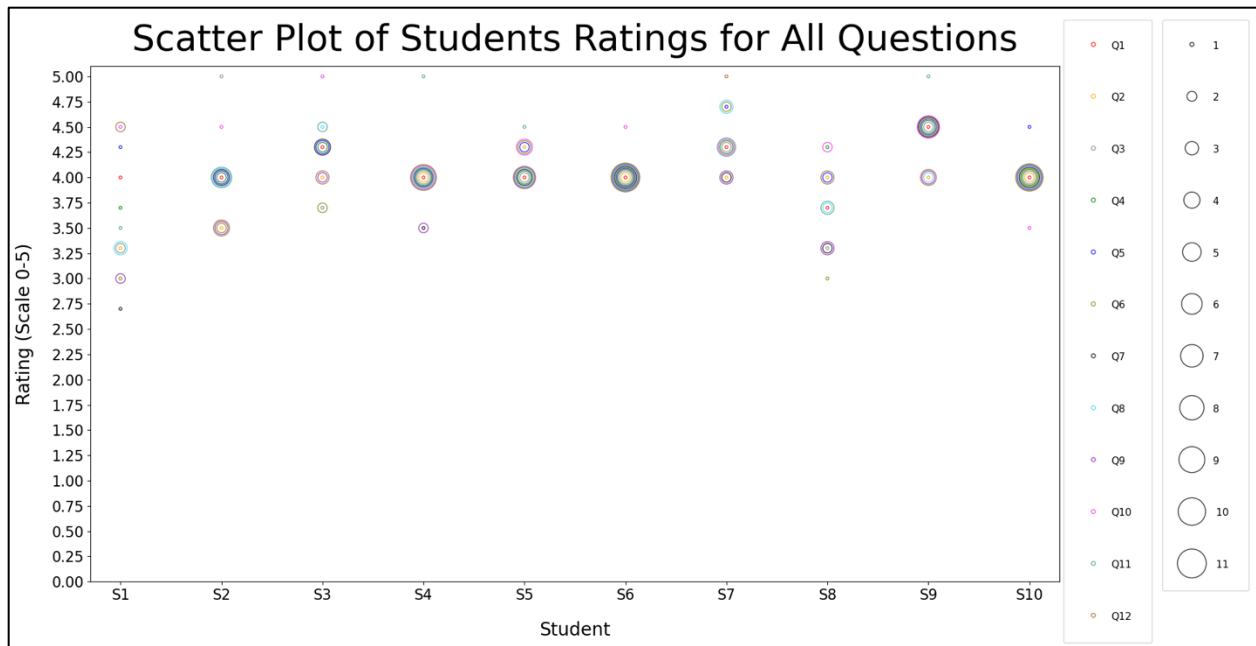
Qualitative Analysis Findings

We conducted post-internship interviews with the interns to further explore their perceptions of K-12 STEM teaching as a potential career. Six participants were interviewed in person. These audio/video interviews were recorded using the research assistant's iPhone and transcriptions were written in Microsoft Word on the research assistant's computer based on the recordings. Four participants were unable to be interviewed in person; instead, recorded interviews were conducted for those interns. Phone interviews were not recorded, but were transcribed by hand during the interview process. Field notes created by the research assistant were also written during all interviews. Transcriptions from interviews were coded to identify common themes in participants' responses. Written data were chunked into common themes of response regarding interns' perceptions of their experience. At times, interns used the same or similar words or phrases such as "challenging" or "low pay," while at other times, different words or phrases such as "rewarding/fulfilling" or "inadequate/couldn't do it" were used. After themes were established, they were organized into positive and negative categories. Two positive themes: 1) sharing a passion for STEM content between intern and student, and 2) student and intern personal growth, and two negative themes: 1) concerns about classroom management, and 2) a lack of interest in the students, were identified prominently in the interview data.

Sharing Passion for Content

Interns reported their positive emotional response when the students understood the material being presented. Interns described the campers' responses as "lighting up" and "getting excited." Multiple interns shared experiences in which the campers naturally engaged in learning activities that were of shared interest with the interns. Experiments involving the dissection of a brain and a cow's eye were mentioned, along with third grade campers discussing the physical laws and limitations of

black holes with one intern. Interns shared surprise and excitement that campers were interested in complex scientific concepts. Interns also shared their feelings of connection with the campers through their common interest in the STEM subjects. Interns' favorable responses to campers' passion for science topics informed us that interns' positive interactions with the campers may influence interns to consider pursuing their own passions for science in the classroom.

Personal Growth

Another positive theme was the interns' own reflections of personal growth in their ability to learn how to explain difficult concepts to the campers. One intern confided,

> This internship opened my eyes to difficulties of teaching, but also how much of an impact teachers can have on the lives of students. I learned how to communicate with people who are not on my same level of knowledge on a subject (Intern 5).

Another intern shared that "the internship taught me how to adapt to all kinds of situations with a variety of knowledge. I had to deliver a variety of knowledge in a truthful, but basic way, and with activities and explanations" (Intern 6). One intern shared that prior to his experience at the KSC he had never worked with such young students before. Intern 2 had the perception that "early elementary kids were just babies, but they are actually little people." This revelation of the complexity of younger learners' thoughts and engagement was affirmed by other interns as well. Having their previous perceptions of teaching challenged and changed was received as a positive experience by the interns. Additionally, interns expressed that knowing earlier in their academic career the prospects for entering into a graduate degree program (MAT degree) and becoming a STEM teacher would have made a positive impact on their interest in entering the teaching profession.

Ultimately, positive responses to participating in the internship indicated interns' perceptions of entering the STEM teaching profession were impacted favorably overall. For example, when we interviewed the interns, two out of ten who responded as "Maybe" in regards to considering the STEM teaching profession on the survey expressed they did actually have an interest in pursuing a K-12 teaching career as an option. One of these interns, who was interested in the field of genetics, indicated she was interested in teaching in the future, and confirmed the transformative nature of the internship experience. The other intern, who was interested in pediatric oncology, decided to enroll the following semester in another education-related internship in which she could work with high school students. She also stated that she had plans to enroll in the MAT program.

Concerns about Classroom Management

One recurring negative theme obtained from the interns' interview data was their dissatisfaction with the behavioral issues of some of the campers. Almost all interns mentioned the difficulty with managing the summer camp classroom. A specific instance of classroom management was shared by one of the interns:

> I was in charge of teaching kids the structure of DNA using candy. I was already nervous before we started. Throughout the entire lesson the kids didn't listen, someone was always trying to eat the candy, someone was always on the verge of crying… I felt bad because I just couldn't make it work (Intern 3).

The struggle for this intern was found in attempting to teach and explain a complex scientific concept - in this case the structure and function of DNA - while also attempting to keep the all the summer camp participants engaged and focused in that module.

Intern 6 shared his perception going into the internship: "I knew [teaching] was time management a lot, and difficult… the internship confirmed these perceptions." A high-anxiety element of the internship experience for several interns was the large number of campers. Finding the ability to keep campers in the large summer camp classrooms engaged proved to be a challenge for several of the interns.

Lack of Youth Interest

An additional negative theme that emerged was the perceived lack of interest in some campers and the interns' struggle with trying to keep them engaged. It is probable that these two negative themes are related. Intern 9 stated that the internship had opened her mind to the possibility of becoming a teacher, sharing: "I enjoyed going to work. My last week I had a huge, wild group of kids. Some groups were really good, others not so much. There are just highs and lows of working with kids for extended periods."

This reporting contradicts the positive finding of some campers' interest in science subjects. These two perceptions offered by the interns are not necessarily mutually exclusive and may even be demonstrated by the same campers on different days, or in relation to different science topics. The interns' lack of previous teaching experience likely led to internalizing and personalizing the campers lack of interest and contributed to a sense that the intern may not be cut out to be a teacher. A lack of interest on the part of the camper correlates to difficulty in classroom management. A camper who is bored will act out in ways which make it difficult for the teacher to effectively continue teaching the rest of the class. The concern about classroom management discussed above is perhaps correlative with a lack of interest.

Discussion

The ten STEM interns who took part in the internship at the KSC's STEM summer camps acquired valuable training and experience. The interns carried out hands-on engaging science activities with Pre-K through 8th grade campers that developed their skills in STEM education. The interns used the engineering design processes, inquiry-based learning, *Next Generation Science Standards* (NGSS, 2013) to conduct these summer camps. The internship was a positive experience for the STEM students, and it increased the interest of several students in the teaching profession, with two of the ten students indicating and initiating plans for pursuing education.

Intern responses showed that the availability of Noyce scholarships would potentially be a motivating factor for some of the STEM students to obtain a teaching certification for a K-12 teaching career. Additionally, interns expressed that knowing earlier in their academic career the prospects for entering into a graduate degree program (MAT degree) and becoming a STEM teacher would have had a positive impact on their interest in entering the teaching profession. As Hutchinson (2012) found, teachers come into the profession from many different points, but knowing about the option for teaching early on is important in the recruitment process.

As stated in the post-internship interviews, several interns indicated that knowing about the options available in STEM teaching earlier in their academic career may have given them reason to consider a career in K-12 STEM teaching. However, these interns felt that they had gone too far in their academic pursuits in their respective STEM fields to change course. From this, we can conclude that having knowledge and internship opportunities in STEM education earlier in their academic career may also have an impact on their interest in a career as a K-12 STEM teacher. Offering an

internship experience in STEM education in the students' first or second year of their undergraduate STEM degree might be optimal in order to cultivate an interest in pursuing an academic career as a K-12 STEM teacher. This is important given the research about recruiting students in their junior or senior years (Moin et al., 2005); allowing students opportunities for teaching in their earlier degree pursuits (i.e. freshman and sophomore years) might contribute to more firm decisions to add an MAT degree in their junior or senior year.

Staff from the KSC indicated a desire to continue this NSF-funded summer internship program as they were impressed by the STEM students' work ethic and eagerness to learn. Interns also acknowledged that they were surprised at learning that they were better at working with younger students than they originally perceived themselves to be. This finding supports other research that suggest the relational aspect of teaching is an important factor in STEM students' desire to go into teaching (Borgerding, 2015). This was particularly noted as the explanatory cause of the change in opinion towards pursuing a career in K-12 STEM teaching by one of the interns who previously had not considered K-12 teaching as a career option. Earlier engagement and positive relationship building with the summer camp participants may be a key to encouraging STEM college students to consider K-12 teaching as a profession.

Interns' concerns about classroom management support similar research that found that more of a focus on content (over classroom management) excites more STEM students about teaching (Borgerding, 2015). We posit that feelings of inadequacy in the area of classroom management could be mitigated by taking the relevant education courses as a part of the students' core STEM curriculum in pursuit of their STEM degree. Since these students had not yet taken any education courses, they had no experience in pedagogy or best practices in inquiry-based instruction that would have supported their knowledge and skills in the area of classroom management. The KSC did provide some training in managing materials and groups, but further coursework and field experiences through a teacher training program would provide more comprehensive knowledge and skill development.

Overall, our research supports that of Wong-Ratcliff and Mundy (2019), Mundy and Ratcliff (2021), and others who have found internships to be an important recruitment tool for STEM majors to consider a career in science and/or mathematics teaching. Beyond recruitment, it is essential to continue to support developing teachers at all points of their training (prior to becoming an education major, during, and beyond into their beginning teaching) (Hutchinson, 2012). Future research is important to longitudinally track interns' career choices over time to see if interns' actually did pursue a degree in teaching (even if it was not their immediate post-graduation career choice).

Implications

Various reports have expressed the nationwide growing shortage of STEM teachers and the challenges faced in increasing the number of qualified STEM teachers in the public schools. There is an urgent national need to improve the quality of STEM education by putting well-qualified STEM teachers in high-need schools. This is so these individuals can teach STEM subjects in an effective way that may lead to more students interested in pursuing a STEM degree in college. We need to ensure that K-12 students receive the STEM training that is essential for future success. Effective STEM teacher preparation provides a vital function in pursuit of an educated and engaged populace. This project comes at a critical time for Kentucky because of the changing student demographic and the state's adoption of the *Next Generation Science Standards* (NGSS, 2013) that is undergirding a strong curriculum in STEM and the development of new statewide performance assessments aimed at improving the educational achievements for all students.

Interns shared key takeaways in their interviews regarding the variables that impacted their willingness to consider a career in STEM teaching. One of the cited reasons for some of the interns to not consider such a career move in K-12 STEM teaching was that they had already decided what

degree they wanted to pursue. In addition, the interns had already spent too much time in their academic STEM degree track and did not want to change their career pursuits at this stage. We conclude that if knowledgeable and informed STEM faculty could promote K-12 STEM teaching as a career in their STEM subjects with freshmen and sophomores, the outcome may prove different. Early exposure to internships also allows time for interested students to continue seeking education related opportunities during college, such as the student who went on to do another education-related internship with high school students. In this way, students have multiple opportunities in education to see if the fit is right for them before embarking in an MAT program.

Options in pursuit of attaining a degree in teaching with a focus on STEM education may include scholarships, accelerated programs, wrap around supports for student success, and targeted field experiences with strong mentor teachers. It is our opinion that these details should be shared with students in STEM fields early on in their undergraduate STEM degree program. Information about options and incentives such as summer internships at the KSC should not only be held as resources in academic advising offices but be readily available to interested STEM students. A forum between education and STEM faculty may need to be established in order for accurate, up-to-date information to be shared between faculty and students, especially in the area of wages and career development.

Kristin Cook (kcook@bellarmine.edu), Associate Professor of Science Education and Associate Dean of the School of Education at Bellarmine University, received her doctorate in Curriculum & Instruction - Science Education and Environmental Sciences from Indiana University. Dr. Cook serves as a professional developer and consultant for K-12 STEAM-focused school reform and project/problem-based learning development. Dr. Cook's research focuses on engaging students and teachers with the community of science through the exploration of socio-scientific inquiry and transdisciplinary STEAM instruction.

Akhtar Mahmood (amahmood@bellarmine.edu) received his Ph.D. in Physics from the State University of New York at Albany in 1998. He is a Professor of Physics at Bellarmine University and founded the Eureka Learning Community for STEM students. He served as the President of the Kentucky Association of Physics Teachers (KAPT) from 2009 – 2012 and is currently serving as a member on the Kentucky Academy of Science's Committee on STEM Education. He was featured among the 20 People to Know in Education and Workforce Development by the Louisville Business Journal in 2017. He served as the PI of an NSF S-STEM project, an NSF Noyce Capacity Building project and is currently serving as the PI of an NSF Noyce Track-1 project.

Cody Nygard (cnygard@bellarmine.edu) is a PhD student in the Education & Social Change program at Bellarmine University. His research revolves around environmental education and connectedness to nature. Cody is an adjunct instructor at Bellarmine, as well as a high school Life Sciences teacher at Walden School, a private, non-religious K-12 school in Louisville, KY.

Mira Gentry-Johnson (mira.gentry-johnson@louisvilleky.gov) received her M.S. in Microbiology and Evolutionary Genetics from the University of Louisville in 2008 and BS in Biology from Centre College. She is the Senior Manager of Education at the Kentucky Science Center. Mira has been in informal education for 16 years.

Mellisa Blankenship (mellisa.blankenship@louisvilleky.gov) received her B.A. in Biology from Indiana University in 2009. She is the Director of Education at the Kentucky Science Center. Mellisa has been in informal education for 17 years.

References

Almy, S. & Theokas, C. (2010). *Not prepared for class: High-Poverty schools continue to have fewer in-field teachers*. Education Trust. https://edtrust.org/resource/not-prepared-for-class-high-poverty-schools-continue-to-have-fewer-in-field-teachers/

American Association of College of Teacher Education (AACTE). (2013). *The changing teacher preparation profession - A report from AACTE's professional data system (PEDS)*. Author. https://aacte.org/2013/03/aacte-releases-first-national-data-report-on-teacher-preparation-profession/

Barth, P., Dillon, N., Hull, J., & Higgins, B.H. (2016). *Fixing the holes in the teacher pipeline: An overview of teacher shortage*. Center for Public Education.

Borgerding, L. (2015). Recruitment of early STEM majors into possible secondary science teaching careers: The role of science education summer internships. *International Journal of Environmental & Science Education, 10*(2), 247-270.

Heron, P. & McNeil, L. (2016). Phys 21: Preparing physics students for 21st century careers, A report by the joint task force on undergraduate physics programs. American Physical Society and the American Association of Physics Teachers.

Hutchinson, I. F. (2012). Addressing the STEM teacher shortage in American schools: Ways to recruit and retain effective STEM teachers. *Action in Teacher Education, 34*(5-6), 541–550.

Kyriacou, C. & Coulthard, M. (2000). Undergraduates' views of teaching as a career choice, *Journal of Education for Teaching,* 26(2), 117-126, https://www.doi.org/10.1080/02607470050127036

Low, E., Ng, P., Hui, C., & Cai, L. (2017). Teaching as a career choice: Triggers and drivers. *Australian Journal of Teacher Education*, *42*(2), 28-46. http://dx.doi.org/10.14221/ajte.2017v42n2.3

Matthews, C. M. (2007). *Science, engineering, and mathematics education: Status and issues* (98-871 STM). Congressional Research Service.

McDonald, C. V. (2017). Who is teaching science in our high schools? Exploring factors influencing pre-service secondary science teachers' decisions to pursue teaching as a career. *Teaching Science: The Journal of the Australian Science Teachers Association*, *63*(4), 33–44.

Moin, L., Dorfield, J., & Schunn, C. (2005). Where can we find future K-12 science and math teachers? A search by academic year, discipline, and academic performance level. *Science Education, 89*(6), 980-1006.

Morrell, P. D., & Salomone, S. (2017). Impact of a Robert Noyce scholarship on STEM teacher recruitment. *Journal of College Science Teaching, 47*(2), 16-21.

Moss, J. D. (2020). "I was told to find what broke my heart and fix it:" College students explain why they want to become teachers. *Cogent Education*, *7*(1). https://www.doi.org/10.1080/2331186X.2020.17342

Mundy, M. & Ratcliff, M. (2021). Attracting Noyce students to become STEM teachers. *Journal of Instructional Pedagogies, 25*, 1-8.

National Research Council. (2010). Preparing teachers: Building evidence for sound policy. National Academies Press.

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

Parsons, S. (2013). Cerritos college employs internships to boost STEM teacher supply. *Community College Week, 25*(12), 9.

President's Council of Advisors on Science and Technology. (2010). *Prepare and inspire: K-12 science, technology, engineering, and math (STEM) education for America's future.* Author. http://www.whitehouse.gov/sites/default/files/microsites/ostp/pcast-stemed-execsum.pdf

Richardson, P. & Watt, H. (2014). *Why people choose teaching as a career: An expectancy-value approach to understanding teacher motivation.* Routledge.

Taimalu, M., Luik, P., Kantelinen, R., & Kukkonen, J. (2021). Why they choose a teaching career? Factors motivating career choice among Estonian and Finnish student teachers. *Journal of the Humanities & Social Sciences, 25*(1), 19–35. https//doi.org/10.3176/tr.2021.1.02

U.S. Department of Education. (2022). *Teaching shortage areas: Nationwide listing.* Office of Postsecondary Education. https://tsa.ed.gov/#/reports

White, S. & Tyler, J. (2014). *Who teaches high school physics? Results from the 2012-13 nationwide survey of high school physics teachers.* American Institute of Physics.

Wong-Ratcliff, M. & Mundy, M. (2019). Recruiting future STEM teachers through summer internship program. *Journal of Education and Human Development, 8*(3), 1-9.

Appendix A

Survey of Interest in Teaching

Name:
Internship Focus/Topics and Reason for Choosing:
Internship Grade Levels Taught:
Current or Completed Major:
College Year: Freshman/Sophomore/Junior/Senior/Graduated/Graduate Student
Career Interest Prior to Internship:
Career Interest After Internship:
Interest in Teaching Career:
      Definitely Interested (takes you to section of Interest)
      Potentially Interested (takes you to section of Potential Interest)
      Not at all Interested (takes you to section of Non-Interest)
Reasons for Interest in Teaching
      Enjoy working with youth
      Other
      Good pay
      Job security Family members work in education
      Want others to have a better education than I had
      Great experience in my own education
      To show others how to do things
      Variety-every day is different
      Passion for a specific subject
      Inspired by my teacher(s) at school
      Want to make a difference/Give back to the community
      Other (please list)
Reasons for Potential Interest in Teaching
      Better Pay
      Learning more about what the job entails
      More opportunities for student loan forgiveness
      Less work outside of school hours
      Safer schools and classrooms
      More opportunities for college scholarships
      More prestige or respect
      More opportunity for career advancement
      More flexibility in how I would do my job
      Other (please list)
Reasons for Non-interest in Teaching
      Requires working too many hours
      There are teachers I don't like
      Lack of school resources and supplies
      I cannot be my own boss
      Too much pressure to "teach to the test"
      Teachers do not get enough respect
      I don't like working with children or young people
      I wouldn't be good at it
      Not much opportunity for career development

It does not initially pay as well as other occupations

Other (please list)

Internship experiences that contributed to your view on teaching:

In what ways did your perceptions of teaching change or not change during the internship?

Most positive aspects of internship:

Most challenging aspects of internship:

If you were offered a scholarship from NSF in the amount of $10,000 in your Junior Year, $10,000 in your Senior Year, and $10,000 towards your teacher certification tuition expense, would you consider obtaining a high school teaching certification for a teaching career at a high school for at least 3 years?

Recommendations to future Noyce interns:

Appendix B

Observational Instrument given by the Kentucky Science Center

Intern Evaluation Form - School's Out Science Camp

Intern Name: _Stash_   Grade _1-2_   Duration (15 min minimum) _4 hours_
Camp Name _Speedometry_   Number of campers _11_   Camp Teacher: _Cana_
Observer _Cana_   Position _Teacher_   Date/Time: _July 12_

Please rate the Intern using the rating scale listed below:

1= **Unsatisfactory:** Performance is consistently unacceptable.
2= **Below Expectations:** Performance fails to meet position requirements on a frequent basis.
3= **Meets Expectations:** Performance is regularly competent and dependable.
4= **Exceeds Expectations:** Performance is routinely above position requirements.
5= **Outstanding:** Performance is consistently superior.

Facilitation Skills	1	2	3	4	5	Comments
Follows Science Center rules and policies with students				✓		
Encourages campers to participate and interact in activities				✓		
Allows students take control of learning but is there to assist with guiding questions.				✓		
Projects positive image and attitude about STEM					✓	
Is courteous and respectful of to all students and staff					✓	
Is motivated and takes initiative			✓			A little hesitant to initiate lessons great help during lesson activities
Confident in engagement			✓			hesitant at first to initiate lesson but supported engagement throughout lesson
Makes the classroom a fun, learning environment				✓		
'Reads' student level of engagement effectively and adjust facilitation accordingly.			✓			
Comfortable with grade level content.				✓		
Any additional content provided from college degree (Engineering, Physics, Math, Chemistry).				✓		
Navigate content delivery through conversation rather than lecture.				✓		

Please circle the words that describe the Intern:

Adaptable	Hard worker	Lazy	Rowdy
(Calm)	(Helpful)	Negative	Rude
Disruptive	Honest	(Polite)	Self-Motivated
Energetic	Insubordinate	(Quiet)	Stubborn
Fun	Intelligent	(Reliable)	Team Player

Best practices observed from Intern's facilitation:

Please note any additional comments:

I think with time and a little practice he
can overcome his hesitation as a class leader

☑ I recommend this Intern for future camp programs.

☐ I do NOT recommend this Intern for future camp programs

College STEM Faculty Teaching Practices: The Influence of a Professional Development

Lloyd M. Mataka [iD]
Lewis-Clark State College

Jon C. Saderholm [iD]
Berea College

Tracy T. Hodge [iD]
Berea College

ABSTRACT

In this paper, we present the influence of a professional development, including a faculty learning community (FLC), on STEM teaching practices at a southern US liberal arts college. We used a quasi-experimental mixed-method design to collect information about teaching practices and the influence of the professional development. The Reformed Teaching Observation Protocol (RTOP) was used to evaluate the faculty teaching strategies. The RTOP ratings indicated that most faculty used some sort of active strategies. Further, faculty who attended the professional development had significantly higher RTOP scores than their counterparts. Results from interviews show that faculty were willing to make changes to their teaching approaches based on the workshop. Results from this study contribute to the knowledge base describing the influence of professional development on STEM college faculty teaching strategies.

Keywords: professional development, STEM, RTOP, FLC, quasi-experiment

Introduction

Colleges need to produce graduates who are independent, able to apply their knowledge, solve complex problems, and work in teams among other attributes (Gijbels et al., 2006). This requires learning environments that focus on the acquisition of these skills, such as reformed teaching and learning approaches, that encourage the use of active teaching techniques for holistic student learning (Piburn & Sawada, 2000). For this reason, evidence-based practice has been encouraged because it emphasizes the importance of research in teaching and learning (Masters, 2018). Researchers have been tasked with informing educators of effective strategies to enhance students' learning (Masters, 2018; Slavin, 2002). According to Georgiou et al. (2020), evidence-based practice is critical because it bridges the gap between research and practice. To this end, some colleges and universities provide professional training for their faculty to be familiar with evidence-based practice. In the present case, a professional development (PD) was designed around participation in a faculty learning community (FLC) that promoted the use of learner-centered strategies to foster student engagement (Harris & Cullen, 2008). In this study, we explore the effect of participation in a PD with a FLC on STEM college faculty's use of learner-centered teaching strategies.

Reformed (Learner-Centered) Teaching Approaches

According to MacIsaac and Falconer (2002), reformed courses are courses that are "taught via the kinds of constructivist, inquiry-based methods advocated by professional organizations and researchers so that these future teachers would be taught as they were expected to teach" (p. 479). Reformed classrooms focus on lesson designs that encourage students' exploration using alternative approaches, use of thought-provoking activities, active classroom communication, respectful group interactions, and students' reflections of the learning material (MacIsaac & Falconer, 2002). According to the Geological Society of North America (GSNA, 2016), reformed teaching entails that faculty shift from teacher-centered to learner-centered approaches that emanate from evidence-based research.

As a result of this paradigm shift, many colleges and universities are now promoting the transformation of faculty teaching strategies from passive, to learner-centered environments. These environments include activities that can be described as active, constructive, and interactive (Chi, 2009). Meaningful learning is encouraged through active interaction between the students and instructors and among the students (Chi, 2009). Teachers in these environments use group work to encourage active questioning from students and discussion among students. Group activities help foster teamwork and camaraderie among students while distributing cognitive load (Chi, 2009; Weinstein, 2002). In a reformed classroom, teachers ask questions that intentionally promote students' conceptual understanding. An effective teacher listens and probes students' understanding (Ruiz-Primo & Furtak, 2007) to challenge the students and make them seek deeper responses to the questions. Effective reformed classrooms also require instructors to understand the students' conceptions before learning a new concept to challenge their understanding (National Research Council, 2000). Kranzfelder et al. (2019) observed that teachers in reformed classrooms guide as students are working and encourage both group and whole-class discussions before summarizing the material to the whole class. Further, MacIsaac and Falconer (2002) encourage using activities that will "anchor student dialogue in a shared, negotiated, and explicit external representation" (p. 484).

Faculty Professional Development

According to Pesce (2015), graduate schools have not focused on preparing students for college teaching jobs. This necessitates a need for faculty PD in higher education. PD programs are vital, as they enable college faculty to adopt new active and research-based teaching practices. For instance, a meta-analysis conducted by Bilal et al. (2019) found that PDs improve pedagogical knowledge and competence among faculty. Based on the meta-analysis, Bilal et al. (2019) encouraged institutions to include PDs in the faculty support system to enhance teaching, leadership, and administrative skills. While faculty welcome PD, Hott and Tietjen-Smith (2018) found that they need incentives to encourage them to get involved in the PD activities (e.g., contribution to tenure portfolio). Pesce (2015) found that faculty were more willing to get involved in PD when they collaborate with colleagues. Therefore, PD programs involving FLC are becoming more popular because they provide a supportive learning environment for participating faculty, and they model the strategies used in active-learning classrooms. FLCs involve a group of faculty members who meet regularly throughout the semester or year to address a topic or actions related to teaching (Richlin & Essington, 2004). Brydges et al. (2013) recommended that a group of 8-12 faculty should meet biweekly for 90 minutes each meeting. Orzech (2021) has suggested that facilitators of FLCs must have a clear vision that is shared by the group and encourage input from the members. Cox (2004) defines FLCs as "a cross-disciplinary faculty and staff group of six to fifteen members who engage in an active, collaborative, year-long program with a curriculum about enhancing teaching and learning and with frequent seminars and activities that provide learning, development, the scholarship of teaching, and community building" (Cox, 2004, p. 8). Cox (2004) identified two types of FLCs: (1)

cohort-based, which is a group of a specific cohort of faculty (e.g., first-year faculty), and (2) topic-based, which focuses on a specific issue that needs to be addressed for effective student learning. For instance, groups can focus on an innovative teaching approach, syllabi redesign, and action research.

A study by Richlin and Essington (2004) provides an overview of the extent of the availability of FLCs in the US and Canada. This study found that about 132 higher education institutions in the US and 308 in Canada have been involved in FLCs. In a study of the outcomes of participation in FLCs at six universities, Beach and Cox (2009) found that instructors involved in an FLC used nontraditional teaching practices in their classes. While traditional teaching approaches (e.g., direct lecture) focus on faculty telling students what to know, nontraditional approaches rely on evidence-based theories that advocate meaningful learning through students' active engagement in the learning process (Freeman et al., 2014). Another study by Brydges et al. (2013) also reported that faculty used nontraditional teaching practices and others planned to use these practices in their subsequent classes after participation in an FLC. Further, Genc and Ogan-Bekiroglu (2004) have argued that faculty who are involved in PD activities are more likely to use nontraditional teaching practices. Cox (2018) indicated that faculty who are involved in an FLC use active learning/teaching techniques such as cooperative and collaborative learning, discussions, and students' centered learning in general. Pelletreau et al. (2018) echoed these sentiments when they found that faculty from different backgrounds were able to teach using active approaches after collaboratively developing a unit on the "potential effect of mutations on DNA replication, transcription, and translation" (p. 1). In this study, the authors also found that students who learned through the developed unit performed significantly better on the exam than those who did not. Hopkins (2020) also found that faculty at a community college believed that PD had an impact on how they designed, taught, and assessed their courses. In another study by Tinnell et al. (2019), faculty indicated that FLCs were responsible for their decision to either begin using or focusing on students' collaborative learning. These authors also found that the FLC enhanced positive faculty collaboration. This can influence students' learning as Schmitz et al. (2018) found that students improved in some aspects of their learning outcomes after their faculty joined an FLC. In a study by Schmidt et al. (2018), faculty believed that FLCs improved their "knowledge and skill-building for creating more equitable learning environments for their students" (p. 99). This was echoed by Hirst et al. (2021) who found that FLCs increased faculty awareness of the needs of students from diverse backgrounds. In the present study, we developed a PD program with an FLC component to encourage the faculty's use of active teaching practices.

Purpose of the Study

The purpose of this study was to explore the science, technology, engineering, and math (STEM) faculty's use of active teaching techniques at a liberal arts college. We also compared the teaching approaches between faculty involved in the PD with a FLC (workshop participant) and those not involved (comparison group). In doing this work, we addressed the following research questions:

- How is participation in a PD project using an FLC associated with changes in faculty use of student-centered approaches to teaching?
- What aspects of participation in an FLC are perceived by STEM college faculty as being important?

Methods

Research Design

This research study is a treatment-control group, quasi-experimental mixed-method design (Creswell, 2003; Tashakkori & Teddlie, 1998). This study describes the teaching practices of STEM college faculty at a southern, rural liberal arts college and explores the extent of the use of active approaches in teaching to determine the influence of a PD project. The Reformed Teaching Observation Protocol ([RTOP], Piburn & Sawada, 2000) was used to structure the ratings of classroom observations of faculty in both the PD and comparison groups. In total, nineteen faculty members were observed, thirteen in the PD and eleven in the comparison group. Additional data were collected from five participants before and after the PD. These members are present in both treatment and control groups and participated in structured interviews designed to expose their knowledge of, and dispositions toward, teaching. Tables 1 and 2 describe the characteristics of the faculty members.

Table 1

Participants and Their Disciplines

Discipline	PD Group	PD Group 2014 Cohort	2015 Cohort	Comparison Group
Biology (BIO)	2	1	1	3
Chemistry (CHM)	3	2	1	3
Computer Science (CSC)	3	2	1	1
Math (MAT)	2	1	1	3
Physics (PHY)	3	2	1	1
Total	13	8	5	11

Table 2

Participants Characteristics

	PD Group ($N = 13$)	Comparison Group ($N = 11$)
Mean teaching years	13.1 ± 7.3	10.8 ± 7.9
Tenured	8	6
Nontenured	5	5
Male	8	9
Female	5	2

Professional Development

The PD was designed and carried out by an invited post-secondary STEM education expert who administered the workshop and a STEM PD expert from the host college. This was a two-stage PD model promoting the adoption of reformed teaching methods in post-secondary STEM classrooms. Faculty self-selected into the project. The project began with a four-day course redesign workshop at the end of the academic year in May 2014 and 2015. The workshops took three full days (9:00 am – 3:30 pm) and one-half day (9:00 am – 12:00 pm). During these summer workshops,

participants were exposed to learner-centered pedagogies (e.g, Fink, 2003) that were designed to target deep learning approaches, while redesigning the syllabus for a course they intended to transform for the following academic year. Workshop themes included student motivation, significant learning goals and objectives, learner-centered assessment, major non-traditional assignments, and objective-assignment alignment. The following year, participating faculty were supported by ongoing participation in a bi-weekly FLC of one and half hours each meeting. Throughout the following academic year, during these regularly scheduled gatherings, faculty discussed the challenges they were facing, learned about potential sources of data describing teaching effectiveness, and watched occasional videos of each other's classroom instruction. Participation in the FLC was encouraged but not mandatory. The majority of FLC meetings were attended by all PD faculty, with one or two members missing an occasional meeting.

The Teaching Context

The courses observed were mainly for students majoring in math and the sciences, which include chemistry, physics, biology, computer science, math, and statistics. Some of the courses such as chemistry, physics, and biology had associated compulsory laboratory components. The lecture component for each course, where the observation took place, had a 70-minute time limit, three days per week. Although most of the observed courses were 100-level entry courses, some were 200-level courses.

Classroom Observations

The classroom observations were rated using the RTOP (Piburn & Sawada, 2000). Observations were documented through running-record notes, which were then rated using the RTOP. See Table 3 for observation frequencies of participants. The RTOP has been widely used as an observational tool for reformed practices in colleges and universities (Lakshmanan et al. 2011; Lund et al. 2015; Park et al. 2011). The RTOP is a five-point Likert scale instrument that has been well validated (Sawada & Piburn, 2000). The instrument has a total of 25 items with a maximum score of 100. The instrument has three major sections: lesson design, knowledge, and classroom culture. The knowledge section is further subdivided into propositional and procedural knowledge, while classroom culture is subdivided into communicative interaction and student/teacher interactions. This makes a total of five RTOP sections. The lesson design section looks at how the lesson aims to engage students, while the propositional knowledge section looks at the content and how it is presented to students for conceptual understanding. The procedural knowledge focuses on the actions students are encouraged to do for conceptual understanding. The communicative interactions focus on teacher/student communications and student/student communications within the lesson. Lastly, the student/teacher interactions show how the teacher's actions encourage active participation of students in the learning process.

Before conducting the observations, the two observers conducted a two-step inter-rater reliability process. The first step involved training to use the RTOP. This was done by the two observers watching three RTOP videos and rating them independently, then comparing the rating to the suggested rating from the RTOP training guide. In the second step the two researchers, who are both science educators, observed five classes simultaneously and rated them separately for comparison. During the comparison, differences were discussed with reference to the training videos to ensure accuracy in the RTOP rating. Afterward, individual researchers observed and rated the rest of the classes.

In total, 19 faculty members were observed. The comparison group had eleven faculty while the PD group had thirteen faculty. Of the thirteen faculty in the PD group, five had been observed as

part of the comparison group in the previous year, before these faculty attended the second workshop and joined the participant group. The observations were conducted concurrently with the ongoing FLC, with the first observations carried out about four weeks into the semester and a second observation occurring late in the term. See Tables 3 and 4 for observation information. To discern which changes in classroom practice were associated with participation in this faculty development program, data were gathered from both PD faculty and a comparison group consisting of non-participating STEM colleagues (see Tables 1 and 2).

Table 3

Observation Frequencies of Various Categories

Category	Participants	Participants with	
		3 observations	2 observations
Fall 2014 semester	12	9	3
Fall 2015 Semester	12	4	8
Treatment	13	7	6
Control	11	6	5

Table 4

Treatment and Control Groups Observations Each Semester

Semester	PD faculty	Comparison faculty
Fall 2014	6	6
Fall 2015	7	5

Data Analysis

Minitab software was used for data analysis. The Kolmogorov-Sminorv (KS) test was used to determine the normality of distribution of RTOP scores. T-tests were used to compare sample means and Cohen's *d* provided the effect size. One-way ANOVA with Bonferroni correction was used to compare means from different sections of the RTOP. Mean RTOP scores of the observations for faculty were used to calculate the difference between the participant group and members of the comparison group.

Interviews

Emails were sent to faculty members who participated in the PD and FLC asking them to volunteer to be interviewed. Six faculty members volunteered for individual interviews and all six participated in the interviews. The semi-structured interviews sought to determine what changes in their practice, if any, had occurred since the workshop. The interviewees answered two questions: (1) what changes have you made to your class this semester and (2) in what ways did the workshop influence those changes.

Results

We conducted a hypothesis of normality to determine if the population this sample is drawn from is normally distributed using all the RTOP scores. The null hypothesis was that RTOP scores of this sample belong to a normal distribution. Using a Komolgorov-Sminorv test, we found $D = 0.11$, $p = 0.99$ for workshop participants ($N = 13$) and $D = 0.20$, $p = 0.72$ for the comparison group ($N = 11$) and $D = 0.2$, $p = 0.4$ for all the participants ($N=19$). Therefore, we accepted our null hypothesis. Using an ANOVA, we found that teaching experience ($F= 1.12$, $p = 0.37$), faculty discipline ($F = 2.63$, $p = 0.08$), and observation cycle did not influence the RTOP scores.

To What Extent Did Faculty Use Reformed Teaching in Their Classrooms?

Each sub-section of the RTOP is scored out of 20 possible points, making the total possible score 100. From all the RTOP observations, the mean score was 63.5 with a range of 36-81. This implies that the faculty in this study generally used reformed teaching approaches (GSNA, 2016). We also investigated the mean scores for different sections of the RTOP. The RTOP sections are Lesson design (LD), Propositional Knowledge (Prop K), Procedural Knowledge (Proc K), Communicative Interactions (CI), and Student/Teacher Interaction (STI) as shown in Figure 1. The mean scores in Figure 1 represent both PD and comparison faculty with the data collected after the PD workshop.

Figure 1

Mean Scores Based on RTOP Sections From 19 Science Faculty
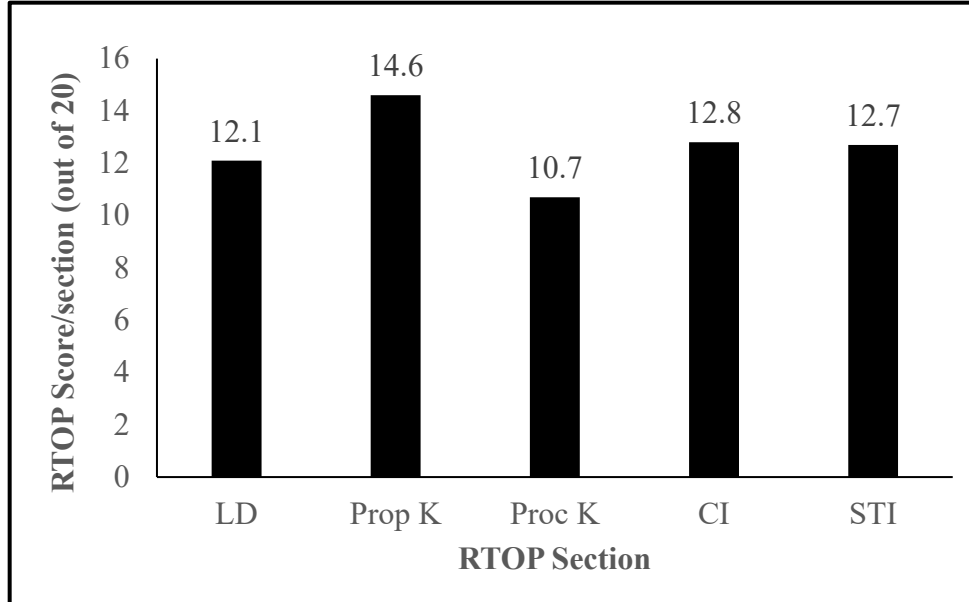
As might be expected, the trend from Figure 1 shows that faculty displayed the highest propositional knowledge and lowest procedural knowledge. We also observed that both sections of the classroom culture have relatively high scores. An ANOVA ($F = 14$, $p = 0.000$) indicated that there was a significant difference in the scores among the sections. Follow-up tests with Bonferroni correction indicated that scores on propositional knowledge were significantly higher than the scores in the other categories. Further, scores on procedural knowledge were significantly lower than scores in the other categories.

What Is the Effect of Workshop Attendance?

Since the *F*-test (*p* = 0.01) showed that variances between the participating faculty and the comparison group were different, Welch's t-test was used. Participating faculty had higher RTOP scores than those in the comparison group. See Table 5 for this information. A significant difference was observed (*t* = 4.03, *p* = 0.000) with a large effect size (*d* = 1.74). Furthermore, when compared against the comparison group, a *t*-test with Bonferroni correction indicated statistical differences in all the RTOP sections between the two groups. See Figure 2 for this information. Although these tests show that participating faculty taught in a more reformed way, they are unable to show that participating in the project affected the difference because there is no pre-test data for most of the faculty members. For that, we looked at five faculty who were in the control group during the first phase of data collection and were in the PD group during the second phase of data collection (Table 5). These participants provided data for pre- and post-workshop. Further these participants represented the disciplines, with one faculty member each, from physics, chemistry, biology, computer science, and mathematics. We conducted a repeated-measures *t*-test of these five faculty members RTOP scores collected before and after they attended the workshop (see Table 6). There was an observed significant improvement on the faculty instructional approach (*t* = 3.40, *p* = 0.03) with a large effect size (*d* = 1.04). Table 7 provides a summary of the observations of one of the five faculty members to demonstrate how their changes were coded on the RTOP.

Table 5

Comparison of RTOP Scores Between Groups

Group	N	Mean	SD	Range	*t*	*p*	*d*	Grand mean
Participants	13	70.4	5.4	60-81	4.03	0.001	1.74	63.5±11.4
Comparison	11	55.3	11.3	36-70				

Table 6

RTOP Repeated Measures t-test of Observations From Five Science Faculty

Observation	N	Mean	SD	*t*	*p*	*d*
Pre	5	63.2	6.3	3.40	0.03	1.04
Post	5	68.5	3.3			

Table 7

Sample Observation Summary Depicting an Observed Change in One Faculty's Approach

Before the workshop	After the workshop
• The class addressed important concepts in the subject	• The class addressed important concepts in the subject
• Students' participation was not encouraged during the lecture time	• The class encouraged students' participation
• There was some group work	• Crossover of students between groups was encouraged.
• Minimal exchange of ideas between teachers and students was observed	• The teacher respected students' ideas: students respected each other's ideas
• The class was mostly teacher talk with some teacher ↔ student interaction	• Student-Student talk was encouraged, although the teacher dominated the discussion.
• Connections between class content with both other content and real-world not observed	• Some connections between content and real world observed.
• Probing of students' understanding was rarely used	• Probing of students' understanding was observed.

Figure 2 shows scores between the PD and the comparison group for different sections of the RTOP. The results show that the PD group had better scores than the comparison group among the different sections of the instrument. Statistical analyses indicated that all these differences were significant at the alpha level of 0.05.

Figure 2

Mean Scores Based on RTOP Sections for Workshop (N = 13) and Comparison (N = 11) Groups After PD
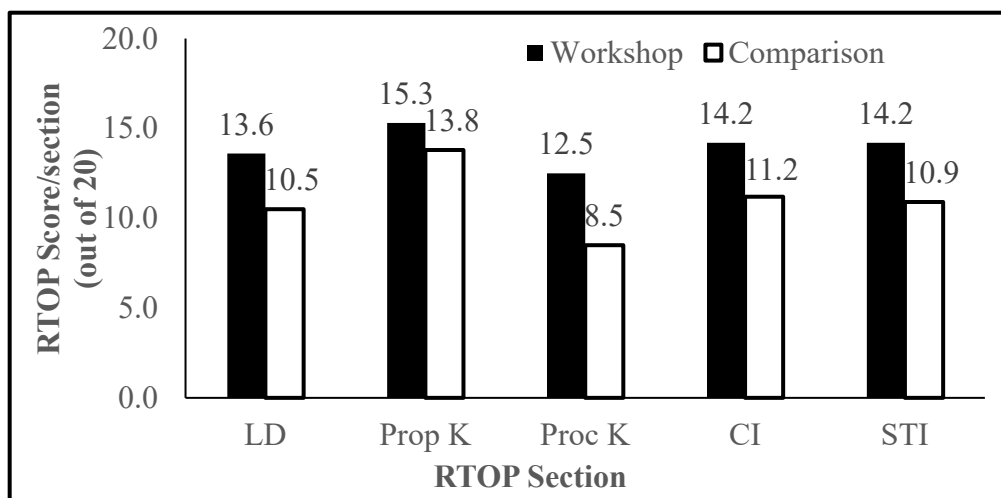
Figure 3 shows differences in RTOP scores in each section, with each value indicating how much the faculty in the PD section scored higher than the faculty in the comparison group. The largest difference in mean section scores was observed between Procedural Knowledge (Proc K) scores seconded by Student-Teacher Interactions (STI). The smallest difference in mean section scores was observed between Propositional Knowledge (Prop K).

Figure 3

Mean Scores Differences Between the Workshop and Comparison Groups for Different RTOP Sections
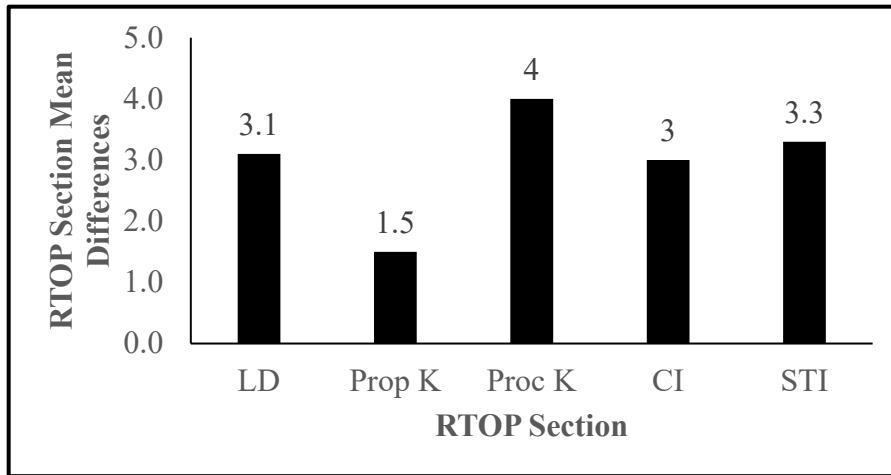
What Was the Effect of the FLC?

Analysis was also done to determine changes in the RTOP scores within the semester to see how the FLC influenced instructional approaches. In this case, the first and last observations were used. For inferential statistics, the repeated measure t-test and its p-value were calculated to determine statistical significance. All these results are presented in Table 8 and for reference, data for the comparison group is also included. Although there was a small increase of 2.5 in the RTOP scores for the PD group between the first and last observation, this increase was not statistically significant. A decrease of 4.3 was observed for the comparison group between the first and last observations, which was also not statistically significant.

Table 8

Analysis of RTOP Mean Scores for First and Last Observation

Measurement	PD Group	Comparison Group
N	13	11
First Observation mean score	68.6	56.9
Second Observation mean score	71.1	52.6
p	0.36	0.27

The Interview Data

The qualitative data answered questions regarding changes to the faculty teaching practice and how the workshop influenced that change. These questions were aimed at supporting results from the quantitative section of this study. Below is a summary of responses from the six participating faculty members. They are identified by an abbreviation for their discipline followed by a designated participant number for the study.

CSC1 redesigned a course as part of an ongoing plan to include more outside assignments and process oriented guided inquiry learning (Simonson, 2019) activities. However, the workshop

provided ideas that helped design some of these classroom activities and deal with the classroom dynamics, such as the placement of students in various groups. CSC1 stated:

> The workshop has given me some new ideas about how I can construct some particular activities. I think I was already kind of pretty far ahead of the curve in terms of active learning. I am making some changes to how I structure teamwork that I think has grown out of the meetings, so they have been valuable.

MAT1 moved the class to a new room that was conducive to facilitating group classroom interactions. He reported that having an appropriate space enabled him to have a kind of classroom that facilitates the kind of activities he has always wanted to do with his students. In this case, the workshop helped him articulate his learning objectives and construct a syllabus that was reflective of what he wanted the students to accomplish in class. The workshop encouraged him to formalize activities that he could not do regularly but now had become a regular part of his course due to the workshop. MAT1 said:

> What I did get out of the workshop is that you can take those one-time things, those things that I do sort of every once in a while, and make them more systematic. That's why I think the workshop has helped me, to take the thing that I do throughout the class and make it an actual part of the class, right, and make it more of an expectation of the students as opposed to some technique that I am doing.

CHM2 reported that he transformed his syllabus and added more active learning assignments to his class based on the workshop. As a result, he attempted to cover all (or most) of Bloom's taxonomy in his learning objectives and classroom activities. He talked about Bloom's taxonomy in terms of alignment of assignments and classroom goals. The workshop changed his mindset about how he viewed himself in the classroom, from a scientist who is teaching, to a teacher who is a scientist. He had more recognition that students were looking at him as a model and not just at their classroom notes. He said:

> One of the things that I have modified a little bit, but it's still work in progress, is changing my mindset from being a scientist who is in the classroom to a teacher who is a scientist, recognizing that students are looking to me for more than just notes … you know, thinking about the taxonomy in order to be able to really address all six areas, I have to approach it differently than an old straight lecture.

BIO1 reported adding some model assignments to improve students' interactions and made some changes to her syllabus as a result of the workshop. Although she had used process oriented guided inquiry learning before, the workshop provided ideas about learning goals and how to effectively use them. She indicated that the transformed learning goals provided support for developing activities and working with the teaching assistant, who also happened to attend the workshop. The workshop provided more confidence in facilitating activities that she already had in mind, but was a little bit reluctant to undertake that road. Furthermore, the FLC was also a great experience for this faculty member because participants shared different experiences in their classrooms. When asked what prompted her to change, she answered:

> Well, this WIDER grant thing … I had similar things, so I knew the premise of it, but it got me to thinking at a better level of trying to pick a couple of … Starting with the learning goals, I mean stuff that I know how to do but it really forced me to fit into it starting with the learning

goals for the class and really thinking about what those really are. And then trying to come up with a couple of activities.

CHM3 reported interspersing his classes between lectures and group activities. He emphasized "student-directed learning" and de-emphasized 'the role of flow of knowledge." Facing pushback from some students who were uncomfortable with having more responsibility, he is still determined to keep on improving because it is worth it. PHY1 believed that working with other faculty members to effect changes in their classrooms is an important experience because participants can affirm each other's practices through a series of meetings. She did not change much in her classroom practice because she has been familiar with active teaching practices before the workshop.

Responses from the interview activities demonstrate that the workshop had an influence on faculty teaching practice both in the classroom and how they prepared for teaching. In summary, the workshop provided confidence to the instructors as they tried to increase students' responsibility in their classrooms, through appropriate learning objectives and activities.

Discussion

In this study, we explored the college faculty's use of active instructional approaches and what influenced their approach to teaching. We used the RTOP to investigate their teaching practices and interviews to determine what influenced their teaching approaches. We further investigated the influence of a PD with a FLC on their instructional approach. According to GSNA (2016), the RTOP cut-off scores are described as follows in Table 9:

Table 9

RTOP Rating Criteria According to GSNA (2016)

Type of Teaching	RTOP Score
Traditional/Teacher centered	0-30
Transitional/Teacher-guided	31-45
Transitional/student-influenced	46-60
Reformed/student-centered	61-100

Our results show that most of the faculty members used reformed teaching approaches. The mean RTOP score on our observations (63.5) is consistent with the mean score of faculty who went through the Arizona Collaborative for Excellence in the Preparation of Teachers (ACEPT) workshop training (Piburn & Sawada, 2000), thus indicating that faculty in this study generally used a reformed teaching approach.

When we compared the two groups of faculty, the PD group (RTOP score = 70.4) used reformed or student-centered approaches while the comparison group (RTOP score 55.3) used a transitional/student-influenced approach based on a GSNA (2016) rating. The comparison tests have shown that faculty who participated in the PD with a FLC had significantly higher RTOP scores than their counterparts who did not participate. Further analysis showed the workshop participants performed better on all sections of the RTOP. Possibly, among other experiences, participation in the workshop may have influenced their teaching approach. From our observations, the faculty involved in the workshop asked more probing questions, required students to use prior knowledge as suggested

by the National Research Council (2000), and provided more opportunities for students to work on their problems than their counterparts. These are issues addressed in the workshop. Further, the workshop addressed teaching for deeper learning, where students are encouraged to take charge of their learning to understand the material better (Haitte, 2012; Lombard, 2018; Marton & Saljo, 1976). This type of learning involves making hypotheses, discussing ideas, and making reflections. More of these characteristics were observed in classes taught by faculty who attended the workshop and were involved in the FLC than in classes of non-participating faculty. Brydges et al. (2013) also observed similar changes after a PD with a FLC in that several faculty members adopted non-traditional teaching styles or planned to adopt them. This also agreed with results by Beach and Cox (2009) and Pelletreau et al. (2018). Further, Cox (2018) has indicated that faculty involved in PD with a FLC generally use active learning techniques, which were observed in this study. The faculty in the control group either used direct lecture or some group discussions mainly to solve textbook problems. Students were not generally encouraged to make predictions, hypotheses, or reflect upon their learning.

In addition, although the faculty in the PD group performed better in all sections of the RTOP, the biggest difference was observed in the procedural knowledge, seconded by student-teacher interactions. This makes sense because the purpose of the workshop, and the FLC that followed, was to change the faculty's mindset about teaching and to guide them towards reformed approaches to teaching. Procedural knowledge focuses on making predictions, engaging in thought-provoking activities, challenging ideas, and reflecting upon the learning process. These skills are important in reformed approaches to learning. Specifically, the workshop addressed these issues by encouraging faculty's use of higher-order objectives, assessments that addressed these objectives, and the classroom activities that would align with the objectives and the assessments. Franklin and Chapman (2012) made a similar observation of physics faculty in workshop classrooms. The smallest difference was observed in propositional knowledge. As Franklin and Chapman (2012) discerned, understanding the fundamental concepts of the lesson is valued in both traditional lecture classes and reformed classes. Therefore, the smallest observed difference in this category makes sense.

A 2.5-point increase in faculty RTOP scores was observed between the first and last observation to determine the impact of the FLC. Although this difference is not significant, it shows that the faculty members did not slide back from using reformed approaches to teaching throughout the semester. This makes a lot of sense when this 2.5-point increase is compared to a 4.3-point decrease in RTOP scores for the comparison group between the first and last observation. It is possible for faculty to slide back and teach using direct lecture, especially towards the end of the semester, when there is so much to cover with very little time to spare. In this case, it is possible that the FLC might have put faculty in the PD group on their toes on using reformed teaching since they already had high RTOP scores at the beginning. Furthermore, the first data set was collected four weeks into the semester when the first FLC meeting was already conducted. Therefore, the first observation may have already reflected the influence of the FLC.

Interviews with the faculty indicated that the workshop helped them to make changes in their teaching practice. These changes included formulating objectives that facilitate more student involvement in the learning process, objectives that aligned with Bloom's taxonomy, and creating activities that aligned with these objectives. Further, the participants reported that the workshop enhanced their confidence in designing and facilitating active learning environments. This may help explain why these participants had higher RTOP scores than the comparison group. Our classroom observations indicated that the PD participants designed their lessons to enhance students' participation. These included providing varied learning experiences, encouraging students to make predictions, encouraging group discussions and communications among students, and encouraging meaningful hands-on experiences and students' reflections. Results from this study concur with those reported by Genc and Ogan-Bekiroglu (2004) who found that teachers who participate in some sort

of PD or training are likely to use student-centered learning. In addition, Tinnell et al. (2019) found that faculty used reformed teaching approaches, such as collaborative learning, after undergoing PD with a FLC. Further, Viskupic et al. (2019) observed that faculty who undergo at least 24 hours of PD are likely to change their teaching approach. In the case of this research, faculty underwent four days of PD, which might have brought about meaningful changes in their teaching approaches.

Limitations of the Study

Three limitations that may affect the results in this study consist of (1) a lack of pre-workshop results for all observed faculty members, (2) a small sample size, and (3) a selection effect. In the first case, not all faculty participating were observed before the workshop. This limits the extent to which we can associate potential changes in teaching practice with participation in the project. However, data from five faculty members who were observed before and after the workshop showed a significant improvement in teaching practice. The second limitation is that the sample size is small, especially for the five members who had both PD and comparison data, which may also affect the results and interpretation of the data. In the third case, there may be a selection effect in that the faculty who chose to go through the faculty development may have been primed to incorporate student-centered teaching even if they had not done so before.

Conclusions and Recommendations

Results from this study have shown that the faculty observed at this liberal arts college have transitional teaching approaches. Further, the workshop participants had higher RTOP scores than the comparison group. These results show that there is potential for further improvement in the use of reformed teaching practices with more PD activities. Studies by Hauck (2012) and Viskupic et al. (2019) showed that faculty who frequently attend PD activities are likely to use reformed teaching practices. Therefore, we propose more PD activities and an expansion of these activities to more faculty at this and other colleges.

This research has significance to the scientific community because we have added to the knowledge about how PD with a FLC influences instructors' use of learner-centered approaches. Results from this study will, therefore, inform science educators and administrators on the importance of appropriate PD for effective undergraduate education. Further, determining what influences faculty's teaching approach is crucial to designing appropriate PD activities that either address any gaps in the teaching practice or reinforce the existing practices.

Lloyd M. Mataka (lmmataka@lcsc.edu) is an Associate Professor of Chemistry at Lewis-Clark State College. He received his MSc from the University of Malawi and PhD from Western Michigan University. He has been involved with both pedagogy and chemistry research and has published in reputable journals.

Jon C. Saderholm (saderholmj@berea.edu) taught high school science and mathematics for 17 years before leaving the classroom to pursue his Ph.D. in curriculum and instruction. He returned to Berea College, his alma mater, in 2007 to prepare future STEM teachers. His research involves the relationship between teacher preparation and professional development. He also directs Berea's STEM outreach center.

Tracy T. Hodge (hodget@berea.edu) is an Associate Professor of Physics and Chair of the Division of Natural Sciences, Nursing, and Mathematics at Berea College. Her current faculty development interests are centered on diversity, equity, and inclusion in the STEM classroom. Her research focuses on ground-based observations of young variable stars.

References

Beach, A., & Cox, M. (2009). The impact of faculty learning communities on teaching and learning. *Learning Communities Journal*, *1*(1), 7–27.

Bilal, A., Guraya, S. Y., & Chen, S. (2019). The impact and effectiveness of faculty development program in fostering the faculty's knowledge, skills, and professional competence: A systematic review and meta-analysis. *Saudi Journal of Biological Sciences, 26*(4):688–97. https://doi.org/10.1016/j.sjbs.2017.10.024.

Brydges, S., Chilukuri, L., Cook, G., Feeley, M., Herbst, M., Tour, E., & Van Den Endel, L. (2013). Building a faculty learning community at a research university. *Currents in Teaching and Learning*, *5*(1&2), 17–35.

Chi, M. T. H. (2009). Active-constructive-interactive: A conceptual framework for differentiating learning activities. *Topics in Cognitive Science*, *1*, 73–105.

Cox, M. D. (2004). Introduction to faculty learning communities. In M. D. Cox & L. Richlin (Eds.), *Building faculty learning communities* (pp. 5-23). Jossey-Bass.

Cox, M. D. (2018). Building and enhancing honors programs through faculty learning communities. In J. Ford & J. Zubizarreta (Eds.), *Breaking barriers in teaching and learning* (pp. 81-113). University of Nebraska-Lincoln Digital Commons.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative, and mixed methods approaches* (2nd ed.). Sage Publications.

Fink, D. L. (2003). *A self-directed guide to designing courses for significant learning.* Jossey-Bass.

Franklin, S. V., & Chapman, T. (2012). Diversity of faculty practice in workshop classrooms. *Physics Education Research Conference 2012 Part of the PER Conference series, 1513*, 130-133.

Freeman, S. Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., & Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *PNAS, 111*(23), 8410-8415.

Genc, E., & Ogan-Bekiroglu, F. (2004). *Patterns in teaching styles of science teachers in Florida and factors influencing their preferences.* Florida State University. http://files.eric.ed.gov/fulltext/ED490781.pdf

Georgiou, D., Mok, S. Y., Fischer, F., Vermunt, J. D., & Seidel, T. (2020). Evidence-based practice in teacher education: The mediating role of self-efficacy beliefs and practical knowledge. *Frontiers in Education, 5,* 1-15. https://www.doi.org/10.3389/feduc.2020.559192

Gijbels, D., van de Watering, G., Dochy, F., & Van den Bossche, P. (2006). New learning environments and constructivism: The students' perspective. *Instructional Science*, *34*(3), 213–226.

Harris, M., & Cullen, R. (2008). Learner-centered leadership: An agenda for action. *Innovation in Higher Education*, *33*(1), 21–28.

Hauck, N. (2012). *Effects of sustained teacher professional development on the classroom science instruction of elementary school teachers. doctoral dissertation* (1343.) [Doctoral dissertation, Utah State University]. Digital Commons. https://digitalcommons.usu.edu/cgi/viewcontent.cgi?article=2324&context=etd

Hirst, R. A., Anderson, K. L., Packard, B. W., Liotta, L. J., Bleakley, B. H., Lombardi, P. J., & Burkholde, K. C. (2021). Faculty learning at the individual and group level: A multi-year

analysis of an interdisciplinary science faculty learning community focused on inclusive teaching and mentoring. *Journal of College Science Teaching, 50*(6), 20-30.

Hopkins, J. E. (2020). *The impact of college faculty perceptions of professional development on course design, teaching, and assessment practices* [Doctoral dissertation, Regent University]. Proquest.

Hott, B. L., & Tietjen-Smith, T. (2018). The professional development needs of tenure track faculty at a regional university. *Research in Higher Education Journal, 35*, 1-12.

Kranzfelder, P., Lo, A. T., Melloy, M. P., Walker, L. E., & Warfa, A. M. (2019). Instructional practices in reformed undergraduate STEM learning environments: A study of instructor and student behaviors in biology courses. *International Journal of Science Education, 41*(14), 1944-1961.

Lakshmanan, A., Heath, B. P., Perlmutter, A., & Elder, M. (2011). The impact of science content and professional learning communities on science teaching efficacy and standards-based instruction. *Journal of Research in Science Teaching, 48*, 534–551.

Lund, T. J., Pilarz, M., Velasco, J. B., Chakraverty, D., Rosploch, K., Undersander, M., & Stains M. (2015). The best of both worlds: Building on the COPUS and RTOP observation protocols to easily and reliably measure various levels of reformed instructional practice. *CBE—Life Sciences Education 14*(2), 1-12. https://www. doi.org/10.1187/cbe.14-10-0168

MacIsaac, D., & Falconer, K. (2002). Reforming physics instruction via RTOP. *The Physics Teacher, 40*, 479-485

Masters, G. N. (2018, August). *The role of evidence in teaching and learning* [Conference presentation]. ACEReSearch Online Conference 2018 - Teaching practices that make a difference: Insights from research. https://research.acer.edu.au/research_conference/RC2018/13august/2

National Research Council. (2000). *How people learn: Brain, mind, experience and school.* National Academy Press.

Orzech, M. J. (2021). Planning, implementing, and assessing an over faculty learning community: A facilitator's lens. *The Open/Technology in Education, Society, and Scholarship Association Journal, 1*(1), 1–13. https://doi.org/10.18357/otessaj.2021.1.1.6.

Park, S., Jang, J.-Y., Chen, Y.-C., & Jung, J. (2011). Is pedagogical content knowledge (PCK) necessary for reformed science teaching? Evidence from an empirical study. *Research in Science Education, 41*(2), 245–260. https://www.doi.org/10.1007/s11165-009-9163-8.

Pelletreau, K. N., Knight, J. K., Lemons, P. P., McCourt, J. S., Merrill, J. E., Nehm, R. H., Prevost, L. B., Urban-Lurain, M., & Smith, M. K. (2018). A faculty professional development model that improves student learning, encourages active-learning instructional practices, and works for faculty at multiple institutions. *CBE—Life Sciences Education, 15*(2), 1-11. https://www. doi.org/10.1187/cbe.17-12-0260

Pesce, J. R. (2015). *Professional development for teaching in higher education: faculty perceptions and attitudes* [Doctoral dissertation, Boston College]. Boston College University Library, eScholarship. https://dlib.bc.edu/islandora/object/bc-ir:104134.

Piburn, M., & Sawada, D. (2000). *Reformed teaching observation protocol (RTOP): Reference Manual* ACEPT Technical Report, No. IN00-3, 1–41. https://mathed.asu.edu/instruments/rtop/RTOP_Reference_Manual.pdf

Richlin, L, & Essington, A. (2004). Overview of faculty learning communities. In M. D. Cox & L. Richlin (Eds.), *Building faculty learning communities: New directions for teaching and learning* (pp. 25–39). Jossey-Bass.

Ruiz-Primo, M. A., & Furtak, E. M. (2007). Exploring teachers' informal formative assessment practices and students' understanding in the context of scientific inquiry. *Journal of Research in Science Teaching, 44*(1), 57–84.

Simonson, S. R. (2019). *POGIL: An Introduction to process oriented guided inquiry learning for those who wish*

to empower learners. Stylus Publishing, LLC.

Schmitz, S., Ebersole, T. E., Morrell, C., & Parker, C. (2018). Measuring the effectiveness of community college faculty learning communities on student performance. *Learning Communities Journal, 10*, 87-108.

Slavin, R. E. (2002). Evidence-based education policies: Transforming educational practice and research. *Educational researcher, 31*(7), 15-21. https://doi.org/10.3102/0013189X031007015

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches.* Sage Publications.

The Geological Society of North America (GSNA). (2016). *Classroom observation project: Understanding and improving our teaching using the reformed teaching observation protocol (RTOP).* National Association of Geoscience Teachers. https://serc.carleton.edu/NAGTWorkshops/certop/vignettes.html

Tinnell, T. L., Ralston, P. A. S., Tretter, T. R., & Mills, M. E. (2019). Sustaining pedagogical change via faculty learning community. *International Journal of Stem Education, 6*(26). https://doi.org/10.1186/s40594-019-0180-5

Viskupic, K., Ryker, K., Teasdale, R. Manduca, C. Iverson, E., Farthing, D., Bruckner, M. Z., & McFadden, R. (2019). Classroom observations indicate the positive impacts of discipline-based professional development. *Journal for STEM Education Research*, *2,* 201–228. https://doi.org/10.1007/s41979-019-00015-w.

Weinstein, D. F. (2002). Duty hours for resident physicians— tough choices for teaching hospitals. *New England Journal of Medicine*, *347*, 1275–1278.

Reasoning at the Intersection of Science and Mathematics in Elementary School: A Systematic Literature Review

Liam Quinn McCashin
University of Alberta

P. Janelle McFeetors
University of Alberta

Mijung Kim
University of Alberta

ABSTRACT

Despite efforts to integrate science and mathematics learning in elementary school through carefully designed activities, students' cognitive processes remain relatively untapped as a possible place of intersection. We believe *reasoning* is a productive co-curricular concept that could lead to meaningful integration. We conducted a systematic literature review of empirical research published over the past 20 years on students' reasoning in both science education and mathematics education. Articles were summarized and examined for their: (1) methodological approach and experimental design; (2) social dimension in the classroom; (3) definition of reasoning and associated structures; and, (4) evidence of students' engagement in reasoning. For each theme, relationships between scientific and mathematical reasoning research were examined for the purposes of finding intersections and discrepancies between the two subject areas. As a result, we suggest the term *STEM reasoning* that embodies the core skills and thought processes across both subject areas to provide an authentic approach to integrating elementary school science and mathematics learning.

Keywords: mathematical reasoning; scientific reasoning; interactions of reasoning; elementary education

Introduction

Reasoning is a complex skill that is critical to problem solving in students of all ages and across a wide range of subject areas. In the context of solving a problem, reasoning stands as the "power that enables us to move from one step to the next" (Wu, 2009, p. 14). Reasoning is what happens in the time between the end of a question and the resulting answers and discussion that follow. Purzer et al. (2015) identified "reasoning processes such as analogical reasoning as navigational devices to bridge the gap between problem and solution" (p. 2) as mental tools and thought systems in place to solve problems within an interdisciplinary context of science, technology, engineering, and mathematics (STEM) education. We understand STEM education to be the integration of two or more of the disciplinary areas (English, 2016) -- our areas of interest are in science and mathematics. Devoid of reasoning, however, mental tools and thinking processes used in problem solving in science and mathematics are merely a set of tools in a toolbox. The ability to reason serves as our neural tour guide. Reasoning functions as the mechanism that helps students determine which thought pathway

to follow - to pick out which pieces of information are critical, how they need to be dealt with, and where to go next.

This reasoning process is an essential skill in the intellectual development of a classroom learner (Alexander, 2017). While forms of reasoning are present in all subject area classrooms, there is a special emphasis on developing reasoning processes in STEM education (Johnson et al., 2021). In fact, reasoning is present to some degree at almost every level of science curricula (Alberta Education, 1996; Next Generation Science Standards, 2013) and mathematics curricula (Alberta Education, 2007/2016; Common Core State Standards Initiative, 2010). Across both subject areas, knowledge-based curricular outcomes are easier to identify and evaluate, while the complex reasoning operations that guide students to their given responses are often implied in curricular outcomes and more difficult to analyze directly. Even for experienced teachers, recognizing and observing the visual signs of reasoning is a process that takes much concentration, awareness, and training (Blanton & Kaput, 2005; Mata-Pereira & da Ponte, 2017; Mercer et al., 1999; Shin, 2020). As science and mathematics educators, we have found that even comparison between science and mathematics curricula can be difficult given the semantic and procedural differences in science and mathematics education. There is also the assumption that since curriculum consumers have a well-formed understanding of reasoning, it is often not explicitly defined or demarcated.

Reasoning in Science Education

Our theoretical perspective on scientific reasoning focuses on the understanding and practice of the relationships between claims and evidence. Encountering a question or puzzling circumstance, children need to examine the context of the question and situation, look for supporting and contrasting evidence, develop, accept, and articulate claims, and justify why the coordination of available claims and evidence has led them to accept certain theories and reject others (Bower, 2009; Varma, 2014). In the problem solving process, scientific reasoning occurs through questions, claims, and evidence to reach conclusions that are critical. Kuhn (1989) explains, "the heart of scientific thinking is the coordination of theories and evidence" (p. 674). Scientific claims or theories stand in relation to actual or potential bodies of evidence, yet, often children make or accept scientific claims not based on evidence but on personal ideas or beliefs, e.g., confirmation bias (Kuhn, 1991; Kuhn & Pearsall, 2000). Further, in "hypothetico-deductive reasoning in which children's working memory accesses and sustains hypotheses from associative memory to be tested and then actively seeks predictions and evidence that follow" (Lawson, 2004, p. 333), it is critical for children to understand how a hypothesis could be tested, evaluated, and sustained (or rejected) based on data sets and evidence that are available. The understanding of correlations between claims and evidence is essential in the scientific reasoning process. Children learn to make, evaluate, and justify claims with evidence as a core value of scientific reasoning. Such scientific reasoning enables students to be critical thinkers as they learn science in school and as they participate in society (Vieira & Tenreiro-Vieira, 2016).

National and international organizations have produced standards emphasizing that critical thinking and induction in problem solving processes are skills that lie at the core of scientific inquiry in science education (National Research Council, 1996; Organisation for Economic Co-operation and Development, 2013). The National Research Council (1996) acknowledges that scientific understanding develops "by combining scientific knowledge with reasoning and thinking skills" (p. 2) and the standards expect students to engage in scientific reasoning as both a process and outcome of learning. Additionally, reasoning in science education has been discussed in complex STEM problem solving contexts by integrating mathematics, causal reasoning, evidence-based evaluation, and argumentation in STEM contexts (Next Generation Science Standards, 2013). As problem solving contexts are complex and integrated with science, mathematics, and technology domains in today's world, reasoning needs to be understood and taught in complex contexts where disciplinary

knowledge, skills, and values are intertwined for decision making and justification. Scientific reasoning cannot be independently processed to solve everyday problems.

Reasoning in Mathematics Education

Our theoretical perspective on mathematical reasoning and experiences that promote students' development of mathematical reasoning is informed by seminal publications in mathematics education and established standards within curriculum frameworks which include reasoning as an essential component. Mathematics education research builds on Polya's (1954) demonstrative and plausible reasoning, with ensuing elaborations of the various forms of reasoning such as transformational (Simon, 1996), metaphoric and analogic (English, 1997), imagistic (Thompson, 1996), and indirect (Brown, 2018). Davydov (1990) determined that children reason as they generalize about concepts and relationships which in turn "enables students to think systematically and to apply rules in concrete situations" (Venenciano & Heck, 2016, p. 23). Brodie (2010) defined reasoning as a way to "develop lines of thinking or argument … to convince … to solve … or to integrate a number of ideas" (p. 7). When students reason "about and with the objects of mathematics" (Brodie, 2010, p. 7), they invoke processes like conjecturing, investigating, representing, analyzing, justifying, refuting, generalizing, and convincing (Mason et al., 2010). Reasoning culminates in students' construction of proofs as an essential feature of the discipline of mathematics (Hanna, 1983; Lakatos, 1976). Mathematical reasoning is critical to the growth of mathematically proficient students (Lannin et al., 2011; White et al., 1998) evidenced in heightened achievement (Nunes et al., 2007). Mathematical reasoning is both a way to learn mathematics by actively constructing understanding and a way to be mathematical as a capacity that is developed over time through explicit use in mathematical contexts.

National organizations have produced standards emphasizing mathematical reasoning. The National Council of Teachers of Mathematics (NCTM, 2000) identified reasoning as a fundamental process for learning mathematics, where students are thinking analytically and systematically as they investigate mathematical phenomena, highlighting the role of conjecturing and justifying. Students are expected to move from informal reasoning to reasoning inductively toward conventional deductive proofs. Subsequently, adaptive reasoning was described as "the glue that holds everything together" (Kilpatrick et al., 2001) within mathematical proficiency. These two documents influenced the "Standards for Mathematical Practice' (Common Core State Standards Initiative, 2010) where five out of eight standards involve reasoning as noticing and expressing patterns and thinking with "logical progression" (p. 6). NCTM has further elaborated on teaching practices which all incorporate and enhance students' reasoning, along with specifically emphasizing the use of "tasks that promote reasoning" (2014, pp. 17-24) and identified the inclusion of reasoning in early childhood and elementary classrooms as transformational in students' development of "deep mathematical understanding" (2020, p. 9).

Exploring Reasoning as the Intersection

As has been demonstrated above, reasoning is often theorized or reasoning standards for student learning are created within domain-specific sites, either situated singularly within science or mathematics. Even in research about reasoning that transcends the boundaries of science and mathematics, one content area is often emphasized over the other. For example, Wasserman and Rossi (2015) used mathematics problems to investigate reasoning approaches of science and mathematics teachers in the context of STEM education. Limited scholarship exists that addresses the possibility of reasoning impacting the relationship between learning in science and mathematics. Research that investigates reasoning within the context of science and mathematics learning simultaneously often focuses on a type reasoning within a specific domain, like model-based reasoning (Lehrer & Schauble,

2000), proportional reasoning (Dole et al., 2012) or socioscientific reasoning (Owens et al., 2019). More promising, some researchers have recognized reasoning as a cognitive process that could cross the disciplinary boundaries of science and mathematics. Alexander (2017) posited that relational reasoning is "a foundational capacity" (p. 8) that could contribute to STEM learning across associated disciplines. Hwang et al. (2020) analyzed reasoning in Trends in International Mathematics and Science Study data and found that there is evidence for reasoning as "the same cognitive practices were utilized in science and mathematics" (p. 534). These studies open up the conversation for reasoning to exist at the intersection of science and mathematics learning within a STEM approach.

Though subject-specific curricula are written independently, the skills and mental processes we aim to develop in science and mathematics are not dissimilar. Reasoning skills are often distinguished in curriculum with unique labels and slightly variant definitions, but the thought processes they draw on are related and overlapping (Pisesky et al., 2018). Especially in an elementary school environment, the potential for teachers to take advantage of the intersection of science and mathematics through reasoning is immense but often not thoroughly utilized. Even if cross-curricular work is attempted, there is often a focus subject and extraneous outcomes from other subject areas used in superficial ways (Babb et al., 2016). Creation of a simple papier-mache science cell, for example, may be the purpose of an activity, while an estimation of the volume of the object could be layered on top. While this type of interdisciplinary project has definite value, it does not integrate the thought processes of science and mathematics. Rather, it takes a mathematics skill and requires its application on top of an existing science project framework that is not necessarily authentic to mathematics.

The aim of this theoretical study was to better understand the convergence points where scientific reasoning and mathematical reasoning intertwine. A primary literature review was conducted to explore how reasoning is researched as a co-curricular concept in the elementary classroom. We will first evaluate science and mathematics education research on reasoning independently to explore their unique methodologies and conceptualizations of reasoning. Through cross-examination of the findings on reasoning in science and mathematics subject areas, we will note the uniqueness and highlight commonalities between the two disciplines. We focus on the question: *What are the critical intersection points offered by reasoning in science and mathematics?*

Methods

We used a systematic review approach (Cooper, 2015) to create our final subset of literature for this study. This systematic process involved the following procedure: (1) established the required search databases and determined a list of appropriate search terms; (2) ran a search and consolidated the resulting articles; and, (3) used an established screening process to narrow down results to a focused subset of the initial article pool. Our final article pool was determined through several iterative rounds of term selection and literature searching. Our primary inclusion criteria were determined through the following conditions. First, literature must have been published between January 2000 and August 2020. Second, included articles needed to be peer reviewed for an academic research journal. Third, articles had to be focused on elementary education. This definition varies from country to country, but generally ranges from preschool or kindergarten to around grades 5 or 6 (depending on the country's range of elementary grades). Studies that contained large grade ranges (i.e., grades 4-12) were eliminated unless there was a clear way to separate out elementary education data.

Two searches were completed per database, one for science and one for mathematics. This was required because of the synonymous (yet differing) language used around reasoning in science and mathematics education. For science, the search terms used were "science" and either "scientific reasoning" or "logical reasoning." For mathematics, the search terms used were "math" and either "mathematical reasoning" or "logical reasoning." Of note, "reasoning (logic)" was a common keyword

that was missed initially and, when employed, produced a large quantity of viable articles for inclusion. Filters were applied to fulfill the peer review requirement and limit dates to the preferred date range. Finally, to specify results to our desired age range, we added search parameters of "elementary," "K-6," "primary" or "early childhood." Of note, we decided to bound our searches with how "reasoning" specifically is used in science and mathematics education because we see it as a broad, integral, cognitive process. We also recognize that terms like "argumentation" are whole domains that are very well-researched in science education (Erduran et al., 2015; Nielsen, 2013), but may be less so in mathematics education (c.f., Staples & Cavanna, 2021; Staples & Newton, 2016), and would have produced unbalanced results between the two disciplines.

These search criteria were applied separately to education databases determined after consultation with an elementary education research librarian. We narrowed our search to two primary engines: ERIC (EBSCOhost) and Proquest. Using these search engines, the following databases were accessed: ERIC, Education Research Complete, Proquest Education Databases, and the Canadian Business & Current Affairs Database. Identical science and mathematics searches were completed using both engines, resulting in 647 total results for science and 542 total results for mathematics. Both lists were exported to Refworks and duplicate articles were removed, leaving 469 unique science education articles and 494 unique mathematics education articles.

A comprehensive screening process was required to eliminate extraneous articles that mention search terms, but are not related to our student focus. Our results were narrowed to short lists for science and mathematics using the following exclusion criteria: (1) Articles needed to be focused on empirical research conducted in a classroom or similar environment. Therefore, literature reviews were excluded. (2) Articles must have been written for academic research journals. Therefore, professional journals or other non-academic publications (even if peer-reviewed) were removed. (3) Articles were required to focus on the reasoning ability of students themselves. Therefore, teacher responses to reasoning, evaluation of existing reasoning skill metrics or examination of reasoning in teachers or preservice teachers themselves were excluded. We did not impose *a priori* definitions of reasoning as there is no single agreed-upon definition in either science education or mathematics education; rather, our aim was to observe how reasoning is conceptualized and operationalized within the empirical literature. Execution of these criteria left our research team with a shortlist of 44 science articles and 43 mathematics articles. A final in-depth screening was completed collaboratively by a team of two faculty members and one PhD student to select the strongest articles that focused most directly on students' reasoning capability. We scanned abstracts and research findings to eliminate any articles that passed through initial filters but did not actually meet our criteria. Many of these new exclusions were eliminated because, though they involved elementary students, they did not focus on student reasoning specifically, such as concentrating on validating a measurement parameter for a diagnostic test or teacher professional development. This final filter left us with 20 science and 21 mathematics articles to proceed with for our analysis.

The final pool of 41 articles was analyzed for critical elements with respect to reasoning by the whole research team. The articles were initially sorted according to their subject area to support immersion in the orientation to students' reasoning within the specific subject area. Then, the following process was used for both science and mathematics articles. The research question or purpose was highlighted to frame analysis of the findings. Research methods were noted because the authors' findings and research consumers' interests rely on these categorizations, including methodological approach, design of the classroom activities (intervention) or measurement tool (descriptive), and grade level. Additionally, the findings of each study were summarized along with salient quotes that represented both students' reasoning and the researchers' interpretations of reasoning. Making brief notes of the findings enabled preliminary analysis of connections across the articles. The research team met to discuss their collective observations that led to emerging themes and issues addressed by articles across science and mathematics. The emerging themes developed out

of both a broad reading of all the articles and a first round of open coding to allow salient ideas to arise in our analysis rather than applying an analytic framework. This emergent process enabled us to discuss elements authors of each article emphasized in their reports.

The synthesis discussion led to a second in-depth round of analysis. In re-reading the articles, we recorded the subject-specific context (e.g., fractions, shape attributes, forces, density), the operational definition of reasoning, the conventional form of reasoning (e.g., inductive, deductive, metaphoric, imagistic), processes of reasoning (e.g., conjecture, hypothesize, generalize, convince), and the social context (e.g., individual students, whole class discussions). The second round of analysis provided finely nuanced insights into the similarities and differences of how reasoning is discussed and researched in science education and in mathematics education. In the second round, we reviewed and compared key ideas that emerged from the first round reading and compared them to develop certain categories to discuss similarities and differences in research on students' science and mathematics reasoning. We focused on what research approach or methods researchers chose to examine students' reasoning, how they perceived and defined reasoning, what types of reasoning researchers focused on, in what context reasoning was researched (e.g., individual cognitive process or social cognitive process, or independent or complex process). We discussed the ideas from the first round of reading and possible categories back and forth to saturate the approach and meanings of students' reasoning in research (Corbin & Strauss, 2008). Four categories were developed to discuss the findings; researchers' approach to examining reasoning, social dimensions, definition of reasoning, and students' engagement in reasoning. We discussed the analysis and developed shared understandings of the categories reported in the results below. Appendices A and B contain the reference lists for science and mathematics education articles, respectively.

Results

In this section, we explore our understanding of how children's scientific reasoning and mathematical reasoning are discussed and researched in the respective subject areas developed through our systematic review of literature published in the past 20 years. Typically, systematic literature reviews in reasoning have been conducted independently in subject areas (for science education, see Engelmann et al. (2016); for mathematics education, see Hjelte et al. (2020) or Jeannotte & Kieran (2017)), but through our exploration we found that critical discussions of interdisciplinarity only occurred as we framed understanding one subject area's conceptualization of reasoning in light of the other area's conceptualization. In this way, ideas and approaches that cursorily appeared to be differences between the disciplines could be better understood as similarities, when viewed through research into children's learning. The differences that remained were generative in enriching both scientific and mathematical reasoning processes rather than inhibiting children's interdisciplinary learning.

We developed tables to provide a structured overview of the results. Because of the substantial list of articles and the two subject areas, each table summarizes the research in one subject area. Table 1 displays research in science education. While maintaining a similar structure for Table 2, which displays the mathematics education research, the subcategories for the "Type of Reasoning" classification differ slightly because of its treatment in the research literature. To succinctly refer to articles in the ensuing results, we have labelled each article with "S" for science or "M" for mathematics, followed by a numeral.

The two overview tables are organized by the categories of general findings discussed below. The four categories include:
1. Researchers' approaches to investigating children's reasoning empirically;
2. Researchers' analysis of a range of social dimensions observed during children's engagement in reasoning;

3. Researchers' working definitions of reasoning, inclusive of domain-general and domain-specific approaches along with conventional forms to structure reasoning statements; and,
4. Researchers' observations and interpretations of children's ways of enacting reasoning in their learning.

In reporting on each of these categories, we first provide an analysis of each of the subject area's findings and then synthesize by interpreting similarities and differences across science education and mathematics education. Our findings provide a foundation to discuss a response to our research question to suggest reasoning as a critical intersection point in integrating science and mathematics learning.

Approaches to Research in Children's Reasoning

Science Education Research

The science articles in the final literature pool (20 total) were diverse in their age range, research methodology, and evidence for reasoning. Target groups varied from pre-kindergarten to grade 7 (using local definitions of elementary education) and were split evenly between early elementary (PK-3, 11 articles) and later elementary learners (grades 4-7, 12 articles) with three articles spanning the full spectrum of elementary learners (S6, S13, S17). Many science articles explored reasoning through the lens of informal play or exploration, often in three- to five-year-old preschool children (S1, S3, S9, S16). Though formal definitions of reasoning are used in the research as they relate to eventual curriculum, these studies access children before they have been formally taught reasoning processes in a classroom setting.

Research in the science group focused on quantitative methodologies, demonstrated through 13 quantitative, five qualitative, and two mixed methods studies. One science article stated that a mixed methods approach was employed in the study; yet, data was presented solely in quantitative manners (S2). For the purposes of this analysis, we have taken the researchers' intentions into consideration and classified it as a mixed methods approach. Most quantitative studies used a pretest-posttest experimental design. The quantitative studies commonly used a range of standardized and/or validated test items over a temporarily extended research period of a few months to a year. Test questions were used to evaluate student reasoning ability (S1, S2, S9 - S14, S17, S19, S20) through analysis of individual test question completion and grading. The majority of the quantitative experimentation focused either on a reasoning-focused lab or class activity, (S1, S3, S6, S9, S16, S19, S20) or a pre-designed individual task or examination (S2, S10, S11, S12, S13, S14, S17) as their treatment. One study appears to be an outlier, as it examined the effect of teacher professional development on student reasoning in the classroom (S4).

Qualitative research in this group used either interviews (S5, S7, S18) or audiovisual recordings (S4, S5, S8, S15) as their means of investigation. Evidence of student reasoning for both data collection designs was qualitatively coded and interpreted. Student statements of reasoning were largely absent in the science reasoning research, though some articles do contain concrete evidence of student voice (S4, S5, S8, S15, S18). Within these articles, the researchers contextualized their investigation of students' reasoning within students' interactions related to actively solving a problem. Data was collected either during the problem solving or in reflective interviews, which resulted in capturing students' expressions of reasoning that were displayed in the articles through extended exchanges.

Mathematics Education Research

With the mathematics articles, age range was similarly diverse, though older elementary learners were emphasized slightly more. There were six articles focused on PK-3 education (M1, M8,

Table 1

Reasoning in Science Education

Article # First Author (Year)	Method			Grade		Social Dimension			Type of Reasoning			Engagement
	Qual	Quan	Mixed	PK-3	4-7	Ind	Small	Class	Process	Concept	Approach	
S1 Bauer (2019)		X		X		X					C	Infer
S2 Chen (2013)			X		X	X			X	X	C E	Conjecture, elaborate, justify, infer, explain, hypothesize, observe, …
S3 Fernbach (2012)		X		X		X					C	Infer, hypothesize, diagnose
S4 Gillies (2013)			X		X			X	X	X	O	Investigate, predict, explain, question, elaborate
S5 Hackling (2015)	X				X			X	X	X	E A	Justify, critique, claim, explain, counterargue, conclude
S6 Hardy (2010)		X		X	X			X	X	X	E	Hypothesize, experiment, justify, generalize, correlate
S7 Hatzinikita (2005)	X				X	X				X	C	Identify, explain
S8 Kim (2019)	X				X		X	X	X	X	E	Analyze, evaluate, justify, explain, claim, argue
S9 Koksal-Tuncer (2018)		X		X		X			X		C	Hypothesize, test, counterargue, refute, revise
S10 Lazonder (2012)		X			X	X				X	C	Hypothesize, experiment, evaluate
S11 Lazonder (2014)		X			X	X				X	C	Predict, justify, hypothesize, infer, question, investigate
S12 Mayer (2014)		X			X	X			X	X	O	Hypothesize, analyze
S13 Osterhaus (2016)		X		X	X	X			X	X	O	Experiment
S14 Osterhaus (2020)		X		X		X			X	X	O	Hypothesize, experiment, analyze
S15 Paparistodemou (2008)	X			X			X		X		E A	Infer, conjecture, generalize, conclude, argue, interpret
S16 Schulz (2007)		X		X		X					C	Explore
S17 Schiefer (2019)		X		X	X	X			X		O	Hypothesize, experiment, theorize, analyze, synthesize
S18 Tytler (2004)	X			X		X			X	X	E	Evaluate, interpret, explore, explain, generalize
S19 Van der Graaf (2015)		X		X		X				X	C	Hypothesize, experiment, evaluate
S20 Van der Graaf (2019)		X			X			X	X	X	C	Hypothesize, experiment, conclude

Note: Qual = qualitative; Quan = quantitative; PK = pre-kindergarten; Ind = individual; NOS = nature of science; C = causal; E = evidence-based; A = argumentation; O = other

Table 2

Reasoning in Mathematics Education

Article # First Author (Year)	Method			Grades		Social Dimension			Type of Reasoning			Engagement
	Qual	Quan	Mixed	PK-3	4-7	Ind	Small	Class	Logic	Content	Form	
M1 Delay (2016)		X		X			X		X		I	Generalize
M2 Depaepe (2007)			X	X				X				
M3 Flegas (2013)			X	X		X			X		D O	Justify, argue, convince, support, explain, organize
M4 Francisco (2010)	X						X	X	X	X	O	Generalize, justify, convince, explain, describe, explore, …
M5 Gurbuz (2016)		X			X	X				X		Justify, explain
M6 Houssart (2006)	X			X			X	X	X			Conjecture, generalize, modify, analyze, discuss, adapt, …
M7 Hughes (2020)		X			X	X				X		Justify, convince, explain, represent, visualize, …
M8 Hunter (2017)	X			X			X	X		X		Justify, convince, explain, reflect, investigate, …
M9 Jurdak (2013)		X			X	X				X		Generalize, justify, explain
M10 Kumpulainen (2003)	X				X		X			X	O	Hypothesize, argue, explain, organize, represent, test, …
M11 McFeetors (2018)	X				X		X		X		I D O	Conjecture, generalize, convince, explain, refute, …
M12 Mercer (2006)			X		X		X	X	X			Hypothesize, discuss, test
M13 Nunes (2007)		X		X		X				X		
M14 Petrovici (2018)		X		X					X		D	
M15 Reid (2002)	X				X		X		X		I D O	Conjecture, generalize, explain, refute, observe, test, …
M16 Saleh (2018)		X			X	X				X		
M17 Sumpter (2015)	X			X			X		X		D O	Argue, represent, conclude, support, challenge, …
M18 Vale (2016)	X			X	X		X	X	X		O	Conjecture, generalize, justify, convince, verify, extend, …
M19 Vandermaas-Peeler (2015)		X		X			X			X		Predict, explain, conclude, explore, evaluate
M20 Wong (2017)		X			X	X			X		D	
M21 Yankelewitz (2010)	X				X		X		X		D O	Hypothesize, justify, convince, argue, analyze, verify, show, …

Note: Qual = qualitative; Quan = quantitative; PK = pre-kindergarten; Ind = individual; I = inductive; D = deductive; O = other

M13, M14, M17, M19), 14 studies looked at grades 4-7 learning, and only one spanned both age groups (M18). The inclusion of early childhood education also saw less prevalence in the mathematics studies, with only three articles (M14, M17, M19) focused on pre-kindergarten years. In contrast, there were ten studies that focused on the last two years of elementary education (M2, M3, M4, M5, M7, M10, M11, M12, M15, M21).

Mathematics articles were evenly distributed between quantitative and qualitative methods. There were nine articles that adopted qualitative research methods (M4, M6, M8, M10, M11, M15, M17, M18, M21), nine used quantitative methods (M1, M5, M7, M9, M13, M14, M16, M19, M20), and three studies used a mixed methods approach (M2, M3, M12). As with science, methodologies varied throughout the mathematics education articles, although a large subset emerged. These studies shared three major characteristics: they used qualitative methodologies (or mixed methods containing a significant qualitative component), relied on audiovisual recording for data collection, and contained many student statements of reasoning (M2, M4, M6, M11, M12, M15, M17, M19, M21). Explicit evidence of reasoning was much more common in the mathematics studies, with some primarily quantitative research even quoting student statements to support their quantitative coding scheme (M3, M5).

Similarities and Differences

Science and mathematics education research differed somewhat in their approach to exploration and evaluation of reasoning ability. Participants' age focused on the lower grades in the science research studies and higher grades in mathematics. This may have been intentional based on the verbalization of reasoning ability that was often of interest in the mathematics group (M3, M4, M6, M10, M11, M15, M17, M18). Similarly, mathematics did not have as much focus on the pre-kindergarten and kindergarten age groups. This may be due to a common focus in the science subset on reasoning in game-like settings or in a play environment (S1, S3, S9, S16). This existed in only two mathematics articles (M14, M17) and one article that focused on science and mathematics (M19).

Methodologically, there was a clear difference in the science subset towards quantitative research (14 quantitative, five qualitative, one mixed), while mathematics had more balanced research approaches (nine quantitative, nine qualitative, three mixed). Science studies tended to focus on evaluation of individual learner reasoning through standardized testing or as a consequence of the evaluation of novel measurement tools. Even in quantitative mathematics experimentation, reasoning item analysis was less of a focus -- articles seemed more interested in correlational factors like intelligence scores, learning disability status, and classroom relationships (M1, M7, M13, M14).

Within the qualitative research, science and mathematics shared a reliance on audiovisual lesson recording and subsequent coding for evidence of reasoning. Interestingly, this did not translate to the presence of student statements of reasoning in the articles' results sections. Concrete evidence of reasoning was present in 14 mathematics articles and only five science articles. This may partially be a consequence of the increased usage of qualitative and mixed methods in the mathematics subset. This may be further evidenced by the usage of student statements of reasoning that were present on the science side. All five studies containing student quotations were all qualitative or mixed analyses and, as a result, had transcript evidence of these reasoning moments (S4, S5, S8, S15, S18).

Social Dimensions of Engagement in Reasoning

Science Education Research

Science articles mainly focused on students' reasoning ability in an individualized manner. Students worked on activities and researchers examined how they used reasoning to solve the

problems. Even if learning occurred in a small group or classroom activity setting (S2, S4, S15, S18), evaluation of student reasoning was still measured through individual testing. In a study by Van der Graaf et al. (2019) (S20), students participated in six inquiry-based lessons and then took pre-, post-, and long-term tests of reasoning skills. Students were engaged in inquiry activities in groups, but their reasoning skills were tested individually. Individual examination occurred for nearly all science articles (exceptions included S5, S6, and S8). Testing was either used as a metric to determine success of an experimental treatment quantitatively (S1-S4, S7, S9-S11, S13, S15, S16, S18, S20) or as a means of item validation for the examination artifact itself (S12, S14, S17, S19). For example, Osterhaus and Koerber (2016) (S14) developed a reasoning testing instrument, the Science Primary School Reasoning Inventory (SPR-I), to measure diverse components of scientific reasoning skills: experimentation, data interpretation, and understanding the nature of science. In this test, students individually worked on problem solving that included inference making based on various conditions.

There were studies (S4, S5, S6, S7, S8, S18) which examined students' reasoning by coding classroom dialogues or conducting interviews. In the Hackling and Sherriff (2015) article (S5), researchers analyzed classroom discourse based on argumentation structure to examine how students provided their reasoning for claims and critiquing and justifying of ideas. In this approach, students' reasoning was understood as a social cognitive process, and was thus examined while students were engaged in social dialogical interactions in classrooms. This approach, however, was uncommon amongst the science subset with only five articles (S4, S5, S7, S8, S18) using a qualitative methodology to analyze group discussions. Hardy et al. (2010)'s article (S6) used collaborative coding of classroom discussions but then translated results into categorical numerical data and analyzed it quantitatively. In these classroom settings, students were engaged in inquiry activities and classroom talk as a whole or small group to discuss their ideas during and for problem solving. Researchers observed and analyzed classroom talk between the teacher and students or students themselves to examine students' reasoning skills.

Mathematics Education Research

Substantial research in mathematics education has resulted in an emphasis on mathematical discourse as an important contributing element to meaningful mathematics learning (Cazden, 2001; Herbel-Eisenmann & Cirillo, 2009; Sfard, 2008). Mathematics articles most commonly used small groups as the unit of analysis for the study. Social engagement was used as a driver for classroom discussion that was then recorded and analyzed for reasoning themes and keywords. The usage of this discussion element was not agreed upon -- most articles used small group or classroom discussion as a tool to elicit reasoning talk (M2, M3, M4, M6, M8, M15), but did not intend to explore the effect of that social dynamic itself. Houssart and Sams (2006) (M6), for instance, detail the importance of group discussion on exposing reasoning thought, but focus their analysis on successful reasoning on game strategies. Depaepe et al. (2007) (M2) even noted the likely effect of social interaction, but concluded that their data was not sufficient to explore the social effect on reasoning. Other studies focused on the importance of the collaborative impact of social interaction on reasoning itself (M1, M3, M10, M11, M12, M17, M18, M21). These were all qualitative studies that analyzed small group or classroom discussion transcripts except for the notable Delay et al. (2016) study (M1). This experiment was unique in its examination of the effect of peer dyads on mathematical reasoning ability. Though there was an implied criticality to the relationship under investigation, this study was quantitative and interactions between dyad pairs were not observed directly.

There was a small subset of mathematics articles that focused on individual reasoning ability (M5, M7, M9, M13, M16, M20). Most of these studies were quantitative experiments (exception M7), with a strong emphasis on exploring correlations to reasoning ability using standardized measurement instruments (M5, M9, M13, M20). Generally, this research demonstrated that students' ability to

reason mathematically impacted their overall achievement in content-specific mathematics topics. The exception article by Hughes et al. (2020) (M7) focused on the importance of critical mathematical vocabulary in communicating reasoning through students' written responses. The researchers noticed that students' written reasoning statements were limited, and researchers reported results through summary of data rather than display of students' responses.

Similarities and Differences

Scientific and mathematical reasoning take dramatically different approaches to understanding the social dimensions of reasoning in the articles we reviewed in this study. The science articles focus mainly on evaluating students' individual reasoning ability. Even in those articles that explore classroom discussion and/or small group environments, reasoning was still often tested at the individual level. This was mostly done in pen and paper examinations or via one-on-one interviews with researchers. Mathematics had a small subset of studies that took this approach, but the majority of mathematics articles focused more on evidence of the reasoning process as it presented itself in group conversation. It was much more common in mathematics to use qualitative methods to record, code, and analyze verbal discussion than it was in science. This may have been a result of the focus on higher elementary grade levels and thus stronger communication abilities. Or it may have tied to outcomes or learning goals in mathematics that focus on the explanation of thought processes.

Defining Reasoning and the Types of Reasoning Invoked in the Research

Science Education Research

Scientific reasoning is defined or discussed with various terms and processes in the articles. Some articles explained scientific reasoning with a specific definition (S2, S3, S6, S7, S9) and some articles explained the elements of scientific reasoning (S4, S5, S8, S10, S11, S12, S13, S14, S17, S19, S20). Some articles did not explain what scientific reasoning was but suggested key ideas such as inference (S15), causal relationship (S16) and coordination of theory and evidence (S18). For instance, Chen and She (2013) (S2) paraphrased definitions of scientific reasoning from various scholars such as "science reasoning skills are the ability to define a scientific question, plan a way to answer the question, analyze data, and interpret results" (National Research Council, 1999, p. 3). Lazonder and Kamp (2012) (S10) explained that "children can start to develop proficiency in the scientific reasoning skills of hypothesis generation, experimentation, and evidence evaluation" (p. 69). Lazonder and Kamp (2012) did not state the definition of scientific reasoning, yet their statements on the components of scientific skills explain what they emphasize as scientific reasoning. Based on the review of the articles, the key ideas to explain scientific reasoning are categorized as follows; causal reasoning (S1, S2, S3, S7, S16), the coordination of theory and evidence (S2, S5, S16, S18), hypothesis, test, and evidence evaluation or analysis (S2, S4, S9, S10, S11, S12, S13, S14, S17, S18, S19, S20), control of variables or cause-effect relationship (S10, S11, S19, S20) and evidence-based, including data-based, information-based, and/or rule-based processes (S2, S5, S6, S8, S15). In the category of causal reasoning, we included articles that emphasized cause and effect relationships, causality, and causal inference. The category of the coordination of theory and evidence also includes claim-evidence relationships. When researchers emphasized the process of hypothesis, test, data analysis, and conclusion, they were categorized in hypothesis, test, and evidence evaluation or analysis. There were also studies that particularly focused on students' understanding of the control of variables as scientific reasoning during inquiry processes. Also, some researchers employed the term evidence-based to emphasize the importance of data, information, theories, rules, and so on as evidence to justify one's claims.

Even though scientific reasoning was explained with different terminologies in different contexts, there were core values of students' scientific reasoning amongst the articles, that is, understanding of the relationship between claim and evidence. In science classrooms, students are engaged in diverse science inquiry activities to develop reasoning and problem-solving skills. In a problem solving process, students attempt to design methods, gather data (or information), and suggest conclusions and solutions to the problems. In the process, students need to understand how the problem, test design, data, and conclusion are related and to justify their conclusion in relation to data and the problem. Thus, it is critical that students make, evaluate, and justify claims based on evidence in scientific problem solving and communication. In the relationships between claim and evidence in scientific contexts, claims include hypothesis, prediction, inference, theory, and conclusion. Evidence includes data, information, knowledge, theory, etc. For instance, students make a claim that the air moves up and expands when it is heated (claim) on the phenomenon of hot air balloons floating (evidence). In scientific explanations, students need to explain how evidence supports, refutes, or revises claims, and that is evidence-based reasoning. As a classroom example, the articles (S2, S4, S6, S8, S9, S10, S11, S12, S15, S18, S20) explain that students are engaged in scientific investigations which include questions, hypothesis-making, testing, and conclusion and in this process, students need to understand the relationship among questions, hypothesis, data from testing, and conclusions. Additionally, the importance of students' causal reasoning is emphasized in controlling variables to test the hypothesis and in justification for the conclusion based on data (evidence). Overall, understanding and explaining how claim and evidence are related, that is, evidence-based reasoning, is presented as the key element of scientific reasoning in the articles.

To examine and discuss students' scientific reasoning skills, some researchers developed students' activities in certain science concepts or topics and some developed activities in general contexts of causal reasoning or relationships. Articles that discussed specific science concepts can be categorized in the content area of physics (S2, S4, S6, S8, S10, S11, S 18, S19, S20), chemistry (S2, S5,S7, S18), biology (S15, S18), and earth and space science (S2). Further, S13 and S14 focused on the Nature of Science. Despite the variety of science topics and content areas, only a half of the articles (S2, S5, S6, S13, S14, S18) mentioned the relationship between conceptual development and scientific reasoning, and the rest used the topics to contextualize students' inquiry and reasoning process in science content areas.

Mathematics Education Research

In mathematics education research, reasoning is of great interest to researchers within a broad range of perspectives on what constitutes reasoning demonstrated through varied definitions, contexts of mathematical topics, and conventional structures. Close to half of the articles do not include a definition of mathematical reasoning (M2, M3, M4, M8, M9, M12, M13, M14, M19), which is not a surprising finding considering Lithner and Palm's (2010) acknowledgement that mathematics educators tend to rely on an "implicit assumption that there is a universal agreement on its meaning" (p. 285). Analyzing the remaining articles that did state the researchers' definition of reasoning does not reveal a universal agreement. Articles that provided an explicit definition of reasoning described it as a process that is a way of thinking that is logical and systematic (M1, M5, M6, M11, M15, M17, M18) that is seen as inherently mathematical. Others described the product of reasoning as a focus in their definitions, rather than a process, as arriving at explanation (M7, M10), conclusion (M16, M21) or solution (M20). Reid (2002) (M15) contains the most deliberate development and discussion of reasoning among this corpus, contributing "*ways* of reasoning…, *needs* to reason…, *formulation* or awareness of reasoning" (p. 6, emphasis in original) to the field's investigation of students' mathematical reasoning. Many of the articles point to a detailed list of specific actions of reasoning to animate their explicit definitions (M5, M11, M15, M17, M18) or as a way to indicate what reasoning

looks like and focus their investigations (M3, M4, M6, M7, M9, M10, M16, M21). These actions will be explored further in the fourth category below.

Situating these conceptualizations of reasoning, researchers explored students' mathematical reasoning within varied contexts and structures. In many of the articles, researchers referred to mathematical reasoning, or logical reasoning, as the focus of their investigation (M1, M3, M4, M6, M11, M12, M14, M15, M17, M18, M20, M21). These authors foregrounded students' actions of logical thinking as they engaged in problem solving tasks that happened to be related to mathematical topics or in domain-general tasks. Examples of domain-general tasks include abstract strategy games (M6, M11, M14) and outdoor play (M17). Focusing on logical thinking, researchers investigated the nature of students' use and/or development of mathematical reasoning. All of the qualitative studies contain thick descriptions of students' statements of reasoning that illustrate both students' actions and the way students structured their mathematical reasoning through conventional forms. These forms included mainly deductive reasoning (M3, M11, M14, M15, M17, M20, M21) and inductive reasoning (M1, M11, M15), but were also supplemented by other forms, such as indirect reasoning (M4, M11), reasoning by cases (M4, M21), or metaphoric and analogic reasoning (M10, M11, M12, M15). A smaller number of articles foregrounded mathematical topics and studied students' achievement within these domain-specific areas given that reasoning occurred. The mathematical topics ranged from reasoning about quantities (M4) to fractions (M4, M7, M16) to computations (M5, M13) to algebra (M9). Overall, these studies did not identify the conventional forms of reasoning, as the focus varied, and most of the studies were quantitative, where students' statements of reasoning were not collected. Additionally, a few articles foregrounded qualities of communicative interaction in social situations discussed above as a context for reasoning (M2, M8, M10, M19), where results focused on the nature of communication rather than qualities of reasoning.

Similarities and Differences

In all the articles in both science and mathematics, researchers established that the investigation and understanding of students' reasoning is critical to pursue to improve learning in elementary school classrooms. Consensus is apparent; and yet, the complexity of how reasoning can be characterized often leads to implicit use of the term "reasoning" or absence of a clear definition by researchers. This commonality between science and mathematics research represents a shared-as-given understanding of reasoning that is held by researchers, even if it is not coordinated or goes so far as to be contested. More often, researchers chose to include a list of ways that students could engage in reasoning to illustrate their stances on what constitutes reasoning, where science emphasized different approaches to reasoning (e.g., causal, theory-evidence coordination, inquiry cycle), while mathematics referred to both the processes (e.g., conjecturing, exploring, analyzing) and product (e.g., explanation, conclusion, solution) of students' reasoning. The illustrations of reasoning point to a difference in science education researchers consistently situating specific moments of reasoning within the broader frame of the scientific inquiry process, while mathematics education researchers employed notions of logical thinking as a broader frame within which they identified processes that lead to products.

One of the ways that researchers in both science and mathematics emphasized the importance of investigating reasoning is highlighted in the contexts of their studies. In both subject areas, almost two-thirds of the research was conducted within domain-general areas related to either science or mathematics. In this way, science researchers were emphasizing students' scientific inquiry process or causal relationships and mathematics researchers were emphasizing students' logical thinking. The remaining articles explore students' reasoning but foregrounded domain-specific content, both in science (e.g., sound, electricity, changes to materials) and in mathematics (e.g., quantities, computations, algebra). While differences in content occur because of the different subject areas, the

similarity in primarily focusing on students' thinking while problem solving and how they reached a solution indicates that the two fields are united in their purpose for understanding students' cognitive processes during moments of reasoning.

Investigating the structures within which students experience and express reasoning was evident in both the science and mathematics reports on research. An intersection between scientific and mathematical reasoning appears in common conventional forms of reasoning such as inductive (mathematics) or evidence-based (science), deductive (mathematics) or causal (science), and several other forms that emphasize the relational aspect drawing on several ideas (science) or representations (mathematics). In science, the coexistence of inductive and deductive reasoning in students' thinking within an inquiry activity occurred more frequently as researchers indicated that experimentation begins deductively as students hypothesize based on known facts or prior experiences, then incorporate inductive reasoning, as they create a claim based on evidence from analyzing data. Mathematics research differed, in that researchers observed and interpreted distinct moments of inductive, deductive, or many other well-defined forms of reasoning in excerpts of students' logical thinking rather than looking at how the different conventional forms of reasoning are coordinated. Despite distinctions in focus related to reasoning structures, there are possible opportunities for reciprocity between subject areas to enhance students' interdisciplinary learning through reasoning: science research could begin to identify more varied forms of reasoning, while mathematics research could begin to coordinate the forms of reasoning.

Students' Engagement in Reasoning Processes

Science Education Research

Teaching science as a process has been emphasized for decades to improve students' scientific reasoning and problem-solving skills (Harlen, 1999). Science process skills include basic process skills (observing, inferring, measuring, using tools and equipment, communicating, classifying, predicting, etc.) and integrated process skills (formulating hypotheses, controlling variables, interpreting data, experimenting, formulating models, etc.). These terms have been widely used to describe students' engagement and actions during scientific inquiry and problem-solving tasks, in addition to the learning objectives that students are expected to develop in science classrooms. Researchers in the articles demonstrated these process skills to explain what process students were engaged in to practice scientific reasoning. These researchers provided students with tasks that engaged basic process skills such as observing and inferring (e.g., S1) and also integrated skills such as experimenting, controlling variables and justification (e.g., S10, 12, 17). When students are engaged in problem solving inquiry tasks, basic process skills are integrated into the complex process skills. For instance, students' observing and using tools and apparatus are all part of the experimenting process. In the hypothesis-verification process, students' engagement in reasoning is more complex, with the process of questioning, hypothesizing, testing, analyzing, and justifying conclusions. In these processes, scientific reasoning, such as the coordination of claim and evidence and justification with evidence, is practiced and developed.

In the articles, many researchers emphasized students' knowledge and skills of evidence-based reasoning and causal relationships, and thus provided a hypothesis-verification process to examine and develop students' scientific reasoning. Starting from questions, students develop, test, and justify hypotheses through controlling variables, analyzing, and justifying with evidence. Students' inquiry process was explained with the verbs of process skills. The verbs can be categorized in the processes of making claims (question, infer, hypothesize, predict, conjecture), testing claims (test, experiment, explore, control variables), analyzing (analyze, interpret, examine, evaluate, synthesize, conclude), and justify (argue and counter argue, critique, refute, generalize, explain, theorize). Some verbs such as

investigate and research pointed to the entire inquiry process, and thus, were not included in this categorization. Some verbs such as explain, theorize, or conclude could be included in various categories such as making claims and justifying.

Mathematics Education Research

A critical focus of reasoning in mathematics education is investigating the processes students invoke that can be categorized as reasoning, and here we refer to as the ways in which students engage in mathematical reasoning. Jeannotte and Kieran (2017) identify these processes as one of two aspects of mathematical reasoning. The processes of reasoning are commonly represented by a broad range of verbs, where the verbs are signifiers of actions taken by students and demonstrate how they go about engaging in mathematical reasoning. Inclusion of statements of students' reasoning or examples of the ways students engaged in processes of reasoning within articles' results followed methodological distinctions: quantitative studies contained no examples (M1, M7, M13, M14, M16, M20) or minimal examples (M5, M9, M19); mixed methods contained varied amounts of student data (M2 had no examples, M12 had limited examples, and M3 had many examples); qualitative studies contained many illustrations of students' engagement in reasoning (M4, M6, M8, M10, M11, M15, M17, M18, M21). Articles that contained many examples tended to focus their investigation on logical thinking (all except M8 and M10) and portrayed the complexity of students' reasoning through many different actions; whereas, articles with limited examples emphasized domain-specific reasoning (except M12) and the product of student reasoning with little variation in actions.

Researchers incorporated students' data and then interpreted it by labeling the action with a particular verb that highlighted the specific reasoning process. Our analysis revealed six clusters of verbs linked thematically; however, it must be noted that these clusters are not mutually exclusive nor applied in a linear fashion in that students' actions are complex and contribute to different purposes at different moments in their problem solving. Conjecturing is giving "a 'conscious guess'" (Lakatos (1976) as quoted in Houssart & Sams (M6), 2006, p. 60), which are "put forth without being considered to be true or false--as something subject to testing" (Reid (M15), 2002, p. 14) and sometimes labeled as a hypothesis or prediction. While an initial conjecture was often made after some mathematical activity, like building and observing a pattern (M10, M15, M18, M21), it can also be revised throughout problem solving to respond to counterexamples (M6, M11, M15).

Investigating (M4, M8, M10, M11, M12, M15, M18, M19) may occur near the beginning of the activity when students explore and play, when observing and acting within the problem with various approaches, and even near the end as they test emerging generalizations. Rather than the reasoning imbedded in investigating remaining ephemeral, researchers highlight the importance of students representing (M3, M4, M6, M10, M11, M17, M21), or tangibly demonstrating with communicative "tools to assist the students' sharing" (Kumpulainen & Kaartinen (M10), 2003, p. 366), to make external and visible their reasoning so that it can be consolidated and interrogated by others. Articles report a wide variation of representations, including show, use manipulatives, gesture, draw diagrams, make images/pictures, organize, describe, discuss, and label. Students used their representations to systematize their investigating, which moves their actions toward analysis (M6, M11, M15, M17, M21), or "consider more completely" (McFeetors & Palfy (M11), 2018, p. 115) strategies and patterns expressed earlier to begin moving to generalizing. Researchers pointed to analysis as a process of mathematical reasoning using verbs like visualize, modify, adapt, specialize, compare, contradict, give counterexamples, and refute.

Students' engagement in the range of activities eliciting mathematical reasoning result, in all the studies, in expressions of generalizing and justifying. Generalizing occurs through students "finding a similarity" and establishing a "a general formula or fact or a meaning of an object or idea" (Vale et al. (M18), 2016, p. 876) and can also be described with the verbs like concluding, relating, and

extending. Examples include creating winning game strategies through repeated success of discrete moves across games (M6, M11), creating a rule from a pattern (M3, M4, M9) or a statement accepted by the group (M15, M17, M18, M19). While generalization is illustrated in just over half the articles, a greater emphasis is placed on students' justification (all articles expect M6, M12) as a process to "construct generic arguments and persuade" and "develop mathematical arguments" (Flegas & Charalampos (M3), 2013, p. 71). The articles seem to advocate for students supporting a solution is more important in the reasoning process than arriving at a solution (generalization) itself. Justifying is represented by other verbs such as convince, argue, explain, support, reflect, and evaluate where the commonality is that the students are using various representations to defend, in a convincing manner, their processes of mathematical thinking and solution to the task as an invocation of mathematical reasoning.

Similarities and Differences

In both science and mathematics articles, students' reasoning is emphasized through processes rather than outcomes of thinking and knowledge. Researchers designed their studies to engage students in the overall process of investigating, researching, and problem solving to understand how students develop their ideas scientifically or mathematically. These researchers characterized the processes of students' thinking and reasoning with various verbs in the categories of testing ideas (predict, hypothesize, experiment, collect data, manipulate materials), reaching solutions (analyze, synthesize, conclude, generalize), and justifying (explain, argue, evaluate, justify), especially if there was a clear emphasis on justification in both areas. For instance, in science education, students need to explain why certain information can be evidence to certain claims (evidence-based reasoning) or how conclusions are justifiable and valid through controlling variables (causal reasoning). In mathematical problem solving, students need to explain how they reached certain solutions (patterning, use manipulatives, gesture, represent with diagrams or images/pictures, organize, explain, discuss, and label) and how their solutions need to be supported deductively as an early process leading toward constructing proofs (conjecture, analyze, synthesize, generalize, justify).

Students' engagement in scientific reasoning is often discussed in the whole process of scientific inquiry. Students practiced scientific inquiry processes, which developed scientific reasoning. Thus, scientific inquiry skills often denoted students' reasoning skills and vice versa. In mathematics articles, the whole process of and the phases of problem solving are not emphasized in the discussion of mathematical reasoning. Certain processes such as generalization or justification are incorporated as one specific skill of mathematical reasoning. As the hypothesis-based inquiry has been emphasized in science education, hypothesizing (conjecturing in mathematics) and related processes such as controlling variables are frequently discussed, which is rarely mentioned in the mathematics articles.

There are the discipline-specific terms and differences between science and mathematics education (see Table 3). Yet, it is also evident that students' reasoning is explained as a process rather than a product and that students understand their meaning-making process of how questions, phenomena, and knowledge are related, interact, and are justified for problem solving. Encountering questions and problems, students are engaged in reasoning processes to develop reasonable and logical explanations and justification to their questions that are represented in a variety of ways. We acknowledge that similarities are evident in the different terms and actions.

Table 3

Reasoning Verbs and Emphasis in Science and Mathematics Education

	Science reasoning & problem solving	Mathematics reasoning & problem solving
Verbs	• question • predict, hypothesize, infer • experiment, manipulate materials, control variables, collect data • analyze, interpret, synthesize • conclude, generalize, • explain, evaluate, examine, argue, justify, theorize	• question • conjecture, predict • investigate, explore, inquire, examine, use manipulatives, systematize • analyze, synthesize, pattern, integrate, modify/adapt, evaluate • generalize, classify, pattern • convince, explain, gesture, represent, justify, refute, argue, discuss, label
Emphasis	• evidence-based reasoning (claim-evidence-justification) • causal and relational reasoning	• logical, systematic thinking using varied conventional structures to present justifications such as inductive, deductive, analogic, imagistic, indirect, metaphoric, by contradiction, etc.

Discussion

Based on the systematic review, similarities and differences in educational approaches to scientific and mathematical reasoning were evident. The terms and pedagogical contexts for students' problem solving were different and diverse according to the disciplinary foci of science and mathematics. Yet, students' reasoning processes were similar as a logical and holistic meaning-making process both in science and mathematics. Once problems are framed and suggested in classrooms, students explore the problems and ways to find answers, develop their answers, and evaluate why their answers are reasonable and justifiable. This process of reasoning in problem solving processes needs to be understood as interdisciplinary, rather than a disciplinary-specific approach belonging exclusively to science or mathematics.

An interdisciplinary approach for science and mathematics learning in elementary education requires students to develop knowledge integration and skills in interdisciplinary problem-solving processes (Honey et al., 2014). Students approach complex STEM questions by seeking, testing, justifying, and negotiating knowledge claims. Reasoning skills are critical throughout STEM problem solving in science and mathematics education. Scientific and mathematical reasoning skills are emphasized within their individual disciplinary domains, and yet, with the recent interest in a STEM-centered approach to teaching (English & King, 2015; Estapa & Tank, 2017; Stohlman et al., 2012), students' reasoning process is increasingly interdisciplinary (Mayes & Gallant, 2018; Tan et al., 2022). In our findings, domain-specific terms and emphases were well defined and researched in science and mathematics education. As examples, skills could include hypothesis testing or scientific method in science, and conjecture and methods of proof in mathematics. Though similar in purpose and implementation, terminology differences make direct comparisons and cross-curricular connections more difficult. Our findings show that both subject areas emphasize students' logical thinking (inductive, deductive, abductive, etc.) to seek out solutions and justify the process and products of their problem solving. When students' reasoning skills are integrated across science and mathematics,

however, direct translation of meaning is not inherently obvious and thus pedagogical discussion needs to follow. A key takeaway from these findings is that the development of students' understanding of the relationships between problems, observations, data, and knowledge claims, rather than focusing on the discipline-specific terms and characteristics, is critical to the advancement of reasoning in science and mathematics education.

We suggest the term 'STEM reasoning' to describe the interdisciplinary nature of reasoning that could be shared by students in science and mathematics classrooms. To occupy that theoretical space, STEM reasoning would have to take on a broad meaning and incorporate both the processes of student engagement and subsequent ways of thinking through problem solving processes. It would aim to identify the finely nuanced facets of reasoning that science and mathematics both offer independently and incorporate them into the process of STEM problem solving. This approach is not, however, intended to focus on the historic integration of science and mathematics -- typically using mathematics formulae as a simple, procedural tool to solve the various substeps of complex science or technological problems (Bosse et al., 2010; Bursal & Paznokas, 2006). Rather, from this literature review we are suggesting that the field can aim for much higher cognitive activities, integrating mathematics-style logic, systems thinking, and methods of proof into approaches of nature-of-science style scientific inquiry.

There is some research to suggest that authentic integration of science and mathematics can be effective (Treacy & O'Donoghue, 2014). Existing definitions for subject-specific reasoning that overlap may converge or disappear through the process of students' problem-solving inquiry. There may also be room for new terminologies that describe complex co-curricular thinking that did not previously occur. Overall, we suggest that STEM reasoning could exist as an umbrella term for the thinking and reasoning that science and mathematics share in their approach to solving problems because of the similarities we identified in students' engagement in reasoning across both disciplines. Students' reasoning and decision-making process in problem solving contexts is integrated with diverse knowledge, skills, and values in problem contexts. Thus, we believe a key takeaway is that a focus and use of STEM reasoning could lead to more authentic and meaningful integration of science and mathematics learning experiences for students.

STEM reasoning in this study focused on the intersection of scientific and mathematical reasoning. There remains an unexplored area in how STEM reasoning may extend to the other area disciplines of technology and engineering. Finding the intersection points between scientific and mathematical reasoning and their engineering (Tan et al., 2022) and technology (Kennedy & Kraemer, 2018) parallels would be critical to formalizing the idea of STEM reasoning. In acknowledging this limitation, we hope to see expansion and clarification of STEM reasoning emerge in future research, consolidating reasoning integration in all aspects of the STEM disciplines. Slavit et al. (2021) recently posited that there is precedent for the consolidation of reasoning skills and abilities in science and mathematics and seem to agree that this happens frequently in practice, but lacks epistemological foundation in the literature. This may include an exploration of the intersection in reasoning with critical thinking and design engineering (Silk et al., 2009; Siverling et al., 2017) or with computational thinking and computer science (Kennedy & Kraemer, 2018; Olabe et al, 2014; Weintrop et al., 2016). Inclusion of these subjects is difficult in this study, however, as relevant literature in both engineering and computational thinking is lacking at the elementary level. Students' reasoning to solve every day-related problems is not a sole discipline-based reasoning but an interdisciplinary logical thinking process to reach out solutions. The demarcation among disciplinary thinking and reasoning is neither present nor meaningful. To understand and develop interdisciplinary STEM reasoning, cross-curricular and problem-based approaches could be useful. This could be approached through cross-pollinated activities involving both science and mathematics curricula. Expanding science and mathematics into STEM integration could occur through curriculum-ambiguous activities that apply

science and mathematics more subtly -- perhaps through broad, project-based learning activities or digital or analog gameplay.

Regardless of the curricular grounding for future research, there is much in the way of methodology and experimental design that can be shifted among science and mathematics. Our findings show that science could gain valuable insight into student reasoning processes from shifting more frequently to a qualitative methodology. Despite the importance of integration of diverse knowledge, skills, and values in collective STEM knowledge building and problem solving, students' reasoning abilities were examined in individual contexts in most of the studies in our review. As STEM problem solving requires integrated, collective, and social domains of learning (Crippen & Antonenko, 2018), students' reasoning needs to be examined and evaluated in interactive and social dimensions. Evidence derived from students' conversations and interactions in the classroom could help bridge the gap from individual student moments of reasoning to the overarching nature of science problem solving. Conversely, our findings show that mathematics researchers may benefit from the application of some of the reasoning inventories developed in science education. Taken together, STEM reasoning research would be invited to include science and mathematics education standards -- both to ground it in previous literature and to fully realize the benefits of their subject counterparts. However, this would need to be expanded to technology and engineering reasoning and problem solving. STEM reasoning could provide a unique approach to STEM education that combines individual disciplinary thinking into a common critical approach to problem solving.

Liam Quinn McCashin (mccashin@ualberta.ca) is a Ph.D. student studying educational technology in the Faculty of Education at the University of Alberta. He is formerly a high school math and science teacher with Edmonton and Elk Island Public schools. Quinn's research explores digital technology integration across K-12 school environments, coding and computational thinking (CT), and games, gamification, and game environments in the classroom.

P. Janelle McFeetors (Janelle.mcfeetors@ualberta.ca) is an Associate Professor in Mathematics Education at the University of Alberta. She is also the Co-director of the Centre for Mathematics, Science and Technology Education (CMASTE). Her research focuses on applying Dewey's theory of experience to designing for and understanding students' mathematical learning and employs emergent methodologies to interpret students' experiences within the complexity of classroom contexts. Her research work spans using commercial games in elementary classrooms, to occasion spatial and logical reasoning, all the way to professional learning program development for post-secondary instructors.

Mijung Kim (mijung.kim@ualberta.ca) is a professor in science education and co-Director of the Centre for Mathematics, Science, and Technology Education at the Faculty of Education, University of Alberta, Canada. Her research interests include science inquiry, dialogical argumentation, and critical and collective reasoning in children's decision making and problem solving process.

References

Alberta Education. (1996). *Science, grade 1 to 6 program of studies*. Author.

Alberta Education. (2007/2016). *Mathematics, kindergarten to grade 9 program of studies*. Author.

Alexander, P. A. (2017). Relational reasoning in STEM domains: A foundation for academic development. *Educational Psychology Review, 29*(1), 1-10. https://www.doi.org/10.1007/s10648-016-9383-1

Babb, A. P. P., Takeuchi, M. A., Yanez, G. A., Francis, K., Gereluk, D., & Friesen, S. (2016). Pioneering STEM education for pre-service teachers. *International Journal of Engineering Pedagogy, 6*(4), 4-11. https://www.doi.org/10.3991/ijep.v6i4.5965

Blanton, M., & Kaput, J. (2005). Characterizing a classroom practice that promotes algebraic reasoning. *Journal for Research in Mathematics Education, 36*(5), 412-446. https://www.doi.org/10.2307/30034944

Bossé, M.J., Lee, T.D., Swinson, M., & Faulconer, J. (2010). The NCTM process standards and the five *Es* of science: Connecting math and science. *School Science and Mathematics, 110*(5), 262-276. https://doi.org/10.1111/j.1949-8594.2010.00033.x

Bower, B. (2009). Think like a scientist: A class of curious sixth-graders arguing over moist, mucky jars may represent the future of science education. *Science News: Magazine of the Society for Science & the Public, 175*(13), 20-23. https://www.doi.org/10.1002/scin.5591751320

Brodie, K. (2010). *Teaching mathematical reasoning in secondary school classrooms*. Springer. https://doi.org/10.1007/978-0-387-09742-8_1

Brown, S. A. (2018). Are indirect proofs less convincing? A study of students' comparative assessments. *Journal of Mathematical Behavior, 49*, 1-23. https://www.doi.org/10.1016/j.jmathb.2016.12.010

Bursal, M., & Paznokas, L. (2006). Mathematics anxiety and preservice elementary teachers' confidence to teach mathematics and science. *School Science and Mathematics, 106*(4), 173-180. https://www.doi.org/10.1111/j.1949-8594.2006.tb18073.x

Cazden, C. (2001). *Classroom discourse: The language of teaching and learning* (2nd ed.). Heinemann.

Common Core State Standards Initiative (CCSSI). (2010). *Common core state standards for mathematics*. National Governors Association Center for Best Practices and the Council of Chief State School Officers. http://www.corestandards.org/Math/

Cooper, H. (2015). *Research synthesis and meta-analysis: A step-by-step approach* (Vol. 2; 5th ed.). Sage.

Corbin, J., & Strauss, A. (2008). *Basics of qualitative research: Techniques and procedures for developing grounded theory* (3rd ed.). Sage. https://doi.org/10.4135/9781452230153

Crippen, K.J., & Antonenko, P.D. (2018). Designing for collaborative problem solving in STEM cyberlearning. In Y. J. Dori, Z. R. Mevarech, & D.R. Baker (Eds.), *Cognition, metacognition, and culture in STEM education: Innovations in science education and technology* (pp. 89-116). Springer. https://doi.org/10.1007/978-3-319-66659-4_5

Davydov, V. V. (1990). *Types of generalization in instruction: Logical and psychological problems in the structuring of school curricula (Soviet studies in mathematics education, Vol. 2)* (J. Teller, Trans.). National Council of Teachers of Mathematics.

Dole, S., Clarke, D., Wright, T., & Hilton, G. (2012). Students' proportional reasoning in mathematics and science. In T. Y. Tso (Ed.), *Proceedings of the 36th conference of the International Group for the Psychology of Mathematics Education* (Vol. 2; pp. 195-202). Psychology of Mathematics Education.

Engelmann, K., Neuhaus, B. J., & Fischer, F. (2016). Fostering scientific reasoning in education -- meta-analytic evidence from intervention studies. *Educational Research and Evaluation, 22*(5-6), 333-349. https://www.doi.org/10.1080/13803611.2016.1240089

English, L. D. (Ed.). (1997). *Mathematical reasoning: Analogies, metaphors and images*. Lawrence Erlbaum Associates.

English, L. D., & King, D. T. (2015). STEM learning through engineering design: Fourth-grade students' investigations in aerospace. *International Journal of STEM Education, 2*(14), 1–18. https://www.doi.org/10.1186/s40594-015-0027-7

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

Erduran, S., Ozdem, Y., & Park, J. Y. (2015) Research trends on argumentation in science education: A journal content analysis from 1998–2014. *International Journal of STEM Education, 2*(1). https://www.doi.org/10.1186/s40594-015-0020-1

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

Hanna, G. (1983). *Rigorous proof in mathematics education*. OISE Press.

Harlen, W. (1999). Purposes and procedures for assessing science process skills, assessment in education. *Principles, Policy & Practice, 6*(1), 129-144. https://www.doi.org/10.1080/09695949993044

Herbel-Eisenmann, B., & Cirillo, M. (2009). *Promoting purposeful discourse: Teacher research in secondary math classrooms*. National Council of Teachers of Mathematics.

Hjelte, A., Schindler, M., & Nilsson, P. (2020). Kinds of mathematical reasoning addressed in empirical research in mathematics education: A systematic review. *Education Sciences, 10*(10), 289-303. https://www.doi.org/10.3390/educsci10100289

Honey, M., Pearson, G., & Schweingruber, H. (Eds.). (2014). *STEM integration in K-12 education: Status, prospects, and an agenda for research*. National Academies Press. https://doi.org/10.17226/18612

Jeannotte, D., & Kieran, C. (2017). A conceptual model of mathematical reasoning for school mathematics. *Education Studies in Mathematics, 96*(1), 1-16. https://www.doi.org/10.1007/s10649-017-9761-8

Johnson, C. C., Peters-Burton, E. E., & Moore, T. J. (Eds.). (2021). *STEM road map 2.0: A framework for integrated STEM education in the innovation age*. Routledge.

Kennedy, C., & Kraemer, E. T. (2018). What are they thinking? Eliciting student reasoning about troublesome concepts in introductory computer science. *Proceedings of the 18th Koli Calling International Conference on Computing Education Research*, 1–10. https://www.doi.org/10.1145/3279720.3279728

Kilpatrick, J., Swafford, J., & Findell, B. (2001). *Adding it up: Helping children learn mathematics*. National Academy Press.

Kuhn, D. (1989). Children and adults as intuitive scientists. *Psychological Review, 96*(4), 674–689.

Kuhn, D. (1991). *The skills of argument*. Cambridge University Press.

Kuhn, D., & Pearsall, S. (2000). Developmental origins of scientific thinking. *Journal of Cognition and Development, 1,* 113-129.

Lakatos, E. (1976). *Proofs and refutations: The logic of mathematical discovery*. Cambridge University Press.

Lannin, J. K., Ellis, A. B., & Elliott, R. (2011). *Developing essential understanding of mathematical reasoning for teaching mathematics in prekindergarten-grade 8*. National Council of Teachers of Mathematics.

Lawson, A. (2004). The nature and development of scientific reasoning: A synthetic view. *International Journal of Science and Mathematics Education, 2,* 307–338.

Lehrer, R., & Schauble, L. (2000). Developing model-based reasoning in mathematics and science. *Journal of Applied Developmental Psychology, 21*(1), 39-48. https://www.doi.org/10.1016/S0193-3973(99)00049-0

Lithner, J., & Palm, T. (2010). Learning difficulties and mathematical reasoning. In B. Sriraman, C. Bergsten, S. Goodchild, G. Pálsdóttir, B. Dahl & L. Haapasalo (Eds.), *The first sourcebook on Nordic research in mathematics education: Norway, Sweden, Iceland, Denmark and contributions from Finland* (pp. 283-298). Information Age Publishing.

Mason, J., Burton, L., & Stacey, K. (2010). *Thinking mathematically* (2nd ed). Prentice Hall.

Mata-Pereira, J., & da Ponte, J. P. (2017). Enhancing students' mathematical reasoning in the classroom: Teacher actions facilitating generalization and justification. *Educational Studies in Mathematics, 96*(2), 169–186. https://www.doi.org/10.1007/s10649-017-9773-4

Mayes, R., & Gallant, B. (2018). The 21st century STEM reasoning. *US-China Education Review B, 8*(2), 67-74. https://www.doi.org/10.17265/2161-6248

Mercer, N., Wegerif, R., & Dawes, L. (1999). Children's talk and the development of reasoning in the classroom. *British Educational Research Journal, 25*(1), 95-111. https://www.doi.org/10.1080/0141192990250107

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

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

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

National Research Council. (1996). *National Science Education Standards.* National Academies Press. https://www.doi.org/10.17226/4962.

National Research Council. (1999). *New strategies for America's watersheds.* National Academies Press. https://www.doi.org/10.17226/6020

Next Generation Science Standards Lead States (NGSS Lead States). (2013). *Next generation sciences standards: For states, by states.* National Academies Press. https://www.nextgenscience.org/search-standards

Nielsen, J. A. (2013). Dialectical features of students' argumentation: A critical review of argumentation studies in science education. *Research Science Education, 43*(1), 371–393. https://www.doi.org/10.1007/s11165-011-9266-x

Nunes, T., Bryant, P., Evans, D., Bell, D., Gardner, S., Gardner, A., & Carraher, J. (2007). The contribution of logical reasoning to the learning of mathematics in primary school. *British Journal of Developmental Psychology, 25*(1), 147-166. https://www.doi.org/10.1348/026151006X153127

Olabe, J., Basogain, X., Olabe, M., Maz, I., & Castao, C. (2014). Solving math and science problems in the real world with a computational mind. *New Approaches in Educational Research*, 3(2), 75-82. https://www.doi.org/10.7821/naer.3.2.75-82

Organisation for Economic Co-operation and Development. (2013). *PISA 2015 draft science framework.* https://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Science%20Framework%20.pdf

Owens, D. C., Herman, B. C., Oertli, R. T., Lannin, A. A., & Sadler, T. D. (2019). Secondary science and mathematics teachers' environmental issues engagement through socioscientific reasoning. *EURASIA Journal of Mathematics, Science and Technology Education, 15*(6), em1693. https://www.doi.org/10.29333.ejmste/103561

Pisesky, A., McFeetors, J., & Kim, M. (2018). Intersections in reasoning within science and mathematics. *Delta-K, 55*(1), 9-18.

Polya, G. (1954). *Mathematics and plausible reasoning: Induction and analogy in mathematics.* Princeton University Press.

Purzer, S., Goldstein, M. H., Adams, R. S., Xie, C., & Nourian, S. (2015). An exploratory study of informed engineering design behaviors associated with scientific explanations. *International Journal of STEM Education, 2*(9), 1-12. https://www.doi.org/10.1186/s40594-015-0019-7

Sfard, A. (2008). *Thinking as communicating: Human development, the growth of discourses, and mathematizing.* Cambridge University Press

Shin, D. (2020). Prospective mathematics teachers' professional noticing of students' reasoning about mean and variability. *Canadian Journal of Science, Mathematics and Technology Education, 20*(3), 423–440. https://www.doi.org/10.1007/s42330-020-00091-w

Silk, E. M., Schunn, C. D., & Strand Cary, M. (2009). The impact of an engineering design curriculum on science reasoning in an urban setting. *Journal of Science Education and Technology, 18*(3), 209–223. https://www.doi.org/10.1007/s10956-009-9144-8

Simon, M. A. (1996). Beyond inductive and deductive reasoning: The search for a sense of knowing. *Educational Studies in Mathematics, 30*(2), 197-210. https://www.doi.org/10.1007/bf00302630

Siverling, E. A., & Suazo-Flores, E., & Mathis, C. A., & Moore, T. J., & Guzey, S. S., & Whipple, K. S. (2017, June). *Middle School Students' Engineering Discussions: What Initiates Evidence-Based Reasoning?* [Paper presentation]. 2017 ASEE Annual Conference and Exposition, Columbus, Ohio. https://www.doi.org/10.18260/1-2--28668

Slavit, D., Grace, E., & Lesseig, K. (2021). Student ways of thinking in STEM contexts: A focus on claim making and reasoning. *School Science and Mathematics, 121*(8), 466–480. https://www.doi.org/10.1111/ssm.12501

Staples, M., & Cavanna, J. (2021). Student argumentation work sample sorting task and teachers' evaluations of arguments. *Mathematics Teacher Educator, 9*(2), 94-109. https://www.doi.org/10.5951/mte.2020.0001

Staples, M., & Newton, J. 92016). Teachers' contextualization of argumentation in the mathematics classroom. *Theory into Practice, 55*(4), 294-301. https://www.doi.org/10.1080/00405841.2016.1208070

Steen, L. A. (1999). Twenty questions about mathematical reasoning. In L. V. Stiff & F. R. Curcio (Eds.), *Developing mathematical reasoning in grades K-12, 1999 yearbook* (pp. 270-285). National Council of Teachers of Mathematics.

Stohlmann, M., Moore, T. J., & Roehrig, G. H. (2012). Considerations for teaching integrated STEM education. *Journal of Pre-College Engineering Education Research, 2*(1), Article 4. https://www.doi.org/10.5703/1288284314653

Tan, A.-L., Ong, Y. S., Ng, Y. S., & Tan, J. H. J. (2022). STEM problem solving: Inquiry, concepts, and reasoning. *Science & Education.* Advance online publication. https://www.doi.org/10.1007/s11191-021-00310-2

Thompson, P. W. (1996). Imagery and the development of mathematical reasoning. In L. P. Steffe & P. Nesher (Eds.), *Theories of mathematical learning* (pp. 267-283). Routledge.

Treacy, P., & O'Donoghue, J. (2014). Authentic integration: A model for integrating mathematics and science in the classroom. *International Journal of Mathematical Education in Science and Technology, 45*(5), 703–718. https://www.doi.org/10.1080/0020739X.2013.868543

Varma, K. (2014). Supporting scientific experimentation and reasoning in young elementary school students. *Journal of Science Education and Technology, 23*(3), 381–397. https://www.doi.org/10.1007/s10956-013-9470-8

Venenciano, L., & Heck, R. (2016). Proposing and testing a model to explain traits of algebra preparedness. *Educational Studies in Mathematics, 92*(1), 21-35. https://www.doi.org/10.1007/s10649-015-9672-5

Vieira, M., & Tenreiro-Vieira, C. (2016). Fostering scientific literacy and critical thinking in elementary science education. *International Journal of Science and Mathematics Education, 14*(4), 659–680. https://www.doi.org/10.1007/s10763-014-9605-2

Wasserman, N. H., & Rossi, D. (2015). Mathematics and science teachers' use of and confidence in empirical reasoning: Implications for STEM teacher preparation. *School Science and Mathematics, 115*(1), 22-34. https://www.doi.org/10.1111/ssm.12099

Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining computational thinking for mathematics and science classrooms. *Journal of Science Education and Technology, 25*(1), 127–147. https://www.doi.org/10.1007/s10956-015-9581-5

White, C. S., Alexander, P. A., & Daugherty, M. (1998). The relationship between young children's analogical reasoning and mathematical learning. *Mathematical Cognition, 4*(2), 103-123. https://www.doi.org/ 10.1007/s10649-015-9672-5

Wu, H. (2009). What's sophisticated about elementary mathematics? *American Educator, 33*(3), 4-14.

Appendix A: Science Education Articles

Bauer, J.-R., & Booth, A. E. (2019). Exploring potential cognitive foundations of scientific literacy in preschoolers: Causal reasoning and executive function. *Early Childhood Research Quarterly, 46*, 275-284. https://www.doi.org/10.1016/j.ecresq.2018.09.007

Chen, C.-T., & She, H.-C. (2015). The effectiveness of scientific inquiry with/without integration of scientific reasoning. *International Journal of Science and Mathematics Education, 20*, 1-20.

Fernbach, P. M., Macris, D. M., & Sobel, D. M. (2012). Which one made it go? The emergence of diagnostic reasoning in preschoolers. *Cognitive Development, 27*(1), 39-53. https://www.doi.org/10.1016/j.cogdev.2011.10.002

Gillies, R. M., & Nichols, K., Burgh, G., & Haynes, M. (2014). Primary students' scientific reasoning and discourse during cooperative inquiry-based science activities. *International Journal of Educational Research, 63*, 127-140. http://www.doi.org/10.1016/j.ijer.2013.01.001

Hackling, M., & Sheriff, B. (2015). Language-based reasoning in primary science. *Teaching Science, 6*(2), 14-25.

Hardy, I., Kloetzer, B., Moeller, K., & Sobdian, B. (2010). The analysis of classroom discourse: Elementary school science curricula advancing reasoning with evidence. *Educational Assessment, 15*(3-4), 197-221. https://www.doi.org/10.1080/10627197.2010.530556

Hatzinikita, V. (2005/2006). Pupils' linear causal reasoning patterns in changes of matter. *International Journal of Learning, 12*(10), 43-54.

Kim, M., & Pegg, J. (2019). Case analysis of children's reasoning in problem-solving process. *International Journal of Science Education, 41*(1), 739-758. https://www.doi.org/10.1080/09500693.2019.1579391

Koksal-Tuncer, O., & Sodian, B. (2018). The development of scientific reasoning: Hypothesis testing and argumentation from evidence in young children. *Cognitive Development, 48*, 135-145. https://www.doi.org/10.1016/j.cogdev.2018.06.011

Lakatos, I. (1976). *Proofs and refutations: The logic of mathematical discovery*. Cambridge University Press.

Lazonder, A. W., & Kamp, E. (2012). Bit by bit or all at once? Splitting up the inquiry task to promote children's scientific reasoning. *Learning and Instruction, 22*(6), 458-464. https://www.doi.org/10.1016/j.learninstruc.2012.05.005

Lazonder, A. W., & Wiskerke-Drost, S. (2015). Advancing scientific reasoning in upper elementary classrooms: Direct instruction versus task structuring. *Journal of Science Education and Technology, 24*(1), 69-77. https://www.doi.org/10.1007/s10956-014-9522-8

Mayer, D., Sodian, B., Koerber, S., & Schwippert, K. (2014). Scientific reasoning in elementary school children: Assessment and relations with cognitive abilities. *Learning and Instruction, 29*, 43-55. http://www.doi.org/10.1016/j.learninstruc.2013.07.005

Osterhaus, C., Koerber, S., & Sodian, B. (2017). Scientific thinking in elementary school: Children's social cognition and their epistemological understanding promote experimentation skills. *Developmental Psychology, 53*(3), 450-462. http://www.doi.org/10.1037/dev0000260

Osterhaus, C., Koerber, S., & Sodian, B. (2020). The Science-P Reasoning Inventory (SPR-I): Measuring emerging scientific-reasoning skills in primary school. *International Journal of Science Education, 42*(7), 1087-1107. https://www.doi.org/10.1080/09500693.2020.1748251

Paparistodemou, E., & Meletiou-Mavrotheris, M. (2008). Developing young students' informal inference skills in data analysis. *Statistics Education Research Journal, 7*(2), 83-106.

Schulz, L. E., & Bonawitz, E. B. (2007). Serious fun: Preschoolers engage in more exploratory play when evidence is confounded. *Developmental Psychology, 43*(4), 1045-1050. https://www.doi.org/10.1037/0012-1649.43.4.1045

Schiefer, J., Golle, J., Tibus, M., & Oschatz, K. (2019). Scientific reasoning elementary school children: Assessment of the inquiry cycle. *Journal of Advanced Academics, 30*(2), 144-177. https://www.doi.org/10.1177/1932202X18825152

Tytler, R., & Peterson, S. (2004). Form "try it and see" to strategic exploration: Characterizing young children's scientific reasoning. *Journal of Research in Science Teaching, 41*(1), 94-118. https://www.doi.org/10.1002/tea.10126

van der Graaf, J., Segers, E., & Verhoeven, L. (2015). Scientific reasoning abilities in kindergarten: Dynamic assessment of the control of variable strategy. *Instructional Science, 43*(3), 381-400. https://www.doi.org/10.1007/s11251-015-9344-y

van der Graaf, J., van de Sande, E., Gijsel, M., & Segers, E. (2019). A combined approach to strengthen children's scientific thinking: Direct instruction on scientific reasoning and training of teacher's verbal support. *International Journal of Science Education, 41*(9), 1119-1138. https://www.doi.org/10.1080/09500693.2019.1594442

Appendix B: Mathematics Education Articles

DeLay, D., Laursen, B., Kiuru, N., Poikkeus, A.-M., Aunola, K., & Nurmi, J.-E. (2016). Friend influence and susceptibility to influence: Changes in mathematical reasoning as a function of relative peer acceptance and interest in mathematics. *Merrill-Palmer Quarterly, 62*(3), 306-333. https://www.doi.org/10.13110/merrpalmquar1982.62.3.0306

Depapepe, F., De Corte, E., & Verschaffel, L. (2007). Unraveling the culture of the mathematics classroom: A video-based study in sixth grade. *International Journal of Educational Research, 46*(5), 266-279. https://www.doi.org/10.1016/j.ijer.2007.10.008

Flegas, K., & Charalampos, L. (2013). Exploring logical reasoning and mathematical proof in grade 6 elementary school students. *Canadian Journal of Science, Mathematics and Technology Education, 13*(1), 70-89. https://www.doi.org/10.1080/14926156.2013.758326

Francisco, J. M., & Maher, C. A. (2011). Teachers attending to students' mathematical reasoning: Lessons from an after-school research program. *Journal of Mathematics Teacher Education, 14*(1), 49-66. https://www.doi.org/10.1007/s10857-010-9144-x

Gurbuz, R., & Erdem, E. (2016). Relationship between mental computation and mathematical reasoning. *Cogent Education, 3*(1), 1-18. https://www.doi.org/10.1080/2331186X.2016.1212683

Houssart, J., & Sams, C. (2008). Developing mathematical reasoning through games of strategy played against the computer. *International Journal for Technology in Mathematics Education, 15*(2), 59-71.

Hughes, E. M., Riccomini, P. J., & Lee, J.-Y. (2020). Investigating written expressions of mathematical reasoning for students with learning disabilities. *Journal of Mathematical Behavior, 58*, 1-14. https://www.doi.org/10.1016/j.jmathb.2020.100775

Hunter, J. (2017). Developing interactive mathematical talk: Investigating student perceptions and accounts of mathematical reasoning in a changing classroom context. *Cambridge Journal of Education, 47*(4), 475-492. https://www.doi.org/10.1080/0305764X.2016.1195789

Jurdak, M. E., & El Mouhayar, R. R. (2014). Trends in the development of student level of reasoning in pattern generalization tasks across grade level. *Educational Studies in Mathematics, 85*(1), 75-92. https://www.doi.org/10.1007/s10649-013-9494-2

Kumpulainen, K., & Kaartinen, S. (2003). The interpersonal dynamics of collaborative reasoning in peer interactive dyads. *The Journal of Experimental Education, 71*(4), 333-370. https://www.doi.org/10.1080/00220970309602069

McFeetors, P. J., & Palfy, K. (2018). Educative experiences in a games context: Supporting emerging reasoning in elementary school mathematics. *Journal of Mathematical Behavior, 50*, 103-128. https://www.doi.org/10.1016/j.jmathb.2018.02.003

Mercer, N., & Sams, C. (2006). Teaching children how to use language to solve math problems. *Language and Education, 20*(6), 507-528. https://www.doi.org/10.2167/le678.0

Nunes, T., Bryant, P., Evans, D., Bell, D., Gardner, S., Gardner, A., & Carraher, J. (2007). The contribution of logical reasoning to the learning of mathematics in primary school. *British Journal of Developmental Psychology, 25*(1), 147-166. https://www.doi.org/10.1348/026151006x153127

Petrovici, C., & Havarneanu, G. (2018). Logical-mathematical game in the context of early education. *Review of Science, Mathematics and ICT Education, 12*(2), 61-82.

Reid, D. A. (2002). Conjectures and refutations in grade 5 mathematics. *Journal for Research in Mathematics Education, 33*(1), 5-29. https://www.doi.org/10.2307/749867

Saleh, M., Prahmana, R. C. I., Isa, M., Murni, M. (2018). Improving the reasoning ability of elementary school student through the Indonesian realistic mathematics education. *Journal on Mathematics Education, 9*(1), 41-54. https://www.doi.org/10.22342/jme.9.1.5049.41-54

Sumpter, L., & Hedefalk, M. (2015). Preschool children's collective mathematical reasoning during free outdoor play. *Journal of Mathematical Behavior, 39*, 1-10. https://www.doi.org/10.1013/j.jmathb.2015.03.006

Vale, C., Widjaja, W., Herbert, S., Bragg, L. A., Loong, E. Y.-K. (2017). Mapping variation in children's mathematical reasoning: The case of 'what else belongs?'. *International Journal of Science and Mathematics Education, 15*(5), 873-894. https://www.doi.org/10.1007/s10763-016-9725-y

Vandermaas-Peeler, M., Massey, K., & Kendall, A. (2016). Parent guidance of young children's scientific and mathematical reasoning in a science museum. *Early Childhood Education Journal, 44*(3), 217-224. https://www.doi.org/10.1007/s10643-015-0714-5

Wong, T. T.-Y. (2018). Is conditional reasoning related to mathematical problem solving? *Developmental Science, 21*(5), e12644. https://www.doi.org/10.1111/desc.12644

Yankelewitz, D., Mueller, M., & Maher, C. A. (2010). A task that elicits reasoning: A dual analysis. *Journal of Mathematical Behavior, 29*, 76-85. https://www.doi.org/10.1016/j.jmathb.2010.02.002

Supporting Pre-schoolers' Acquisition of Geometric Knowledge Through Mind Mapping

Türker Sezer 🆔
Bolu Abant Izzet Baysal University

Özgül Polat 🆔
Marmara University

ABSTRACT

Mind mapping refers to the use of a specific graphic organizer to support learning. This paper describes the effect of mind mapping on pre-schoolers' geometric learning. Using a pre- and post-test control group quasi-experimental model, researchers found that the use of mind maps resulted in a statistically significant difference in geometry learning for pre-schoolers (mean age = 65.0 months). These results are discussed in terms of their ramifications for pre-school geometry education, as well as for the use of mind maps with pre-school children.

Keywords: young children; early geometry teaching; mind mapping; preschool education

Introduction

Researchers have studied the development of children's geometric thinking since the 1950s. In 1967, Piaget and Inhelder described three sequential and developmental stages of geometric learning (topological, projective, and euclidean) (Piaget & Inhelder, 1967). Since the early 1980s, research on children's geometric thinking has been profoundly influenced by van Hiele's theory of developmental stages (Olkun & Toluk Uçar, 2007). In this theory, the first level of geometric thinking is visualisation, in which the child uses visual reasoning based on the appearance of whole shapes without thinking about separate components (Levenson et al., 2011). For example, when children are asked why they know a shape is a rectangle, they may say, "*It looks like a window*" rather than talking about sides or angles. The second level is the analytical (analysis) at which a shape is described according to its properties. Properties and components of the shapes are analysed, and properties and rules are discovered experimentally. At this level, children may identify right angles and opposite sides being the same length when discussing a rectangle. At the third level (informal deduction), the properties of the shapes are organised logically and hierarchically. Students move from one trait to another. They also use their definitions to distinguish shapes by referring to these characteristics (van Hiele, 1986, 1999). So, for instance, students at this stage understand that a square is also a rectangle, since the definition of a rectangle is a quadrilateral with four right angles. At the formal deduction level, students can understand the meaning and importance of proof based on axioms, theorems, and definitions, and develop an understanding of it (Olkun & Toluk Uçar, 2007). At the last level (rigor), students construct theorems in different hypothetical systems, analyze, and compare these theorems (Fuys et al., 1988), and form abstract deductions (Usiskin, 1982).

For many years and throughout the world, teachers have struggled to teach geometry effectively, and student achievement has been poor (Abdullah & Zakaria, 2013; Alex & Mammen,

2012; Bal, 2014; Fidan & Türnüklü, 2010). For example, elementary and middle school students in the United States were shown to have difficulty learning basic geometric concepts and solving geometric problems, and are not ready for more complex geometric concepts and test-based skills (Clements et al., 2001; Szinger, 2008). The results of research conducted in Turkey also have similar findings. For example, in a study involving 1,270 primary school students (4th, 5th, 6th, and 7th grade), it was determined that 64.5% of the students were at van Hiele's visual level (Bal, 2014). Similarly, in another study involving 1,644 students (5th grade), it was revealed that 47.9% of the students were at the visual level (Fidan & Türnüklü, 2010).

Problems in Early Geometry Education

Like many areas of mathematics, geometric knowledge builds on itself. Early learning can lead to later achievement. Research reveals that children who received preschool education outperformed those who did not (Fidan & Türnüklü, 2010) and high-quality early education positively affects children's geometric thinking in later grades and on international assessments like the Program for International Assessment (Gamboa & Krüger, 2016; Pholphirul, 2017; Usta & Demirtaşlı, 2018). However, if young students do not learn important geometrical ideas, they have difficulty advancing at high levels (Siew et al., 2013).

Unfortunately, researchers have documented widespread problems in early geometry education internationally (Sunzuma & Maharaj, 2019). Researchers from the United States, Scotland, Israel, and South Africa have found that pre-service teachers exhibited low levels of geometric content knowledge (*US*-Clements, 2003; *UK-Scotland*-Fujita & Jones, 2006a, 2006b; *Israil*-Markovits & Patkin, 2020; *South Africa*-van der Sandt, 2007). In one US study, many teacher candidates could only reach van Hiele level 1, and they were not supported in reaching level 2, the descriptive/analytic level (Clements, 2003). These problems were also documented in US pre-school and primary school teachers, not just teacher candidates (Sarama & Clements, 2009). This was also seen to negatively affect the geometry content and techniques that American teachers taught to their students (Clements & Sarama, 2011). According to the results of the studies held in the UK, Scotland, Israel, Zimbabwe, and Malaysia, teachers also had various problems in terms of geometry skills (see Fujita & Jones, 2006a, 2006b; Markowitz & Patkin, 2020; Sunzuma & Maharaj, 2019). For instance, Markowitz & Patkin, (2020) claimed that some of the teachers had negative attitudes towards geometry, that they had insufficient knowledge of content, that they had difficulty in using the correct mathematical language, and they were influenced by the appearance of the figures. A Malaysian study found that teachers don't take into account the individual differences in students' geometric thinking; and they assume that all students have the same cognitive visual-spatial ability and level of understanding (Abd Wahab et al., 2014). Other researchers documented deficiencies invisualizing and solving the geometric figures stemming from a lack classroom practice, lack of appropriate teaching methods and materials, lack of prior knowledge about basic geometric concepts, insufficient teacher-student interaction, lack of motivation in the classroom, and unsuitable learning environment (Chaudhary, 2019). Researchers have also documented problems with early geometry teaching, including a lack of problem-solving (Sulistiowati et al., 2019), a lack of collaborative learning (Chianson et al., 2010), insufficient technology support (Clements & Sarama, 2011; Rohendi et al., 2018) and not allocating time for geometry as much as the other topics in mathematics (such as counting) (Moss et al., 2015; Sarama, 2002). Even parents neglect geometry in favor of numeracy and counting (Zippert & Rittle-Johnson, 2020).

Research in Turkey mirrors the international research described above (see Kandır et al., 2017; Korkmaz & Şahin, 2020; Inan & Temur-Dogan, 2010; Zembat et al., 2013). For instance, as a result of a current study, it was discovered that although prospective teachers were able to partially identify the mistakes made by the children related to the figures in short stories; they had difficulty in

suggesting strategies to address those mistakes (Korkmaz & Şahin, 2020). In another study, preschool teachers were unable to use appropriate mathematical language and couldn't develop open-ended questions (Kandır et al., 2017). In addition to this, some Turkish teachers have insufficient content knowledge (Inan & Temur-Dogan, 2010), use ineffective methods and techniques (Zembat et al., 2014), and do not focus on children's participation in the activities (Pekince & Avcı, 2016). Furthermore, materials used for the teaching of pre-school geometry are also problematic - most sources do not present examples that are different from the prototypes (Aslan & Aktaş-Arnas, 2007).

Quality early interventions can support young children in developing foundational geometric knowledge (van Hiele, 1986). In 1999, Clements and colleagues argued that the development of geometric thinking is not determined by age, but can be supported with quality instruction. They suggested that observation, measurement, drawing, and modelling can all support geometric learning for young children. These authors also argued that teachers needed to give examples of shapes that displayed the range of the category (e.g., triangles that did not have a base that was horizontal) so that children did not acquire limited images of those shapes (Clements et al., 1999).

Experimental studies have been performed in Turkey in the last decade to promote geometry skills in children 4-6 years old (Gecu-Parmaksiz & Delialioğlu, 2020; Kesicioğlu & Alisinanoğlu, 2014; Kılıç & Şahin, 2019; Korkmaz, 2017; Öngören, 2008; Şen, 2017). First of all, technology support such as using computer-assisted educational programs and augmented reality-based virtual manipulatives were shown to be effective in early geometry education (Gecu-Parmaksiz & Delialioğlu, 2020; Kesicioğlu & Alisinanoğlu, 2014). In addition, the use of concrete manipulatives such as Montessori materials and Froebel gifts also positively affected children's knowledge of geometry (Öngören, 2008; Şen, 2017), as did inquiry-based activities and geometry educational programs developed by researchers (Kılıç & Şahin, 2019; Korkmaz, 2017). These research results showed that when young children's knowledge of shape and space was supported by qualified instructional aids, significant improvements were made.

Unfortunately, there are some problems in the widespread use of the methods and techniques used in all these experimental studies by preschool teachers. For example, the technical knowledge of the teachers and the technological equipment of the school may not be sufficient to carry out technology-supported studies. Similarly, using Montessori and Froebel materials and research-based programs requires costly training. Therefore, it is clear that there is a need for easily accessible, applicable, and cost-effective methods and techniques by teachers to support young children's knowledge of geometry.

Mind Mapping

Mind map refers to a powerful and unique graphic organiser that dates back to the 1960s and agitates all the functions of the brain (words, imagination, numbers, cause, rhythm, pictures, lists, details, colours, and spatial awareness) (Buzan & Buzan, 2009; Buzan et al., 1999). In a mind map, the main object of study (main text) is placed in the center of the page, and then various ideas related to the main text are placed as sub-texts at the end of colored main branches. Sub-branches are then drawn off of the subtext to show new subsidiary ideas. The most complex relationships between texts are formed by adding numbers, images, arrows, or connecting elements to mind maps (Buzan & Buzan, 2002). Relevant information is simplified with keywords. This method can be used as a learning, study, and application tool (Rustler, 2012). In recent years, mind maps have been used to support students in developing a new idea, taking notes to recall the existing information easily, establishing links to understand a complex concept, and developing memory (Buzan & Buzan, 2002).

From a theoretical point of view, mind mapping is based on "*The Dual Coding Theory*" (Paivio, 1991), and "*Conjoint Retention*" (Kulhavy et al., 1985). The basic assumption of both theoretical perspectives is that there are two separate cognitive subsystems that people use to process information:

the verbal and pictorial cognitive subsystems. Verbal information is processed and encoded only in the verbal system, while pictorial information is processed and encoded in both systems. Effective instruction supports students in encoding information in both systems (Paivio, 1991), and the use of visual images in instruction provides especially powerful support, as they are stored and processed in both the verbal and pictorial cognitive subsystems (Sadoski & Paivio, 2013). Therein lies the rationale for mind maps - they support students in creating and storing information in both cognitive systems (Merchie & Van Keer, 2016).

Furthermore, mind mapping is compatible with constructivism. When students create a mind map, they actively connect new information to what they already know (Dhindsa & Anderson, 2011). Mind maps also support students in developing metacognition. They are meta-learning tools (for detailed information, see Buzan, 2020), such as concept maps, and Ven diagrams (see Novak & Gowin, 1984) that allow students to discover how they learn (Ismail et al., 2010; Merchie & Keer, 2016; Stull & Mayer, 2007).

Recent research has shown that mind mapping leads to effective results in reading, writing, mathematics, science, problem-solving, creative thinking, and memory development (Abi-El-Mona & Adb-El-Khalick, 2010; Brinkmann, 2003a; Inayah & Argawati, 2019; Irman, 2019; Khatimah & Rachman, 2018; Mardiyah et al., 2018; Widiana & Jampel, 2016). Most of these studies were conducted with participants in secondary school and higher education levels, and were focused on reading and writing. A small, but increasing, number of studies have been done with preschool-age children, showing that mind mapping supported learning in math, science, language, and critical thinking (Polat et al., 2022; Polat & Aydın, 2021; Polat & Aydın, 2020; Polat et al., 2017; van der Wilt et al., 2019). Taken together, these studies show that mind mapping is an important technique that allows the children to interact directly with objects and events, and supports discussing different points of view, generating questions, searching and analyzing information (Polat & Aydın, 2020). Mind mapping also allows young children to visually represent themes/concepts/objects and to associate them with other themes/concepts/objects. This supports deeper understanding and more long-lasting and accessible memory (Polat, 2021) as well as developing their imagination and memory (Polat et al., 2022). This study extends this work on mind mapping with pre-school children into geometry.

Early Education in Turkey

For children to enter primary school in Turkey, they must be 69 months old in September of that year (https://sgb.meb.gov.tr). Preschool education is not compulsory and differs according to age groups and types of institutions. There are independent kindergartens (36-69 months), kindergartens in primary schools (57-69 months), and nurseries (3-36 months) in a variety of educational institutions. There are also private children's clubs (mixed age), and kindergartens affiliated with religious institutions (36-68 months). The Turkish Preschool Education Program was updated in 2013 and its objectives are to support all areas of children's development, to help students acquire good habits, to prepare them for primary school, to provide a common educational environment for children from diverse backgrounds, and to support their Turkish language development. The program is child-centered, flexible, eclectic, balanced, play-based, and spiraling. It also supports creativity and learning by suggesting a stimulating environment suitable for children (learning centers, etc.). The program emphasizes family involvement, family education, and counseling. Additionally, it asks teachers to use a multidimensional assessment (e.g., portfolio) approach. Finally, the program encourages teachers to organise large group (whole class), small group (collaboration), and individual (considering individual differences) activities in a balanced way, to plan activities in an integrated way, and to carry out activities adapted for children with disabilities (if any, is an inclusion student) (MoNE, 2013).

The Turkish Preschool Education Program describes learning goals for geometry that are aligned with the theories of Piaget and van Hiele, as well as the National Council of Teachers of Mathematics (NCTM) standards (NCTM, 2000; Piaget & Inhelder, 1967; van Hiele, 1986, 1999). Pre-school children are expected to learn about direction and position in space, as well as simple shapes (circle, triangle, square, rectangle, ellipse, edge, corner). The program suggests teaching this content through game-based activities that support the active participation of young children (MoNE, 2013).

Although the philosophy, theoretical foundations, framework, methods, and techniques of the program are clear, several problems have arisen in early geometry teaching in Turkey (Fırat & Dinçer, 2018; Inan & Temur-Dogan, 2010; Sezer & Güler-Öztürk, 2011; Zembat et al., 2013). However, because of the limited scope and small sample sizes of these studies, it is difficult to make general claims about early geometry teaching in Turkey. Again, this study should help the field move towards making more general claims about pre-school geometry teaching and learning in Turkey; claims that could then inform more effective instruction.

Present Study

Educational and psychological studies show that children learn information about shapes when they have active experiences; young children change their judgments about shapes as they gain rich experience with geometry (Olkun & Toluk Uçar, 2007). Therefore, creating a teaching process that supports intuitive thinking in children, uses explanatory, informal teaching as a base from which to start, utilizes concrete materials to support learning, and ensures active participation is vital for the gaining of geometric concepts (Siew et al., 2013; van Hiele, 1986). In particular, children need research-based educational experiences that provide opportunities to attend to and identify properties of shapes and transformations of those shapes (Clements, 2003).

Mind-mapping would seem to be an appropriate teaching method for supporting pre-school students in learning geometry because it supports both verbal and visual presentations of knowledge and helps children create effective mental images (see Kulhavy et al., 1985; Paivio, 1991). Mind-mapping also provides opportunities for discussion, question generation, information search and analysis, imagination, concretization, and association, all of which align with the recommendations for teaching geometry described above (see Polat & Aydın, 2020; Polat et al., 2022). Finally, mind-mapping has been shown to effectively support geometric learning in older students (Bütüner, 2006).

In sum, and according to scholars, mind mapping which affects learning and teaching in a positive way (Liu et al., 2014) will have an influence on introducing geometry skills if prepared and implemented based on the developmental features of pre-schoolers. In addition, it is noteworthy that the number of mind mapping studies conducted with preschool children is quite low in the literature. For the above reasons, this study aims to examine the effect of mind mapping studies on the geometry skills of 5-6-year-old children. A quasi-experimental design with a pre-test and post-test, as well as a treatment and control group were utilized for this study. Hypotheses to be tested for this purpose are as follows:

1.   There is no significant difference between the pre-test scores for geometric knowledge from the experimental group and the control group.
2.   There is a significant difference between the pre-test and post-test scores of the experimental group after the mind mapping studies are applied.
3.   There is a significant difference between the pre-test and post-test scores of the control group after the Preschool Education Program is applied.
4.   There is a significant difference between the post-test scores for geometric knowledge from the experimental group and the control group.

Method

Participants

The study group consisted of two classes, totaling 29 children (age 5-6 years), from a preschool in a city located in the western part of Turkey's Central Anatolian Region. One class served as the experimental group and the other served as the control.

There were 14 children in the experimental group (six girls, eight boys, with a mean age of 64.35 months) and 15 children in the control group (eight girls, seven boys, with a mean age of 65.6 months). In the experimental group, three children had preschool education for one year, and 11 children were in their second year of preschool education. In the control group, four children were in their first year and 11 children were in their second year. Children in the experimental and control groups were also similar in terms of their mother's age, parental education, and parental income.[1]

Measures

The research data were obtained using the "Early Geometry Skills Test-EGST" developed by the researcher to determine the geometric thinking skills of children aged 5-7. Several skills were included in the test. These skills are *recognizing typical examples of basic 2D shapes* (MoNE, 2013; NCTM, 2000), *recognizing typical, atypical, and invalid examples of basic 2D shapes* (Clements et al., 1999; Clements, 2003), *drawing basic 2D shapes* (Clements, 2003; Ho, 2003; NCTM, 2000), *recognizing the edges and corners of basic 2D shapes* (MoNE, 2013; NCTM, 2000), *placing (rotating) basic 2D shapes in the appropriate space* (Clements et al., 1999; Clements, 2004), *making shapes by combining different shapes, making shapes with sticks* (Clements et al., 1999; Clements, 2003), *figure-ground relationships* (NCTM, 2000)*, recognizing 3D shapes (intuitively)* (MoNE, 2009), *pattern* (MoNE, 2009; Olkun & Uçar-Toluk, 2007), *perspective taking* (NCTM, 2000), *construction with blocks* (Clements, 2003; Fisher et al., 2013), and *prediction of a surface of shapes 3D* (Yetkin & Dascan, 2010). Secondly, for content validity, 13 expert opinions were solicited (four from the Department of Mathematics, two from the Department of Classroom Teaching, four from the Department of Preschool Education, one from the Department of Measurement and Evaluation, and two experts who are classroom teachers). The *content validity ratio* (CVR) was calculated for each item of the test. Eighteen items that were lower than the determined values (0.54) were excluded from the test. In addition, it was found that the value of the *content validity index* (CVI) for the entire test had a higher value than the value of the *content validity ratio* (CVI = 0.65 > CVR = 0.54). Therefore, the results of the content validity value showed that the test was valid (Lawshe, 1975). Thirdly, the Frostig Visual Perception Test (r=.670, p< 0.05) and Early Number Test (r=.432, p< .05) were used for the criterion validity. After that, the item discrimination results found that there was a significant difference between 27% of the lower group and 27% of the upper group. In addition, statistically significant differences were found between age groups (5, 6, and 7). Finally, the reliability of the EGST was examined. The reliability coefficient of the test was KR20, 0.853. Pearson's correlation coefficient between the two test halves was 0.697. In addition, the Pearson test-retest correlation coefficient of the test was 0.898.

The test is comprised of 42 items. The scoring of 37 items is on a true-false basis (correct=1, incorrect=0). Scoring for the other five items is calculated by the total score (maximum score 8 for Item four, Item five, and Item seven maximum score of six for Item six, and maximum score of four

---

[1] Most of the mothers of the children in the experimental group are within the age range of 30 to 34 years (n = 7) and have university degrees (n = 10). Similarly, most of their parents are within the age range of 30 to 34 years (n = 12) and have university degrees (n = 12). Furthermore, most of the families had a medium-income level (n = 10). In the control group, most of the children's mothers are within the age range of 35 to 39 years (n = 8) and have university degrees (n = 12). On the other hand, most of their parents are within the higher age range of 40-44 (n = 6) and have university degrees (n = 11) and their families have middle-income (n = 12).

for Item 39). The lowest score that can be obtained from the test is zero, and the highest score is 71 (Sezer & Güven, 2016, 2019). The skill groups of the test items are given below in Table 1.

Table 1

Content of Early Geometry Skills Test

Skill Group	Item No
Recognizing typical examples of basic 2D shapes	(1)
	(2)
	(3)
Recognizing typical, atypical, and invalid examples of basic 2D shapes	(4)
	(5)
	(6)
	(7)
Drawing 2D basic shapes (copying)	(8)
	(9)
	(10)
Recognizing the edges and corners of basic 2D shapes	(11)
	(12)
Placing (rotating) 2D basic shapes into the appropriate space	(13)
	(14)
	(15)
Making new shapes by combining different shapes (imaginarily)	(16)
	(17)
Decomposing a shape	(18)
	(19)
Making new shapes by combining different shapes (with the wooden material)	(20)
	(21)
Making new shapes with the sticks	(22)
	(23)
	(24)
Figure-ground relationships	(25)
	(26)
	(27)
Recognizing 3D shapes (intuitively)	(28)
	(29)
	(30)
	(31)
	(32)
	(33)
Pattern	(34)
	(35)
	(36)
	(37
Perspective-taking	(38)
Building with blocks	(39)
Predicting a surface of 3D shapes	(40)
	(41)
	(42)

Experimental procedure

Preparation for the Experiment: After an initial information meeting with school administration, faculty, and parents, two teachers and the parents of their students volunteered to participate in the research project. Four days of mind mapping training was given to the two participating teachers. This training included theoretical underpinnings of mind maps, examination and analysis of example mind maps, creation of mind maps, and methods for conducting mind map studies with children. Lots were

drawn to determine which of the two classes would be the control group and which would be the experimental group.

Classroom teachers and advisors received training on administering the EGST and demonstrated that they were able to do so. They then administered the pre-test to study participants. Students in the experimental group engaged in mind-mapping exercises with the researchers once a week for four weeks. These exercises were designed to familiarize the students with mind maps, and focused on the themes of toys, animals, babies, and cars. The relevant theme (e.g., animals) was written in the middle of the page and the children were asked to think about the related ideas and then to illustrate their thoughts. Under each picture created by the children, labels or short explanations were written by the teachers and researchers.

Treatment: In November, December, January, and February children in both the experimental and control groups participated in the same activities prepared by the teachers based on the achievements and indicators in the Ministry of National Education Preschool Education Program. Children in the experimental group created individual mind maps on the second Friday of each month. These children created mind maps with the theme circle, triangle, square, and rectangle (see Figure 1).

Figure 1

Example of the Square-themed Mind Mapping Exercise
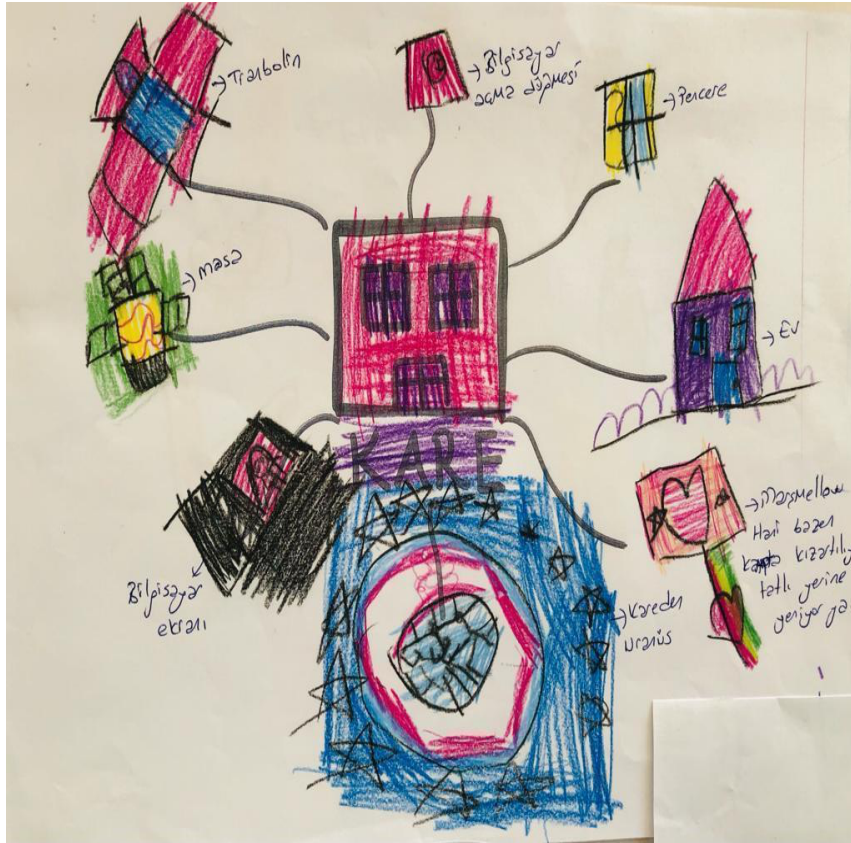
Note: Figure 1 shows a mind map drawing of a 65-month-old boy with a square theme (*Turkish spellings of drawings in parentheses*). In the drawing, there are a table surface (masa yüzeyi), a trampoline (trambolin), a computer power button (bilgisayar açma düğmesi), a window (pencere), a house (ev), a marshmallow (lokum/marşmelov), and a computer screen (bilgisayar ekranı). There is also the planet Uranus, designed using squares in the drawing.

After creating their new mind map, children were given their previous mind map and asked if they wanted to change or add anything to their old map. (e.g., after creating a mind map for the triangle, they were given their previous mind map of the circle and given a chance to revise it.) This allowed children to reinforce previously learned information by repeating it, as well as provided opportunities to add new information, to establish relationships between previous and new knowledge, and to restructure relationships. From March to May, this process was repeated with pentagons, hexagons, and ellipses (see Figure 2). At the end of the experiment, each child had produced 7 mind maps.

At the end of the experiment, children were given the EGST again. After the post tests were administered, mind mapping studies were carried out with the control group starting in the 3rd week of May and into June. These mind mapping exercises followed the same steps as those in the experimental group but were accelerated to take place over a shorter period of time. These maps were not subject to evaluation in the research process and were created with the sole aim of providing equal learning opportunities to all research participants.

Figure 2

Example of the Hexagon-themed Mind Mapping Exercise

Note: Figure 2 shows a mind map drawing of a 63-month-old girl with a hexagonal theme (Turkish spellings of drawings in parentheses). In the drawing, there is a cat (kedi), a mushroom (mantar), a television (televizyon), a sign (tabela), a toy (oyuncak), and a snowman (kardan adam).

Data Analysis

In evaluating the research data, the children's personal information and EGST scores were first analysed and converted into data sets. The data sets were then analysed using a statistical package program. Finally, a Shapiro-Wilk test was used to determine if the data were normally distributed. The normality results are presented in Table 2.

Table 2

Normality Test

Group	Shapiro-Wilk		
	Statistic	**df**	***p***
Exp. group: pre-test	.888	14	.077
Cont. group: pre-test	.932	15	.290
Exp. group: post-test	.944	14	.475
Cont. group: post-test	.944	15	.441

In Table 2, it was concluded that the distribution was normal (p> .05). Therefore, parametric tests were used to analyse the data that were found to have a normal distribution. The Independent group *t*-test was performed for the pre-tests and an analysis of covariance (ANOVA) for the post-tests. In addition, the effect size was calculated and reported.

Results

The means of the EGST scores before and after the test of the experimental and control groups were examined and the results are presented in the tables below.

Table 3

Independent Samples t-test Result of Pre-tests

Group	*n*	*x̄*	*sd*	*t*	*df*	*p*
Experiment	14	42.15	4.39	-.931	27	.360
Control	15	43.73	4.79			

In Table 3, according to the results of the independent samples *t*-test conducted between the pre-test scores of the experimental group (*M* = 42.15, *SD* = 4.39) and control group (*M* = 43.73, *SD* = 4.79) obtained from the EGST, no significant difference was found *t* (27) = 4.30, *p*= .360. These findings showed that children in the experimental and control groups prior to the application (mind mapping) were at a similar level in terms of geometry knowledge.

Table 4

Descriptive Statistics for Pre- and Post-tests

	Pre-test		Post-test		Total	
	X̄	**sd**	**X̄**	**sd**	**X̄**	**sd**
Experiment	42.14	4.39	55.21	4.62	48.68	7.10
Control	43.73	4.79	50.13	4.25	46.10	5.51
Total	42.97	4.59	52.59	5.07	47.78	6.81

These results show that the children in the experimental group (pre-test: *M* = 42.14, *SD* = 4.39; post-test: *M* = 55.21, *SD* = 4.62) and in the control group (pre-test: *M* = 43.73, *SD* = 4.79; post-test: *M* = 50.13, *SD* = 4.25) showed higher levels of geometry skills in the post-tests. A 2x2 ANOVA with test (pre-test, post-test) and group (experiment, control) was performed to determine the effect

of variables on the geometry knowledge of young children. The assumptions required were then tested and homogeneity of variance $F = .149$, $df1 = 3$, $df2 = 54$, $p = .930$ were found to be adequate for the two-way ANOVA. The findings were given in the following table.

Table 5

Two-way ANOVA Results

Source	SS	df	MS	F	η_p^2
Intercept	132396.112	1	132396.112	6483.307***	.992
Group	44.112	1	44.112	2.160	.038
Test	1372.736	1	1372.736	67.222***	.555
Group * Test	161.150	1	161.150	7.891**	.128
Error	1102.738	54	20.421		
Total	135037.000	58			

**p < .01. ***p < .001.

There was no significant difference according to the group $F (1, 54) = 2.160$, $p = .147$, $\eta_p^2 = .038$. On the contrary, it was determined that the effect of the test variable was significant $F (1, 54) = 67.222$, $p = .000$, $\eta_p^2 = .555$. In addition, it was understood that the main effect between group and test variables on EGST scores was significant $F (1, 54) = 7.891$, $p = .007$, $\eta_p^2 = .128$. In the Post Hoc test, the difference in the test variable was found to be in favour of the post-tests. Furthermore, when the main effect between group and test variables was examined, it was determined that there were significant differences in favour of the post-tests of the experimental group. In Figure 3, the interaction graph presented shows the change in the EGST scores of the experimental and control groups as a function of the measurement time.

Figure 3

Interaction Graph of Group and Time Variables
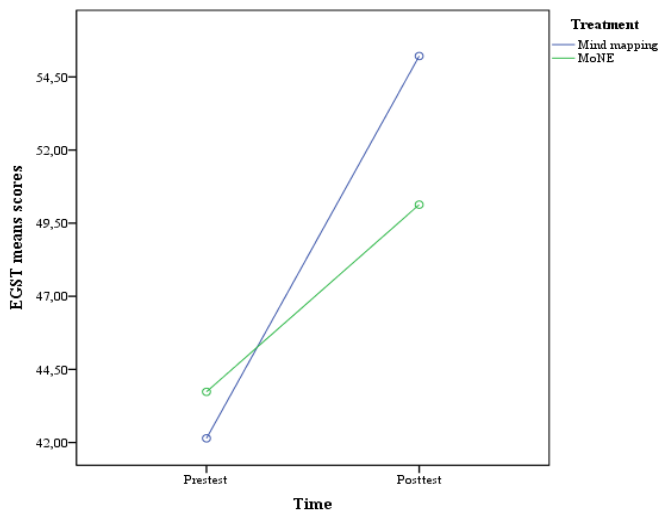
When Figure 3 is examined, it is seen that there is a change in the pre and post-test scores of the experimental and control groups over time. The post-test scores of the children in the experimental group were significantly higher than the post-test scores of the children in the control group.

Results and Discussion

One significant result of this study is that even absent mind mapping, the Ministry of Education Preschool Program supported the geometry learning skills of the children in the control group considerably. For some of the skills measured by the EGST, both groups experienced gains and there were no statistically significant differences in those gains. These skills included recognizing typical examples of basic 2D shapes, recognizing edges and corners of basic 2D shapes, rotating, creating shapes by combining different shapes and sticks, and the relationships between form and ground, recognition of 3D shapes (intuitively), patterning, taking perspective, and estimating a surface of 3D shapes. These findings serve as evidence that theoretical underpinnings of the Pre-school education program in geometry are sound, as well as the constructivist facets of the program, including its child-orientation, flexibility, play-based, emphasis on creativity and exploration, and participation of family.

However, despite the geometric learning demonstrated by the control group, the results of this quasi-experimental study show that mind mapping does have a statistically significant, positive effect on the geometry learning of pre-school children. When compared with a control group who was not taught using mind-mapping, pre-school children who were taught using mind-mapping did better on a post-test of geometry learning. The effect size was seen as medium ($\eta_p^2=.128$; Cohen, 1988) and 13% of the improvement in the experimental group is explained by the application of mind maps.

The EGST scores indicated that the mind mapping exercise contributed to children's learning of how to draw figures, how to distinguish typical, atypical, and invalid examples of the shapes, and how to develop a new shape from different shapes and develop various structures with the blocks. This helps us understand *how* mind mapping supports geometric learning for pre-school children. One way mind mapping supports pre-school children is that it provides opportunities to draw shapes. Mind mapping supports students' drawing skills because students begin the mind-mapping process by drawing the shapes. This is consistent with other studies highlighting the contribution of drawing shapes by children in early geometry education (see Clements, 2003; Ho, 2003; NCTM, 2000).

A second way mind mapping assists pre-school children's geometric learning is by supporting the development of relational thinking. This was seen in the experimental group's better scores in distinguishing between typical, atypical, and invalid examples of shapes. Mind mapping supports relational thinking by supporting students in presenting and processing visual and verbal information together, supporting the creation of mental images (see Kulhavy et al., 1985l; Paivio, 1991), providing opportunities for active participation, discussion, generating questions, searching and analyzing the information, concretization, association (see Polat & Aydın, 2020; Polat et al., 2022), hierarchical organization of the information (concretization of complex relations) (see Buzan & Buzan, 2002). Mind mapping also supports the focus of the brain on details, nourishes associations, develops memory, and eases recall with its multi-dimensional structure (Farrand et al., 2002).

Finally, mind mapping has supported the child's ability to develop a new shape from different figures and various structures with the blocks. The ability to create geometric shapes from different figures and obtaining shapes from one figure to another has been described as a skill group and a significant conceptual area within the field of geometry (NCTM, 2000). Moreover, this important skill is in no way trivial for young learners. In a study, conducted by Copley (2000) it was recorded that 4-year-olds were surprised when they observed that two triangles were obtained after the square was cut by turning it diagonally, and that a square was formed after these triangles were brought together again. Thus, it can be hypothesized that using mind maps supported students' imaginative knowledge, which, in turn, contributed to their relatively greater success on a range of geometric tasks.

In addition to these specific supports, it can be hypothesized that mind mapping provides more general support for students learning geometry. Research results have shown that when

conducting geometry studies with young children, play, concrete materials, and the use of technology have positive effects on the acquisition of geometry knowledge (Fisher et al., 2013; Keren & Fridin, 2014; Kılıç & Şahin, 2019). Furthermore, when the activities supported thinking skills such as classification and comparison, it was concluded that they positively affected the early teaching of geometry (Kılıç & Şahin, 2019; Parks, 2015; Şen, 2017). Mind mapping provides opportunities for fun, an exchange of ideas, comparison, classification, teacher-child interaction, and problem-solving, all elements of effective geometry pedagogy for pre-school children. Another reason mind-mapping may have supported pre-schoolers' geometry learning may have to do with the nature of pre-schoolers' minds. The learning capacity, memory, and attention of young children are limited (Berk & Meyers, 2015). Mind mapping supports understanding by reducing the cognitive burden of learning through visual materials (Schwamborn et al., 2011). It emancipates the learner from unnecessary cognitive burdens (Stull & Mayer, 2007). Hence it ensures the effective use of the limited capacity of memory (Anderson, 2011). Mind maps also work by forcing students to attend to relationships between new information and what they already know (Dhindsa & Anderson, 2011), and by drawing on a wide range of skills such as attention, coordination, reasoning, thinking, analyzing, creativity, imagination, memory, planning, integration, visuality, hearing, and kinesthetics into use (Wang et al., 2010).

This study also builds on the work of researchers who have studied mind mapping with older students. Mind mapping affected language skills (Hsin-Yi, 2015), critical thinking skills (Polat & Aydın, 2020), and school readiness (Polat et al., 2017) in young children. In addition, mind mapping contributed to the teaching of geometry in studies conducted with different grades (high schools) (Kariyana & Sonn, 2016; Loc & Loc, 2020). Kariyana and Sonn (2016) explained that the positive development in students' knowledge of geometry, with the effect of mind maps, is that since learning and teaching become more animated, student motivation increases and it's easier to remember what they learned. Additionally, mind maps can develop new ways for students to take initiative while learning and enhance their creativity and thinking skills (Loc & Loc, 2020). Furthermore, it can be said that mind mapping, which makes mathematics enjoyable (Entrekin, 1992), develops mathematical reasoning (Ayal et al., 2016), and mathematical creativity (Vijayakumari & Kavithamole, 2014) is an effective technique for young children to learn geometry. Mind maps as instructional aids help young children learn because they are manipulable and meaningful (Clements & Sarama, 2007).

Mind mapping is a powerful technique that activates many brain functions (Buzan & Buzan, 2002). It supports words, imagination, numbers, causes, rhythm, pictures, lists, details, colours, and spatial awareness effectively (Buzan et al., 1999), facilitates learning, and ensures permanence (Buzan & Buzan, 2002). All of these happen by allowing learners to visualize the relationships of mathematical objects in a structured way (Brinkmann, 2003b). Mind mapping allows for the support of many skills (thinking about themes, imagination, emotions, attention, identifying the possible causes, reasoning, analyzing, creativity, memory, coordination, reading, writing, visualization, planning, and application) (Wen-Cheng et al., 2010) facilitates the organization of geometric concepts as hierarchies and categories via colours, shapes and pictures, visualization of the relationships between shapes and their features through brainstorming, active participation, and interaction of children (Budd, 2004). Mind mapping, which is based on a constructivist approach, caused children to expand their conceptual schemes. It enabled the children to use their personal experiences, emotions, and intuitions as well (Dhindsa & Anderson, 2011; Ültanır, 2012).

All of the above research indicates that there is great possible potential for the effective use of mind mapping with pre-school children. The claim can be made that this experiment shows that mind maps do, indeed, support pre-school geometry learning. This aligns with van Hiele's contention that geometric learning, although developmental, is not constrained by age, and can be supported and accelerated through the use of effective pedagogy (van Hiele, 1999). According to van Hiele, it is essential for children to explore, discuss, and solve problems in order to learn new ideas. In this

context, it can be suggested that the children in the experimental group had opportunities to do exactly that.

Conclusion

In summary, it was concluded that mind mapping is a technique that can be used in early childhood and a practice that enhances the geometry skills of young children. The results of this study made contributions to the literature both regarding mind mapping and early geometry teaching. In addition, it was understood that the constructivist-based curriculum, which has been applied in preschools in Turkey since 2013, promotes early geometry skills as well. However, the size of the study, and the relatively narrow context, make generalization of the obtained results difficult. More research is needed to test if these findings may hold true for larger and more diverse groups of participants in a wider range of settings.

Another implication of this study is the possibility that geometric thinking levels and skills, which are specified separately from each other, may be considerably less static and discrete. Students often exhibit different van Hiele levels on different questions, and the levels of their thinking show dynamism and continuity, which echoes previous research studies (Burger & Shaughnessy, 1986; Clements & Battista, 1992; Fuys et al., 1988). This indicates that continued research is needed to refine and clarify van Hiele levels and other aspects of children's geometric thinking.

Türker Sezer (sezer_t@ibu.edu.tr) is an assistant professor of Preschool Teacher Education in the Education Faculty at the Bolu Abant Izzet Baysal University of Central Bolu in Turkey. He has national and international experience with early math teaching and preschool teacher education. His major research interests include (1) preparing and supporting pre-service and in-service preschool teachers to teach children under six years, and (2) exploring and comparing teaching strategies and mathematical performance of children in an international context. His recent publications can be found in Participatory Educational Research and the International Journal of Psychology and Education Studies.

Özgül Polat (ozgul.polat@marmara.edu.tr) is an associate professor of Preschool Teacher Education in the Ataturk Education Faculty at the Marmara University of Central Istanbul in Turkey. She has been working in the Preschool Education Program Writing and Dissemination Commission of the Ministry of National Education since 2002. She is also one of the authors and consultants of the 0-18 Age Family Education Programs of the General Directorate of Lifelong Learning. In addition, she is taking part in many national and international projects and has many books on various topics such as Preparation for Primary Education and Primary Education Programs, Reading and Writing Awareness, Visual Reading, Sound and Writing Education, Preschool Education Programs, Sensory Education, Multiple Intelligence Applications, Family Education, Multicultural Education, Children's Capacity Building, etc. Furthermore, she has her signature as a writer, editor, and consultant in many resources on preschool and primary school through first grade education. Her recent publications can be found in Education and Science, Thinking Skills and Creativity, and Early Child Development and Care.

References

Abd Wahab, R., Abu, M. S. B., Bin, A. H., Abdullah, N. A. B. A., & Mokhtar, M. B. (2014, December). The improving of cognitive visual-spatial using google SketchUp in overcoming learning difficulties in geometry. In *International Education Postgraduate Seminar 2014* (pp. 275-285).

Abdullah, H. A., & Zakaria, E. (2013). Enhancing students' level of geometric thinking through van Hiele's phase-based learning. *Indian Journal of Science and Technology*, 6(5), 4433-4446.

Abi-El-Mona, I., & Abd-El-Khalick, F. (2008). The influence of mind mapping on eighth graders' science achievement. *School Science and Mathematics, 108*(7), 298-312. https://doi.org/10.1111/j.1949-8594.2008.tb17843.x

Alex, J. K., & Mammen, K. J. (2012). A survey of South African grade 10 learners' geometric thinking levels in terms of the van Hiele theory. *Anthropologist, 14*(2), 123-129.

Anderson, O. R. (2011). Brain, mind, and the organization of knowledge for effective recall and application. *Learning Landscapes, 5*(1), 45-61. https://ww.doi.org/10.1080/09720073.2012.11891229

Aslan, D., & Aktaş Arnas, Y. (2007). Okul öncesi eğitim materyallerinde geometrik şekillerin sunuluşuna ilişkin içerik analizi [Content analysis of the presentation of geometric shapes in preschool education materials]. *Çukurova Üniversitesi Sosyal Bilimler Enstitüsü Dergisi, 16*(1), 69-80.

Ayal, C. S., Kusuma, Y. S., Sabandar, J., & Dahlan, J. A. (2016). The Enhancement of mathematical reasoning ability of junior high school students by applying mind mapping strategy. *Journal of Education and Practice, 7*(25), 50-58.

Bal, P. A. (2014). Predictor variables for primary school students related to van Hiele geometric thinking. *Journal of Theory and Practice in Education, 10*(1), 259-278.

Berk, L. E., & Meyers, A. B. (2015). *Infants and Children: Prenatal through Middle Childhood* (8th edition). Pearson.

Brinkmann, A. (2003a). Mind mapping as a tool in mathematics education, *Mathematics Teacher, 96*(2), 96-101. https://www.doi.org/10.5951/MT.96.2.0096

Brinkmann, A. (2003b). Graphical knowledge display – mind mapping and concept mapping as efficient tools in mathematics education, *Mathematics Education Review, 16*, 35-48.

Budd, J. W. (2004). Mind maps as classroom exercises. *The Journal of Economic Education, 35*(1), 35-46. https://www.doi.org/10.3200/JECE.35.1.35-46

Burger, F. W., & Shaughnessy, J. M. (1986). Characterizing the van Hiele levels of development in geometry. *Journal for Research in Mathematics Education, 17*(1), 31-48. https://www.doi.org/10.5951/jresematheduc.17.1.0031

Bütüner, S. Ö. (2006). *Açılar ve üçgenler konusunun ilköğretim 7. sınıf öğrencilerine VEE diyagramları ve zihin haritaları kullanılarak öğretimi [Teaching of angles and triangles by using mind maps and vee diagrams to primary school 7th class]*, [Unpublished master thesis], Balıkesir Üniversitesi.

Buzan, T. (2020). *Zihin haritalama: Evrendeki en güçlü düşünce aracını öğrenmek ve kullanmak için tam rehber* (Ç., Yasemin Bayraktar). Olimpos Yayınları. (2018).

Buzan, T., & Buzan, B. (2002). *How to mind map*. Thorsons.

Buzan, T., & Buzan, B. (2009). *The mind map book: Unlock your creativity, boost your memory, change your life*. Pearson

Buzan, T., Dottino, T., & Israel, R. (1999). *The brainsmart leader*. Gower Publishing.

Chaudhary, R. (2019). *A study on difficulties in learning geometry at secondary level students* [Unpublished master's thesis], Tribhuwan University.

Chianson, M. M., Kurumeh, M. S., & Obida, J. A. (2010). Effect of cooperative learning strategy on students' retention in circle geometry in secondary schools in Benue State, Nigeria, *American*

Journal of Scientific and Industrial Research, 2(1), 33-36.
https://www.doi.org/10.5251/ajsir.2011.2.1.33.36

Clements, D. H. (2003). Geometric and spatial thinking in early childhood education. In D. H.
Clements, J. Sarama, & A. DiBiase (Eds.), *Engaging young children in mathematics: Standards for early childhood mathematics education* (pp. 267-297). Routledge.

Clements, D. H., & Battista, M. T. (1992). Geometry and spatial reasoning. In D. A. Grouws (Ed.),
Handbook of research on mathematics teaching and learning (pp. 420-464). Macmillan.

Clements, D. H., & Sarama, J. (2007). Effects of a preschool mathematics curriculum: Summative
research on the Building Blocks project. *Journal for Research in Mathematics Education*, *38*(2),
136-163. https://www.doi.org/10.2307/30034954

Clements, D. H., & Sarama, J. (2011). Early childhood teacher education: The case of geometry.
Journal of Mathematics Teacher Education, *14*, 133-148. https://www.doi.org/10.1007/s10857-011-9173-0

Clements, D. H., Battista, M. T., & Sarama, J. (2001). Logo and geometry. In Ema Yackel (Ed.),
Journal in Research Mathematics Education, Monograph Series, 10 (pp. 177), NCTM.

Clements, D. H., Swaminathan, S., Hannibal, M., & Sarama, J. (1999). Young children's concepts of
shape. *Journal for Research in Mathematics Education*, *30*(2), 192-212.
https://www.doi.org/10.2307/749610

Cohen, J. C. (1988). *Statistical power analysis for the behavioral sciences* (2nd edition.). Lawrence Earlbaum
Associates.

Copley, J. V. (2000). *The young child and mathematics.* National Association for The Education of
Young Children.

Dhindsa, H. S., & Anderson, O. R. (2011). Constructivist-visual mind map teaching approach and
the quality of students' cognitive structures. *Journal of Science Education and Technology*, *20*(2),
186-200. https://www.doi.org/10.1007/s10956-010-9245-4

Entrekin, V. S. (1992). Mathematical mind mapping. *Mathematics Teacher, 85*(6), 444-445.

Farrand, P., Hussain, F., & Hennesy, E. (2002). The efficacy of the 'mind map' study technique.
Medical Education, *36*, 426–431. https://www.doi.org/10.1046/j.1365-2923.2002.01205.x

Fidan, Y., & Türnüklü, E. (2010). İlköğretim 5. sınıf öğrencilerinin geometrik düşünme düzeylerinin
bazı değişkenler açısından incelenmesi [Examination of 5th grade students' levels of
geometric thinking in terms of some variables]. *Pamukkale Üniversitesi Eğitim Fakültesi Dergisi*,
27, 185-197.

Fırat, Z. S., & Dinçer, Ç. (2018). Okul öncesi öğretmenlerin doğal matematiksel dil kullanımlarının
incelenmesi [Examining the use of mathematical language of preschool teachers]. *Abant İzzet
Baysal Üniversitesi Eğitim Fakültesi Dergisi, 18*(2), 895-914.
https://www.doi.org/10.17240/aibuefd.2018..-396525

Fisher, K. R., Hirsh-Pasek, K., Newcombe, N., & Golinkoff, R. M. (2013). Taking shape:
Supporting preschoolers' acquisition of geometric knowledge through guided play. *Child
development*, *84*(6), 1872-1878. https://www.doi.org/10.1111/cdev.12091

Fujita, T., & Jones, K. (2006a). Primary trainee teachers' knowledge of parallelograms. *Proceedings of
the British Society for Research into Learning Mathematics, 26*(2), 25–30.

Fujita, T., & Jones, K. (2006b). Primary trainee teachers' understanding of basic geometrical figures
in Scotland. In J. Novotná, H. Moraová, M. Krátká & N. Stehlíková (Eds.), *Proceedings 30th
conference of the international group for the psychology of mathematics education (PME30)* (Vol. 3, pp.
129–136). Prague, Czech Republic.

Fuys, D., Geddes, D., & Tishler, R. (1988). The van Hiele model of thinking in geometry among
adolescents. *Journal for Research in Mathematics Education. Monograph, 3*, 1-196.
https://doi.org/10.2307/749957

Gamboa, L. F., & Krüger, N. (2016). Does the contribution made by early education to later academic achievement differ in Latin America?: PISA 2009-2012. *CEPAL Review*, *118*, 82-96.

Gecu-Parmaksiz, Z., & Delialioğlu, Ö. (2020). The effect of augmented reality activities on improving preschool children's spatial skills. *Interactive Learning Environments*, *28*(7), 876-889. https://www.doi.org/10.1080/10494820.2018.1546747

Hsin-Yi, W. (2015). The effects of incorporating mind mapping with picture-book reading instruction on the reading comprehension ability of 1st grade elementary school students. *International Journal of Humanities, Arts and Social Sciences*, *1*(4), 147-155. https://www.doi.org/10.20469/ijhss.20003-4

Inan, Z. H., & Dogan-Temur, O. (2010). Understanding kindergarten teachers' perspectives of teaching basic geometric shapes: A phenomenographic research. *ZDM Mathematics Education, 4*(5), 457-468. https://www.doi.org/10.1007/s11858-010-0241-1

Inayah, R., & Argawati, N. O. (2019). Nurturing students' writing narrative interest through mind mapping and cooperative integrated reading and writing. *Indonesian EFL Journal, 5*(2), 121-130. https://www.doi.org/10.25134/iefljv5i2.1781.

Irman, I. (2019). The Effectiveness of mind mapping techniques in counseling of enhancement ability memory students in learning. *Ta'dib, 22*(2), 131-140. http://dx.doi.org/10.31958/jt.v22i2.1424

Ismail, M. N., Ngah, N. A., & Umar, I. N. (2010). The effects of mind mapping with cooperative learning on programming performance, problem solving skill and metacognitive knowledge among computer science students. *Journal of Educational Computing Research*, *42*(1), 35-61. https://www.doi.org/10.2190/EC.42.1.b

Kandır, A., Türkoğlu, D., & Gözüm, A. İ. C. (2017). Assessment of Turkish 2013 Preschool Education Program of The Ministry of National Education (MONE) in terms of mathematical activities. In R. Efe, I. Koleva, E. Atasoy & V. Kotseva (Eds.), *Current Trends in Educational Sciences* (pp. 393-406). ST. Kliment Ohridski University.

Kariyana, I., & A. Sonn, R. (2016). Teaching methods and learners' concept formation, development, and integration in geometry: Assessing the relationship. *International Journal of Educational Sciences*, *12*(1), 75-88. https://www.doi.org/10.1080/09751122.2016.11890415

Keren, G., & Fridin, M. (2014). Kindergarten social assistive robot (KindSAR) for children's geometric thinking and metacognitive development in preschool education: A pilot study. *Computers in Human Behavior, 35*, 400-412. https://www.doi.org/10.1016/j.chb.2014.03.009

Kesicioğlu, O. S., & Alisinanoğlu, F. (2014). Doğrudan öğretim modeline göre hazırlanan eğitim uygulamalarının okul öncesi çocuklarının geometrik şekil öğrenmelerine etkisinin incelenmesi. *Ordu Üniversitesi Sosyal Bilimler Enstitüsü Sosyal Bilimler Araştırmaları Dergisi, 4*(8), 71-77.

Khatimah, K., & Rachman, D. (2018). Mind mapping vs semantic mapping: Which technique gives EFL learners more benefits in reading comprehension?. *Journal of English Educators Society, 3*(2), 165-176. https://www.doi.org/10.21070/jees.v3i2.1498

Kılıç, M., & Şahin, F. T. (2021). Okul öncesi geometri eğitim programının çocukların geometri becerilerine ve şekilsel yaratıcılıklarına etkisi [The effect of preschool geometry education program on children's geometry skills and figural creativity]. *Ondokuz Mayıs Üniversitesi Eğitim Fakültesi Dergisi, 40*(1), 231-256.

Korkmaz, H. İ. (2017). *Doğal açık alanlarda uygulanan sorgulama temelli etkinliklerin çocukların geometrik ve uzamsal düşünme becerilerine etkisi [The effects of inquiry-based activities implemented in natural outdoors on children's geometric and spatial thinking skills],* [Unpublished doctoral dissertation], Hacettepe Üniversitesi.

Korkmaz, H. I., & Şahin, Ö. (2020). Preservice preschool teachers' pedagogical content knowledge on geometric shapes in terms of children's mistakes. *Journal of Research in Childhood Education*, *34*(3), 385-405. https://www.doi.org/10.1080/02568543.2019.1701150

Kulhavy, R. W., Lee, J. B., & Caterino, L. C. (1985). Conjoint retention of maps and related discourse. *Contemporary Educational Psychology*, *10*(1), 28-37. https://www.doi.org/10.1016/0361-476X(85)90003-7

Lawshe, C. H. (1975). A quantitative approach to content validity. *Personnel psychology*, *28*(4), 563-575.

Levenson, E., Tirosh, D., & Tsamir, P. (2011). *Preschool geometry. Theory, research, and practical perspectives.* Sense Publishers.

Liu, Y., Zhao, G., Ma, G., & Bo, Y. (2014). The effect of mind mapping on teaching and learning: A meta-analysis. *Standard Journal of Education and Essay, 2*(1), 17-31.

Loc, N. P., & Loc, M. T. (2020). Using mind map in teaching mathematics: An experimental study. *International Journal of Scientific & Technology Research, 9*(4), 1149-1155.

Mardiyah, I., Suhito, S., & Safa'atullah, M. F. (2018). Analysis of mathematical reasoning ability of junior high school students of grade VII viewed from cognitive style on problem based learning with mind mapping. *Unnes Journal of Mathematics Education, 7*(2), 122-128. https://www.doi.org/10.15294/ujme.v7i2.24883

Markovits, Z., & Patkin, D. (2020). Preschool in-service teachers and geometry: Attitudes, beliefs, and knowledge. *International Electronic Journal of Mathematics Education*, *16*(1), em0619. https://www.doi.org/10.29333/iejme/9303

Merchie, E., & Van Keer, H. (2016). Mind mapping as a meta-learning strategy: Stimulating pre-adolescents' text-learning strategies and performance?. *Contemporary Educational Psychology*, *46*, 128-147. https://www.doi.org/10.1016/j.cedpsych.2016.05.005

MoNE. (2009). İlköğretim Matematik Dersi Öğretim Programı (1–5. Sınıflar) [Elementary Mathematics Curriculum]. T.C. Milli Milli Eğitim Bakanlığı, Talim ve Terbiye Kurulu Başkanlığı, Ankara.

MoNE. (2013). Okul Öncesi Eğitim Programı [Preschool Education Program]. T.C. Milli Eğitim Bakanlığı, Temel Eğitim Genel Müdürlüğü, Ankara.

Moss, J., Bruce, C., & Bobis, J. (2015). Young children's access to powerful mathematics ideas: A review of current challenges and new developments in the early years. In L. D. English & D. Kirshner (Eds), *International handbook on mathematics education* (pp. 153–190). Routledge.

NCTM. (2000). *Principles and Standards for School Mathematics.* Author.

Novak, J. D. & Gowin D. B. (1984). *Learning how to learn.* Cambridge University Press.

Olkun, S., & Toluk-Uçar, Z. (2007). *İlköğretimde etkinlik temelli matematik öğretimi [Activity-based mathematics teaching in primary education]*. Maya Akademi.

Öngören, S. (2008). *Okul öncesi eğitim kurumlarına devam eden 4-5 yaş grubu çocuklarına geometrik şekil kavramı kazandırmada Montessori eğitim yönteminin etkililiği [The effectiveness of Montessori education method in gaining the concept of geometric shape to 4–5-year-old children attending pre-school education institutions.]*. [Unpublished master thesis], Selçuk Üniversitesi.

Paivio, A. (1991). Dual coding theory: Retrospect and current status. *Canadian Journal of Psychology, 45*(3), 255–287. https://www.doi.org/10.1037/h0084295.

Parks, A. N. (2015). *Exploring mathematics through play in the early childhood classroom.* Teachers College.

Pekince, P., & Avcı, N. (2016). Okul öncesi öğretmenlerinin erken çocukluk matematiği ile ilgili uygulamaları: Etkinlik planlarına nitel bir bakış [Preschool teachers' applications related to early childhood mathematics: A qualitative look at activity plans]. *Kastamonu Eğitim Dergisi*, *24*(5), 2391-2408.

Pholphirul, P. (2017). Pre-primary education and long-term education performance: Evidence from Programme for International Student Assessment (PISA) Thailand. *Journal of Early Childhood Research*, *15*(4), 410-432. https://www.doi.org/10.1177/1476718X15616834

Piaget, J., & Inhelder, B. (1967). *The child's concepts of space*. Routledge & Kegan Paul.

Polat, Ö. (2021). Effects of mind mapping studies on the creativity skills of 60 to 72-month-old children. *Education and Science*, *46*(207), 21-42. https://doi.org/10.15390/EB.2021.9582

Polat, Ö., & Aydın, E. (2020). The effect of mind mapping on young children's critical thinking skills. *Thinking Skills and Creativity*, *38*, 100743. https://doi.org/10.1016/j.tsc.2020.100743

Polat, Ö., Küçükoğlu, E. K., & Aydın, E. (2022). Analysis of the effect of mind mapping studies on the acquisition of the concept of time among 60–72-month-old children. *Psycho-Educational Research Reviews*, *11*(1), 254-269. https://www.doi.org/10.52963/PERR_Biruni_V11.N1.16

Polat, O., Yavuz, E. A., & Tunc, A. B. O. (2017). The effect of using mind maps on the development of maths and science skills. *Cypriot Journal of Educational Sciences*, *12*(1), 32-45. https://www.doi.org/10.18844/cjes.v12i1.1201

Rohendi, D., Septian, S., & Sutarno, H. (2018, February). The use of geometry learning media based on augmented reality for junior high school students. In *IOP conference series: Materials science and engineering* (Vol. 306, No. 1, p. 012029). Manado, North Sulawesi, Indonesia: IOP Publishing.

Rustler, F. (2012). *Mind mapping for dummies*. John Wiley & Son.

Sadoski, M., & Paivio, A. (2013). *Imagery and text: A dual coding theory of reading and writing*. Routledge.

Sarama, J. (2002). Early childhood corner: Listening to teachers: Planning for professional development. *Teaching Children Mathematics, 9*(1), 36–39. https://www.doi.org/10.5951/TCM.9.1.0036

Sarama, J., & Clements, D. H. (2009). *Early childhood mathematics education research: Learning trajectories for young children*. Routledge.

Schwamborn, A., Thillmann, H., Opfermann, M., & Leutner, D. (2011). Cognitive load and instructionally supported learning with provided and learner-generated visualizations. *Computers in Human Behavior*, *27*(1), 89-93. https://www.doi.org/10.1016/j.chb.2010.05.028

Şen, P. A. (2017). *Froebel armağanlarının okul öncesi eğitim kurumuna devam eden 60-72 aylık çocukların geometri becerilerine etkileri [Effects of Froebel Gifts on geometry skills of 60 - 72 months old pre-school children]*. [Unpublished master thesis], Okan Üniversitesi.

Sezer, T., & Güler-Öztürk, D. S. (2011). The effects of drama in helping five-year-old children acquire the concepts of number and operation. *Educational Research, 2*(6), 1210-1218.

Sezer, T., & Güven, Y. (2016). Erken Geometri Beceri Testinin geliştirilmesi [The development of Early Geometry Skills Test]. *Atatürk Üniversitesi Kazım Karabekir Eğitim Fakültesi Dergisi*, (33), 1-22.

Sezer, T., & Güven, Y. (2019). 5-7 yaş grubu çocukların geometri becerilerinin incelenmesi [Investigation of geometry skills of 5-7 years old children]. *Erken Çocukluk Çalışmaları Dergisi, 3*(2), 514–540. https://www.doi.org/10.24130/eccd-jecs.1967201932187

Siew, N. M., Chong, L. C., & Abdullah, M. R. (2013). Facilitating students' geometric thinking through van Hiele's phase-based learning using tangram. *Journal of Social Sciences*, *9*(3), 101-111. https://www.doi.org/10.3844/jsssp.2013.101.111

Stull, A. T., & Mayer, R. E. (2007). Learning by doing versus learning by viewing: Three experimental comparisons of learner-generated versus author-provided graphic organizers. *Journal of educational psychology*, *99*(4), 808-820. https://www.doi.org/10.1037/0022-0663.99.4.808

Sulistiowati, D. L., Herman, T., & Jupri, A. (2019). Student difficulties in solving geometry problem based on Van Hiele's thinking level. *Journal of Physics: Conference Series*, *1157*(4), 042118. https://www.doi.org/10.1088/1742-6596/1157/4/042118

Sunzuma, G., & Maharaj, A. (2019). In-service teachers' geometry content knowledge: Implications for how geometry is taught in teacher training institutions. *International Electronic Journal of Mathematics Education*, *14*(3), 633-646. https://www.doi.org/10.29333/iejme/5776

Szinger, I. S. (2008). The evolvement of geometrical concepts in lower primary mathematics (parallel and perpendicular). *Annales Mathematicae et Informaticae, 35*, 173–188.

Ültanır, E. (2012). An epistemological glance at the constructivist approach: Constructivist learning in Dewey, Piaget, and Montessori. *International Journal of Instruction*, *5*(2), 195-212.

Usiskin, Z. (1982). *van Hiele levels and achievement in secondary school geometry: CDASSG project*. National Institution of Education.

Usta, H. G., & Demirtaşlı, R. N. (2018). An international comparison according to PISA 2012 mathematical literacy Turkey and Finland. *Electronic Turkish Studies*, *13*(11), 1389-1420. http://dx.doi.org/10.7827/TurkishStudies.13377

van der Sandt, S. (2007). Pre-service geometry education in South Africa: A typical case?. *IUMPST: The Journal, 1*, 1–9.

van der Wilt, F., Boerma, I., van Oers, B., & van der Veen, C. (2019). The effect of three interactive reading approaches on language ability: An exploratory study in early childhood education. *European Early Childhood Education Research Journal*, *27*(4), 566-580. https://www.doi.org/10.1080/1350293X.2019.1634242

Van Hiele, P. M. (1986). *Structure and Insight: A Theory of Mathematics Education.* Academic Press.

Van Hiele, P. M. (1999). Developing geometric thinking through activities that begin with play. *Teaching Children Mathematics, 5*(6), 310-316. https://www.doi.org/10.5951/TCM.5.6.0310

Vijayakumari, K., & Kavithamole, M. G. (2014). Mind Mapping: A tool for Mathematical Creativity. *Guru Journal of Behavioral and Social Sciences*, *2*(1), 241-246.

Wang, W. C., Lee, C. C., & Chu, Y. C. (2010). A brief review on developing creative thinking in young children by mind mapping. *International Business Research*, *3*(3), 233-238.

Wen-Cheng, W., Chung-Chieh, L., & Ying-Chien, C. (2010). A brief review on developing creative thinking in young children by mind mapping. *International Business Research*, *3*(3), 233-238.

Widiana, I. W., & Jampel, I. N. (2016). Improving students' creative thinking and achievement through the implementation of multiple intelligence approach with mind mapping. *International Journal of Evaluation and Research in Education, 5*(3), 246-254.

Yetkin, D., & Daşcan, Ö. (2010). *İlköğretim Programı (1-5 Sınıflar).* Anı Yayıncılık.

Zembat, R., Sezer, T., Koçyiğit, S., & Balcı, A. (2014). Preschool teachers' views on geometry in preschool period. In M, Yasar, O. Ozgun and J. Galbraith (Ed.), *Contemporary Perspectives and Research on Early Childhood Education* (pp. 153-163), Cambridge Scholars Publishing.

Zippert, E. L., & Rittle-Johnson, B. (2020). The home math environment: More than numeracy. *Early Childhood Research Quarterly*, *50*, 4-15. https://www.doi.org/10.1016/j.ecresq.2018.07.009

ICRSME
International Consortium for Research in
Science & Mathematics Education

A Comparative Study of Fifth-grade Mathematics Textbooks Used in Turkey and Singapore

Zehra Toprak
Independent Researcher

Mehmet Fatih Özmantar
Gaziantep University

ABSTRACT

This study reports on comparing the presentation of mathematical content, types of responses required of students, conformity with the curriculum outcomes, and potentially confusing expressions in the subjects of percentages, triangles, and quadrilaterals of fifth-grade mathematics textbooks used in Turkey and Singapore. Document analysis of the two mathematics textbooks indicated that both books remained committed to achieving the curriculum outcomes, that the content of the Singapore textbook was more simply organized, and that the Singapore textbook included more visual representations and multiple solution strategies. The analysis also showed that the questions requiring numerical responses were common in both textbooks, while the Singapore textbook included relatively more questions requiring both response and explanation. The findings also indicated that the Singapore textbook was comparably more error-free; however, the Turkey textbook involved several difficulties concerning syntactic errors, ambiguity, unrealistic task presentation, redundant operation, unclarity in mathematical purposes, and weak mathematical context. We discuss the usefulness and intelligibility of mathematics textbooks regarding established cultural and philosophical traditions and trust and respect for textbooks.

Keywords: comparative study, curriculum outcomes, mathematics textbooks, content presentation

Introduction

Mathematics textbooks are viewed as powerful instructional tools that could act as sources for teaching activities and instructional ideas and hence, provide a particular version of curricular content in a specific sequence (Reys et al., 2004). In so doing, textbooks provide a vision on what and how teachers should teach and how students will learn (Dole & Shield, 2008; Fan & Zu, 2007). Schmidt et al. (1997) define textbooks as micro-organizers of in-class activities, emphasizing their effects on teachers and teaching. This realization led many to focus their attention on understanding the links between teachers and textbooks. Research in this direction produced compelling evidence that mathematics textbooks significantly determine pedagogical approaches adopted by teachers relevant to the delivery of certain contents (Haggarty & Pepin, 2002). For example, Fan and Zhu (2007) state that many mathematics teachers use textbooks as a source while structuring their teaching approach. In addition, the research provides evidence that textbooks affect teaching strategies (Fan & Kaeley, 2000), and points to the similarities between textbook style and teacher's instructional style (Krammer, 1985). Research findings, however, do not necessarily indicate that teachers readily follow and always comply with the textbook prescriptions. It is well established that teachers are selective while choosing

textbooks and even adapting certain parts from within the book (Love & Pimm, 1996; Remillard, 2005). Brown (2009) argues that affordances and constraints of textbooks affect teachers' utilization and instructional beliefs and pedagogical understandings shape their reliance on and selection of the content from within the text. The author also emphasizes that this bidirectional relationship had a significant effect on instructional practices in the classroom.

Considerable research attention has also been paid to theorize the connections between the curricular documents and textbooks. For instance, Valverde et al. (2002) classify curriculum into four categories: intended, enacted, potentially implemented, attained and, they define textbooks as the potentially implemented curriculum within this framework. For them, by functioning as mediators between intended and enacted curriculum, mathematics textbooks contribute to the attainment of instructional goals through strategies, activities, and applications. Such a view accentuates textbooks as de facto curriculum (Harwood, 2016), and as materials that bridge formal curriculum goals with in-class teaching (Schmidt et al.,1997). In effect, textbooks could be viewed as envoys that largely influence students' mathematics learning opportunities (Stylianides, 2009) and tools that enable the feasibility of curriculum.

Effects of textbooks on students' learning is another research focus in this area. Remillard (2018), for example, states that textbooks are designed to guide, support, and facilitate learning. On the other hand, Macintyre and Hamilton (2010) provide evidence that selection of content and its presentation have a potential effect on students' participation in mathematical activities and hence their achievement. In addition, Sievert et al. (2019) conclude that mathematics textbooks have a significant impact on students' skills to develop creative strategies, and choose and use appropriate solution methods. Mathematical skills to be acquired by students can be present at different levels in textbooks. For example, one textbook may emphasize multiple solutions while another on problem-solving strategies. Sharing this observation, Valverde et al. (2002) note that the features that make textbooks different from each other can create different mathematical expectations and cause variations in students' mathematical performances and achievement levels.

Research studies have also been conducted to gain insights into the features and components of effective textbooks with regard to student learning and achievement. A typical mathematics textbook includes exercises, problems, tasks, activities, worked examples, strategies, definitions, models, and representations designed to enable learning (Remillard, 2018). The researchers relate each of these components and how they are employed in the creation of the content to the effectiveness of textbooks on student learning and achievement. That is, mathematics textbooks, with their pedagogical potential, and physical and structural features, provide affordances and constraints for students' learning opportunities (Hadar, 2017; Törnroos, 2005; van den Ham & Heinze, 2018). For instance, instructional tasks that encourage students to use multiple solution strategies, use multiple representations, and make explanations are found to be positively correlated with student achievement (Stein & Lane, 1996).

Further, the research establishes links between the textual features and students' attitudes toward and understanding of the content. For example, Zevenbergen's study (2001) offers evidence indicating that students need more help while structuring the information in a text-intensive textbook. With a focus on structural features of textbooks, Fan (2010) puts particular importance on consistency, intelligibility, and error-freeness of the content as discriminating characteristics of effective textbooks. Czeglédy and Kovács (2008) point to the necessity of simple, clear, and concise language and design features that make textbooks easy to follow. O'Keeffe (2013, p.8) relates these features to the notion of readability which is defined as "a number of factors which influence the reader, including interest and motivation, legibility of the print, complexity of the words and sentences in relation to the ability of the reader". The research insights shared hitherto suggest that structural and physical features such as intensity of text, the simplicity of the language, the level of error-freeness, readability, and intelligibility of the texts impacted student learning.

Comparative Studies on Textbooks

The textbook can be considered a cultural and historical artifact that mediates and is mediated by teaching and learning traditions by its nature (Rezat, 2006). Cultural and social values affect the development of textbooks, which influences the cultural and social values of future generations (Fan, 2013). Having been produced with the influence of social, cultural, and historical values, textbooks are unique to particular societies in which they are developed. That is, textbooks have features that make them different from those produced by other societies or nations. Such peculiar features of textbooks are described as 'textbook signature' (Charalambous et al., 2010). Comparisons of mathematics textbooks from different countries are essential in coming to know particular textbook signatures and, in turn, shed some light on what societies can learn from each other in teaching and learning mathematics. This line of thinking constitutes one of the motives for researchers to perform comparative analyses of textbooks used in different countries (e.g., Kang, 2014; Takeuchi & Shinno, 2020). The researchers of comparative textbook studies aim to understand and explain the cross-national differences and similarities between learning opportunities presented by different mathematics textbooks (Charalambous et al., 2010), and how content selected by the textbooks vary across countries (Takeuchi & Shinno, 2020). By producing knowledge on instructional practices within particular cultural and educational traditions, comparative studies also allow and encourage the best instructional ideas to travel across borders (Liu, 2019).

A few comparative studies have recently been carried out in Turkey. In this regard, the Turkey textbook has been compared with, for example, that of the USA (Kar & Işık, 2015), Japan (Acar, 2019), and Canada (Kul et al., 2018). One of the main drivers of research on mathematics textbooks in Turkey is Turkish students' consistently low level of achievement in tests such as the Trends in International Mathematics and Science Study (TIMSS) and the Programme for International Student Assessment (PISA). This achievement landscape seems to have prompted research on Singapore's mathematics textbooks, a country with consistently high achievement in these tests. However, the existing comparative research on mathematics textbooks of Turkey and Singapore predominantly focus on the concept of cognitive demand (Engin, 2015; Özgeldi & Esen, 2010; Reçber, 2012; Yılmaz, 2018) and questions (Özer, 2012). Erbaş and Alacacı (2009) compared sixth and seventh-grade mathematics textbooks used in Turkey, the US, and Singapore. These authors found out that the number of activities in mathematics textbooks used in Turkey is higher, textbooks of Turkey and the US included comparably more activities that are connected with real life, Singapore textbooks adopted a more abstract approach, and Turkey textbooks lack significant definitions of concepts. This study also demonstrated that Turkey textbooks have deficiencies in worked examples; which included less modeling compared to Singapore textbooks, while exercises are more enriched through visual aids in both Singapore and the US textbooks. In another study, Erbaş et al. (2012) examined sixth-grade mathematics textbooks used in Turkey, the US, and Singapore in terms of text density, visual elements, organization, weight and the number of topics, and presentation of topics. These authors found that textbooks used in Singapore are distinguished due to their low text density, use of rich visual elements, smooth organization, clarity, and simplicity. A study with a similar focus comparing Turkey 10th grade mathematics textbooks and Singapore mathematics textbooks by Sağlam (2012) showed that less complex problems that require recognition of knowledge/concepts and solution of one-step problems are more common in Singapore textbooks, while Turkey textbooks are richer in terms of student-centered activities.

Although studies comparing Singapore and Turkey mathematics textbooks produced significant results, these studies largely fall short in specifying structural aspects that need to be improved in textbooks used in Turkey. One such inadequacy is related to the approach adopted by textbooks in presenting the content. Hence, comparative studies are needed on textbook approaches adopted to organize and present the mathematical content. Another issue that demands further

consideration is textbooks' conformity with curriculum outcomes and potentially confusing expressions in textbooks. One might tempt to think that there is less need for research on textbooks' conformity with curriculum outcomes, and on potentially confusing expressions. This is on account of the assumption that textbooks are normatively developed based on the curriculum outcomes with expectations of rigorous proofreading and a quality assurance process in the course of textbook development. However, in an examination of two secondary school ninth-grade mathematics textbooks, Akkaya (2016) observed that certain curricular standards and acquisitions were ignored and further determined that scientific and syntactic errors were present.

Furthermore, teachers interviewed in Hıdıroğlu's (2016) study mentioned some errors in a different fifth-grade mathematics textbook and stated that these errors formed barriers for learning. Similarly, Çakır (2009) evaluated a fifth-grade mathematics textbook in line with the teachers' views. His findings revealed that most of the teachers in this research thought that the fifth-grade mathematics textbook contained mathematical errors and redundant details. These findings point to the significance of analyzing mathematics textbooks in terms of their conformity with curriculum outcomes and the potential content errors.

Instructional tasks that encourage students to explain how they arrive at an answer or to justify their solutions are found to be important for learning more than simply asking for numerical answers to a question (Stein & Lane, 1996). It is also stated that these tasks in textbooks do not necessarily lead students to engage in mathematical explanations and justifications; yet, such tasks have a genuine potential to serve these purposes (Stein et al., 1996). Examination of the types of responses required is essential in developing our current understanding of the degree to which textbooks encourage students to make explanations; this is partly because solving a mathematical problem also calls for an explanation of the solution process (Li, 2000). Questions that require students to find as well as explain and justify a solution are instrumental for long-lasting and meaningful learning (National Council of Teachers of Mathematics [NCTM], 2000).

On the basis of our considerations shared hitherto, this paper aims to perform a comparative analysis of fifth grade mathematics textbooks employed in Turkey and Singapore. The comparative analyses have been performed regarding four particular dimensions:

1. Content presentations
2. Type of responses required of the students
3. Conformity with the curriculum outcomes
4. Potentially confusing expressions

With these analyses, this study attempts to contribute to textbook research by designating the similarities and differences between the mathematical competencies, learning opportunities, and learning cultures that Turkey and Singapore consider valuable for their students. The findings and results of this study are also expected to yield theoretical and practical insights that textbook developers, textbook evaluators, policy-makers, and teachers may find helpful for reflection purposes while employing textbooks as part of mathematics instruction.

Methods

This study adopted the document analysis technique that requires a systematic examination and interpretation to understand and make sense of data in print and digital materials (Bowen, 2009). Document analysis was used to compare fifth-grade mathematics textbooks of Turkey and Singapore. The reason for the fifth-grade selection was that this grade corresponds to a transitional stage in both countries. In Turkey, the fifth grade is the transitional stage from primary to secondary school levels. In Turkey, a common practice during primary education is that the classroom teachers teach students

in all subject fields, including mathematics. However, at the secondary level, mathematics instruction is performed by mathematics teachers who are trained and specialized in this area. In Singapore, fifth grade is the transition stage where the subject-based banding approach begins to be practiced. As part of this practice, either foundational mathematics curriculum (intended for those weaker in their foundations), or the standard mathematics curriculum (designed to further students' mathematical accumulation) is followed. Hence this, in a sense, reflects a transition from a single curriculum followed in the initial four years of education to the implementation of a selective program starting from the fifth grade. In this study, we considered the standard mathematics curriculum for Singapore. Turkey's textbook analyzed was *Ortaokul Matematik Ders Kitabı 5* [Middle School Mathematics Textbook 5] (Cırıtcı et al., 2017), the only textbook developed and distributed for fifth-grade after the introduction of the updated mathematics curriculum. Among the other five available mathematics textbooks in Singapore, we selected Singapore's *My Pals are Here 5* (Kheong et al., 2017) because it was one of the most commonly used textbooks.

Mathematics Textbooks Used in Turkey and Singapore

Singapore's Standard Mathematics Syllabus covers numbers and algebra, measurement and geometry, and statistics. Turkey's Middle School Mathematics Curriculum similarly covers numbers and operations, geometry and measurement, and data processing for fifth grade. By placing mathematical problem-solving at the center, Singapore's mathematics curriculum adopts a pentagonal instructional framework that includes five interrelated components: metacognition (monitoring one's thinking, self-regulation of learning), process skills (reasoning, communication, connections, thinking skills and heuristics, applications and modeling), concepts (numerical, algebraic, geometrical, statistical, probabilistic, analytical), skills (numerical calculation, algebraic manipulations, spatial visualization, data analysis, measurement, estimation, use of mathematics tools) and attitudes (beliefs, interest, appreciation, confidence, perseverance). On the other hand, Turkey's mathematics curriculum focuses on problem-solving, process skills, mathematical modeling, building connections, emotional skills, psychomotor skills, information, and communication technology skills.

The current curriculum in Singapore was first introduced in 2013 and has been regularly revised every six years. The Singapore textbook analyzed in this study was the third edition of My Pals are Here 5 released in 2017. The mathematics education curriculum was revised in Turkey in 2017 and, different from the previous one, teaching particular manners and certain values have been selected as the primary foci. The values are embedded in every subject matter curriculum, including mathematics, and teachers are supposed to convey eight core values to students during their instructions: justice, friendship, honesty, self-control, patience, respect, love, responsibility, patriotism, and helpfulness (Talim ve Terbiye Kurulu Başkanlığı [TTKB], 2017). Also, various skills and competencies were specified by considering the EU Competency Framework, National Education Quality Framework, and 21st-century skills.

Singapore Ministry of Education (MoE) assigns the task of textbook development to private publishers to increase diversity (Kaur et al., 2015). To allow school principals, department heads, grade level coordinators, and subject heads to make more informed decisions, Singapore MoE publishes an Approved Textbook list for primary and secondary schools every year in August. It demands schools to choose from the list based on their needs (MoE, 2017). Since the mid-1990s, textbooks have been developed by private publishers in accordance with the mathematics curriculum developed by MoE and are approved by an evaluation committee affiliated with MoE. Publishers are asked to revise the textbooks based on the feedback received from the committee and resubmit them for approval (Fan, 2010).

Similarly, in Turkey, the approval of the central education office must be obtained during the textbook development process. Textbooks are first pre-evaluated by the Directorate of Education

Council. Once approval is secured, the process of panel examination and evaluation of the draft textbook begins. Criteria and rubric to be used in examining and evaluating the textbooks are developed and announced by the Directorate of Education Council. Textbooks are graded by the panelists based on these rubrics. Later, a detailed panel report is drafted with the rationales for the given grade. Based on this report, a final decision is made for a textbook for publication by the central government, which is then freely distributed to all students at the appropriate grade level.

Textbook Analysis Framework

In our attempt to carry out a comparative analysis of Turkey and Singapore textbooks, we initially performed a similarity check for the contents of both books. The headings were crosschecked, and corresponding titles with the mathematical substance were examined in detail. Our examination determined two particular topics with considerably higher matches in terms of substance in both textbooks: (1) percentages and (2) triangles and quadrilaterals. Table 1 presents a brief outline for each topic in both textbooks. Our comparative analyses focused on these two topics due to high content correspondences between both textbooks.

Table 1

Two Mathematical Topics Employed to Perform Comparative Analyses of Textbooks

Turkey			Singapore		
Headings	Learning Domains	Content Outline	Headings	Learning Domains	Content Outline
Percentages	Numbers and Operations	Percent, percentages as decimals/fractions, comparison of percentage expressions, finding the percentage of a quantity	Percentages	Numbers and Algebra	Percent, expressing percentages as fractions/decimals, expressing fractions/decimals as percentages, finding the percentage of a quantity
Triangles and Quadrilaterals	Geometry and Measurement	Polygons; types of triangles and quadrilaterals; the measures of the interior angles of triangles and quadrilaterals; finding unknown angles in triangles and quadrilaterals	Triangles and Quadrilaterals	Geometry and Measurement	Classifying triangles; the sum of the angles in a triangle; right-angled, isosceles and equilateral triangles; finding unknown angles; triangle drawings; Parallelogram, rhombus, and trapezoid; finding unknown angles; quadrilateral drawings

We comparatively analyzed the textbook contents in the two topics based on the following dimensions: (1) presentation of the content, (2) types of responses required of students, (3) conformity with curriculum outcomes, and (4) potentially confusing expressions. Table 2 shows the analytical framework used for each dimension.

Table 2

Analytical Framework

Presentation of Mathematical Content	Type of Responses	Conformity with Curriculum Outcomes	Potentially Confusing Expressions
o How is the subject organization scheme?	o Response only	o Does the textbook fully cover the relevant curriculum outcomes?	o Error resulting from repetition
o How is the subject taught?	o Response and explanation		o Syntactic errors
o Are multiple solution strategies provided?	o Explanation only	o Does the textbook digress from the relevant curriculum outcomes?	o Ambiguity
			o Unrealistic tasks
o Are visual representations used?			o Unclarity in mathematical purposes
o Is technology use encouraged?			o Redundant procedures
			o Weak mathematical context

We utilized type of responses required of students and presentation of the content dimensions of the textbook analysis framework developed by Charalambous et al. (2010) in the process of development of our analysis framework shown in Table 2. We added the first two items in the presentation of mathematical content. We also compared the content of subjects in the textbooks and outcomes when examining the level of textbook's conformity with curriculum outcomes. At this point, it should be noted that, unless otherwise stated, the term curriculum refers in this paper to nationally intended curriculum. Potentially confusing expressions emerged as a result of our content analysis of the textbooks. Further details of the data analysis procedure are provided for each dimension below.

Presentation of Mathematical Content

To develop an organization scheme pertaining to the presentation of the content, we first examined the introduction sections that included the explanations given by the authors regarding how the textbooks must be used. We then examined the sequence of each construct located in the introduction sections from the first subject to the final subject and schematized them. We changed the names of some constructs based on their intended purpose to facilitate readers' understanding further. For example, we changed the construct 'Learn' in the Singapore textbook and the construct 'Let us Do it Together!' in Turkey's textbook into "worked example". We presented it as such in the content organization scheme. Also, in this dimension, we carried out a holistic examination of the subject of percentages in each textbook and reported on the presentation of the topic. We later analyzed this report in accordance with how the subject of percentages was taught. We analyzed the

use of visual representations, encouragement for technology use, the use of multiple solution strategies by focusing on worked examples, and problems/tasks/drill-and-practice questions assigned to students.

Type of Responses Required

We determined each question as the unit of analysis. Although multiple questions were numbered and itemized under each question, we considered each itemized question as a separate one. We coded the questions based on whether they required students to give a response only, both a response and explanation, and an explanation only. More specifically, a question that required a verbal or numerical response was coded as *response only*, a question that required students to explain how they had found a solution in addition to requiring a verbal/numerical response was coded as *response and explanation*, and a question that required students to give an explanation only without asking them to give any verbal/numerical response was coded as *explanation only*.

Conformity with Curriculum Outcomes

To understand if the textbooks covered the curriculum outcomes, and if they went beyond these intended outcomes, we examined the mathematics curriculum outcomes for subjects common in both textbooks (percentages, triangles, and quadrilaterals), the explanations for these outcomes, and the presence of each outcome, and its explanation in the relevant sections of each textbook. To this end, by following the sequence adopted by the textbooks, we carried out a comparative examination of common subjects and the relevant curriculum outcomes and explanations.

Potentially Confusing Expressions

During the analyses, we also tried to note down any expressions that could confuse students. For this, we selected reading texts, activities, games, research-thinking sections, information boxes, speech-thought bubbles, worked examples, and each question students were expected to solve. Through a line-by-line examination of the texts, we identified and categorized any potential issues such as errors induced by repetition, redundancy, and ambiguity that could confuse students. We also referred to the relevant literature, which helped us make sense of the observed errors, and allowed us to come up with categories for potentially confusing expressions. Among the particular studies that inspired us towards this direction were Adler (2000), Sullivan et al. (2003), and Fan (2010). The categories of confusing expressions employed in this study are listed in Table 3 with explanations and particular illustrative examples.

Reliability of the Findings

We invited an expert not directly involved in our study and asked her to audit our data analysis procedures. We first informed the expert on the purpose, method of our study, and the details of our analytical framework. We then asked the expert to conduct an analysis by using one-third of our data. The expert and one of the researchers carried out an independent analysis of the data, and later met for debriefing. The inter-coder reliability was measured as 100 % for the dimension of conformity with curriculum outcomes, 96% for the type of response required, and 90% for the dimension of confusing expressions. We further discussed the codes on which there was a disagreement and reached a consensus. The remaining sections of the book were also analyzed by a researcher based on the discussed and agreed-upon points. Further, analyses were iteratively revisited, the findings were compared with previous results, and, thus, a final version of the findings was established.

Table 3

Coding Scheme for Potentially Confusing Expressions

Category	Explanation	Example
Error resulting from repetition	An unnecessary repetition of words, numbers or expressions	∠PQR and ∠RSO are right right angles (Kheong et al., 2017, p.135). The word "right" in this statement was written twice.
Syntactic errors	Grammatical errors in the verbalizations	With stadium's shopping center and its lounges becoming the center of our city (Cırıtcı et al., 2017, p.165). The grammatical errors create difficulties in understanding the real intention, which could have been better expressed as "The stadium is becoming the center of our city with its shopping center and lounges."
Ambiguity	Ill-defined situations in which it is hardly possible to corroborate the plausibility of an answer or to make sense of the problem demand	Is there any building around you that is still under construction? If so, how much (i.e., in percentage) of this building is completed? (Cırıtcı et al., 2017, p.165) It is difficult for students to provide a sensible answer to the question as students cannot know the final product of the buildings before they are fully constructed. In addition, in order to check the plausibility of the answer, the teacher must also know the building under construction.
Unrealistic task	Assignment of a duty impossible to realize due to uncertainties involved in a situation and/or the scope	Select a species of endemic plant and animal and then state its percentage in relation to the number of species of that specific animal (Cırıtcı et al., 2017, p.181). In this task, students were asked to state, in percent, the number of one endemic plant and of one animal but there are uncertainties involved in the duty (i.e., percent of endemic species in comparison to what?). The scope is also problematic in that it is hardly possible to determine the number of plants or animals in particular species (unless there are particular records held by authorities).
Unclarity in mathematical purposes	Failure or hardships in making sense of what mathematical purpose a shared text/request serves	Is there any mathematical symbol drawn from the symbol of percentage? Search it! (Cırıtcı et al., 2017, p.166). This question may intend to encourage students to do research. However, it is hard to understand what kinds of answers students were expected to come up with; not clear if such a symbol existed; and if it did, how it was related to the symbol and the concept of percentage.
Redundant procedure	Unnecessary or irrelevant operations, steps, or procedures shared in a solution	There is 25% discount applied to televisions in a store. Let's find out how much discount applied to a TV with a price of 800₺. $800 \div 100 = 8$ $8 \times 25 = 200$ ₺ discount will be made. The discounted price of the television is $800 - 200 = 600$₺. (Cırıtcı et al., 2017, p.166) Students were asked to find the amount of discount. Although the amount was found, the textbook continued the solution and although not asked, the discounted price was found. Therefore, the worked example included a redundant solution step.
Weak mathematical context	Insufficiencies in creating a context for mathematical discussions or considerations	What should we be careful about so that endemic living species like these butterflies can continue their lives? (Cırıtcı et al., 2017, p.178) This question intends to increase student participation but the mathematical context it tries to develop is weak.

Findings

We present the findings under four separate headings: presentation of mathematical content, type of responses, conformity with curriculum outcomes, and potentially confusing expressions.

Presentation of Mathematical Content

In this part, we present the schemes regarding the organization by wholly studying both textbooks as a result of our analysis of the sequence of constructs, such as worked examples, questions, activities, and games and developed schemes indicating the organization of the subjects (see Figure 1). We then explain the presentation of the content by analyzing the way the subject is taught, multiple solution strategies, visual representations, and the use of technology in the subject of percentages.

Figure 1

Textbooks' Organization Scheme of Subjects
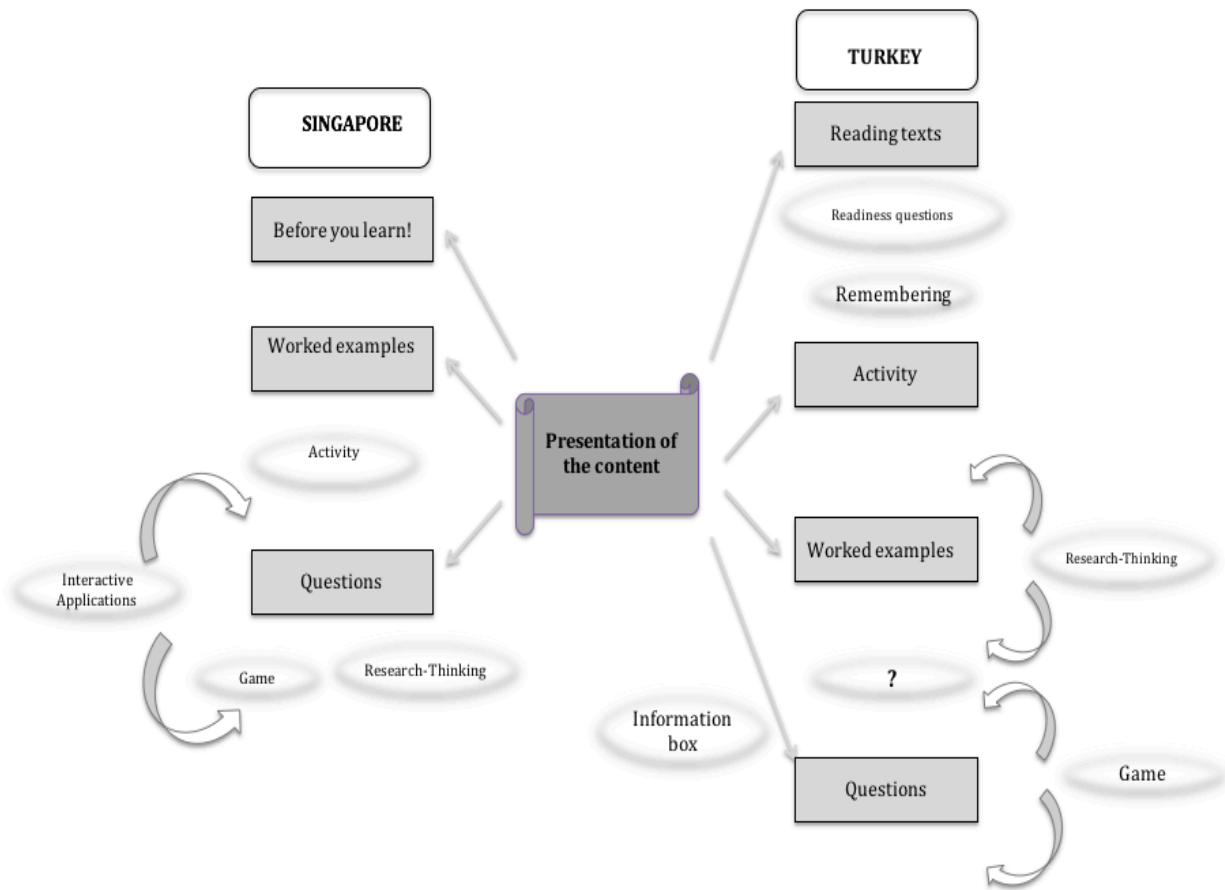
Constructs in bold in the figure above were those that had always been used, and constructs in light color were those used at intervals. Both textbooks included activities, games, worked examples, questions, and research-thinking parts. Different from Turkey's textbook, previous learning was presented under the worked examples in the Singapore textbook. In comparison with the Singapore textbook, the Turkey textbook included reading text and questions aimed at increasing readiness, while the Singapore textbook included 'Before You Learn' and interactive application parts. The Singapore textbook concluded the unit with unit evaluations, unit summary, and non-routine problems after presenting the subject. The Singapore textbook also adopted a much simpler organizational structure and used more subheadings.

Each unit in the Singapore textbook began with a colorful cover that was relevant to mathematical content in the unit and in which cartoon characters spoke and invited students to think. The covers have taken mathematical context as its focus and had designs that students could find appealing. They included numbers that must have enabled students to carry out procedures easily. The mathematical content to be taught was introduced through speech bubbles. Also, the classes to be taught were given, and the main mathematical idea of the unit was presented under the 'Big Idea' heading. The Turkey textbook did not make efforts to create a mathematical context, except in the headings of the subjects to be taught in the colored unit covers.

The introduction of subjects through revisiting students' prior knowledge was periodically evident in both textbooks. Each subject started with a reading text in the Turkey textbook. These reading texts were expected to increase students' interest and invite students to think about mathematical concepts. Some questions related to these texts were also posed. Each class started with a question under 'the Before You Learn' heading in the Singapore textbook. Students were expected to think on the question, develop connections between concepts and visual representations, and occasionally make explanations. Lower text density and the use of plain language were also distinguishing features of the Singapore textbook.

Both textbooks chose to teach the content through reliance on worked examples. Reading texts were followed by an activity that was followed by worked examples in the Turkey textbook. Worked examples were presented in a row, and then several questions were posed altogether in a row. On the other hand, the Singapore textbook first presented a worked example and then a relevant question to consolidate learning. Turkey's textbook included two child cartoon characters (one girl, one boy) with speech bubbles. These speech bubbles contained procedural knowledge, additional knowledge, and a repetition of the expression included in the solution. The Singapore textbook, in contrast, had six real characters (three girls, three boys) depicted as Singaporean kids. They were presented in thought/speech bubbles, which included steps/models for the problem-solving process, a question that can facilitate a specific solution step, and a question to elicit students' learning, information, or alternative solution processes. While Singapore's textbook included an interactive application for each subject accessible under the heading 'app-tivity', the Turkey textbook lacked such a practice. To delve more into the differences between the two textbooks, Table 4 presents the instructional structure of the percentage topic in both textbooks.

Regarding the presentation of mathematical content, we focused on worked examples and problems/tasks/drill-and-practice questions assigned to students to designate the use of visual representations, encouragement for technology use, and the use of multiple solution strategies. To this direction, we initially determined worked examples and questions/problems about percentages, which are presented in Table 5. In the Turkey textbook, there were 12 worked examples and 58 questions/problems. In the Singapore textbook, 13 worked examples along with 89 questions/problems were presented.

Table 4

Instructional Structure of Percentage Topic

Turkey	Singapore
The meaning of the concept of percentage was given after presentation of the reading text, game, and worked example.	The meaning of the concept of percentage was given as soon as the unit started under the heading 'Big Idea' in the unit cover that explained the unit and the lessons to be taught.
Solution strategies in worked examples were not clearly separated from one another.	Solution strategies in worked examples were presented in separate sections under the headings Method 1, Method 2, or in speech bubbles. Sometimes 3 different strategies were presented.
100-square grids were used for the connections between percentage and fractions.	100-square grids and percentage scale were used for the connections between percentage and fractions. A linear scale was used for the connections between percentage and decimals.
A scarce use of models was present.	A frequent use of models is present. In the student directed questions, ready-given models were included, the solution path was divided into steps and gaps were left, that is, guided tasks were presented.
Worked examples on conversion of fractions to percentages were carried out through fractions with a denominator lower and higher than 100.	Worked examples were presented through selection of fractions with a denominator lower and higher than 100 and selection of fractions, the denominators of which, cannot result in 100 through reduction and expansion.
Trio conversions, the conversion of fractions to decimals, and conversion to percentages was given in addition to dual conversions.	Students were asked to convert percentage either to fractions or decimals, and to convert decimals or fractions to percentages, which meant there were only dual conversions.
Discount and wage raise related problems were presented, while there were no interest rates and tax questions.	Discount, wage raise, and tax related problems were presented. The meanings of interest rates, discount, and tax were given in speech bubbles.
Strategy of estimation of the solution before starting the solution process was used only in one problem, and after the solution was completed and estimated, the real solution was not compared.	Polya's four-step problem solving strategy that intends to answer the questions such as: What is given? What is expected? How can I solve it? Is my solution logical? was used only in one problem. However, the justification of the solution was not sought.

Table 5

Examples and Questions Presented in the Percentage Section of the Books

Types	Turkey	Singapore
Worked examples	12	13
Questions/problems	58	89
Total	70	102

Multiple solutions, visual representation, and use of technology in tasks on percentages are shown in Table 6. Since the intensive use of models in the Singapore textbook draws attention, our analyses also included this. As seen in Table 6, the Singapore textbook gives more space for multiple

solutions, visual representation, modeling, and technology than the Turkish textbook. The high use of visual representations in the Singapore textbook (32%) and modelings are two distinct characteristics. The Turkey textbook gave much less space to multiple solutions (4%) and technology (1%).

Table 6

Multiple Solutions, Visual Representation, and Use of Technology in Percentage Section

	Turkey (*n*=70)	**Singapore** (*n*=102)
Use of visual representations	18 (26%)	33 (32%)
Use of modellings	9 (13%)	33 (32%)
Use of multiple solutions	3 (4%)	7 (6%)
Use of technology	1 (1%)	12 (12%)

We can briefly summarize the main differences in the presentation of percentages as follows. While the Singapore textbook chose to introduce the subject with a mathematical question that required students to think and to explain their thinking, the Turkey textbook chose to use reading texts with social contexts to introduce the subject. Worked examples were given by wholesale in a row, and questions were posed altogether in the Turkey textbook. On the other hand, the Singapore textbook prioritized the immediate consolidation of learning by introducing a similar problem right after giving a worked example. The Singapore textbook is different from the Turkey textbook in terms of including a higher number of and more diverse visual representations, more intense use of modelings, utilization of multiple solution strategies in the worked examples, and promotion of technology use. Provision of modelings of problems posed to students, the guidance provided to students by leaving blanks in the problem-solution process, presentation of the subjects in a more simplified organization, less and more cogent explanations, and low text density were among the other salient features of the Singapore textbook.

Type of Responses Required

This part presents findings related to the type of responses for problems in the percentages, triangles, and quadrilaterals. The type of responses both textbooks required of students were examined based on three different categories: response only (RO), response and explanation (RE), explanation only (EO) (see Table 7). It could be observed that the majority of the questions in both textbooks only ask the students for answers and less for explanations. In addition, it has been determined that the questions in the Turkey textbook require only answers at a relatively higher rate. In contrast, the questions in the Singapore textbook encourage students to make more explanations.

Conformity with Curriculum Outcomes

We have observed that both textbooks remained committed to the relevant curriculum outcomes; they covered every target outcome and did not go beyond these outcomes.

Table 7

Type of Responses Required of Students in the Textbooks

	Turkey			Singapore		
	Percentages (n= 58)	Triangles (n=47)	Quadrilaterals (n=33)	Percentages (n=89)	Triangles (n=75)	Quadrilaterals (n=77)
RO	57 (98%)	44 (94%)	32 (97%)	80 (90%)	44 (59%)	56 (73%)
RE	1 (2%)	2 (4%)	1 (3%)	8 (9%)	21 (28%)	20 (26%)
EO	-	1 (2%)	-	1 (1%)	10 (13%)	1 (1%)

Note. RO: response only; RE: response and explanation; EO: explanation only

Potentially Confusing Expressions

The findings of the confusing expressions observed in both textbooks within the scope of two analyzed sections are presented in Table 8, where we shared the frequencies of the determined incidents. Compared with Singapore's, the volume and diversity of confusing expressions were strikingly higher in Turkey's textbook. While we observed one minor error resulting from the repetition in the Singapore textbook, seven instances could mislead or confuse the users of the Turkey textbook. Unclarity in mathematical purposes was the most common difficulty observed in the Turkey textbook.

Table 8

Potentially Confusing Expressions in the Textbooks

Potentially Confusing Expressions	Turkey (*f*)	Singapore (*f*)
Error resulting from repetition	1	1
Syntactic error	1	-
Ambiguity	1	-
Unrealistic tasks	1	-
Redundant operation	1	-
Weak mathematical context	1	-
Unclarity in mathematical purposes	3	-

Discussion

The first focus of our comparative examination was on the presentation of mathematical content. Findings that are relevant to this aspect highlight that the Singapore textbook utilizes more subheadings and adopts a simpler presentation style. The Singapore textbook presents a more comfortably capacious layout with lower text density and more space on the pages. This observation corroborates the findings by Erbaş et al. (2012), who pinpointed low text density and a simpler form as two prominent features of the Singapore textbook. Given that low text density can serve to limit the time needed for completion of instruction and tasks (Morrison et al., 1988) and considering that students need more assistance when faced with higher text density (Zevenbergen, 2001), it could be argued that the Singapore textbook can enable more effective instruction.

Another significant difference between the two textbooks regarding the presentation of the content is related to definitions and explanations of the concepts commonly used in daily life. For instance, the Turkey textbook does not explain the concept of "discount" when presenting discount problems in the percentage topic. On the other hand, if students do not fully comprehend the meaning of these concepts, the definitions of concepts such as tax, interest (the meaning of income and loans), and discount were provided in speech bubbles in the Singapore textbook. The complexity of the words and sentences is one of the most important factors affecting the readability of textbooks (O'Keeffe, 2013). In this respect, we might argue that the Singapore textbook appears to be more concerned with readability. Furthermore, the Singapore textbook also aims at intelligibility (Czeglédy & Kovács, 2008; Fan, 2010) with its simple organizational structure, low writing density, spacious page structure, frequently used visual representations, models, and plain language. One reason for using plain English might be because this textbook is prepared for non-native Singaporean student speakers of English (Wang-Iverson et al., 2009). That Turkey's textbook fails to present explanations on technical terms and the use of a more intense language might generate some issues for students whose mother tongue is not Turkish and who have not yet developed a good command of the Turkish language.

An essential feature of the Singapore textbook was related to the importance attached to the use of multiple solution strategies (even occasional presentation of three different strategies), a feature lacking in the Turkey textbook. Employment of multiple solution strategies is an important feature that characterizes Singapore's mathematics education approach (Soh, 2008) in supporting the development and diversification of learning opportunities (Yoong & Hoe, 2009). The Turkey textbook's failure to include multiple solutions might be limiting students from developing multiple approaches to problems, and from developing skills that could otherwise enable them to handle problems from diverse perspectives.

Singapore textbooks use modelings and visual representations more often. This approach might serve various purposes. Modelings in mathematics education in Singapore are used as tools in analyzing and processing the information and developing logical steps needed for problem-solving (Kaur et al., 2015), in visualizing the problem (Ng & Lee, 2009), and in concretization (Fong & Lee, 2009). This is a well-established practice in Singapore mathematics education, which follows through a three-stage developmental trajectory known as concrete-pictorial-abstract (Kaur et al., 2015). In this approach, ideas are first introduced with concrete materials or practical activities followed by visual illustrations, which are eventually translated into abstract mathematical conceptualizations (see Ginsburg et al., 2005; Jaciw et al., 2016). The effect of this approach on textbook production could easily be observed in many Singapore mathematics textbooks (Soh, 2008). Yoong and Hoe (2009) argue that, with this approach, real-life and abstract ideas are better connected, and concepts and skills are further developed.

Singapore textbook developers frequently use visual representations and modelings due to the positive effects on the development of mathematical understanding (Arcavi, 2003; Naroth & Luneta, 2015). Bora and Ahmed (2019) argue that mathematical modeling supports students' creative and original problem-solving skills, improving their self-perception and self-esteem. Bahmaei (2011) similarly discusses the impact of modeling on enhancing students' knowledge and skills, and helping them grow into more self-confident mathematical thinkers. In the tasks/problems presented in the Singapore textbook, pre-developed modelings are presented to students right after worked examples. Also, spaces related to problem-solving steps and some hints are provided to guide students more effectively. Such an approach comes with pros and cons. For example, the guided tasks adopted by the Singapore textbook may limit students' creativity and cognitive efforts during the problem-solving process.

On the other hand, guided tasks can serve many significant purposes. They could prevent students from developing incorrect models, minimize frustration by allowing them to discover the solution paths, build their self-confidence and self-efficacy, and provide further scaffolding for

struggling students. Turkey textbook's avoidance of guided tasks might avert certain limitations, yet this avoidance could also result in certain disadvantages for students to develop self-confidence and self-efficacy. Given that the trade-offs involved in guided tasks have not been fully resolved, future research on the strategic integration of guided tasks into mathematics textbooks by considering the cultural fabric of different communities could help develop a more informed view.

Singapore has been formally using modelings at the elementary school level since 1983 (Ng & Lee, 2009). Other basic approaches adopted by Singapore textbooks include the employ of modelings in visualizing a problem, using pictorial representations in displaying the connections, and developing strategies for problem-solving (Soh, 2008). Textbooks also include modelings and diagrams to show basic concepts, non-routine tasks, use additional diagrams and pictures to display different thinking methods, and include extra exercises, guided practices, and problems that become more difficult (Ginsburg et al., 2005). These approaches that have been identified with Singapore textbooks are also defined as "textbook signature" for these books (Charalambous, et al., 2010). The Singapore textbook also includes modelings that conform to this signature through visual representations, multiple solution strategies, guided practices, and non-routine tasks. It then becomes evident that Singapore textbooks, including the one analyzed in our study, are developed based on an established instructional philosophy of mathematics education, and that deep historical traditions shape textbooks' content and structural organizations. These traditions seem to evolve into certain educational visions and produce pedagogic approaches that have been tried over the years and have been effective (see also Oates, 2014).

On the other hand, the Turkey textbook appears to suffer from the lack of similar established approaches on mathematics education and the lack of a fundamental philosophy that could have, otherwise, guided the content and structural organization of the textbook. The presence of certain guiding instructional approaches on the content and presentation of a textbook may contribute to the development of an established culture within textbooks. Such established approaches may also potently serve as prevention and mitigation of wasted labor and sources by eliminating the need to initiate new instructional pursuits in every book development process and allowing new textbook and material development under the guidance of certain corroborated principles. That the first edition of the Singapore textbook we examined in this study was dated 2005, the second edition was dated 2008, and the third edition was dated 2017, which underscores the protection of labor needed for textbook development and the refocusing of efforts on improving the existing textbook. Rather than using a textbook for a while and then pulling it out from the market, we argue for the instrumentality of an approach for sustaining, updating, and improving the quality of an existing textbook following emerging instructional innovations, an approach evident in the Singapore textbook.

In the percentage topic, the Turkey textbook showed the estimation of percentages through the use of a calculator only in one worked example, while the Singapore textbook asked students to use a calculator in 12 questions. The Singapore textbook encouraged calculator use in tasks that include big numbers (e.g., 20% of 7.500), and procedures that are difficult to carry out without a calculator (e.g., 7% of 699). Research evidence on the use of calculators in mathematics instruction indicates benefits, such as the potential to assist students to formulate calculations, to interpret these calculations (Ruthven, 2009), in developing concepts, carrying out exercises, improving problem-solving, saving time (Van de Walle, 2007), and in developing positive attitudes (Och & Indoshi, 2011). Calculators are also supported in mathematics education due to their potential in increasing the mathematical expertise of students by making the mathematical tasks more accessible and, thus, enabling students to better cope with complex tasks (Ruthven, 1996). Also, the use of calculators in mathematics textbooks is argued to be an effective instructional practice when mathematical tasks do not aim not to find a calculation-based solution, but to solve a complex problem (Thompson & Sproule, 2000). Therefore, the use of calculators can be further encouraged in the problems that include larger numbers when the aim is to discover the potential links in the given data.

Regarding the second focus of our study, which is the type of responses required of students, both textbooks predominantly presented questions/problems expecting students to give only a numerical response. However, the Singapore textbook tended to elicit more explanations from the students than that of Turkey. Mathematical explanations encourage students to think on mathematical ideas, open these ideas for further discussions, review, and update. All these features contribute to students' mathematical development and support students' communication in mathematics (NCTM, 2000). Providing students with the opportunities to explain their solutions in mathematics textbooks is significant in the social construction of mathematics inside the classroom (see Yackel, 2001). Given the possible effects of explanations on students' mathematical growth and understanding, we support the further inclusion of more tasks that require students to make mathematical explanations.

The use of tasks that require only numerical responses seems to be a global practice. As a matter of fact, Li (2000) similarly showed that most problems included in both the US' and China's mathematics textbooks required only numerical responses from students. The author reported that only 19 % of the problems in the US textbook expected students to make an explanation for solutions and that no problem in the Chinese textbook required explanation for any solution. Törnqvist's (2019) study that examined Sweden's and China's textbooks and the study on Kosova's and Albania's conducted by Vula et al. (2015) also confirmed the common practice of questions that require numerical response only. In examining eighth grade Singapore and Turkey mathematics textbooks, Özer (2012) also showed the predominant use of tasks requiring numerical responses in both textbooks. However, Özer (2012) found that Turkey textbooks at that grade required comparably more explanations from students.

Regarding the third focus of our study (conformity with the curriculum outcomes), it has been determined that the textbooks remain loyal to the country's mathematics curriculum, include each target acquisition, and do not go beyond the acquisitions. This could be because compliance with the officially issued curriculum documents is considered an important criterion for the textbook production process in both countries and that care was given to realize the expected compliance. In fact, compliance to the national set of standards devised by the official curriculum is an important condition (Fan, 2010; TTKB, 2015) employed to evaluate the adequacy of textbooks in both Turkey and Singapore.

As to this study's fourth and final focus, our observations of potentially confusing expressions in percentages, triangles, and quadrilaterals subjects (only one in the Singapore textbook and nine in the Turkey textbook with seven different categories) are also interesting. The presence of only one error caused by the repetition of the same word twice in the Singapore textbook is a testament to higher standards for textbook development and a more rigorous evaluation process. This is supported by Kaur et al.'s (2015) argument on meticulous mathematics textbook preparation and the approval process in Singapore. Fan's (2010) contention that publishers seriously consider the input given by the approval committee and carefully revise the book accordingly also helps explain our result. Our study has categorized and defined seven difficulties that potentially reduced the textbook's instructional and practical value, intelligibility, and readability: repetition, syntactic error, ambiguity, unrealistic tasks, unclarity in mathematical purposes, redundant operation, and weak mathematical context. The Turkey textbook we examined in this study included all these types of difficulties, and others (e.g., Çakır, 2009; Hıdıroğlu, 2016) also reported similar observations in different fifth-grade mathematics textbooks. Questions that have unclarity in mathematical purposes, procedures that are redundantly lengthened by the continuation of further solutions, tasks that aim at increasing interest but have weak mathematical context, unrealistic tasks, and ambiguous expressions might increase students' cognitive load and diminish their motivation for learning. Such potentially confusing expressions might also decrease users' trust in textbooks and compromise respect for the textbooks.

Turkey's textbook examined in this study is distributed to millions of students and thousands of teachers free of charge. It is evident that expressional issues present in this book influence students'

learning opportunities and teachers' interpretation of the content. This might negatively affect the productive application of the content presented through the book. Research focusing on the use of mathematics textbooks by teachers in Turkey (e.g., Özmantar et al., 2017) showed that although textbooks are distributed at no cost, teachers commonly refrain from using these textbooks. While explaining their reluctance to use Turkish textbooks, teachers often referred to the difficulties related to the ones observed and reported in our study. Based on analysis of potentially confusing expressions that emerged from our study and the prior research, we can arguably suggest that expressional issues form a barrier for the achievement of outcomes set by the textbooks and have the potential to influence readability negatively. We also wish to point out the need for further studies to develop organizational and systematic structures to detect such issues in the textbooks before they are published.

Zehra Toprak (zehratoprak0@gmail.com) worked as a middle school mathematics teacher for nine years. She received her PhD degree in mathematics education from Gaziantep University, Turkey. Her research focuses on mathematics textbooks, mathematical reasoning, and activity-based mathematics teaching.

Mehmet Fatih Özmantar (ozmantar@gantep.edu.tr) is a professor of mathematics education at Gaziantep University, Turkey. He earned his PhD from Leeds University, in the UK. His research focuses on mathematical task design, mathematics teacher education, socio-cultural theories of teaching and learning, program development, and historical investigations as well as ethical issues involved in instructional practices.

References

Acar, G. (2019). *Türkiye ve Japonya 4, 5 ve 6. sınıf matematik ders kitaplarındaki problem yapılarının karşılaştırmalı olarak incelenmesi* [Comparative analysis of structure of problems in Turkey and Japan 4, 5 and 6. grade mathematics textbooks] [Master's Thesis, Erciyes University]. Yüksek Öğretim Kurulu National Theses Repository.

Adler, J. (2000). Conceptualising resources as a theme for teacher education. *Journal of Mathematics Teacher Education, 3*(3), 205–224.

Akkaya, G. (2016). *Ortaöğretim dokuzuncu sınıf matematik ders kitaplarının öğretim programına uygunluğu açısından incelenmesi* [The appropriateness of the ninth grade mathematics textbooks regarding the high school mathematics curriculum] [Master's Thesis, Akdeniz University]. Yüksek Öğretim Kurulu National Theses Repository.

Arcavi, A. (2003). The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics*, *52*(3), 215-241.

Bahmaei, F. (2011). Mathematical modelling in primary school, advantages and challenges. *Journal of Mathematical Modeling and Application*, *1*(9), 3-13.

Bora, A., & Ahmed, S. (2019). Mathematical modeling: An important tool for mathematics teaching. *International Journal of Research and Analytical Reviews*, *6*(2), 252-256.

Bowen, G. A. (2009). Document analysis as a qualitative research method. *Qualitative Research Journal*, *9*(2), 27-40.

Brown, M. W. (2009). The teacher–tool relationship: Theorizing the design and use of curriculum materials. In J.T. Remillard, B.A. Herbel-Heisenmann & G. M. Lloyd (Eds.), *Mathematics teachers at work: Connecting curriculum materials and classroom instruction* (pp. 17-36). Routledge.

Charalambous, C. Y., Delaney, S., Hsu, H. Y., & Mesa, V. (2010). A comparative analysis of the addition and subtraction of fractions in textbooks from three countries. *Mathematical Thinking and Learning, 12*(2), 117-151.

Cırıtcı, H., Gönen, İ., Kavas, D., Özarslan, M., Pekcan, N., & Şahin, M. (2017). *Ortaokul matematik ders kitabı 5* [Middle School Maths textbook 5]. Milli Eğitim Bakanlığı Yayınları [The Ministry of National Education Press].

Czeglédy, I., & Kovács, A. (2008). How to choose a textbook on mathematics?. *Acta Didactica Napocensia*, *1*(2), 16-30.

Çakır, İ. (2009). *İlköğretim 5. Sınıf matematik ders kitaplarının öğretmen ve öğrenci görüşleri doğrultusunda değerlendirilmesi* [The evaluation of the fifth grade mathematics textbooks of the primary education according to the views of the teachers and students] [Master's thesis, Çukurova University]. Yüksek Öğretim Kurulu National Theses Repository.

Dole, S. & Shield, M. (2008). The capacity of two Australian eighth-grade textbooks for promoting proportional reasoning. *Research in Mathematics Education*, *10*(1), 19-35.

Engin, Ö. (2015). *Türkiye 7. sınıf matematik ders kitabındaki etkinliklerin bilişsel istem düzeylerinin program ve farklı ülkelerle karşılaştırılması* [A comparison of the cognitive demand levels of tasks in 7th grade mathematics textbook in Turkey with those in the national curriculum and the textbooks of other countries] [Master's thesis, Ankara University]. Yüksek Öğretim Kurulu National Theses Repository.

Erbaş, A. K., & Alacacı, C. (2009, October 10). *6 ve 7. sınıf Türk matematik ders kitaplarının Amerikan ve Singapur ders kitapları ile karşılaştırmalı bir analizi* [A comparative analysis of 6th and 7th grade Turkish mathematics textbooks with American and Singaporean textbooks]. http://users.metu.edu.tr/erbas/resources/textbooks_executive_summary.pdf

Erbaş, A. K., Alacacı, C., & Bulut, M. (2012). A comparison of mathematics textbooks from Turkey, Singapore, and the United States of America. *Educational Sciences: Theory and Practice*, *12*(3), 2324-2329.

Fan, L. (2010, July). *Principles and processes for publishing textbooks and alignment with standards: A case in Singapore.* https://www.apec.org/Publications/2010/07/Replicating-Exemplary-Practices-in-Mathematics-Education-among-APEC-Economies-July-2010

Fan, L. (2013). Textbook research as scientific research: Towards a common ground on issues and methods of research on mathematics textbooks. *ZDM Mathematics Education, 45*(5), 765-777.

Fan, L., & Kaeley, G. S. (2000). The influence of textbooks on teaching strategies: An empirical study. *Mid-Western Educational Researcher, 13*(4), 2-9.

Fan, L., & Zhu, Y. (2007). Representation of problem-solving procedures: A comparative look at China, Singapore, and US mathematics textbooks. *Educational Studies in Mathematics*, *66*(1), 61-75.

Fong, N. S., & Lee, K. (2009). Model method: A visual tool to support algebra word problem solving at the primary level. In W.K. Yoong, L.P. Yee, B. Kaur, F.P. Yee, & N.S. Fong (Eds.), *Mathematics education: The Singapore journey* (pp.169-203). World Scientific Publishing.

Ginsburg, A., Leinwand, S., Anstrom, T., & Pollock, E. (2005, January 28). *What the United States can learn from Singapore's world-class mathematics system (and what Singapore can learn from the United States): An exploratory study.* American Institutes for Research. https://files.eric.ed.gov/fulltext/ED491632.pdf.

Hadar, L. L. (2017). Opportunities to learn: Mathematics textbooks and students' achievements. *Studies in Educational Evaluation*, *55*, 153-166.

Haggarty, L., & Pepin, B. (2002). An investigation of mathematics textbooks and their use in English, French and German classrooms: Who gets an opportunity to learn what?. *British Educational Research Journal*, *28*(4), 567-590.

Harwood, N. (2016). What can we learn from mainstream education textbook research?. *RELC Journal*, *48*(2), 264–277.

Hıdıroğlu, Ç. N. (2016). *Ortaokul 5. sınıf matematik dersi öğretim programının kesirler ünitesinin değerlendirilmesi* [Evaluation of fractions unit of secondary school 5th grades' mathematics curriculum] [Master's Thesis, Pamukkale University]. Yüksek Öğretim Kurulu National Theses Repository.

Jaciw, A. P., Hegseth, W. M., Lin, L., Toby, M., Newman, D., Ma, B., & Zacamy, J. (2016). Assessing impacts of math in focus: A "Singapore math" program. *Journal of Research on Educational Effectiveness*, *9*(4), 473-502.

Kang, J. H. (2014). A cross-national comparative study of first-and fourth-grade math textbooks between Korea and the United States. *Curriculum and Teaching Dialogue*, *16*(1-2), 91-108.

Kar, T., & Işık, C. (2015). Türk ve Amerikan yedinci sınıf matematik ders kitaplarının tamsayılarla toplama ve çıkarma işlemleri üzerinden karşılaştırılması [Comparison of Turkish and American seventh grade mathematics textbooks in terms of addition and subtraction operations with integers]. *Education and Science*, *40*(177), 75-92.

Kaur, B., Soh, C. K., Wong, K. Y., Tay, E. G., Toh, T. L., Lee, N. H., Ng, F. S., Dindyal, J., Yen, Y. P., Loh, M. Y., Tan, H. C. J., & Tan, L.C. (2015). Mathematics education in Singapore. In S. J. Cho (Ed.), *The proceedings of the 12th international congress on mathematical education: Intellectual and attitudinal challenges* (pp. 311-316). Springer.

Kheong, F. H., Soon, G. K. & Ramakrishnan, C. (2017). *My pals are here: Maths 5B.* Marshall Cavendish Education.

Krammer, H. P. (1985). The textbook as classroom context variable. *Teaching and Teacher Education, 1*(4), 273-278.

Kul, Ü., Sevimli, E., & Aksu, Z. (2018). A comparison of mathematics questions in Turkish and Canadian school textbooks in terms of synthesized taxonomy. *Turkish Journal of Education*, *7*(3), 136-155.

Li, Y. (2000). A comparison of problems that follow selected content presentations in American and Chinese mathematics textbooks. *Journal for Research in Mathematics Education*, *31*(2), 234-241.

Liu, Y. (2019, May). *A comparison study of using origami as a teaching tool in middle-school mathematics class in North America and China.* https://scholar.uwindsor.ca/cgi/viewcontent.cgi?article=1097&context=major-papers.

Love, E., & Pimm, D. (1996). 'This is so': A text on texts. In A. J. Bishop, K. Clements, C. Keitel, J. Kilpatrick & C. Laborde (Eds.), *International handbook of mathematics education* (Vol. 1, pp. 371-409). Springer.

Macintyre, T., & Hamilton, S. (2010). Mathematics learners and mathematics textbooks: A question of identity? Whose curriculum? Whose mathematics?. *The Curriculum Journal, 21*(1), 3-23.

Ministry of Education (2021, October 18). *Approved textbook list.* https://www.moe.gov.sg./education-in-sg/approved-textbook-list

Morrison, G. R., Ross, S. M., & O'Dell, J. K. (1988). Text density level as a design variable in instructional displays. *Educational Technology Research and Development*, *36*(2), 103–115.

Naroth, C., & Luneta, K. (2015). Implementing the Singapore mathematics curriculum in South Africa: Experiences of foundation phase teachers. *African Journal of Research in Mathematics, Science and Technology Education*, *19*(3), 267-277.

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

Ng, S. F., & Lee, K. (2009). The model method: Singapore children's tool for representing and solving algebraic word problems. *Journal for Research in Mathematics Education*, *4*(3), 282-313.

Oates, T. (2014, November).*Why textbooks count.* https://www.cambridgeassessment.org.uk/Images/181744-why-textbooks-count-tim-oates.pdf

Och, J. P., & Indoshi, F. C. (2011). Challenges and benefits of using scientific calculators in the teaching and learning of mathematics in secondary school education. *Journal of Media and Communication Studies*, *3*(3), 102-111.

O'Keeffe, L. (2013). A framework for textbook analysis. *International Review of Contemporary Learning Research-An International Journal*, *2*(1), 1-13.

Özer, E. (2012). *Türkiye 8. sınıf matematik konularına göre Türkiye, Singapur ve ABD kitaplarındaki soruların karşılaştırmalı analizi* [*Based on the topics in Turkey's eight grade mathematics curriculum a comparative analysis of the exercises in the textbooks of US, Singapore and Turkey*] [Master's Thesis, Ankara University]. Yüksek Öğretim Kurulu National Theses Repository.

Özgeldi, M., & Esen, Y. (2010). Analysis of mathematical tasks in Turkish elementary school mathematics textbooks. *Procedia-Social and Behavioral Sciences*, *2*(2), 2277-2281.

Özmantar, M. F., Dapgın, M., Çırak Kurt, S., & İlgün, Ş. (2017). Mathematics teachers' use of source books other than textbooks: Reasons, results and implications. *Gaziantep University Journal of Social Sciences, 16*(3), 741-758.

Remillard, J. T. (2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211-246.

Remillard, J. T. (2018). Examining teachers' interactions with curriculum resource to uncover pedagogical design capacity. In L. Fan, L. Trouche, C. Qi, S. Rezat, J. Visnovska (Eds.), *Research on mathematics textbooks and teachers' resources* (pp. 69-88). Springer.

Reçber, H. (2012). *Türkiye 8. sınıf matematik ders kitabındaki etkinliklerin bilişsel düzeylerinin programdakilerle ve ülkeler arası karşılaştırılması* [An international comparison of level of cognitive demands of the tasks in Turkey's 8th grade mathematics curriculum and textbook] [Master's Thesis, Ankara University]. Yüksek Öğretim Kurulu National Theses Repository.

Reys, B. J., Reys, R. E., & Chavez, O. (2004). Why mathematics textbooks matter. *Educational Leadership*, *61*(5), 61–66.

Rezat, S. (2006). A model of textbook use. In J. Novotná, H. Moraová, M.Krátká & N. Stehlíková (Eds.), *Proceedings of the 30th conference of the international group for the psychology of mathematics education* (pp. 409-416). Charles University in Prague.

Ruthven, K. (2009). Towards a calculator-aware number curriculum. *Mediterranean Journal for Research in Mathematics Education*, *8*(1), 1-14.

Ruthven, K. (1996). Calculators in the mathematics curriculum: The scope of personal computational technology. In A. J. Bishop, M.A.K. Clements, C. Keitel-Kreidt, J. Kilpatrick & C. Laborde (Eds.), *International handbook of mathematics education* (pp. 435-468). Kluwer Academic Publishers.

Sağlam, R. (2012). *A comparative analysis of quadratics in mathematics textbooks from Turkey, Singapore, and the International Baccalaureate Diploma Programme* [Master's Thesis, Bilkent University]. Yüksek Öğretim Kurulu National Theses Repository.

Schmidt, W. H., McKnight, C. C., & Raizen, S. A. (1997). *A splintered vision: An investigation of U.S. science and mathematics education.* Kluwer Academic Publishers.

Sievert, H., van den Ham, A. K., Niedermeyer, I., & Heinze, A. (2019). Effects of mathematics textbooks on the development of primary school children's adaptive expertise in arithmetic. *Learning and Individual Differences, 74*, 1-13. https://www.doi.org/10.1016/j.lindif.2019.02.006

Soh, C. K. (2008). An overview of mathematics education in Singapore. In Z. Usiskin & E. Willmore (Eds.), *Mathematics curriculum in pacific rim countries-China, Japan, Korea, and Singapore: Proceedings of a conference* (pp. 23-36). Information Age Publishing.

Stein, M. K., & Lane, S. (1996). Instructional tasks and the development of student capacity to think and reason: An analysis of the relationship between teaching and learning in a reform mathematics project. *Educational Research and Evaluation, 2*(1), 50-80.

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*(2), 455-488.

Stylianides, G. J. (2009). Reasoning-and-proving in school mathematics textbooks. *Mathematical Thinking and Learning, 11*(4), 258-288.

Sullivan, P., Zevenbergen, R., & Mousley, J. (2003). The contexts of mathematics tasks and the context of the classroom: Are we including all students?. *Mathematics Education Research Journal, 15*(2), 107-121.

Takeuchi, H., & Shinno, Y. (2020). Comparing the lower secondary textbooks of Japan and England: A praxeological analysis of symmetry and transformations in geometry. *International Journal of Science and Mathematics Education, 18*(4), 791-810.

Talim ve Terbiye Kurulu Başkanlığı (2017). *Müfredatta yenileme ve değişiklik çalışmalarımız üzerine* [On our curriculum renewal and change studies]. https://ttkb.meb.gov.tr/meb_iys_dosyalar/2017_07/18160003_basin_aciklamasi-program.pdf

Thompson, A. D., & Sproule, S. L. (2000). Deciding when to use calculators. *Mathematics Teaching in the Middle School*, *6*(2), 126-129.

Törnqvist, C. (2019). *A comparative study of the probability content in upper secondary school textbooks in Sweden and China.* http://muep.mau.se/handle/2043/27639

Törnroos, J. (2005). Mathematics textbooks, opportunity to learn and student achievement. *Studies in Educational Evaluation*, *31*(4), 315-327.

Valverde, G. A., Bianchi, L. J., Wolfe, R. G., Shmidt, W. H., & Houang, R. T. (2002). *According to the book: Using TIMSS to investigate the translation of policy into practice through the world of textbooks.* Kluwer Academic Publishers.

Van de Walle, J. A. (2007). *Elementary and middle school mathematics: Teaching developmentally*. Pearson Education.

Van den Ham, A. K., & Heinze, A. (2018). Does the textbook matter? Longitudinal effects of textbook choice on primary school students' achievement in mathematics. *Studies in Educational Evaluation, 59,* 133-140. https://www.doi.org/10.1016/j.stueduc.2018.07.005

Vula, E., Kingji-Kastrati, J., & Podvorica, F. (2015). A comparative analysis of mathematics textbooks from Kosovo and Albania based on the topic of fractions. In K. Konrad & N. Vondrová (Eds.), *Proceedings of the ninth congress of the European society for research in mathematics education* (pp.1759-1765). Charles University in Prague, Faculty of Education and ERME.

Yackel, E. (2001, July). *Explanation, justification and argumentation in mathematics classrooms.* https://files.eric.ed.gov/fulltext/ED466631.pdf.

Yılmaz, Ş. (2018). *Bilişsel talep düzeylerine göre 8. sınıf matematik ders kitaplarındaki etkinliklerin kalitelerinin belirlenmesi* [Determination of the quality of activities in the 8th grade math textbooks according to cognitive demand levels] [Master's Thesis, Gaziantep University]. Yüksek Öğretim Kurulu National Theses Repository.

Yoong, W. K. & Hoe, L. N. (2009). Singapore education and mathematics curriculum. In W. K. Yoong, L.P. Yee, B. Kaur, F.P. Yee & N. S. Fong (Eds.), *Mathematics education: The Singapore journey* (pp.13-47). World Scientific Publishing.

Wang-Iverson, P., Myers, P., & Lim, E.W.K (2009). *Beyond Singapore's mathematics textbooks.* https://www.aft.org/sites/default/files/periodicals/wang-iverson.pdf.

Zevenbergen, R. (2001). Changing contexts in tertiary mathematics: Implications for diversity and equity. In D. Holton & M. Artigue (Eds.), *The teaching and learning of mathematics at university level: An ICMI study* (pp. 13-26). Kluwer Academic Publishers.

Elementary Pre-service Mathematics Teachers' Noticing Teaching Practices Based on TIMSS Video Study

Emine Gül Çelebi-Ilhan 🄳
TED University

Duygu Ören-Vural 🄳
Kocaeli University

ABSTRACT

This study investigates elementary pre-service mathematics teachers' noticing in a video-based activity within the context of a mathematics teaching methods course. To this end, we analyzed what and how pre-service mathematics teachers noticed in the Public Release Third International Mathematics and Science Study (TIMSS) videos. Data obtained from pre-service teachers' written analysis of classroom teaching videos were then analyzed. The results indicate that pre-service teachers primarily focused on *discourse* followed by *task* and *learning environment* aspects of the classroom teaching presented in the videos. While pre-service teachers' noticing was primarily descriptive and evaluative in a general sense, there was less interpretive level noticing. It was interesting that most comments belonging to the interpretive level appeared in the *task* category. Additionally, pre-service mathematics teachers generally noticed the elements of instruction related to teachers (i.e., the teacher's role, movement, and pedagogy) at descriptive and evaluative levels, but were less attentive to how specific classroom events, tasks, or content were connected with students' learning. Our findings suggest that exposing pre-service mathematics teachers to multiple contexts and diverse mathematical content during the analysis of video cases of classroom practices in collaborative environments supports noticing specific aspects within classroom practice.

Keywords: video-based analysis, TIMSS video study, pre-service mathematics teacher noticing, teacher education

Introduction

Given that teaching requires a range of expertise across multiple domains, it is a complex process. In this complex everyday practice, teachers receive a lot of "sensory data" that forms through their participation in the classroom (Star & Strickland, 2011). This element of teaching has a cyclical nature: once a teacher pays attention to an event or component of the practice, that may give rise to another component that the teacher interprets and responds to accordingly. Within areas of expertise, this process or act is called noticing, which is a relatively new theoretical construct within the field of teacher education (Sherin, 2004).

In its everyday use, the term "noticing" implies other acts such as attending, observing, acknowledging, or discovering. It is simply paying attention to things in the world, although we don't notice everything we see. Research with both pre- and in-service teachers concluded that teachers should be supported in learning to notice noteworthy classroom phenomena and students' mathematical thinking (Santanaga, 2011; Sherin & Han, 2014; Star & Strickland 2008; van Es et al.,

2017). That is, they should practice noticing in order to attend to, interpret, and reflect on important components of teaching and learning (McDuffie et al., 2013). This type of reasoning and awareness of actions and important events in the professional environment is typical to expert teachers. Hence, to support pre-service teachers (PSTs) in their path to becoming experts, learning to notice is crucial; it can also be considered a fundamental component of teacher knowledge and practice among many others. As such, this study investigates what elementary mathematics PSTs marked as being noteworthy in different classroom videos from the Third International Mathematics and Science Study (TIMSS) Video Study, and explores how these were inferred in a written report from a methods course. Based on the findings, recommendations for teacher education research are then offered.

Pre-service Teachers' Noticing

A strong theoretical body of knowledge and skills is necessary, albeit not sufficient, to become an expert in any area including teaching. Expert teachers are more accurate and sensible in terms of recognizing patterns, and making sense of such patterns, within their practice (Berliner, 2001), whereas novice teachers are less qualified in this regard. A developmental perspective of teacher learning shows that noticing, coupled with other skills that are related to the quality of teaching and learning, can be expanded (Matos et al., 2009; Putnam & Borko, 2000).

Although considerable research on noticing has focused on in-service teachers, a growing number of studies are now incorporating the concept of teacher noticing in teacher education programs as a means to prepare PSTs for their future teaching careers (Gupta et al., 2018). For example, Aydeniz and Doğan (2016) investigated Turkish PSTs' noticing of student responses, reporting that PSTs could easily attend to and evaluate mistakes in students' responses, but were not able to notice students' reasoning. In another study, McDonald (2016) compared PSTs' noticing with that of expert teachers. Two groups of teachers were asked to reflect on a video assigned to them by considering interesting moments. The findings revealed important differences between the comments of the two groups; while PSTs were more focused on the teachers, expert teachers attended more to the students. A similar study was conducted by Jacobs et al. (2011) to compare in-service teachers and PSTs. The researchers found that the more experience teachers had with students' thinking, the more engaged they were in noticing tasks. Based on this finding, Jacobs et al. (2011) concluded that focusing on children's understanding can influence teachers' decision-making. Other studies have indicated that PSTs have the ability to develop noticing abilities throughout teacher education programs (Star et al., 2011). However, few studies have focused on understanding the depth of PSTs' noticing.

As new technologies are prevalent in all areas of our lives, using videos has become one of the most salient components in teacher education and professional development over the past two decades. Classroom teaching videos provide their viewers an environment to capture classroom scenes as if they were present in the classroom, even if they were not (Brophy, 2004). This environment can simulate real-time teaching in such a way that the process of identifying and extracting important elements from a classroom video reflects the process of teaching (Seago, 2004). In addition, the use of videos appears to be an effective tool for helping PSTs to learn to notice and teach. Research supports the use of videos in facilitating PSTs to focus on important aspects of classroom practice and develop the necessary analytical skills to examine teaching and learning mathematics (Sanatanaga & Guarino, 2011; Star & Strickland, 2008; van Es & Sherin, 2006). The use of videos in teacher preparation, such as microteaching, lesson analysis, and giving a model for expert teaching (Sherin, 2004) is varied. One common way of incorporating videos into PST education is having video clubs, where teachers or PSTs watch themselves or their colleagues in order to analyze features of their practices. This line of research has found that participants of such video clubs develop their noticing abilities in a range of ways, from chronologically describing the events to a more interpretive level of noticing based on specific moments (van Es & Sherin, 2002; 2010). For example, van Es and Sherin

(2002) found that after using video analysis software, teachers began to capture noteworthy events from their classrooms, and more frequently offered specific evidence in discussions or their own interpretations of these events. Similarly, Males (2017) investigated what secondary mathematics PSTs found noteworthy when watching and discussing video lessons taught by their peers across two semesters. She found that PSTs comments began to focus more on student thinking or actions rather than teacher talk and actions towards the end of the second semester.

The second line of work with video involves viewing and discussing video cases in order to produce a conversation around the aspects of teaching and learning. For instance, Star and Strickland (2008) examined videos as a tool to improve PSTs' ability to observe classroom practices. With this aim, they used two videos from US Public Release TIMSS videos, where participants were asked to complete pre- and post-assessment instruments after watching. Assessment questions were created by the research team with a focus on the observed features of the classrooms, lessons, and teachers' practice. These features were grouped into five categories: the classroom environment, classroom management, tasks, mathematical content, and communication. Participants' abilities to notice classroom features such as the classroom environment, mathematical content of the lesson, and teacher and student communication during a lesson improved. In another study, van Es et al. (2017) grouped topics that PSTs noticed from a short video clip as mathematical content and learning goals, students' thinking, pedagogies for making thinking visible, and classroom discourse norms. These studies analyzed the aspects of teaching or learning, or selected categories, that were frequently noticed by PSTs.

Another line of research has stressed the importance of providing guidance to teachers or PSTs when analyzing videos. For example, in a study on Learning to Learn from the Mathematics Teaching Project, which used extensive videos over two courses, Santanaga and Guarino (2011) reported that video-based activities support PSTs' ability to notice student thinking, to notice teacher moves that make student thinking visible, and to reason about the instructional decisions that can impact student learning. Their findings also highlighted that the guidance provided to teachers, or prospective teachers, when analyzing the videos is crucial for learning and reflecting from teaching. Similarly, McDuffie et al. (2013) argued that participating in video analysis activities supports teachers to notice students' mathematical thinking. Moreover, the prompts used, the video choices, and the opportunities to repeat the activity enabled PSTs to improve their levels of noticing (i.e., developing from merely describing to interpreting). Also, the depth of teachers' foci changed from teacher moves, to classroom interactions and teachers' effects on learning.

Overall, researchers have emphasized that in order to learn to notice, PSTs need support, and teacher education programs should provide explicit opportunities for PSTs to practice these skills (Sherin et al., 2011; Sims & Walsh, 2009; van Es, 2011). Therefore, to support PSTs' development of noticing, teacher educators should comprehend what PSTs notice and investigate the depth of their noticing. Based on this premise, this study aimed to understand the noticing that PSTs have while working with classroom teaching videos.

Teacher Noticing Framework

There are many perspectives and proposals on the key aspects of noticing in the literature. For example, Ball (2011) described three properties that noticing includes: observing, realizing, and attending to. Similarly, van Es and Sherin (2002) highlighted the following three aspects of noticing: "(a) identifying what is important or noteworthy about a classroom situation; (b) making connections between the specifics of classroom interactions and the broader principles of teaching and learning they represent; and (c) using what one knows about the context to reason about classroom events" (p. 573). Mason (2011) stressed that "noticing is a movement or shift of attention" (p. 45). Other researchers have based their conceptions of teacher noticing on one particular topic, such as students'

mathematical thinking. For example, Jacobs et al. (2010) described professional teacher noticing (of students' mathematical thinking) as a set of three interrelated skills: attending, interpreting, and deciding how to respond.

Although researchers' different conceptualizations emphasize different aspects of noticing, there seem to be some commonalities between them. It is seen that attending to what is happening in the classroom and making connections between different elements in classroom situations is critical (Stahnke et al., 2016; van Es & Sherin 2002). Based on this, teachers reason about and interpret what is happening in their classrooms (Li & Superfine, 2018) and decide how to respond to them (Gupta, et al., 2018). As seen, all these concepts relate to teacher noticing and shape teacher movement (Sherin et al., 2011). Understanding how and what teachers and PSTs notice will shed light on their current and future practice, which can better prepare these teachers (Jacobs et al., 2010; Star & Strickland 2008; van Es, 2011).

In this study, we use van Es and Sherin's (2002; 2010) definition of noticing, which centers around three key aspects: 1) the ability to attend to what is significant in a complex situation; 2) using contextual knowledge to reason about important events; and 3) the ability to make connections between specific events and the broader principles they represent. Through her work on video clubs and their impact on teachers thinking and practice, van Es (2011) generated a framework on teacher noticing, which focused on the influence of video cases in teaching teachers to notice the features of students' mathematical thinking based on (i) what teachers notice and (ii) how teachers notice. The category of "what teachers notice" describes both whom and which issues to focus on noticing, and it is comprised of different levels of noticing. The category of how they notice distinguishes the analytical stance of teachers' responses (the extent to which teachers' responses are descriptive, interpretive, and evaluative) and the depth of their noticing. In this study, we first identified a general framework describing the issues that PSTs' instruction should focus on. Based on Stigler et al.'s (1999) three key questions to analyze classroom teaching (lesson environment, content, and the ways that content is studied) and Star and Strickland's (2008) observation categories, we chose the following focus areas: *task (T)*, *classroom discour*se (D), and *learning environment (LE)*. Then, we focused on the analytical stance and specificity of what PSTs found noteworthy.

A mathematical task is defined as a problem or set of problems that focuses students' attention on a particular mathematical idea (Stein et al., 1996). Hiebert et al. (2003) examined all seven countries that participated in the TIMSS 1999 video study and found that 80% of the time spent in eight-grade mathematics lessons was dedicated to working on the mathematical task. The discourse aspect is also an inevitable part of mathematics lessons. Sherin et al. (2008) found that classroom discourse is a process through which groups of individuals communicate. Although differing in degree and quality, in each of TIMSS classroom teaching videos, there were interactions between teachers and students, and among students (Hiebert et al., 2003). Lastly, the aspect of the learning environment includes physical and non-physical elements forming the environment in which learning takes place. Researchers defined the learning environment as consisting of pedagogy, technology, and physical space (Cleveland & Fisher, 2014), and then highlighted its importance in educational settings (Clarke, 2001; Lippman, 2007). As these three aspects were considered important in TIMSS classroom practices videos, we used them to structure PSTs' attention.

Researchers also reported that lenses help novice teachers focus on particular aspects of teaching instead of being overwhelmed by many complexities within the classroom (Roller, 2016). For this reason, these three areas (task, discourse, learning environment) were used to provide a focus and to support PSTs' noticing skills when analyzing the videos. Moreover, it should be noted that the quality of mathematical understanding occurring in the classroom is very much dependent on the quality of the mathematical tasks used (Chapman, 2013), how they are used, and the quality and type of mathematical discourse provided in the classroom (Nathan & Knuth, 2003; Van Zoest & Enyart, 1998); these two aspects influence and are influenced by learning environments (Anderson & Walberg,

2003). Furthermore, in the research field on mathematics-specific pedagogy, two aspects, mathematics tasks and discourse, are highlighted in recent mathematics teaching reform (Borko et al., 2000). The National Council of Teachers of Mathematics (NCTM, 2007) proposed that teaching expertise requires teachers to constantly analyze mathematical tasks and discourse to understand what students learn and develop their teaching practice. Therefore, how PSTs notice these aspects is critical for their future teaching (Borko, 2004; Mason, 2002). In several studies, PSTs' difficulty to attend key aspects of mathematics teaching is noted. For example, in Star and Strickland's (2008) study, PSTs attended to the classroom environment and management issues more than the mathematical content and communication. In another study that replicated Star and Strickland's initial work, Star et al. (2011) obtained a similar result. In these studies, the aspects focused on were readily determined. In contrast, this study investigates what related aspects PSTs will notice about these three aspects.

This study aims to contribute to the existing literature by engaging PSTs in a video-based noticing activity within the context of the mathematics teaching methods course. As such, we used TIMSS video cases as a tool for understanding PSTs' noticing abilities in contexts, such as *task*, *classroom discourse,* and *learning environment*, to produce a conversation around the various aspects of teaching and learning. In contrast to other studies that used TIMSS video cases, PSTs' noticing ability was identified based on the three key aspects that Stigler (1999) proposed and on van Es and Sherin's (2011) framework.

In this study, we address the following question: How and what do PSTs notice in terms of the task, learning environment, and discourse in TIMSS videos of classroom practice? Using the framework developed by van Es and Sherin (2011), we described the classroom features that PSTs found noteworthy, their analytical stances, and how these were evidenced based on their explicit comments on videos in written reports.

Methods

Study Context and Data Collection

In this study, a descriptive qualitative case study design (Yin, 2018) was applied, in which PSTs' written comments were defined as the unit of analysis. In this way, we aimed to understand what and how PSTs collectively notice regarding the three aspects (task, learning environment, discourse) of teaching practice represented in the TIMSS video cases. The study took place in the context of an elementary (grades 5 to 8) mathematics teaching methods course that focused on mathematical communication in the classroom. The course broadly focused on issues of general literature regarding mathematical communication, mathematics and language, mathematical classroom discourse, mathematical symbols and terminology, and multiple representations. A total of 34 PSTs who were all juniors—third-year students of elementary mathematics education majors in a foundation university in Turkey—participated in this study. PSTs took this course as part of a mathematics education research course and they were required to read journal articles, reports, or monographs each week and to write at least one question regarding the topic before each lesson. First, these lessons started with a brief introduction given by one of the researchers, and then, PSTs were asked to identify their questions and comments about the resources before discussing the theme of the week.

After PSTs were introduced to the theoretical concepts regarding mathematical communication, we used classroom teaching videos to develop PSTs' noticing skills in mathematical practice. We used videos since many studies report the benefits of using this medium to support PSTs' learning in general (e.g., Santagana, 2011; Star & Strickland, 2008), and fostering PSTs' abilities to notice in particular. Furthermore, studies also showed that PSTs develop noticing abilities when they analyze and discuss videos in a systematic and structured way (e.g., Kleinknecht & Gröschner, 2016). The influential TIMSS video cases were chosen to expose them to different cultural practices of teaching mathematics. Different videos, including various mathematical content from seven countries,

were randomly assigned to PST groups made up of four or five people (see Table 1). Then, one TIMMS video from each country was selected by the instructor of the course. We chose these videos to cover all learning areas of middle school mathematics in the Turkish National Curriculum: Numbers and Operations, Data analysis and Measurement, Geometry, and Algebra.

Table 1

Description of TIMSS Video Lessons

Group ID	Topic	# of Group Members	Description of each lesson
1	Polygons	5	An eighth-grade mathematics lesson focused on deriving the sum of interior angles in a polygon. It is the first session in a series of four lessons that work toward the more advanced concepts of polygons.
2	Data collection and representation	4	An eighth-grade mathematics lesson focused on simulation and data collection. It is an extension lesson following a nine-lesson unit of work focused on statistics.
3	Exponents	4	An eighth-grade mathematics lesson focused on operations with numbers raised to certain powers and calculating the area of a triangle and a trapezoid. It is the second lesson in a unit of work focused on raising numbers to powers.
4	Changing shape without changing area	5	An eighth-grade mathematics lesson focused on two-dimensional geometry; in particular, the areas of triangles between parallel lines.
5	Surface area	5	An eighth-grade mathematics lesson in which students work on a textbook chapter about factoring. The students either work independently or in small groups, with occasional explanations by the teacher.
6	Introducing algebra	5	An eighth-grade mathematics lesson focused on terms and variables. It is the first lesson on this topic.
7	Secants and tangents	5	An eighth-grade mathematics lesson focused on the measurement of angles formed by secants and tangents intersecting with a circle. This is the fourth lesson in a six-lesson series on this topic.

To complete their assignments, PSTs analyzed their video cases virtually via a cloud-based working environment; group work was completed using online video meetings, chats, or private group channels. PSTs were asked to present their analyses online, while the rest of the students had access synchronously or asynchronously. Then, they were asked to submit a written report of analyses for their video cases. As Choy (2016) argued, providing an explicit focus to teachers would support them in noticing specific details of instruction or practice. Hence, PSTs were given three main areas: *Task, Discourse, Learning Environment,* and guiding questions (see Appendix) to focus on during their analyses of video cases, both for their oral presentations and their written reports.

Data Analysis

Stigler et al.'s (1999) three dimensions of classroom practice (organization of the lesson environment, kinds of mathematical tasks, and the processes of instruction) were used to describe the three broad themes—*Task*, *Discourse*, and *Learning Environment*—for the analysis of PSTs' written reports on video cases. In describing these main themes, we also used Star and Strickland's (2008) observation categories, which included classroom environment, classroom management, tasks, mathematical content, and communication. These were used as initial themes in the coding scheme to describe what PSTs noticed as components of instruction from their cases (see Table 2 for the description of each of these themes).

Table 2

Main Areas Used to Code Written Reports of TIMSS Video Analyses

Themes	Descriptions
Task	Includes the types of tasks selected by the teacher, their level of appropriateness for students, and how the task was implemented by the teacher
Discourse	Includes the students and teachers' roles in supporting classroom discourse, the resources, and tools used for supporting classroom discourse, and the way teacher responds to students' (wrong) answers
Learning Environment	Includes the teachers' management of time, physical space, and social relations in the classroom, as well as teachers' instructional group strategies and expectations of their students

The data analysis was conducted in three stages (see Figure 1); data obtained from the first and second stages were used to answer the research question "what do PSTs notice," whereas data obtained from the second stage was used to answer the research question "how do PSTs notice."

In the first stage, all data were combined in one document to organize coding. Themes in the written reports of the groups were then determined. Since data came from groups' written analyses of their video cases, the idea units were identified in the data first (Jacobs & Morita, 2002). The idea unit was defined in terms of a "distinct shift in focus or change in the topic" (Jacobs et al., 1997, p. 13), which has a meaningful and identifiable focus. In this regard, we identified an idea unit as a sentence or a paragraph that has a meaningful chunk of information and a certain focus. As we divided the written reports into idea units, we labeled each idea unit reflecting a clear focus on one of the three themes (task, discourse, and learning environment). In this way, all idea units were associated with one of the themes. Two researchers (the instructor and the second author of the study) coded each idea unit as a comment of the PSTs. Each researcher applied these themes to each comment separately, and then met to discuss and agree on the terms. Interrater reliability was obtained by determining the idea units in each group's report, as well as the number of idea units that were agreed upon; the rate of agreement was 95%.

In the second stage, we continued with our open coding process by specifically naming and categorizing the data (Strauss & Corbin, 1990). By further examination of the idea units associated, we identified what PSTs noticed, and then sorted the resulting categories into sub-themes. Similar to the process applied at stage 1, two researchers coded the data separately and then met to reach a consensus. Interrater reliability was determined for sub-themes and agreement on the placement of categories; the results were 85%, 90%, and 90%, respectively. Four sub-themes (teacher pedagogy,
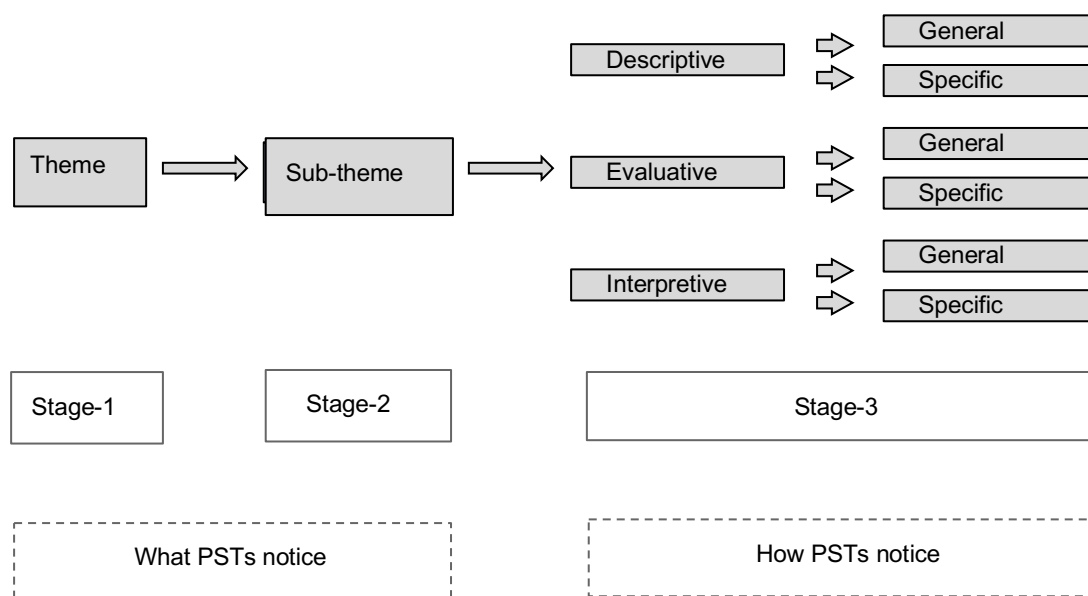
teacher role, content, and student) were related to the task theme, five sub-themes (teacher role, use of materials and resources, students' role and participation, teacher move, and communication) were related to the discourse theme, and four sub-themes (classroom management, classroom climate, resources, and physical space) were related to the learning environment theme at this stage. See Table 3 for the descriptions of all sub-themes and sample comments on each sub-theme.

At the third stage, we identified how PSTs noticed when analyzing their video lessons. At this stage, another cycle was employed for the data coded at the second stage to identify the stance and specificity developed by van Es and Sherin (2010). Stance is defined as the analytical level of PSTs comments—Descriptive, Evaluative, Interpretive—in the idea units. In accordance with van Es and Sherin's (2010) framework, a statement was coded as Descriptive if it described events that occurred; it was coded as Evaluative if the PSTs commented on what was good or bad or should have been done differently; and it was coded as Interpretive if PSTs tried to make inferences about what they observed. In accordance with van Es and Sherin's (2010) framework, a statement was coded as Descriptive if it portrayed simplified observations of events that occurred in the video case; it was coded as Evaluative if the PSTs' comments were judgmental of what they observed (either good or bad) or if commenting on what should have been done differently; and it was coded as Interpretive if PSTs tried to make inferences of, and connections between, the features that they observed.

At the same time, data was also coded in terms of its level of specificity. In line with van Es and Sherin (2010), specificity was determined by the level of details included in the PSTs' comments. A statement was coded as General if it focused on the whole class or involved broad generalizations, and it was coded as Specific if it was specific to a particular event, idea, individual, or issue (such as one student's idea or a particular teacher's movement). Once again, two researchers coded the idea units according to the stance and the level of detail; specificity was coded separately, and a consensus was reached for each case through discussion. The interrater agreement in the coding stance and specificity was 90% and 95%, respectively. See Table 4 for sample comments on each stance and specificity level.

Figure 1

Stages of Coding

Table 3

Descriptions of Sub-themes and Sample Comments

Sub-themes	Codes	Descriptions	Examples
Teacher pedagogy	T1	How the teacher implemented the task	Tasks teacher selected mostly involved procedural understanding.
Teacher role	T2	How teacher used the task to deliver the instructions	Teacher expected students to solve the problem by using the questions and answers method.
Content	T3	Mathematical content of the task	We thought this lesson had rich tasks as students collected data and made representations of the data.
Student	T4	How the task supports students' learning	The tasks did not allow students to make connections to algebraic concepts in a conceptual way.
Teacher's role	D1	What teachers do to support classroom discourse	The teacher took communicative and explanatory roles to create meaningful mathematical conversations in the classroom.
Use of materials and resources	D2	How tools and materials support classroom discourse	The purpose for the use of materials (computer, board, ruler, and cards) is mathematical and action oriented.
Students' role and participation	D3	What students do to support classroom discourse	Students are active and productive in seeking answers to the questions asked by the teacher. This creates an environment for discussion.
Teacher's move	D4	Teachers talk to plan, open, and support classroom discourse	Despite incorrect answers from the students, the teacher explained the problem and asked them again after giving hints.
Communication	D5	The use of mathematical symbols and language to support classroom discourse	There were deficiencies in the teacher's use of mathematical symbols and language. Responding to a student's question, the teacher just says to substitute zero in the formula.
Classroom management	LE1	How teachers organize class (i.e., pace, timing, disciplinary issues)	The teacher is usually in front of the blackboard, not near the students. Some students are not interested in the lesson. Despite this, the teacher can maintain silence in the classroom.
Classroom climate	LE 2	Classroom atmosphere and the way in which the teacher and students interact	The instruction is given in a classroom environment that considers individual differences and is formed because of students' orientations
Resources	LE 3	Resources (i.e., handouts and technological tools) in the learning environment	We see that there are wallboards in the physical environment of the classroom and that calculators and dice are used as learning materials
Physical space	LE 4	Physical features of the classroom.	The seating arrangement in the classroom is traditional.

Table 4

Examples From PSTs Comments About Stance and Specificity

Stance	Specificity	Examples
Descriptive	General	Teacher first expected students to think individually, then she wanted students to discuss in groups.
	Specific	The teacher opened up the problem of angles of the circle on the smartboard, and gave them time for the solution. Students tried to solve the problem individually.
Evaluative	General	Technology is used mathematically and that's the way it should be used.
	Specific	The teacher seems to be good at teaching, but he has some problems in reflecting that to the students. For instance, he responds to a student's question by telling them to solve the problem by using a formula.
Interpretive	General	One thing that caught our attention was the seating arrangement. It enabled the teacher to make eye contact and provided a space that can be used during the activities so the teacher can manage the classroom well.
	Specific	The teacher displays an insistent and investigative attitude to create meaningful mathematical conversations in the classroom. Students who could not solve the problem were directed to the clue cards so as to prevent them from disengaging from the lesson. Students who were able to solve the problem were asked to explain their strategy what they have done, without saying if it is true or false. In this way, the teacher guides the students to talk in mathematical terms.

Results

The analyses of PSTs' written reports are presented under the main categories of *Task, Discourse,* and *Learning environment.* First, we provide what PSTs noticed in their written analyses of cases. In doing that, we first begin describing the number of idea units that PSTs noticed according to our three main themes. Second, we provide results of "how" they noticed, which included the level and specificity of PSTs noticing and the focus of their comments. The analyses include both quantitative results regarding the frequency of the sub-themes addressing what PSTs noticed (with respect to each main theme), as well as how they noticed (by providing the frequencies of specificity and the stance as the analytical level of these comments under each main theme).

For all seven groups of PSTs, we identified 187 idea units in the written reports. Table 5 includes the number of comments regarding each theme, the specificity level, and the levels of the comments.

Table 5

Frequency of the Themes, Specificity, and Levels of the PSTs' Comments

	Themes			Specificity		Levels		
	Task	Discourse	Learning Env.	General	Specific	Descriptive	Evaluative	Interpretive
n	65	94	28	157	30	71	62	54
Total		187			187		187	

Among all the comments, approximately half (50.3 %) were about *Discourse,* which was the most noticed theme by the PSTs. The *Task* theme consisted of one-third of all comments (34.8%), and the least noticed theme was *Learning Environment* (14.9%). Considering the level of specificity, it was found that only 16.1% of PSTs' comments were specific (that is, based on specific issues focusing on a particular individual or event), while the majority of comments were written about the whole class or generalizations of events (83.9%). Looking at the distribution of PSTs' different levels of noticing, it was found that interpretive comments were rated the lowest. 29.6% of PSTs' comments were interpretive. Overall, the number of descriptive comments (37.9%) was higher than the comments in the evaluative (33.2%) and interpretive (28.9%) categories. In the sections below, the results of PSTs noticing are presented per theme (*Task*, *Discourse*, and *Learning Environment*).

Task

PSTs' comments about the *Task* were conceptualized based on the type of tasks, level of appropriateness of these tasks for the students, and the way teachers implemented these tasks. PSTs' comments indicating noticing related to the *Task* (T) accounted for 35% of all idea units (Table 5); those centered on how the task was implemented by the teacher (T1) accounted for 45% of the T codes; those on the teacher's role in delivering the task (T2) accounted for 21.5 % of all T codes; and those on mathematical aspects (T3) accounted for 21.5% of all T codes. The level of appropriateness of the task to the students and student thinking (T4) was the least noticed aspect in the *Task*, accounting for 8% of all T codes (Table 6).

Table 6

Frequency of the PST Task Codes

Sub-themes	Teacher pedagogy (T1)	Teacher role (T2)	Content (T3)	Students (T4)
n	29	14	14	8

Considering how PSTs noticed the *Task*, the comments indicating noticing were categorized based on the level of noticing, namely whether it was descriptive, evaluative, or interpretive. As seen in Table 7, PSTs' interpretive comments about *Tasks* were more than the evaluative and descriptive ones, accounting for 40.0%, 32.3%, and 27.7%, respectively. These results indicate that PSTs comments about the *Task* aspect of the video cases were mostly interpretive, as they tried to make sense of the type of task selected by the teacher and the implementation of the task in the classroom (mainly focusing on the teacher and the teacher's role, and less on the student aspects including the level of appropriateness of the task or the student thinking processes); this finding was further supported by the qualitative analysis of written excerpts corresponding to the *Task* theme.

Table 7

How PSTs Noticed the Task

	Descriptive				Evaluative				Interpretive			
	General		Specific		General		Specific		General		Specific	
	n	%	n	%	n	%	n	%	n	%	n	%
	15	23.1	3	4.6	17	26.2	4	6.1	16	24.6	10	15.4
Total	n		%		n		%		n		%	
	18		27.7		21		32.3		26		40	

Comments indicating general features were more common than those indicating specific details related to *Task*, which was a finding that was consistent at every level (descriptive, evaluative, or interpretive) and accounted for 76% of all T comments. Among the three levels, interpretive-specific comments were more common than specific comments at other levels. This result indicates that most of the PSTs' specific comments were at an interpretive level; although PSTs were largely noticing general features of the lessons, they made specific references to particular events, ideas, individuals, or issues, particularly when they noticed interpretively. In line with this result, in the report from group (7), PSTs wrote about *teacher pedagogy*:

> *When the student gave the wrong answer to the question, instead of making him think about the correct answer, the teacher immediately wrote the correct answer by erasing the answer given by the student, or when the student was speaking but did not know the correct answer, the teacher called another student to solve the question.* This effectively *prevented students from discovering and learning.*

In group (5), however, the way in which PSTs identified and interpreted how the teacher used the task was by inferring based on the *role of the teacher*. As such, it was coded as being interpretive and general:

> The teacher *manages the task in such a way that there are huge differences in time allowing for one-to-one dialogues ... might be since students are moving at the same pace, and they were on different parts of the task.*

On the other hand, analysis of the PSTs' written reports revealed many evaluative comments as well. For instance, what could have been done to ensure that the task could support students'

learning; the *student* aspect of the task was considered at an evaluative level. Since there was no reference to a specific student or event, this comment was coded as being general rather than specific. In this respect, PSTs in group (5) commented that:

> *The task is appropriate for the students' level, but it did not have an engaging context as in the other video cases such as students who love pizza or the number of chocolate chips in the cookies. The context is not engaging students, it merely consists of algebraic equations and a problem related to surface area."*

Here, group (5) noticed the *content* of the task with respect to algebraic thinking and they evaluated it as being "not a good task," since the context was not engaging. They reported this using a general comment that illustrated which mathematical concepts were involved in the tasks of the lesson, and the type of knowledge they called for:

> *The task involved mathematical concepts such as variables, charts, coordinate planes, and graphic representations. The problems used as tasks contained in a book required procedural knowledge.*

In this comment, group (5) shared simple and general observations about the mathematical concepts of the task and the type of problems found in the textbook; hence, their response was descriptive and general.

Discourse

Discourse was conceptualized as the tools, resources, and roles of the teachers and students in supporting discourse within the classroom. Moreover, how the teacher responded to students' wrong answers was also included under the *Discourse* theme. The findings show that there was a high frequency of comments where PSTs focused on discourse, which accounted for more than 50% of all the codes (Table 5). Within the *Discourse* (D) codes, PSTs noticing centered around the roles of the teacher (D1) (29% of codes) and students (D2) (18% of all D codes), followed by the tools and materials used for supporting classroom discourse (D3) (13 % of all D codes), actions used by the teacher to plan, open up and support the classroom discourse (D4) (20% of all D codes) and the identified symbols and language that were used (D5) (20 % of all D codes) (see Table 8).

Table 8

Frequency of the Discourse (D) Codes

	Teacher role (D1)	Students' roles & participation (D2)	Use of materials & resources (D3)	Teacher moves (D4)	Communication (D5)
n	27	17	12	19	19

These results indicate that PSTs were attentive to aspects related to classroom discourse (such as *students' roles, the teacher's discursive moves, the teacher's role, and the ways that materials or resources were used*), as well as *communication* patterns within the classroom context. Although less frequent, the PSTs also evaluated these aspects regarding the classroom discourse but the integration of these observations and evaluations in order to reason about the discourse took place less often.

In regards to discourse, PSTs' noticing was mainly descriptive (45% of all D codes), followed by evaluative (33% of all D codes) and interpretive (21% of all D codes) (see Table 9).

Table 9

How PSTs Noticed Discourse

	Descriptive				**Evaluative**				**Interpretive**			
	General		Specific		General		Specific		General		Specific	
	n	%	n	%	n	%	n	%	n	%	n	%
	40	42.5	3	3.2	30	31.9	1	1.1	14	14.9	6	6.4
Total	n		%		n		%		n		%	
	43		44.8		31		32.9		20		21.3	

Similar to noticing in *Task* theme, PSTs noticed general aspects more often than specific events, cases, or students in the case of *Discourse* (89% of all D codes). This was the case at each stance or level of noticing because PSTs' comments were primarily general, irrelevant of whether they were descriptive, evaluative, or interpretive (Table 9). However, among the three levels, interpretive-specific comments were most common, while evaluative-specific comments were least common. This result indicates that PSTs were most often making specific references to particular events, ideas, individuals, or issues when they integrated their noticing to reason about those particular features. In line with this result, in the report from group (3), the PSTs wrote:

> *The teacher displays an insistent and investigative attitude to create meaningful mathematical conversations in the classroom. Students who cannot find a solution are directed to the clue cards so that they do not disengage from the lesson. Students who find a solution are asked to explain the strategy they used, without saying whether it was true or false. In this way, the teacher guides the students to talk mathematically.*

Here, PSTs were attentive to the supportive *role of the teacher,* and these observations were connected with how the teacher responded to students' solutions; this enabled students to create meaningful mathematical conversations by specifically focusing on solving the problem. At times, PSTs also evaluated their observations of classroom discourse as being either good or bad in the general sense. In relation to this, group (7) made the following comment on the *students' role* in discourse:

> *The students' job was to answer the questions, take notes, and communicate with their teacher. Since there is no discussion context in the classroom, students do not interact with each other, but only with the teacher. Students are not responsible for their own learning, but this responsibility rather belongs to the teacher herself.*

Here, PSTs in group (7) also made observations about *students' roles* within classroom discourse, but their comments were judgmental of the fact that this was not beneficial for students. PSTs also made general observations about the teacher's *use of materials or resources* to model mathematical

situations and concepts, as well as how specific technologies were used. For instance, group (2) commented on a video case that dealt with data analysis and representation:

> *We saw that the materials used to support in-class discourse are a calculator, dice, ruler, and no technological devices.*

Although the tools and resources were described as supporting classroom discourse, there was neither an evaluation on the use of those tools nor any reasoning about how these tools connected to the instruction; therefore, this comment was coded as descriptive and general.

Another category that PSTs commented on related to *Discourse* was the *teacher moves,* which supported opening up and extending classroom discourse. In one example from group (4), PSTs noticed the teacher's moves in a general manner during a lesson that examined how changes in the shape of a triangle did not lead to a change in its area:

> *Despite the students' incorrect answers, the teacher explained the problem to the students and asked it again after giving clues. The teacher did not directly give the students the answer. This allows students to find and see the answers themselves by asking questions about the problem.*

Here, PSTs made connections about how the teacher acts to open up and facilitate discourse by portraying statements about the whole class in general. Hence, these comments were coded as interpretative and general.

PSTs also noticed that mathematical communication patterns exist with respect to classroom discourse. For example, in group (3), the PSTs noticed aspects of *communication* within the context of teaching operations, with numbers raised to powers, by giving a specific example from the lesson:

> *The teacher used the mathematical symbols and language deficiently sometimes. At one time, a student directed a question at the teacher, but the teacher answered by asking the student to apply the formulas and replace them with a zero. In this regard, the teacher should have solved the problem in front of the class and expressed the problem or solution more clearly on the board. He needs to talk more about the mathematical concept and comment more on mathematics.*

Similarly, this comment was a detailed observation about how the teacher answered a student's question. Based on this observation, PSTs evaluated the way the teacher responded as being "deficient." Yet, they also explained the reason behind that judgment as a need for more talk on mathematical concepts; hence this comment was categorized as interpretive and specific.

Learning Environment

The Learning Environment was conceptualized based on the way a teacher manages time, physical space, and social relations within the classroom context. Furthermore, it involved the teacher's expectations of his/her students and the instructional group strategies employed during the class. Based on the frequencies of *Learning Environment* (LE) codes, the least amount of noticing was characterized as this theme (15% of all codes) (see Table 6). The findings indicate that higher frequency codes were classroom climate (LE1) and classroom management (LE2) issues (32% of all LE codes), followed by the teacher's arrangements of the physical space (LE3) (21% of all LE codes) and the organization and utilization of resources (LE4) (14% of all LE codes) within the classroom (see Table 10).

Based on the findings, PSTs' comments about the learning environment were mainly descriptive and evaluative (each accounted for 35.7% of all LE codes), whereas the rest of the

comments were interpretive (28.6% of all LE codes) (see Table 11). Furthermore, the results indicate that PSTs' noticing in the learning environment centered around how the teacher organized the classroom, used physical space and resources, and the state of the classroom environment. PSTs' noticing also indicated making sense of one or more of the specific aspects of the learning environment, while making meaningful connections between them.

Table 10

Frequency of Learning Environment Codes

Sub-themes	Physical space (LE3)	Classroom management (LE1)	Classroom climate (LE2)	Resources (LE4)
n	6	9	9	4

In parallel to the findings on *Discourse*, PSTs' comments primarily centered around the general picture of the *Learning Environment* (90.3% of all LE codes); at all levels of noticing, describing non-specific details about an event or situation was the most common kind of comment (90% of all LE codes). The results also showed that only a few specific comments about the learning environment were made by PSTs (10.7% of all LE codes). Once again, there was a higher frequency of specific comments about the learning environment at the interpretive level (7.1% of all LE comments) than at the other levels (3.6% for descriptive and none for the evaluative level) (see Table 11).

Table 11

How PSTs Noticed the Learning Environment

	Descriptive				Evaluative				Interpretive			
	General		Specific		General		Specific		General		Specific	
	n	%	n	%	n	%	n	%	n	%	n	%
	9	32.1	1	3.6	10	35.7	0	0	6	21.4	2	7.1
Total	n		%		n		%		n		%	
	10		35.7		10		35.7		8		28.6	

One example of specific and interpretive noticing can be seen in the following comment by group (3) on *classroom management*:

> *The teacher paid attention to her students individually. For instance, she instantly noticed a student at the back asking for help from a friend and asked what the matter is. Generally, the class is teacher-centered, and the teacher is successful at managing the classroom.*

PSTs commented on how the teacher interacted with the students and gave a specific example from the video case that supported their claim. Similarly, group (5) noticed the arrangement of the *physical space* by the teacher, by noting the seating arrangement and discussing the advantages of this for students and the teacher without paying attention to any specific event. This comment, which was coded as interpretive-general, stated the following:

> *One of the most important parts that caught our attention is the students' seating arrangement, because the seating arrangement has various advantages for students and teachers during and in between lessons. These advantages include the ability of the teacher to make eye contact with the students during the lesson, to provide a space that the students can use comfortably in activities and drama studies, and allowing the teacher to easily control the students during a lesson.*

Another category that PSTs noticed frequently was *classroom climate* (see Table 10). A group of students (group 1) noted that the tone of classroom atmosphere and the students' actions were controlled by the teacher in an evaluative and general way:

> *In the video, we see that the course is progressing in a way that is under the control of the teacher. Here, we notice the teacher's instructions, that is, students are dependent on the teacher's instructions.*

PSTs also noticed that *resources* were an important aspect of the learning environment, as they can be used by the teacher and the students to support learning and teaching. The use of time was also considered a *resource* that relates to the learning environment. Group (3) made a comment reflecting on the use of handouts, instructional strategies, and time. This group also provided an analysis of their observations and drew conclusions based on specific evidence from the video case; hence, this comment was categorized as an interpretive-specific comment:

> *The teacher passed out the handouts quickly, called upon some of the students to speak and some of them to work on the board, and created small groups for all students to find the correct answer. This showed that he used the time effectively. Another indicator of this is that he finished all of the work he brought for use in the given time of that class.*

Moreover, in this comment, PSTs analyzed the social norms of the classroom, whereby students were given a voice; they also emphasized the teacher's use of instructional group strategies, which reflected the teacher's expectations of his students.

Discussion, Conclusions, and Suggestions

Investigating teachers and PSTs' noticing in complex classroom situations is very important, as what is happening in the classroom is an outcome of what teachers do, and what teachers do in the classroom is an outcome of what they notice (Star & Strickland, 2007). In qualified teaching, the outcome is students having a deep and meaningful understanding (Lampert, 2010); consequently, the efforts on investigating teachers' noticing skills and developing those skills contribute to teachers' teaching expertise and the quality of their teaching (Jacobs et al., 2010; Sherin et al., 2011). Given this context, efforts in this regard can facilitate students learning.

In this study, we examined PSTs' noticing in the context of elementary mathematics classes. Specifically, we aimed to describe the classroom features that PSTs identified as noteworthy (i.e., what PSTs notice), their levels of analysis, and what evidence they presented in their explicit comments on videos, which were taken from their written group reports. In order to describe how and what PSTs

found noteworthy about the different classroom features, PSTs watched TIMSS videos selected from seven countries. In line with the work of Santanaga and Guarino (2011), which proposed that video representations should be used with purpose and guidance, we chose these specific videos based on two main criteria: 1) they taught mathematics content and 2) the teaching aligned with the outcomes of this course.

Our results indicate that in their analysis of the videos, PSTs most frequently commented about the *Discourse* (50%) aspect of classroom teaching (i.e., the role of teachers, teacher movements, communication in the classroom, students' role and participation, and the use of materials and resources) in their reports. Moreover, PSTs' noticing was mostly descriptive (38%) and evaluative (33%), as well as general (86%). As such, their evidence was less focused on specific issues (16%) (i.e., a particular issue or an event). The reason that the *Discourse* category was most frequently noted could be because the course content primarily focused on communication within mathematics lessons. A striking finding in the distribution between the subcategories of *Discourse,* is that PSTs were more attentive to *Discourse* in terms of the teachers' roles than the students' roles; they were also very attentive towards components of communication within the classroom and the use of materials and resources. This finding is in accordance with the conclusions drawn from the existing literature. For example, in a study conducted by van Es et al. (2017), researchers found that PSTs were most attentive to overall features of classroom interactions and teachers' roles in classroom discourse at the beginning of their study. After their study, PSTs begin to develop abilities to attend to norms of communication between teachers and students, and focus more on how students' participation in discourse affects their thinking and learning. Another important finding of van Es et al.'s (2017) study was about the materials and resources, which PSTs were less likely to be attentive to.

The second most frequently noted category by PSTs was *Task* (34%). The distribution of subcategories showed that PSTs focused more on aspects related to the teacher (i.e., teacher pedagogy and role, how it is implemented by teachers, and the teachers' roles in delivering the content of the task) than on student aspects (such as how students experienced the task) and content aspects (the content of the task). This finding is in accordance with Males (2017), who found that throughout the two semesters of her study, PSTs were more focused on teachers and teachers' actions; additionally, comments that involved student talk and students' thinking were more frequent in the second semester than the first semester. Moreover, our findings support Star and Strickland's (2008) findings that PSTs have difficulties attending to the mathematical content of the lesson when they are not specifically observing the content, unless specific attention is devoted to teachers' choice of representation, the accuracy of the content, and the problems that are used. Indeed, in our study, the content of the instruction was not explicitly asked in the guiding questions; therefore, the result obtained about content might also be due to guiding questions. Still, as a content-related aspect, *Task* was the second most frequently noted category following *Discourse,* which shows that these two categories are the most prominent features of the reform movement in pedagogy specific to mathematics (Borko et al., 2000).

Our findings indicated that *Learning Environment* was the least noticed category by PSTs. Though classroom management and classroom climate as subcategories constituted the majority of what PSTs noticed about the *Learning Environment*, resources, and the physical space were less frequently noticed. This finding is also consistent with prior studies (Males, 2017; Star & Strickland, 2008), which found that PSTs were not frequently noticing factors of the physical environment such as the arrangement of desks or other resources in the classroom, but also showed greater attention to the classroom management and climate (which involves time management and social relations, instructional group strategies, and the teacher's expectations of the students). Our result could also be explained by the content of the course, as the course primarily focused on communication in mathematics lessons. A similar finding was also proposed by Males (2017), who stated that students noticed communication issues more frequently than challenges related to classroom management and

the environment; however, in contrast to this view, we also found that frequent attention was paid to *Tasks*. This could be because we took mathematical content as a subcategory of *Task* and combining these two topics might have resulted in a higher frequency of the *Task*. Furthermore, since the learning environment seems to refer to learning and student achievement (Fraser, 1994; McRobbie & Fraser, 1993) more than teaching, the fact that it was less frequently noted is aligned with the findings of other studies that demonstrated that PSTs are more likely to notice teaching aspects rather than student aspects (Aydeniz & Doğan, 2016; McDonald, 2016).

In addition to what PSTs found noteworthy, we also examined analytical levels of PSTs' noticing and the types of evidence they provided in their written reports. In this way, we found that PSTs' noticing was mostly descriptive. Although descriptive and evaluative comments were more common than interpretive comments in this case, the distribution was different to the *Task* category. PSTs' interpretive level comments were more than evaluative and descriptive. These results could be because PSTs are more familiar with the *Task* aspect of a lesson than *Discourse* or *Learning Environment* based on their own experience as elementary students or based on the experience and knowledge they gained throughout the teacher education program. The results on the levels of specificity of PSTs' comments indicate that PSTs mostly noticed general features related to *Task*, *Discourse,* and the *Learning Environment*. This result is supported by the results of other studies, which focused on developing PSTs' ability to notice and found that PSTs tend to comment on the general features of teachers and learning, rather than a specific description about the content or students' understanding (Amador et al., 2016; Jacob et al., 2010). Another noteworthy finding from our study is that PSTs generally noticed specific features of each category when making interpretive comments. It can be understood that PSTs referred to a specific event or action while making interpretations about a teacher or students' actions in relation to the categories of *Task*, *Discourse,* or *Learning Environment*.

The ability to notice is of critical importance in teaching, as highlighted by reform movements that propose student-centered approaches to mathematics lessons (NCTM, 2014). PSTs' awareness of instructional elements taking place in classrooms can inform their future teaching. Therefore, it is crucial for PSTs to have the ability to learn and practice to make sense of significant events and interactions taking place (Ball & Cohen, 1999; Lampert, 2010). Here, "to learn in practice" is understood as "learning in practice from a study of practice" (Jaworski, 2009), based on *social practice theory* (Lave & Wenger, 1991); situating teachers and PSTs' learning as both individuals and as a group is important as they engage in the practice of teaching mathematics. Our findings revealed that elementary mathematics PSTs are attentive to instructional elements related to teachers (i.e., teacher's role, movements, and pedagogy) with a high level of descriptive and evaluative noticing. They can explain what is going on in the classroom (description), and sometimes make judgments about the quality of teachers' actions (evaluation). However, the level of notice showing PSTs' reasoning about how a specific classroom event, task, or content is connected with students' learning (interpretation) is less common. Additionally, these comments are less often based on the evidence presented in the video, which suggests that PSTs may need help to become more responsive to such instructional elements. Furthermore, PSTs may be inexperienced in justifying their evaluations and making reflections. Because of this, PSTs also need support in making "productive reflections" (Davis, 2006) to become more responsive to students' learning, as reform efforts have proposed (Santagana & Guarino, 2011). PSTs should be given an explicit guide or focus to develop their ability to notice productively (Choy, 2016). For example, student-related aspects and learning environments were the least noticed elements in our study; thus, a specific course focusing on these aspects can be designed to enhance PSTs' ability to notice. Also, as shown in this study's findings, PSTs' comments were specific when they made an interpretation. This implies that when encouraged to justify their reasoning using specific examples from the classroom instruction, PSTs may make more connections and integrations with their observations to specific features of instructions (van Es et al., 2017). In this sense, teacher education should deliberately focus on the development of PSTs' ability to notice

(Sherin & van Es, 2005), which will support them in not only attending, but also making sense of, these important events regarding students' thinking; this can also ensure that their future teaching decisions can be shaped accordingly. According to previous studies, PSTs can learn to notice early, namely during their methods course (Amador et al., 2016; Roth McDuffie et al., 2014) or in specifically designed courses that provide multiple contexts and a variety of mathematical content, as this can help them to become experts in attending to important elements of instruction, as well as to develop a "responsive approach to instruction" (van Es et al., 2017). Such contexts can include video cases of classroom practices or collaborative environments in which PSTs can share their ideas and learn from others.

Emine Gül Çelebi-Ilhan (gul.celeki@tedu.edu. tr) is an assistant professor in the Elementary Mathematics Education Department at TED University; in Ankara-Turkey. She has worked as a mathematics teacher at the middle and high school levels, and as a researcher at the Turkish Ministry of National Education for over fifteen years. Her research interests are mathematics teachers' noticing, mathematical discourse - language, and teacher learning.

Duygu Ören-Vural (duygu.oren@kocaeli.edu.tr) is currently working in Kocaeli University, Mathematics and Science Education Department in Education Faculty. Her special field of interests are in-service and pre-service mathematics teacher education, affective variables in mathematics teaching, and professional development programs for teachers.

References

Amador, J. M., Carter, I. & Hudson, R. A. (2016) Analyzing preservice mathematics teachers' professional noticing, *Action in Teacher Education, 38*(4), 371-383, https://www.doi.org/10.1080/01626620.2015.1119764

Anderson, G. & Walberg, H. (2003). Learning environments. In H. Walberg (Ed.), *Evaluating educational performance: A sourcebook of methods, instruments and examples.* (pp. 81-98). McCutchan Pub.

Aydeniz, M., & Dogan, A. (2016). Exploring pre-service science teachers' pedagogical capacity for formative assessment through analyses of student answers. *Research in Science & Technological Education*, *34*(2), 125–141.

Ball, D. L. (2011). Foreword. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. xx-xxiv). Routledge.

Borko, H. (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3–15.

Brophy, J. (Eds.) (2004). *Using video in teacher education.* Elsevier.

Choy, B. H. (2016). Snapshots of mathematics teacher noticing during task design. *Mathematics Education Research Journal. 28*(3), 421-440. https://www.doi.org/10.1007/s13394-016-0173-3

Clarke, H. (2001, September, 13-15). *Building education: The role of the physical environment in enhancing teaching and learning.* [Paper presentation] British Educational Research Association Annual Conference, University of Leeds. Leeds, UK.

Cleveland, B. & Fisher, K. (2014). The evaluation of physical learning environments: a critical review of the literature. *Learning Environment Research, 1,* 1–28. https://www.doi.org/10.1007/s10984-013-9149-3

Fraser, B. J. (1994). Research on classroom and school climate. In D. L. Gabel (Ed.), *Handbook of research on science teaching and learning* (pp.493-541). MacMillan

Gupta, D., Soto, M., Dick, L., Broderick, S., & Appelgate, M. (2018). Noticing and deciding the next steps for teaching: A cross-university study with elementary pre-service teachers. In G. J. Stylianides & K. Hino (Eds.), *Research advances in the mathematical education of pre-service elementary teachers – An international perspectiv*e (pp. 261-275). Springer.

Hiebert, J., Gallimore, R., Garnier, H., Givvin, K.B., Hollingsworth, H., Jacobs, J., Chiu, A.M.-Y., Wearne, D., Smith, M., Kersting, N., Manaster, A., Tseng, E., Etterbeek, W., Manaster, C., Gonzales, P., & Stigler, J. (2003). *Teaching mathematics in seven countries: Results from the TIMSS 1999 video study* (NCES 2003-013). U.S. Department of Education. National Center for Education Statistics. https://nces.ed.gov/pubs2003/2003011.pdf

Jacobs, J.K.,Yoshida, M.,Stigler J. M., & Fernandez, C. (1997). Japanese and American teachers' evaluations of mathematics lessons: A new technique for exploring beliefs. *The Journal of Mathematical Behavior. 16*(1), 7–24. https://doi.org/10.1016/S0732-3123(97)90004-3

Jacobs, V. R., Lamb, L. L., & Philipp, R. A. (2010). Professional noticing of children's mathematical thinking. *Journal for Research in Mathematics Education, 41*(2), 169–202.

Jacobs, V. R., Lamb, L. L., Philipp, R. & Schappelle B. P. (2011) Deciding how to respond on the basis of children's understanding. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: seeing through teachers' eyes* (pp. 97–116). Taylor and Francis.

Jacobs, J. K., & Morita, E. (2002). Japanese and American teachers' evaluations of videotaped mathematics lessons. *Journal for Research in Mathematics Education, 33*(3), 154–175. https://www.doi.org/10.2307/749723

Jaworski, B. (2009). Learning in and from practice. In R. Even & D.L. Ball (Eds.), *The professional education and development of teachers of mathematics* (pp. 143-148.) Springer.

Kleinknecht, M., & Gröschner, A. (2016). Fostering preservice teachers' noticing with structured video feedback: Results of an online- and video-based intervention study. *Teaching and Teacher Education, 59*, 45-56. https://dx.doi.org/10.1016/j.tate.2016.05.020

Lampert, M. (2010). Learning teaching in, from, and for practice: What do we mean? *Journal of Teacher Education, 61*(1–2), 21–34. https://www.doi.org/10.1177/0022487109347321

Lave, J., & Wenger, E. (1991). *Situated learning: Legitimate peripheral participation* (1st ed.). Cambridge University Press.

Li W., & Castro Superfine A. (2018). Understanding the work of mathematics teacher educators from a knowledge of practice perspective. In G. Stylianides & K. Hino (Eds.), *Research advances in the mathematical education of pre-service elementary teachers. ICME-13 monographs.* Springer. https://www.doi.org/10.1007/978-3-319-68342-3_19

Males L.M. (2017). Using video of peer teaching to examine grades 6–12 preservice teachers' noticing. In E. Schack, M. Fisher, & J. Wilhelm (Eds.), *Teacher noticing: Bridging and broadening perspectives, contexts, and frameworks. Research in mathematics education.* Springer. https://www.doi.org/10.1007/978-3-319-46753-5_6

Mason, J. (2002). *Researching your own practice: The discipline of noticing.* Routledge Falmer.

Mason, J. (2011). Noticing: Roots and branches. In M. G. Sherin, V. R. Jacobs, & R. A. Phillipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 35–50). Routledge.

Matos, J.F., Powell, A., Sztajn, P., Ejersbø, L., & Hovermill, J. (2009). Mathematics teachers pofessional development: Processes of learning in and from practice. In R. Even & D.L. Ball (Eds.), *The professional education and development of teachers of mathemati*cs,*New ICMI study series,* (pp. 167-184). Springer. https://www.doi.org/10.1007/978-0-387-09601-8_19

McDonald, S.P. (2016). The transparent and the invisible in professional pedagogical vision for science teaching. *School Science and Mathematics, 116* (2), 95-103. https://www.doi.org/10.1111/ssm.12156

McDuffie, A.R., Foote, M.Q., Drake, C., Turner, E.E., Aguirre, J.M., Bartell, T.G., & Bolson, C. (2014). Use of Video Analysis to Support Prospective K—8 Teachers' Noticing of Equitable

Practices. *Journal of Mathematics Teacher Education*, 17, 245-270. https://doi.org/10.5951/mathteaceduc.2.2.0108

McRobbie, C. J., & Fraser, B. J. (1993). Associations between student outcomes and psychosocial science environment. *Journal of Educational Research*, *87*(2), 78-85.

Nathan, M. & Knuth E. (2003). A study of whole classroom mathematical discourse and teacher change. *Cognition and Instruction, 21*(2), 175-207.

Putnam, R. T., & Borko, H. (2000). What do new views of knowledge and thinking have to say about research on teacher learning? *Educational Researcher, 29*(1), 4–15. https://www.doi.org/10.3102/0013189X029001004

Roller, S.A. (2016). What they notice in video: a study of prospective secondary mathematics teachers learning to teach. *Journal of Mathematics Teacher Education 19*, 477–498. https://www.doi.org/10.1007/s10857-015-9307-x

Roth McDuffie, A., Foote, M. Q., Bolson, C., Turner, E. E., Aguirre, J. M., Bartell, T. G., et al. (2014). Using video analysis to support prospective K-8 teachers' noticing of students' multiple mathematical knowledge bases. *Journal of Mathematics Teacher Education, 17*(3), 245–270. https://www.doi.org/10.1007/s10857-013-9257-0

Santagata, R., & Guarino, J. (2011). Using video to teach future teachers to learn from teaching. *ZDM, 43,* 133-145. https://doi.org/10.1007/s11858-010-0292-3

Seago, N. (2004). Using video as an object of inquiry for mathematics teaching and learning. In J. Brophy (Ed.), *Using video in teacher education* (pp. 259-286). Elsevier.

Sherin, M. G. (2004). New perspectives on the role of video in teacher education. In J. Brophy (Ed.), *Using video in teacher education* (pp. 1–28). Elsevier.

Sherin, M. G., & Han, S. (2004). Teacher learning in the context of a video club. *Teaching and Teacher Education*, *20*(2), 163-183. https://doi.org/10.1016/j.tate.2003.08.001

Sherin, M. G., Jacobs, V. R., & Philipp, R. A. (2011). Situating the study of teacher noticing. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 1–13). Routledge.

Sherin, M., & van Es, E. (2005). Using video to support teachers' ability to notice classroom interactions. *Journal of technology and teacher education, 13*(3), 475-491.

Sims, L., & Walsh, D. (2009). Lesson study with preservice teachers: Lessons from lessons. *Teaching and Teacher Education, 25*(5), 724–733. https://doi.org/10.1016/j.tate.2008.10.005

Stahnke, R., Schueler, S., & Roesken-Winter, B. (2016). Teachers' perception, interpretation, and decision-making: a systematic review of empirical mathematics education research. *ZDM*, 48(1–2), 1–27.

Stein, M. K., Grover, B. W., & Henningsen, M. (1996). Building student capacity for mathematical thinking and reasoning: An analysis of mathematical tasks used in reform classrooms. *American Educational Research Journal, 33*(2), 455-488.

Star, J. R., & Strickland, S. K. (2008). Learning to observe: Using video to improve preservice teachers' ability to notice. *Journal of Mathematics Teacher Education, 11,* 107-125. https://doi.org/10.1007/s10857-007-9063-7

Star, J. R., & Strickland, S. K. (2011). Reflections on the study of teacher noticing. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: seeing through teachers' eyes* (pp. 66–78). Routledge.

Star, J. R., Lynch, K. H., & Perova, N. (2011). Using video to improve mathematics' teachers' abilities to attend to classroom features: A replication study. In M. Sherin, V. Jacobs, & R. Philipp, (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes,* (pp. 117–13). Routledge.

Stigler, J.W., Gonzales, P., Kawanaka, T., Knoll, S., & Serrano, A. (1999). *The TIMSS videotape classroom study: Methods and findings from an exploratory research project on eighth-grade mathematics instruction in Germany, Japan, and the United States.* (NCES 1999-074). U.S. Department of

Education. National Center for Education Statistics
https://nces.ed.gov/pubs99/1999074.pdf

Strauss, A., & Corbin, J. M. (1990). *Basics of qualitative research: Grounded theory procedures and techniques.* Sage.

van Es, E.A. & Sherin, M.G. (2002). Learning to notice: Scaffolding new teachers' interpretations of classroom interactions. *Journal of Technology and Teacher Education, 10(4), 571-596.*

van Es, E. A., & Sherin, M. G. (2006). How different video club designs support teachers in "learning to notice". *Journal of Computing in Teacher Education*, *22*(4), 125-135. http://dx.doi.org/10.1080/10402454.2006.10784548

van Es, E. A., & Sherin, M. G. (2010). The influence of video clubs on teachers' thinking and practice. *Journal of Mathematics Teacher Education*, *13*(2), 155-176. https://doi.org/10.1007/s10857-009-9130-3

van Es, E. A. (2011). A framework for learning to notice student thinking. In M. Sherin, V. Jacobs & R. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 134-151). Routledge.

van Es E.A., Hand V. & Mercado, J. (2017). Making visible the relationship between teachers' noticing for equity and equitable teaching practice. In E. O. Schack, M. H. Fisher, & J. A. Wilhelm (Eds.), *Teacher noticing: Bridging and broadening perspectives, contexts, and frameworks.* (pp. 251–270). Springer International Publishing.

Van Zoest, L., & Enyart, A.(1998) Discourse of course: encouraging genuine mathematical conversations. *Mathematics Teaching in the Middle School, 4*(3), 150-7.

Appendix

Guiding questions for the written video case analysis is presented below.

Task

What types of tasks are selected by the teacher?
What is the level of appropriateness of the task for students?
How did the teacher implement the task?

Discourse

What are the roles of the teacher & students in supporting classroom discourse?
What are the resources and tools used for supporting classroom discourse?
How does the teacher respond to students' (wrong) answers?

Learning Environment

How does the teacher manage time, physical space and social relations in the classroom?
What instructional group strategies are employed in the classroom?
What are the teacher's expectations for her students?