Volume 27, Number 2 • Summer 2023
ISSN: 1087-3430

Electronic Journal for Research in
Science & Mathematics Education

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

ICRSME

EJRSME
Electronic Journal for Research in Science & Mathematics Education

EDITORS
Mark A. Bloom, Science Education, *Dallas Baptist University*
Sarah Quebec Fuentes, Mathematics Education, *Texas Christian University*

MANAGING EDITOR
Morgan Jansing, *Texas Christian University*

COPY EDITOR
Audrey Meador, *West Texas A&M University*

ASSOCIATE EDITORS
Daniel Alston, *University of North Carolina at Charlotte*
Lisa Ann Borgerding, *Kent State University*
Danxia Chen, *Dallas Baptist University*
Matthew Perkins Coppola, *Purdue University*
Kimberly Figaro, *Dallas Baptist University*
Kathryn Green, *Clarke County School District*
Hayat Hokayem, *Texas Christian University*
Gabriela Jonas-Ahrend, *Paderborn University*
Ravinder Koul, *The Pennsylvania State University*
Stefanie Livers, *Missouri State University*
Eryn Maher, *Georgia Southern University*
Cherie McCollough, *Texas A&M University – Corpus Christi*
James A. Mendoza Álvarez, *University of Texas at Arlington*
Erin Peters-Burton, *George Mason University*
Christine Schnittka, *Auburn University*
Byung-In Seo, *Chicago State University*
Julie Westerlund, *Texas State University*
Robert Wieman, *Rowan University*
Dawn Woods, *Oakland University*

PUBLISHER
ICRSME, International Consortium for Research in Science & Mathematics Education

Cover designed by Dusty Crocker

CONTENTS

Volume 27 No. 2 | Summer 2023

RESEARCH / EMPIRICAL

ALL RIGHTS RESERVED
© International Consortium for Research in Science & Mathematics Education (ICRSME)
Electronic Journal for Research in Science & Mathematics Education (EJRSME)

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 2, i-vii

The Intricacies of the STEM Teacher Shortage

Sarah Quebec Fuentes
Texas Christian University

Mark A. Bloom
Dallas Baptist University

Introduction

When conducting an Internet search on the *national teacher shortage* in the popular media, words such as *crisis*, *colossal*, and *critical* appear. However, the nature of the teacher shortage in the United States is also complex (Barnum, 2022; McVey & Trinidad, 2019).[1] Factors including region of the country, district-type, and subject areas must be considered. For instance, states have consistently reported a teacher shortage in STEM subject areas (e.g., AAEE, 2023; McVey & Trinidad, 2019). The present editorial is the opening of a series addressing our role as science and mathematics teachers and teacher educators in addressing the STEM teacher shortage. We start by presenting the problem. The following sections share data and literature that substantiate the shortage of teachers in STEM subject areas as well as related issues including declining enrollment in teacher education programs, perceptions of the teaching profession, race/ethnicity of the teacher population, and out-of-field teachers.

Teacher Shortage

Evidence shows a teacher shortage in STEM subject areas in the United States. Over the last two decades, states have identified shortages of teachers in Mathematics and Science (McVey & Trinidad, 2019). In the American Association for Employment in Education (AAEE, 2023) report, school districts conveyed *Considerable Shortages* of qualified applicants in all STEM subject areas (Mathematics, General Mathematics and Science, Physics, Chemistry, Biology, and Earth/Physical Science) for the Academic Year (AY) 2022-2023 as well as over the previous eight academic years (with a transition from mainly *Some Shortage* to mainly *Considerable Shortage*). This finding held true across most regions of the United States (with three regions reporting *Some Shortage* in some of the STEM subject areas). Data of teacher shortages for the AY 2022-2023 from the U.S. Department of Education (U.S. DOE, 2023) documents 31 states with shortages of mathematics teachers and 34 states with shortages of science teachers.

Further, school districts overall reported that 25% of newly hired teachers did not have traditional preparation, and 80% of school districts described *having enough candidates for open positions* as a *Big Challenge* in the hiring process (AAEE, 2023). For the AY 2022-2023, the majority of public schools with a vacancy reported *Somewhat Difficult* or *Very Difficult* for the anticipated level of difficulty in filling positions in Computer Science, Mathematics, Biology or Life Sciences, and Physical Sciences (IES, 2022).

---

[1] We want to acknowledge that the editorials focus on data and literature reflecting the situation in the United States, which may differ from other countries (e.g., Fray & Gore, 2018; Han et al., 2018).

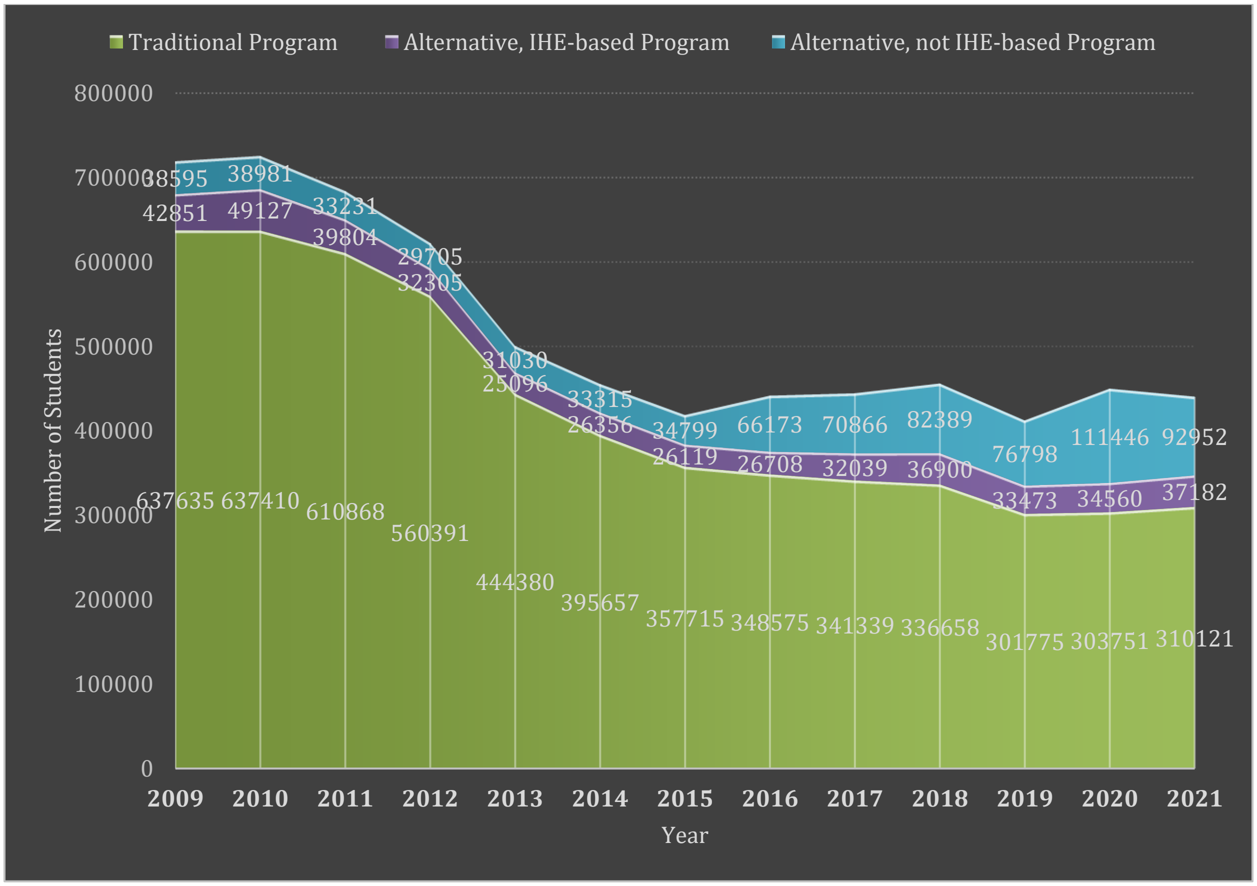
© 2023 International Consortium for Research in Science & Mathematics Education (ICRSME)

Declining Enrollment in Teacher Education Programs

One possible reason for the teacher shortage could be that enrollment in Teacher Education Programs (TEP) in the United States has been declining (AAEE, 2023; U.S. DOE, 2022a). Figure 1 shows enrollment by program type (Traditional Programs; Alternative, Institute of Higher Education [IHE] Based Programs; and Alternative, not IHE-Based Programs). Across all program types, enrollment has decreased by 39% for AY 2008-2009 through AY 2020-2021. Specifically, enrollment has decreased in Traditional Programs by 51% and in Alternative IHE-Based Programs by 13%. In contrast, enrollment in Alternative not IHE-Based Programs has increased by 141% (U.S. DOE, 2022b). Much of this growth is due to the enrollment in Alternative not IHE-Based Programs in Texas (e.g., 75% of the total enrollment in AY 2019-2021) as well as the growth in enrollment in Texas. For additional data by state, refer to the report, *Preparing and Credentialing the Nation's Teachers* (U.S. DOE, 2022a), and associated data sets (U.S. DOE, 2022b).

Figure 1

Teacher Education Program Enrollment from 2008 to 2021 by Program Type (U.S. DOE, 2022b)



Similar trends exist across STEM fields in the United States. Table 1 presents the percent change (between AY 2012-2012 and AY 2018-2019) in program completers in STEM subject areas by the three program types (U.S. DOE, 2022a). Overall, the percent change shows a decrease in program completers for all STEM subject areas except for Computer Science. In particular, the percent change indicates a decrease across all STEM subject areas in Traditional Programs and across most STEM subject areas in Alternative not IHE-Based Programs. Alternatively, the percent change demonstrates an increase in Alternative IHE-based Programs in all subject areas apart from mathematics. The findings of the AAEE (2023) report likewise align, with Colleges and Universities reporting a *Considerable Shortage* in qualified candidates in STEM degree programs (Mathematics, General Mathematics and Science, Physics, Chemistry, Biology, and Earth/Physical Science).

Table 1

Change (and Percent Change) in Program Completers in STEM Subject Areas from 2012 to 2019 (U.S. DOE, 2022a)

	Subject Area						
Program Type	**Mathematics**	**General Science**	**Biology**	**Chemistry**	**Physics**	**Earth Science**	**Computer Science**
Traditional	-47774 (-42%)	-1212 (-30%)	-918 (-28%)	-295 (-29%)	-149 (-27%)	-202 (-36%)	-15 (-40%)
Alternative, IHE-Based	-95 (-8%)	141 (26%)	108 (21%)	45 (26%)	11 (12%)	43 (80%)	6 (300%)
Alternative, not IHE-Based	34 (2%)	570 (46%)	-123 (-23%)	-105 (-49%)	-60 (-41%)	-46 (-42%)	12 (67%)
Total	-4835 (-34%)	-501 (-9%)	-933 (-21%)	-355 (-25%)	-198 (-25%)	-205 (-28%)	3 (6%)

Perceptions of the Teaching Profession

The prestige of a profession "can be understood to mean the reputation and social standing that the profession holds in society as well as the respect and authority workers are afforded as professionals" (Kraft & Lyon, 2022, p. 7). In their examination of the state of the K-12 teaching profession in the U.S. over the last 50 years, Kraft and Lyon (2022) found that the prestige of the teaching profession experienced a decline in the 1970s, a rise in the 1980s, and a steady rate for 20 years prior to the most recent decline starting around 2010. Currently, the perception of the teaching profession is at its lowest.

The present trend pertaining to the prestige of the teaching profession has implications for student interest in pursuing a career in education. Since the public's view of professions influences students' choice of career (Christensen et al., 2019; Han et al., 2018; Kraft & Lyon, 2022), student interest in becoming a teacher has followed a similar pattern as that of the prestige of the profession. Since 2010, interest in the teaching profession has decreased to its lowest level in the last half century (Bartanen & Kwok, 2022; Kraft & Lyon, 2022).

Students' career interests are also influenced by family perceptions of professions (Christensen et al., 2019, 2022; Kraft & Lyon, 2022). With respect to the teaching profession, Christensen et al. (2022) found that one of the factors that predicted *whether parents believe teaching would be the best career option for their children* (p. 9) was parent perception of societal respect for teachers. Notably, in a Phi

Delta Kappan (PDK) Poll (2022), 62% of the adults surveyed indicated that they would not want their child to become a public school teacher in their community. In 2018, this percent was 54%, and all previous surveys the percent was less than 50%.

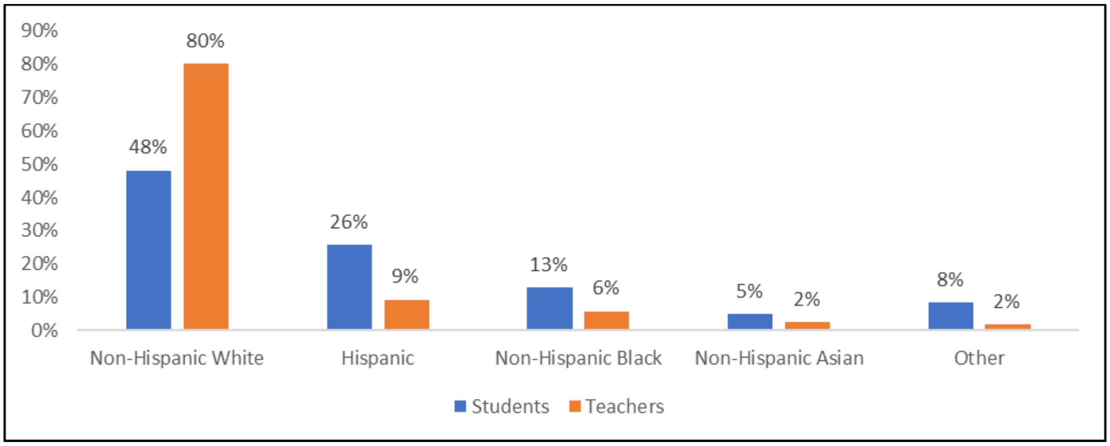
Race/Ethnicity of the Teacher Population

TEPs have made efforts to recruit and retain diverse teachers of color as they recognize the benefits that they bring to the education of students of color (Sleeter et al., 2014). Teachers whose race/ethnicity matches their students are better able to serve as mentors and role models and can better understand their students' cultural backgrounds (Egalite et al., 2015). As early as the pre-Kindergarten years, research has noted improved learning outcomes for students when they are matched with teachers who share their race/ethnicity (Downer et al., 2016). For instance, Gershenson et al. (2021) found that black students who had just one black teacher by the third grade are 13% more likely to graduate high school and 19% more likely to enroll in college than their same-race/same-school peers. The same study revealed that no such trend was detected among white students.

Despite the evidence that a diverse teacher workforce is better for a diverse student body, Figure 2 shows that a great disparity still exists between the racial/ethnic make-up of U.S. K-12 public school students (Fabina et al., 2023) and the corresponding teacher workforce (Taie & Lewis, 2022). In 2021, non-Hispanic white students made up only 48.1% of the total enrolled students, yet 79.9% of K-12 teachers were white. In all other racial/ethnic groups, teachers were underrepresented. Hispanic students comprised 26% of enrolled students and 9% of teachers were Hispanic; Black students made up 13% of enrolled students and 6% of teachers were Black; and Asian students accounted for 5% of student enrollment and 2% of teachers were of Asian descent. These numbers closely match the diversity among science teachers. Zippia Career Expert (2023) used a database of 30 million profiles to estimate the diversity among U.S. Science teachers and found the most common ethnicity of Science and Mathematics teachers to be white (72% Science, 72% Mathematics), followed by Hispanic (12% Science, 12% Mathematics), Black (8% Science, 8% Mathematics), and other (4% Science, 4% Mathematics).

Figure 2

Percent of Students and Teachers in K-12 Schools in 2021



On top of the racial disparity, teachers of color are also reported to be leaving the profession at a higher rate than their white counterparts. A study by the Massachusetts Department of Elementary and Secondary Education (DESE, 2019) found an attrition rate of 12.5% among white teachers in AY

2016-2017, but a higher rate of 17% for Latinx teachers and 24% for Black teachers. This trend appears across the country; Steiner and Woo (2021) found that one in four teachers indicated they planned to leave the profession by the end of AY 2020-2021, and that number was almost 50% among Black teachers.

Out-of-Field Teachers

This problem of teachers assigned to subjects that do not align with their area of expertise spans many subject areas but is especially challenging for Mathematics and Science teachers. As far back as 1990, the National Center for Education Statistics (NCES, 1996) found that among public school Mathematics students, approximately 25% were taught by Mathematics teachers who lacked even a minor in Mathematics or Mathematics Education. Similarly, 39% of public school Biology students had teachers who lacked even a minor in Biology or Life Science. The problem was even worse for Physical Science with 56% of public school Physics students in classrooms taught by teachers who lacked even a minor in Physics, Chemistry, Geology, or Earth Science.

The problem has not improved over time. Shaw et al. (2019) indicated that between 2003 and 2016, the number of out-of-field teachers in Chemistry and Physics classes has continued to increase. Among Mathematics teachers, only 58% were teaching in-field. Among those teaching out-of-field, 14% not only lacked an in-field degree, but also lacked in-field certification (Shaw et al., 2019). Further, Nixon (2017) identified that out-of-field teaching was much more prevalent among new teachers as compared to seasoned, long-term teachers. Their data revealed that among 100 new Science teachers, more than 60% were required to teach Science subjects for which they were not adequately prepared.

Conclusion

In the present editorial, we have acknowledged the ongoing shortage of teachers in general and the fact that current enrollment in TEPs do not indicate an end to this shortage. We have further identified the problems of a general decline in the public's perceptions of the teaching profession, a lack of diversity among K-12 teachers, and teachers assigned to subjects outside of their area of expertise. In the next editorial, we plan to discuss various reasons that have been attributed to the trends presented herein. We will then use the foundational knowledge from the first editorials to elaborate on practical ways in which we, as Mathematics and Science teachers and teacher educators, can contribute to building the STEM teacher workforce.

References

American Association for Employment in Education (AAEE). (2023). *Educator supply and demand report 2022-23*. https://aaee.glueup.com/event/2022-2023-educator-supply-demand-report-digital-access-78774/

Barnum, M. (2022, August 11). Is there a national teacher shortage? Here's what we know and don't know. *Chalkbeat*. https://www.chalkbeat.org/2022/8/11/23300684/teacher-shortage-national-schools-covid

Bartanen, B., & Kwok, A. (2022). *From interest to entry: The teacher pipeline from college application to initial employment century* (EdWorkingPaper No. 22-535). Annenberg Institute at Brown University. https://doi.org/10.26300/hqn6-k452

Christensen, S., Davies, R. S., Harris, S. P., Hanks, J., & Bowles, B. (2019). Teacher recruitment: Factors that predict high school students' willingness to become teachers. *Education Sciences, 9*, 1-12. doi:10.3390/educsci9040282

Christensen, S., Davies, R. S., Larsen, R. A. A., Harris, S., Hanks, J., & Bowles, B. (2022). Parental perceptions of the teaching profession: Factors that predict parental encouragement of students to enter the teaching profession. *Education Sciences, 12*, 1-13. https://doi.org/10.3390/educsci12110734

DESE (2019). *Educator Workforce Diversity*. Massachusetts Board of Elementary and Secondary Education. 5/21/2019 BESE Regular Meeting: Educator Workforce Diversity - Massachusetts Board of Elementary and Secondary Education

Downer, J. T., Goble, P., Myers, S. S., & Pianta, R. C. (2016). Teacher-child racial/ethnic match within pre-kindergarten classrooms and children's early school adjustment. Early Childhood Research Quarterly, 37(4). https://doi.org/10.1016/j.ecresq.2016.02.007.

Egalite, A. J., Kisida, B., & Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, *45*, 44-52. doi.org/10.1016/j.econedurev.2015.01.007

Fabina, J., Hernandez, E. L., & McElrath, K. (2023). *School enrollment in the United States: 2021*. American Community Survey Reports, U.S. Census Bureau. School Enrollment in the United States: 2021 (census.gov)

Fray, L., & Gore, J. (2018). Why people choose teaching: A scoping review of empirical stufies, 2007-2016. *Teaching and Teacher Education, 75*, 153-163. https://doi.org/10.1016/j.tate.2018.06.009

Gershenson, S., Hart, C., Hyman, J., Lindsay, C., & Papageorge, N. W. (2021). *The long-run impacts of same-race teachers* (Working paper 25254). National Bureau of Economic Research. https://www.nber.org/papers/w25254

Han, S. W., Borgonovi, F. & Guerriero, S. (2018). What motivates high school students to want to be teachers? The role of salary, working conditions, and societal evaluations about occupations in a comparative perspective. *American Educational Research Journal, 55*(1), 3-39. DOI: 10.3102/0002831217729875

Ingersoll, R. M. (1999). The problem of underqualified teachers in American secondary schools. *Educational Researcher*, *28*(2), 26-37. https://doi.org/10.3102/0013189×028002026

Institute of Education Sciences (IES). (2022). *Staffing* [Data set]. School Pulse Panel. https://ies.ed.gov/schoolsurvey/spp/

Kraft, M. A., & Lyon, M. A. (2022). *The rise and fall of the teaching profession: Prestige, interest, preparation, and satisfaction over the last half century* (EdWorkingPaper No. 22-679). Annenberg Institute at Brown University. https://doi.org/10.26300/7b1a-vk92

NCES (October, 1996). Out-of-Field Teaching and Educational Equality. National Center for Education Statistics. https://nces.ed.gov/pubs/web/96040.asp

Nixon, R. S., Luft, J. A., & Ross, R. J. (2017). Prevalence and predictors of out-of-field teaching in the first five years. *Journal of Research in Science Teaching, 54*(9), 1197-1218. https://doi.org/10.1002/tea.21402.

Phi Delta Kappan (PDK) International. (2022). *PDK poll of the public's attitudes toward the public schools.* https://pdkpoll.org/2022-pdk-poll-results/

Shaw, L., Jannuzzo, C., Hassan, T., Gadidov, B., Ray, H. E., & Rushton, G. T. (2019, September 25,). Diagnozing the current state of out-of-field teaching in high school science and mathematics. *Plos One.* https://doi.org/10.1371/journal.pone.0223186

Sleeter, C. E., Neal, L. I., & Kumashiro, K. K. (2014). *Diversifying the teacher workforce: Preparing and retaining highly effective teachers*. Routledge.

Steiner, E. D., & Woo, A. (2021). *Job-related stress threatens the teacher supply: Key findings from the 2021 State of the U.S. Teacher Survey.* Rand Corporation. https://api.semanticscholar.org/CorpusID:242350561

Taie, S., & Lewis, L. (2022). *Characteristics of 2020-2021 public and private K-12 school teachers in the United States: Results from the National Teacher and Principal Survey first look* (NCES 2022-113). U.S. Department of Education, National Center for Education Statistics. https://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2022113

U.S. Department of Education (DOE). (2022a). *Preparing and credentialing the nation's teachers: The secretary's report on the teacher workforce.* Office of Postsecondary Education, Title II Higher Education Act. https://title2.ed.gov/Public/OPE%20Annual%20Report.pdf

U.S. Department of Education (DOE). (2022b). *Enrollment, by state, by program type* [Data set]. Office of Postsecondary Education, Title II Higher Education Act. https://title2.ed.gov/Public/DataTools/Tables.aspx

U.S. Department of Education (DOE). (2023). *Teacher shortage areas* [Data set]. https://tsa.ed.gov/#/home/

Zippia (2023). Science Teacher Demographics and Statistics [2023]: Number of Science Teachers In The US (zippia.com). https://www.zippia.com/science-teacher-jobs/demographics/

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 2, 1-24

ICRSME
International Consortium for Research in
Science & Mathematics Education

Teachers as Learners: Outdoor Elementary Science

Sarah J. Carrier
North Carolina State University

M. Nils Peterson
North Carolina State University

Aimee B. Fraulo
North Carolina State University

Laura M. Romeo
North Carolina State University

Kathryn T. Stevenson
North Carolina State University

ABSTRACT

Young learners benefit from learning and engaging in nature, but nature-based outdoor learning impacts on teachers are less understood. The present study examined elementary school teachers' learning experiences with science instruction in outdoor natural settings. An Outdoor Science Education (OSE) program partnered with schools in the southeastern U.S. to provide standards-aligned outdoor science lesson expeditions for fifth-grade students. Teachers reported their participation in expeditions improved their own content knowledge in science and contributed to their confidence teaching science and teaching in the outdoors. In addition, teachers reported their perceptions of students' performance in their science classes and on standardized science tests. The teachers described their perceptions about how the expeditions improved their students' engagement for learning and behavior in outdoor settings when compared to indoor science instruction. They discussed how these changes were most obvious for students who struggled with traditional classroom instruction. Using survey and interview data, we discuss the potential for teacher learning related to outdoor science instruction.

Keywords: outdoors, elementary, science, student engagement

Introduction

Learning science in the outdoors may be particularly effective for young learners, but student success depends on elementary science teachers' instructional practices (National Science Teachers' Association [NSTA], 2020). Outdoor science instruction has shown potential to enhance motivation, behavior, and content mastery for many students (Carrier, 2013; Carrier et al., 2014; Szczytko et al., 2018; Cheng & Monroe, 2012), and outdoor experiences can also support students' critical thinking skills (Ampuero et al., 2015). When science instruction occurs in nature, students can connect directly with science's cross-cutting concepts such as patterns in nature, life cycles, energy flow, and form and function (Cooper, 2015; Fägerstam & Blom, 2013; Next Generation Science Standards [NGSS] Lead States, 2013). Here we define nature as plants, animals, and other objects and organisms in the outdoors not made by people. Further, science instruction in nature supports interdisciplinary learning where students connect science with literacy, mathematics, and social studies (Eick, 2012; Junker & Jacquemin, 2016; McMillan & Vasseur, 2010). Linking science with other subject areas is known to be an effective instructional practice; connects content to students' lives (National Research Council

© 2023 International Consortium for Research in Science & Mathematics Education (ICRSME)

[NRC], 2000; 2004), motivates students to engage in their learning (Banilower et al., 2010), and prepares students for careers in science (Feinstein et al., 2013). While the benefits of outdoor learning experiences for students are well-documented (Cooper, 2015; Harvey et al., 2020; Rios & Brewer, 2014; Wistoft, 2013), less is known about how outdoor learning experiences impact teachers, particularly K-5 teachers who often have limited science backgrounds (Smith, 2020). Here we examine teachers' participation in one outdoor science education (OSE) program's support of science instruction at elementary schools in the southeastern United States. Our research questions asked:

1. What are the relationships between elementary school teachers' participation in OSE experiences led by naturalists and their perceptions of their own and their students' science content knowledge?
2. What are the relationships between elementary school teachers' participation in OSE experiences led by naturalists and their perceptions of science instruction and outdoor science instruction?
3. What are teachers' perceptions of their students' engagement in science and behaviors following their participation in an OSE program?

We begin with a sociocultural frame, then provide an overview of the documented need to support elementary teachers' science instruction, outdoor learning, and teachers' expectations of student learning. Each of these components positions our study's focus on *teachers as learners.*

Theoretical Framework

The present study is situated in Vygotsky's (1978) sociocultural framework that identifies learning as an interactive phenomenon by which people learn through their social interactions. Humans develop cognitively through social interactions in guided learning, where participants co-construct knowledge with other learners (McLeod, 2014). According to Vygotsky, learning occurs through interactions with others who possess varying degrees of knowledge. Those with a higher level of knowledge are referred to as a "More Knowledgeable Other" (MKO). The MKO may model behavior or vocabulary that guides other learners. The MKO does not necessarily need to be a formal teacher but may be a peer or a tool (in the case of a computer tutorial program or device) that holds more knowledge about a topic than the learner. When members of a group learn and develop together, they enact Vygotsky's concept of MKO (Abtahi, 2017; McLeod, 2014; Moalosi, 2013).

As with young learners, sociocultural experiences also support adult learners (Wang, 2018; Wang et al., 2010). Even though teachers are generally viewed as the MKO relative to their students, they can simultaneously share the role of learners with their students. When teachers possess science content knowledge and explore ways to share this knowledge with their students, teachers' own learning continues as they hone their instructional practices. Further, science instruction reforms note that "learning science involves learning a system of thought, discourse, and practice—all in an interconnected and social context—to accomplish the goal of working with and understanding scientific ideas" (NRC, 2012, p. 252). In this study, we focus on teachers as learners and examine the relationships of their participation in outdoor expeditions with their perceptions of their students and of their own science content knowledge and instruction practices.

Literature Review

The marginalization of science instruction in elementary schools is well-documented and persistent. Data from the *2018 National Survey of Science and Mathematics Education* (Plumley, 2019) suggest that only 18% of primary and 26% of intermediate elementary students receive science instruction on all or most days in a school year. Teachers report challenges that inhibit science instruction beyond time limitations that include elementary teachers' low self-efficacy in science and few opportunities for teacher professional development.

In a recent national survey examining trends in science education (Smith, 2020), only 31% of elementary teachers reported they felt well-prepared to teach science. During their teacher preparation, only 34% of elementary teachers participated in science content coursework in life, physical, and earth science as recommended by NSTA (2012). A comparison of survey data from 2012 to 2018 found that elementary teachers' beliefs about science instruction continue to align with traditional instruction rather than reform-based teaching practices. These data indicate a lack of teacher exposure to professional opportunities to learn and rehearse effective science instructional practices to support their students (Smith, 2020).

While such findings identify a need for teachers' professional development in science, less than 60% of elementary teachers reported participating in science content or science methods professional development in the preceding three years, and 24% had never participated in science-related professional development (Smith, 2020). Such opportunities for professional development in science vary widely and are offered at varying frequencies. When professional development focuses on content knowledge and includes active learning, teachers report positive increases in their knowledge, skills, and changes in their classroom practices (Garet et al., 2001).

As teachers build their content knowledge, their knowledge of teaching practices can also expand beyond traditional instruction to include reform instruction (Feiman-Nemser, 2008). This content knowledge (CK) and the practice of teaching, or pedagogical knowledge (PK), contribute to teachers' pedagogical content knowledge (PCK) (Loughran et al., 2001; Shulman, 1986). Teachers with solid content knowledge are often also pedagogically skilled at incorporating effective instructional practices such as rich questioning during student engagement in science activities (Anderson, 2002; Anderson & Helms, 2001; Kennedy, 1998). Baumert and Kunter's (2013) model of professional competence includes CK, PCK, and PK, as well as teachers' beliefs about and motivation to teach. Park and Oliver (2008) developed an organizational tool of observable components of PCK that includes teachers' "(a) orientations to science teaching, (b) knowledge of students' understanding in science, (c) knowledge of science curriculum, (d) knowledge of instructional strategies and representations for teaching science, and (e) knowledge of assessments of science learning" (p. 264). Their pentagonal model supports the dynamic nature of teacher development of PCK.

Helping teachers shift from traditional teacher-centered instruction to reform-based student-centered teaching demands professional development that provides models and supports ongoing teacher rehearsals with this pedagogy (Desimone et al., 2002). In one study that documented teachers' struggles with this shift, teachers designed outdoor spaces for their students as part of their professional development. These teachers reflected on their challenges in shifting from teacher-centered instruction to student-centered instruction (Catapano, 2005). Desimone (2009) identified features of effective professional development that include active learning, coherence with classroom instruction, content focus, collective participation with PD developers and other teachers, and opportunities to practice. While the OSE in the present study is structured for student learning, we identify features that served as professional development for teachers, including active learning in the outdoors (Glackin, 2016).

Outdoor Learning

Researchers have identified the complexity of learning features that inform best practices for student learning. In *How People Learn II,* the roles of environmental, social, and cultural influences are described:

> Learning is a dynamic, ongoing process that is simultaneously biological and cultural. Attention to both individual factors (such as developmental stage; physical, emotional, and mental health; and interests and motivations), as well as factors external to the individual (such as the environment in which the learner is situated, social and cultural contexts, and opportunities available to learners), is necessary to develop a complete picture of the nature of learning (National Academies of Science, Engineering, and Medicine [NASEM], 2018, p. 9).

Outdoor science instruction has the potential to support students' physical, socio-emotional, and mental health developmental factors, and situating learning in the outdoors can motivate student learning. Researchers have found that outdoor instruction can improve students' science test scores, support their applications of science content to real-world situations, increase student enthusiasm and interest in science, and enhance emotional and cognitive development (Dillon et al., 2006; Lieberman & Hoody, 1998; Kahn & Kellert, 2002; Malone, 2008). Such enhanced cognitive and affective outcomes (Eaton, 2000) have further been found to improve student attitudes and comfort in the outdoors (Carrier, 2009; Cheng & Monroe, 2012) and encourage students' critical thinking skills (Hungerford & Volk, 1990). In an analysis of 18 research articles on outdoor learning, Ayotte-Beaudet et al. (2017) found that "when outdoor teaching and indoor teaching are coordinated to complement one another, a more positive effect on learning can be expected than when compared to indoor teaching alone" (p. 5355). These studies showed that students can make connections beyond the classroom when teachers combine outdoor experiences with indoor science instruction. Such shared social connections enhance students' conceptual knowledge by providing students authentic applications for their new learning.

In addition, students' direct connections to nature have been found to increase their interest in and motivation to learn science (Lindemann-Matthies, 2005; Skinner and Chi, 2012; Zoldosova & Prokop, 2006). Other researchers have found that sustained experiences in nature can impact young children's attitudes about nature that extend into adulthood (Ewert et al., 2005; Wells & Lekies, 2006). By facilitating children's connection to nature, such experiences have potential to enhance students' lasting interest in science.

Despite the affordances that outdoor contexts provide for learning about topics such as weather patterns or life cycles (Glackin, 2016; Rickinson et al., 2004), teachers rarely situate their instruction in the outdoors. Teachers have reported obstacles to outdoor instruction that include their limited awareness of outdoor activities, perceptions of a disconnect with instructional standards, time constraints, lack of administrative support, classroom management concerns, transportation costs, limited resources, and safety risks (Lock, 2010; Palavan et al., 2013). As teachers learn to negotiate these challenges, they can provide young students opportunities to engage with phenomena in outdoor settings and build a strong base for students' future learning.

Elementary Science Instruction

Including authentic science experiences in primary grades helps children learn and are the context for the present study. Decades of research on teaching and learning science have revealed that young children's scientific reasoning skills are more sophisticated than previously thought (NRC, 2007). Teachers can capitalize on these findings by providing students with early and rich science learning opportunities that focus children's scientific reasoning skills and help shape their experiences and prolong their interest in science. While student interest in science can persist through elementary school (DeWitt & Osborne, 2007), studies have found that their interest frequently wanes as they enter middle school (Archer et al., 2010). Children's early interest in science is a strong indicator of their potential to pursue science studies and careers (Maltese & Tai, 2010), emphasizing the need for teachers to nurture in students an enduring interest and engagement in science. Significantly, teachers' expectations can influence students' motivation and performance (Florea, 2007) and play a critical role in capturing and nurturing their students' interest in science. Because students' K-5 experiences frequently establish a trajectory for their future learning, teachers' expectations and support are critical for all students, including those traditionally underserved in science classrooms and careers. Zargona and colleagues (2017) identified a need for more research on strategies for meeting the needs of all students in inclusion classrooms, and these needs include teachers' positive expectations.

Expectation Bias

Much research has focused on how teachers' expectations of students from academically stereotyped populations may diminish the quality of instruction they deliver to these students (Weinstein et al., 2004). Expectation theories identify the impact of teachers' expectations on student achievement (Cooper, 2000). McKown and Weinstein (2002) found that African American students were more vulnerable to negative teacher expectations than their European American peers. Other studies have identified that teacher expectations related to students' gender and ethnicity can influence student performance (Retelsdorf, 2015; Rong, 1996; Timmermans et al., 2015). Relatedly, teacher expectations for students identified with learning disabilities have been found to impact their ratings of students' personal attributes and achievements (Rubie-Davies, 2006; Sorhagen, 2013; Woodcock & Vialle, 2011). When teachers' interactions with their students extend beyond the traditional classroom experiences, there is potential to broaden teachers' perceptions of their students' potential. In her work with schoolyard learning, Feille (2013) noted that when instruction expands beyond the classroom walls, teachers and students share direct experiences in the outdoors that can positively impact teacher-student relationships. Situating some science lessons in the outdoors has potential to reach students who may not be well served by traditional instruction. The range of experiences possible in the outdoors offers opportunities for teachers and students to engage in multiple learning modalities making learning accessible to a broad range of students (Eick, 2012; Harris, 2018; Rios & Brewer, 2014).

Methodology

This study is part of a larger examination of an OSE program that focused primarily on student outcomes (Szczytko, 2018; Stevenson, 2021). The present study extends this work to examine the teacher participants' perceptions of the OSE program's impact on their students' science learning and on their own learning of science content and science instruction. A mixed methods embedded design (Creswell, 2014) was employed in this study. All participants were recruited from the OSE program's database (N = 65).

The primary data source presented here is qualitative semi-structured interviews (see Appendix A) with teachers (n = 30) regarding their experiences in the OSE and their perceptions of their students during and after these experiences. In order to supplement the qualitative data, we decided to also include survey data to quantify the response categories and identify patterns and tendencies with the corresponding qualitative data. We sent a survey request the same sample population of the 65 OSE teachers, and to protect teacher identity of those who responded (n = 33), we administered an anonymized five-question targeted quantitative survey (see Appendix B). Embedded designs are appropriate for these analyses as the additional quantitative data supplements the primary qualitative data to provide additional evidence and aid in data interpretation (Creswell & Plano Clark, 2011). This protocol was approved as exempt from the ethics review process by the Institutional Review Board for the Protection of Human Subjects in Research (NC State University IRB # 12084IRB).

Study Context

The context of this study highlights a partnership between one private, non-profit OSE program and participating schools, teachers, and students in a southeastern US state. The partners' mutual goals focused on supplementing fifth-grade teachers' science instruction in the classroom and engaging students in outdoor learning experiences. The teachers' main goals were to enhance their students' science learning, and the OSE program's primary mission was to excite students about the wonders of science in nature situated in outdoor expeditions. These complementary goals positioned the OSE program to guide students and their teachers through a host of outdoor activities directly aligned with the state's science standards.

Schools apply to participate in the OSE program, and once approved, instructors called "naturalists" meet with teachers before their first expedition. During these meetings, the naturalists and teachers schedule multiple expeditions across the school year, and together they collaboratively align their presentation topics with state science standards and teachers' schedules. In the state where the study was conducted, as in many other US states, fifth grade is the only elementary school grade where science is assessed; thus, the emphasis on the state's science standards. The OSE naturalists developed outdoor activities to facilitate students' engagement in outdoor activities while also supporting teachers' goals by addressing the state standards-based science content that is tested. The grade 5 standards include ecosystems, weather, landforms, and life cycles that align well with outdoor learning. The OSE naturalists' science content activities exposed students and teachers to science practices as they asked students to generate and interpret evidence, and they scaffolded students' efforts to form explanations of the natural world (NRC, 2012). Their instructional models further facilitated students' connections to cross-cutting concepts such as patterns, cause and effect, structure and function, and systems (NGSS Lead States, 2013).

The naturalists offer each partner school up to nine expeditions per school year, and most schools average around five expeditions per year. At least one of the expeditions takes place in the schoolyard. More frequently, students and teachers travel to nearby natural areas with the naturalists for a day of science learning and engagement in the outdoors. During the expeditions, the classroom teachers serve in various roles. Sometimes teachers choose to observe the lessons or serve as chaperones to monitor student behaviors, while other times, teachers actively participate in activities along with their students.

Participants

Participants in this study were recruited from the OSE program's participant database. The database consisted of 65 teachers whose schools have engaged in expeditions with the OSE program. Each teacher was individually contacted via email to request their participation in an interview. Thirty teachers agreed to participate in a phone interview (46% response rate); of these 27 respondents identified as female, and three identified as male. The teacher participants had participated in this OSE program for between one to 12 years, with an average of 3.75 years. A separate email with an active link to the survey was sent to the same 65 teachers from the original database to request participation. Of the 65 teachers, 33 teachers anonymously responded to the online survey (50.8% response rate). In an effort to help teachers feel comfortable in honestly sharing their experiences, we prioritized protecting teacher identity; thus, we were not able to align survey data with interview responses. All teacher names are pseudonyms.

Data

Qualitative

We conducted semi-structured interviews with each teacher participant (See Appendix A). Each interview lasted between 25 to 45 minutes, and interviews were audiotaped and transcribed. The interviews were conducted by phone, with the majority administered by the second author, who at that time was a graduate research assistant. Before the start of each interview, the study's aims and the interviewee's rights were reviewed. We obtained informed consent from all participants. The interview process followed a specific sequence of open-ended questions guided by Vygotsky's (1978) sociocultural learning theory that asked the teachers their perceptions about how participating in the OSE programs with the naturalists impacted their learning about science concepts, instructional strategies, comfort in teaching science, and teaching in the outdoors (see Appendix A). The base questions were supplemented with follow-up questions that allowed the teachers to elaborate and clarify their responses.

Quantitative

The online survey (see Appendix B) consisted of response items using a five-point Likert-type scale. The survey questions aligned with research questions to identify patterns and support the qualitative findings. The teachers were asked their perceptions about to what extent the OSE supported their own learning, their comfort teaching science, their comfort teaching science in the outdoors, and their perceptions of their students' science learning and experience in the outdoor expeditions. Additionally, teachers were asked to report on their frequency and levels of participation in OSE expeditions. The findings provided a more holistic view and context for the qualitative descriptions from the semi-structured interviews.

Data Analysis

The interviews were recorded and transcribed, and the transcriptions were coded using a priori coding themes (Creswell & Plano Clark, 2011) derived from the research questions and positioned within Vygotsky's framework of sociocultural learning as teachers participated the OSE along with their students. Additional themes (See Table 1) emerged from the teachers' responses and contributed to the developing theme of teachers as learners.

As recommended by Hsieh & Shannon (2005), interview data were used to construct valid inferences to understand the teachers' perceptions of the OSE experiences. The data analysis occurred in several phases using an inductive process. Three researchers continuously discussed the coding procedures throughout all data analysis phases to assure consistency and provide trustworthiness and reliability of the data to ensure the integrity of the research process (Nadin & Cassell, 2006). The researchers initially read and coded the same four randomly selected transcribed interviews, then compared for coding decisions. The researchers discussed code interpretations and, when necessary, the codes were reviewed and revised until interrater reliability (Krippendorf, 2011) reached 90% agreement. In the next phase of data analysis, the remaining 26 transcribed interviews were randomly assigned to be coded independently by the three researchers. Survey data were analyzed, and descriptive statistics regarding teachers' perceptions of their learning and student learning were calculated using Stata 16 software. Correlation coefficients were calculated using Pearson's R to estimate correlations between teachers' perceptions of their own engagement and learning and that of their students.

Table 1

Coding Themes and Sample Quotes

Coding Themes	Description	Sample Quotes
Science Content Knowledge	Teachers' perceptions of their and their students' science content knowledge	Also, my own knowledge of local botany and wildlife and things like that, I have learned.
Instructional Strategies - Science	Teachers' views about science instruction	I just love the fact that it's a way for our kids to connect hands-on to our environment and to ideas about conservation and things that are going to be lasting impacts...they got to create animals out of natural materials
Outdoor Teaching and Learning	The outdoors as a setting for teaching and learning	It's one thing to read about ecosystems, [but another] to be able to actually be out and see the actual things, roll over dead logs and then find out what's living underneath them
Teachers' Views of Students	Teachers' descriptions of their students in outdoor learning.	They were focused, they were excited. They listened very well. They were willing to do more and participate more outside than versus inside of the everyday paper-pencil.
Teacher as Learner	Teachers' descriptions of their own learning	Science itself I can teach, but being able to connect it to the real world in front of them- that was a little bit harder for me - so I think that improved.

Findings

The teachers' voices from interviews and survey responses are organized around the coding themes that emerged through the inductive analysis process. These findings include teacher interview data and patterns from aggregated survey data.

Teachers' Science Content

The science content of the OSE activities addressed the state's science standards for grade 5. In the interviews, the teachers described OSE lessons on forces and motion, aquatic ecosystems, terrestrial ecosystems, weather, energy and matter, and heat transfer. Despite the fifth-grade teachers' familiarity with the science content topics, 23 of the teachers described their learning or "relearning" science content by the OSE, both within and beyond the standards topics. Darcy felt that her participation in the outdoor program activities built her "own knowledge of local botany and wildlife," and Valerie explained her own learning from the naturalists, "I feel like I learned a lot just from hearing from another science expert." Some teachers referenced elementary teachers' preparation as generalists and their limited backgrounds in science. Alexandra described teachers learning science content from the naturalists' instruction strategies:

> I definitely think it's [OSE] empowering to a lot of teachers. I have a lot of teachers that are on my team that don't know a lot about science. It's been really helpful to them. They've [said], 'I had no idea that heat transfers in that way. I had no idea about all these [concepts], about the plants.'

Abby described connecting the outdoor experiences to her indoor science lessons and her learning along with her students:

> As a teacher and getting outside - and actually participating in the field and the experiments - to me, it gives the students and myself so much more actual hands-on than we can do in the classroom…Then we bring it back into the classroom and even today during science, we were talking about one of our earlier expeditions because we always bring it back. We were talking about fungi and talking about how we saw it on our trip. It keeps bringing everything around and it's interesting and we all learn from it.

In addition to teachers' descriptions of their own learning, many described their perceptions of their students' learning science content.

Teachers' Perceptions of Students' Science Content

Jaden and Darcy used their students' performance on standardized science tests as measures of students' science content knowledge. Both attributed students' strong test scores to their participation in the OSE activities. Jaden explained, "The year that we did [OSE], I had the top 20% highest scores in the state…My kids were all identified in some way with some label [learning or behavior disability]…everyone did well in science." Darcy said, "We noticed that our scores did go up and then we also saw [when] another school in the county dropped it [OSE program], their scores went down and ours were still high."

Data from the teacher surveys (see Table 2) also indicated that while the teachers felt that participation in the OSE program supported students' science learning, it also supported teachers' learning.

There was a significant positive correlation between teachers' perceptions of students' learning and their own learning ($r(29) = .47, p <.01$). Approximately half of the teachers (45.5%) stated they learned "a great deal," 30% reported that they felt their learning was "only moderately" impacted by their participation in the OSE, while 6% (two teachers) reported that they felt that the experience did not support their learning at all. During the interviews, teachers who reported learning little from the OSE experience explained their already strong sense of self-efficacy in their science content knowledge and knowledge of science instruction. The teachers' reflections of their students' learning followed clear patterns, with twenty-five (75.8%) of the teachers reporting that the OSE supported their students' learning "a great deal." Only one teacher (3%) reported that the OSE "only moderately" supported their students' learning, and none of the teachers felt that their students failed to learn from the expeditions. In addition to the teachers learning science content, many teachers described learning from watching the naturalists' teaching methods. These methods included OSE naturalists asking students to design and build models, connecting interactive activities to classroom experiences, and extending instruction beyond the four walls of the classroom, as we examine in the next section.

Table 2

Descriptive Data from Surveys

Survey Item	Mean	SD	Variance
How often do you participate in the expeditions with your students?	4.6	0.9	0.8
How comfortable do you feel:			
teaching science to your students?	4.6	0.7	0.6
teaching science through outdoor experiences?	4.2	0.7	0.5
To what degree has OSE supported:			
comfort level in teaching science?	3.7	1.2	1.5
your students' learning?	4.7	0.5	0.3
your learning?	4.0	1.1	1.2

Note: Data were collected in a 5-point Likert-type scale of 1 = "Not at all" to 5 = "A great deal"

Instruction Strategies

The teacher participants described various instruction strategies they learned from observing the naturalists working with their students that aligned with their standards-based science instruction. Valerie said, "I feel like I learned a lot in the way that they presented information and conducted activities. I feel like it gave me maybe just a better love of science and how to engage the students in science." Relatedly, Ellen pointed out that situating instruction in the outdoors helps students connect science to life outside the classroom. "Giving students the opportunity to explore outside, they can see [science] in real-life. Real-life applications." Robert described similar views of science instruction in outdoor settings:

> It's one thing to read about ecosystems, but to be able to actually be out and see the actual things, roll over dead logs and then find out what's living underneath them, and because when we're studying about decomposers, we can actually see them live and experience them. I think it just adds a lot to the learning process and they appreciate that opportunity.

He went on to describe his learning instruction strategies from the naturalists, "Just seeing how they structure it and how they tie in those standards with the activities. It lets you see a model for how it can be."

In addition to student engagement in the outdoor lessons, the teachers described the OSE expeditions as confidence-building experiences for their teaching and learning in the outdoors. Lindsay explained, "Prior to this experience, I had never taken children outside for learning except an occasional field trip...I feel much more confident now taking them on expeditions outside." Bob appreciated that the OSE naturalists capitalized on "teachable moments" when he said:

> When things get pushed by their [students'] wonderings and understandings beyond what is just that objective…they're [naturalists] able to roll with what's interesting to them and pushing that envelope rather than just stopping straight at the content.

Twenty-three of the 30 teachers who were interviewed described how being outside with the naturalists alongside their students influenced their instruction and their learning strategies for authentic science activities and "hands-on" learning. Eight of the teachers valued science instruction that connected science with students' lives and used terms like "real world" and "real life" connections. Fifteen of the teachers described the interdisciplinary instruction potential of outdoor experiences. Ashley explained that the outdoor experiences shed new light for her about "how everything is interconnected."

Learning in the Outdoors

Many of the teachers described learning the benefits of situating instruction in the outdoors. Jamie elaborated:

> I think they see that science is not just in the classroom...I think that's one of the biggest benefits that the students get because they're able to make the connection that you don't have to be in a classroom to learn science or to gain new knowledge about science.

Cassidy's perception of science instruction in the outdoors addressed both cognitive and affective contributions of situating science in the outdoors, saying, "I am just reassured that this [the outdoors] is a good environment to learn. There are learning opportunities that we wouldn't have in the classroom. It definitely sparks their interest and their curiosity."

During the interviews, the teachers often spoke about how the OSE experiences expanded their own cognitive and affective connections to science teaching. Jaden explained that the outdoor experiences positively impacted her feelings about teaching science, "This program has tightened my comfort in teaching science. I love it even more." Darcy also described her learning the critical instruction practice of connecting to students' lives. "Science itself I can teach, but being able to connect it to the real world in front of them, that was a little bit harder for me, so I think that improved." Several of the teachers described their experiences with the naturalists taking on the role of the MKOs in expeditions. Darcy explained:

> When they [OSE program naturalists] teach in the woods, they are not only teaching the curriculum, but they're also bringing in local issues like invasive species that I didn't necessarily know, so that way I can kind of connect it with something around them a little bit better. Then actually applying--like when we did simple machines, and they're [the students] applying it for [designing] the bear bag [with pulleys] to hang food and stuff.

Sharon also described her learning science instruction strategies that connect to students' lives, to nature, and to classroom learning:

> They're [OSE program naturalists] actually tying it back to what the students are learning in the classroom. I just love the fact that it's a way for our kids to connect hands-on to our environment and to ideas about conservation and things that are going to be lasting impacts...they got to create animals out of natural materials that were found in the woods, and then they got to talk about the adaptations that they had from living in that area.

In describing how participation in the OSE program impacted her instruction strategies, Jamie elaborated on her role as *learner*:

> I've gained so much information from them and the activities that they provide as well as just their knowledge base. I guess I'm almost a student to some degree because I'm taking notes and trying to capture their activities that they do so we can recreate them in the classroom.

Relatedly, Lindsay gave a specific example of how participation in the OSE with her students impacted her learning about science instruction.

> It has definitely helped my teaching. I had mentioned to the OSE instructor that I was struggling, and my students were struggling, with understanding motion graphs. When she [naturalist] did her activity, everywhere that we walked that day we paused occasionally, and she had a student collect the data and created a graph and graphed our own motion.

Jamie and Lindsay both described that they learned science instruction practices from the MKO naturalists that tap into students' creative thinking and active learning. While many variables influence teachers' perceptions of their science teaching and outdoor instruction, the teachers' learning expanded beyond their perceived knowledge of science and instruction strategies. In the following section, we document teachers' descriptions of learning more about their students.

Teachers' Perceptions of Their Students' Behaviors and Engagement in Science

In addition to the OSE supporting teachers learning of effective science instruction through connecting to students' lives in the outdoors, the teachers also reported that the experiences altered their perceptions of students' cognitive and affective connections to their learning. Jaden explained how her perception of one student changed following the OSE, "Last year we had a kid who was in an alternate behavior classroom. The deal was that he could go on the expedition on a trial basis. This kid was probably one of the best-behaved kids out there!"

Margo commented on the potential for outdoor experiences to engage all students in science. "I do honestly think that it could encourage students to love science more, especially some of those girls who have started to get away from liking science and math." The teachers often described seeing their students in new and different ways. As Reese explained, "We get to take a step back and view our students in a completely different environment, with a different teacher and we can just learn more about them that way." In addition to seeing their students' interactions with different teachers, Cassidy described the transition of student behaviors in the outdoor settings saying, "Some of the students who don't speak up in class, they were speaking up and were on it. They speak up in the outdoors, but not in the classroom." Lacey had similar observations about her students "They were focused, they were excited. They listened very well. They were willing to do more and participate more outside than inside with the everyday paper and pencil."

Other teachers spoke directly about students who have been identified as learning or behaviorally challenged. Jamie noticed that in the outdoors, her students "who are not able to sit in a classroom setting without a behavior distraction...those students are engaged, they're willing to write...they're willing to attempt more in the participation than they are in the classroom." Similarly, Robert mentioned the positive impact on "students with learning disabilities, again, gets them out of a textbook into the world where they're able to see the concepts instead of struggling through reading about them." Margo noted, "I think that makes it easier for kids that have short attention spans to stay with what's happening." The teachers highlighted their students' affective engagement with science content when learning collaboratively in the outdoors.

Limitations of the Study

One limitation of the study is that the teachers self-selected by volunteering to participate in the interviews. Selection bias may express itself in many ways, but in this study, it seems most likely to occur in a form where the teachers most excited about learning in nature responded. Thus, our findings may provide better inference for contexts with highly engaged teachers. Further, self-reports of the teachers' experiences offer a snapshot of the teachers' perceptions of their science and outdoor instruction. In order to ensure teachers' privacy, we used an anonymous online survey, so we could not determine which of the teachers who responded to the survey were also interviewed, thus inhibiting triangulation of data. While the data presented here are specific to the targeted OSE program participants, it is essential to consider the potential value of teachers' participation in OSEs to inform teacher development and teacher education.

Discussion

While the OSE program in this study was designed to enhance fifth-grade student science learning, our research questions focused on teachers' perceptions of the OSE program's impact on their and their students' science content knowledge, science instructional practices, outdoor instruction, and learning about their students.

Science Content Knowledge

Our first research question asked, "What are the relationships between elementary school teachers' participation in OSE experiences led by naturalists and their perceptions of their own and their students' science content knowledge?"

The teachers' survey responses and interviews were consistent in their perceptions that both their and their students' science content knowledge grew from participation in the OSE expeditions. Alexandra discussed how elementary teachers must learn to teach all subjects and thus can lack specific science content knowledge as documented in national surveys (Plumley, 2019; Smith, 2020). Alexandra described how her content knowledge about heat transfer grew as she learned along with her students. The teachers' survey responses also revealed their critical recognition of how the OSE program activities were directly linked to their grade level science content standards.

Our focus on the teachers' learning connects with research on professional development for teachers. The OSE program featured some core characteristics of effective professional development, such as active learning, content focus, and collective participation (Darling-Hammond, 2005; Desimone, 2009; Desimone et al., 2002; Fischer et al., 2018; Garet et al., 2001; Fishman et al., 2003), but as Kennedy (2016) pointed out in her review of 28 studies on professional development, there are other factors beyond the design of professional development that influence instructional practices. We offer suggestions for addressing other factors in our implications.

The collective participation of naturalists, teachers, and students aligns with Vygotsky's (1978) sociocultural framework to examine the social connections as teachers in this study accompanied their students and participated together in the outdoor expeditions and further aligns with more recent learning sciences research (NASEM, 2018). While the OSE program focused on student learning, both the teachers and students shared the role of learners with OSE naturalists who served as MKOs as they modeled science instruction strategies to facilitate students' science learning in the outdoors. Interestingly, the data provide many examples of ways that the OSE program significantly also supported the teachers learning science content. Lindsay explained that she learned about forces and motion content from the OSE naturalist, and she further gained knowledge about instructional practices from the naturalists' instructional modeling. The OSE naturalists in the present study partnered closely with teachers and also served as MKOs as they presented outdoor learning activities that aligned with the prescribed science standards and supported teachers' indoor science instruction. This critical and participatory design enhanced the experience for the teachers, and the OSE program helped maximize the impact on teachers' perceptions of both student *and* teacher learning (Ayotte-Beaudet et al., 2017).

Science Instruction and Outdoor Science Instruction

The second research question asked, "What are the relationships between elementary school teachers' participation in OSE experiences led by naturalists and their perceptions of science instruction and outdoor science instruction?"

In addition to teachers' perceptions of their increased science content knowledge, the OSE program also supported teachers' science instructional practices, including outdoor instruction. The teachers described how their knowledge of content (CK) and pedagogies (PK) expanded and contributed to their PCK as they learned new and different instruction strategies for communicating science with their students (Park & Oliver, 2008; Shulman, 1986). In addition, this strategic alignment of science content knowledge, science standards, and instructional practices addressed many of the teachers' perceived barriers to outdoor instruction.

Through their participatory collaborations with the naturalists and observations of naturalists' instruction, the teachers in our study described learning how the outdoors can be an authentic setting to teach science standards in ways that apply to students' lives beyond the classroom. The teachers in this study recognized that situating science content and instructional practices in the outdoors connects science with students' lives. For example, one teacher described an activity that challenged students to consider themselves as campers who need to protect their food from potential bears. This lesson moved learning about forces and motion from the teachers' traditional classroom instruction as they asked students to apply these concepts to design a pulley system in tree branches to elevate their food off the ground. Another example of an activity that teachers and students experienced this collaborative learning context reinforced the concept of heat transfer. In this activity, students were asked to apply concepts of heat transfer in their design of a camping stove to make tea out of pine needles.

The teachers' descriptions of their students' engagement in content-focused science revealed the impact of teachers' learning new science instruction practices (Desimone et al., 2002; Glackin, 2016) and the resulting student motivation and interest in science (Dillon et al., 2006; Kahn & Kellert, 2012; Lindemann-Matthies, 2005; Zoldosova & Prokop, 2006). Our findings support research on the effectiveness of outdoor learning as an instructional strategy. Rickinson and colleagues (2004) recommended that programs "use a range of carefully structured (outdoor) learning activities and assessments linked to the school curriculum" (p. 7). They also identified the potential for school grounds and outdoor community projects to impact students' motivation to learn. Importantly, they further describe that, through these outdoor experiences, students may "develop more positive relationships with each other, with their teachers, and with the wider community" (p. 6).

Teachers' Perceptions of Students

The third research question asked, "What are teachers' perceptions of their students' engagement in science and behaviors following their participation in an OSE program?"

Most notably, interview and survey data reveal multiple ways that teachers in this study learned more about their students. Although teachers reported how participating in the OSE with the naturalists helped them learn science content, science instruction strategies, and about using the outdoors as a setting for teaching and learning, the teachers' descriptions of their increased understanding of their students as they learned together in the outdoors may be the most relevant impact of their OSE experience. As Vygotsky's theory suggests, social interaction is a fundamental component of cognitive and social development. Human development and learning are guided by social interactions as participants co-construct knowledge and engage in their environment within a social-ecological system (McLeod, 2014). These experiences provided the teachers an important lens to recognize how outdoor learning can target students' potential and allow students the opportunity to assume the role of the MKO with their peers and their teacher, especially students whose traditional classroom experiences seemed stagnant. Providing teachers with the opportunity to expand learning beyond the classroom exploring outdoor science with their students can motivate both students and their teachers.

The relationship building that can occur during OSE experiences can influence teachers' assumptions about students in ways that have potential to ultimately improve teachers' instruction, improve student outcomes, and challenge teachers' expectation biases (Cooper, 2000; Florea, 2007). Teachers' relationships with their students impact the learning experiences for both. Woodcock and Vaille (2010) explain, "The way in which teachers perceive the students' behaviour can influence their future expectations and responses to students" (p. 178).

Like Jordan, many of the teachers in this study were surprised by some of their students' positive engagement in the science content and learning in the outdoors, especially those students who had previously been identified as having behavioral or learning challenges. Teachers like Reese described their students' positive engagement in the outdoor setting that they had not seen in the indoor science classroom (Dillon et al., 2006; Lieberman & Hoody, 1998; Malone, 2008); these experiences indicate the potential to influence teachers′ expectations and alter bias of students' learning potential (Woodcock & Vialle, 2011). Another teacher, Margo, confronted her expectation biases about girls and science when she said, "I do honestly think that it [outdoor learning] could encourage students to love science more, especially some of those girls who have started to get away from liking science and math."

In this study, as teachers learned science content along with their students, they described an enhanced appreciation of their students' abilities. Two examples where teachers described learning from their students were when OSEs asked students to design a pulley to elevate food to protect it from bears and when their students developed strategies to make their own tea. In addition to seeing the naturalists as MKOs, the teachers in this study were also able to see their students assume MKO roles with their peers and, at times, with their teachers as they shared their learning in the sociocultural context of reciprocal learning (Vygotsky, 1978). This dynamic illustrates the important MKO role as a dynamic interactive/fluid process as the naturalists, students, and teachers take turns as MKO. This fluidity seldomly occurs in traditional interactions with students and teachers in a science classroom and can promote student-centered instruction (Feiman-Nemser, 2008; Smith, 2020).

Shared learning experiences can expand teachers' expectations and attitudes about their students (Rubie-Davies, 2010; Weinstein, 2002). The significance of teachers' increased awareness and appreciation of their students' potential for learning and engagement highlights the benefits of outdoor science experiences to encompass the kinesthetic, cognitive, and affective domains of learning (Honebein & Honebein, 2015). These experiences further support teacher reflection and integration as they build their PCK (Park & Suh, 2019). Though the benefits of OSE for teachers are understudied and arguably underappreciated in practitioner settings, this research suggests how OSE programs can support both teacher development and student outcomes.

Implications

Although this study's findings are specific to this OSE program, these data provide important insights for both teacher and student learning. For decades, elementary science instruction has been marginalized by accountability measures focused on other domains, notably reading and mathematics (Smith, 2020), yet the need for a scientifically literate society is more relevant than ever. Teacher educators and school administrators are responsible for supporting teachers' science instruction and students' science learning. One of the notable implications is that OSE providers can join teacher educators and schools to expand OSE and other science professional development supports for teachers. The school community primarily serves students, so it must maintain a keen interest in and devotion to resources to supporting teachers' professional development in science.

The experience of teachers' participating in OSE presented here have the potential to inform other educators about strategies to access authentic learning opportunities that connect to the natural world and students' lives. Such experiences reinforce sociocultural collaborations with other learners, such as local OSE programs, informal educators in museums or nature centers (Tran & King, 2011), and science specialist teacher leaders (Herbert et al., 2017). Providing teachers with student-focused teaching models in authentic settings has potential to expand teachers' science content knowledge, enhance their instruction strategies that connect to the natural world and to students' lives, build their confidence in science and science teaching, and, importantly, also help teachers learn more about their students. As teachers, students, and collaborators connect with nature and learn together in the outdoors, they begin to build a rich community of learners.

The authors received no financial support for the research, authorship, and/or publication of this manuscript. The authors would like to thank Muddy Sneakers, Inc. for their school collaborations introducing students to the wonders of the natural world through experiential outdoor education.

Sarah J. Carrier (sjcarrie@ncsu.edu) is Professor of Elementary Science Education at North Carolina State University. Her research has focused on teachers' developing identities as science teachers, environmental education, and outdoor learning. As a former elementary school teacher who spent countless hours teaching in the outdoors and learning along with her students, she guides current and future teachers to embrace the outdoors as a setting for teaching and learning.

Aimee B. Fraulo (abfraulo@gmail.com) is a research scientist and outdoor educator. Her research focuses on the neuroscience and socioemotional benefits of outdoor learning strategies in classroom teaching and child development. She has taught in both traditional and Montessori schools and currently works as a researcher in the agricultural and ecological sciences.

Kathryn T. Stevenson (ktate@ncsu.edu) is an associate professor in the Department of Parks, Recreation & Tourism Management at NC State University. Her research and teaching centers on children and nature and environmental education. In particular, she is interested in understanding how time in natural areas provides benefits to young learners as well as how youth-led conversations can lead to community-scale change.

M. Nils Peterson (mnpeters@ncsu.edu) is a Professor of Fisheries, Wildlife, and Conservation Biology at North Carolina State University. His research focuses on unravelling the drivers of environmental behavior in contexts of endangered species conservation, wildlife crime, environmental education, environmental conflict, and environmental policy making.

Laura M. Romeo (lmmmayer@ncsu.edu) is an experienced teacher educator, instructional designer, curriculum specialist, and former elementary teacher. Her passion for working within the education field began as a young girl in her mother's Kindergarten classroom and has grown with each new experience, both professionally and personally. Laura is devoted to providing teachers with more effective instructional practice and contributing to their love and further pursuit of education in her

current work efforts. Her research interests include Universal Design for Learning, Social and Emotional Learning, preservice teachers, instructional design and development, and equitable teaching practices.

References

Abtahi, Y. (2017). The more knowledgeable other: A necessity in the zone of proximal development? *For the Learning of Mathematics*, *37*(1), 35-39. https://doi.org/10.1080/14794802.2017.1390691

Ampuero, D., Miranda, C. E., Delgado, L. E., Goyen, S., & Weaver, S. (2015). Empathy and critical thinking: Primary students solving local environmental problems through outdoor learning. *Journal of Adventure Education & Outdoor Learning*, *15*(1), 64-78. https://doi.org/10.1080/14729679.2013.848817

Anderson, R. D. (2002). Reforming science teaching: What research says about inquiry. *Journal of Science Teacher Education, 13*(1*)*, 1–12. https://doi.org/10.1023/A:1015171124982

Anderson, R. D., & Helms, J. V. (2001). The ideal of standards and the reality of schools: Needed research. *Journal of Research in Science Teaching, 38* (1), 3–16. https://doi.org/10.1002/1098-2736(200101)38:1<3::AID-TEA2>3.0.CO;2-V

Archer, L., DeWitt, J., Osborne, J., Dillon, J., Willis, B., & Wong, B. (2010). "Doing" science versus "being" a scientist: Examining 10/11-year-old schoolchildren's constructions of science through the lens of identity. *Science Education*, *94*(4), 617-639. https://doi.org/10.1002/sce.20399

Ayotte-Beaudet, J. P., Potvin, P., Lapierre, H. G., & Glackin, M. (2017). Teaching and learning science outdoors in schools' immediate surroundings at K-12 levels: A meta-synthesis. *Eurasia Journal of Mathematics, Science and Technology Education*, *13*(8), 5343-5363. https://doi.org/10.12973/eurasia.2017.00833a

Banilower, E., Cohen, K., Pasley, J. & Weiss, I. (2010). *Effective science instruction: What does research tell us?* Second edition. RMC Research Corporation, Center on Instruction.

Baumert, J., & Kunter, M. (2013). The COACTIV model of teachers' professional competence. In *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 25-48). Springer, Boston, MA.

Bloom, M. A., Holden, M., Sawey, A. T., & Weinburgh, M. H. (2010). Promoting the use of outdoor learning spaces by K-12 in-service science teachers through an outdoor professional development experience. In *The inclusion of environmental education in science teacher education* (pp. 97-110). Springer, Dordrecht.

Carrier, S.J., Thomson, M.M., Tugurian, L.P., & Stevenson, K.T. (2014). Elementary science education in classrooms and outdoors: Stakeholder views, gender, ethnicity, and testing. *International Journal of Science Education*, *36*(13), 2195-2220.

Carrier, S.J., Tugurian, L.P. & Thomson, M.M. (2013). Elementary science indoors and out: Teachers, time, and testing. *Research in Science Education*, *43*(5), 2059-2083.

Carrier, S. (2009). The effects of outdoor science lessons with elementary school students on preservice teachers' self-efficacy. *Journal of Elementary Science Education*, *21*(2), 35-48.

Catapano, S. (2005). Teacher professional development through children's project work. *Early Childhood Education Journal*, *32*(4), 261-267. https://doi.org/10.1007/s10643-004-1428-2

Cheng, J.C.H. & Monroe, M. C. (2012). Connection to nature: Children's affective attitude toward nature. *Environment and Behavior*, *44*(1), 31–49. http://doi.org/10.1177/0013916510385082

Chawla, L. (1998). Significant life experiences revisited: A review of research on sources of environmental sensitivity. *The Journal of Environmental Education*, *29*(3), 11-21. https://doi.org/10.1080/00958969809599114

Christidou, V. (2011). Interest, attitudes, and images related to science: Combining students' voices with the voices of school science, teachers, and popular science. *International Journal of Environmental and Science Education*, *6*(2), 141-159.

Cooper, H. M. (2000). Pygmalion grows up. In P. K. Smith & A. D. Pellegrini (Eds.), *Psychology of education: Major themes* (pp. 338–364). RoutledgeFalmer https://doi.org/10.3102/00346543049003389

Cooper, A. (2015). Nature and the outdoor learning environment: The forgotten resource in early childhood education. *International Journal of Early Childhood Environmental Education*, *3*(1), 85-97. https://files.eric.ed.gov/fulltext/EJ1108430.pdf

Creswell, J. W. 2014. *Research Design: Qualitative, Quantitative, and Mixed Methods*. Approaches. 4th ed. SAGE Publications.

Creswell, J. W., & Plano Clark, V. L. (2011). Choosing a mixed methods design. *Designing and conducting mixed methods research*, *2*, 53-106.

Darling-Hammond, L. (2005). Teaching as a profession: Lessons in teacher preparation and professional development. *Phi Delta Kappan*, *87*(3), 237-240. https://doi.org/10.1177/003172170508700318

Desimone, L. M. (2009). Improving impact studies of teachers' professional development: Toward better conceptualizations and measures. *Educational Researcher*, *38*(3), 181-199.

Desimone, L. M., Porter, A. C., Garet, M. S., Yoon, K. S., & Birman, B. F. (2002). Effects of professional development on teachers' instruction: Results from a three-year longitudinal study. *Educational Evaluation and Policy Analysis*, *24*(2), 81-112. https://doi.org/10.3102/01623737024002081

DeWitt, J., & Osborne, J. (2007). Supporting teachers on science-focused school trips: Towards an integrated framework of theory and practice. *International Journal of Science Education*, *29*(6), 685-710. https://doi.org/10.1080/09500690600802254

Dillon, J., Rickinson, M., Teamey, K., Morris, M., Choi, M. Y., Sanders, D., & Benefield, P. (2006). The value of outdoor learning: Evidence from research in the UK and elsewhere. *School Science Review*, *7*(320), 107–112.

Eaton, D. (2000). Cognitive and affective learning in outdoor education. *Dissertation Abstracts International - Section A: Humanities and Social Sciences*, *60*(10-A), 3595. https://tspace.library.utoronto.ca/handle/1807/12600

Eick, C. J. (2012). Use of the outdoor classroom and nature-study to support science and literacy learning: A narrative case study of a third-grade classroom. *Journal of Science Teacher Education*, *23*(7), 789-803. https://doi.org/10.1007/s10972-011-9236-1

Ewert, A., Place, G., & Sibthorp, J. (2005). Early-life outdoor experiences and an individual's environmental attitudes. *Leisure Sciences*, *27*(3), 225-239. https://doi.org/10.1080/01490400590930853

Fägerstam, E., & Blom, J. (2013). Learning biology and mathematics outdoors: Effects and attitudes in a Swedish high school context. *Journal of Adventure Education & Outdoor Learning*, *13*(1), 56-75. https://doi.org/10.1080/14729679.2011.647432

Feille, K. (2013). Getting outside: Three teachers' stories of using the schoolyard as an integrated tool for elementary teaching. *Electronic Journal of Science Education*, *17*(3), 1-17. http://files.eric.ed.gov/fulltext/EJ1188358.pdf

Feiman-Nemser, S. (2008). Teacher learning: How do teachers learn to teach?. In *Handbook of research on teacher education* (pp. 696-705). Routledge.

Feinstein, N. W., Allen, S., & Jenkins, E. (2013). Outside the pipeline: Reimagining science education for nonscientists. *Science*, *340*, 314-317. https://doi.org/10.1126/science.1230855

Fischer, C., Fishman, B., Dede, C., Eisenkraft, E., Foster, B., Frumin, K., Lawrenz, F., Levy, A., & McCoy, A. (2018). Investigating relationships between school context, teacher professional development, teaching practices, and student achievement in response to a nationwide science reform. *Teaching and Teacher Education, 72(*1), 107-121. https://doi.org/10.1016/j.tate.2018.02.011

Fishman, B. J., Marx, R. W., Best, S., & Tal, R. T. (2003). Linking teacher and student learning to improve professional development in systemic reform. *Teaching and teacher education*, *19*(6), 643-658. https://doi.org/10.1016/s0742-051x(03)00059-3

Garet, M. S., Porter, A. C., Desimone, L., Birman, B. F., & Yoon, K. S. (2001). What makes professional development effective? Results from a national sample of teachers. *American Educational Research Journal*, *38*(4), 915-945. https://doi.org/10.3102/00028312038004915

Glackin, M. (2016). 'Risky fun 'or 'Authentic science'? How teachers' beliefs influence their practice during a professional development programme on outdoor learning. *International Journal of Science Education*, *38*(3), 409-433. https://doi.org/10.1080/09500693.2016.1145368

Glesne, C. (2016). *Becoming qualitative researchers: An introduction.* (5th edition). Pearson.

Haider, M., & Yasmin, A. (2015). Significance of scaffolding and peer tutoring in the light of Vygotsky's theory of Zone of Proximal Development. *International Journal of Languages, Literature and Linguistics*, *1*(3), 2015. https://doi.org/10.18178/ijlll.2015.1.3.33

Harris, F. (2018). Outdoor learning spaces: The case of forest school. *Area*, *50*(2), 222–231. https://doi.org/10.1111/area.12360

Harvey, D. J., Montgomery, L. N., Harvey, H., Hall, F., Gange, A. C., & Watling, D. (2020). Psychological benefits of a biodiversity-focused outdoor learning program for primary school children. *Journal of Environmental Psychology*, *67*, 101381.

Herbert, S., Xu, L., & Kelly, L. (2017). The changing roles of science specialists during a capacity building program for primary school science. *Australian Journal of Teacher Education*, *42*(3), 1-21. https://doi.org/10.14221/ajte.2017v42n3.1

Honebein, P. C., & Honebein, C. H. (2015). Effectiveness, efficiency, and appeal: Pick any two? The influence of learning domains and learning outcomes on designer judgments of useful instructional methods. *Educational Technology Research and Development*, *63*(6), 937-955. https://doi.org/10.1007/s11423-015-9396-3

Hoody, L.L. (1996). *The effects of environment-based education on student achievement*. State Education & Environment Roundtable [SEER]. http://www.seer.org/pages/research/educeff.pdf

Hsieh, H. F., & Shannon, S. E. (2005). Three approaches to qualitative content analysis. *Qualitative Health Research*, *15*(9), 1277-1288. https://doi.org/10.1177/1049732305276687

Hungerford, H. R., & Volk, T. L. (1990). Changing learner behavior through environmental education. *The Journal of Environmental Education*, *21*(3), 8-21. https://doi.org/10.1080/00958964.1990.10753743

Junker, C. R., & Jacquemin, S. J. (2016). Bridging the gap: Surveying interdisciplinarity in the environmental literature classroom. *ISLE: Interdisciplinary Studies in Literature and Environment*, *23*(2), 395-411. https://doi.org/10.1093/isle/isw038

Kahn Jr, P. H., & Kellert, S. R. (Eds.). (2002). *Children and nature: Psychological, sociocultural, and evolutionary investigations*. MIT press.

Kennedy, M. M. (1998). Education reform and subject matter knowledge. *Journal of Research in Science Teaching, 35*(3), 249–263. https://doi.org/10.1002/(SICI)1098-2736(199803)35:3<249::AID-TEA2>3.0.CO;2-R

Kennedy, M. M. (2016). How does professional development improve teaching? *Review of Educational Research, 86*(4), 945–980. https://doi.org/10.3102/0034654315626800

Krippendorff, K. (2011). Agreement and information in the reliability of coding. *Communication Methods and Measures, 5*, 93-112.

Lieberman, G. A., & Hoody, L. L. (1998). *Closing the achievement gap: Using the environment as an integrating context for learning*. State Education and Environment Roundtable. https://files.eric.ed.gov/fulltext/ED428943.pdf

Lindemann-Matthies, P. (2005). 'Loveable' mammals and 'lifeless' plants: How children's interest in common local organisms can be enhanced through observation of nature. *International Journal of Science Education*, *27*(6), 655-677. https://doi.org/10.1080/09500690500038116

Lock, R. (2010). Biology fieldwork in schools and colleges in the UK: An analysis of empirical research from 1963 to 2009. *Journal of Biological Education*, *44*(2), 58-64.

Loughran, J., Milroy, P., Berry, A., Gunstone, R., & Mulhall, P. (2001). Documenting science teachers' pedagogical content knowledge through PaP-eRs. *Research in Science Education*, *31*(2), 289-307. https://files.eric.ed.gov/fulltext/ED442631.pdf

Malone, K. (2008). *Every experience matters: An evidence based review of the role of learning outside the classroom on the development of the whole young person*. Farming and Countryside Education.

Maltese, A.V., & Tai, R.H. (2010) Eyeballs in the fridge: Sources of early interest in science. *International Journal of Science Education, 32*(5), 669-685. https://doi.org/10.1080/09500690902792385

Mannion, G., Fenwick, A., & Lynch, J. (2013). Place-responsive pedagogy: Learning from teachers' experiences of excursions in nature. *Environmental Education Research*, *19*(6), 792-809. https://doi.org/10.1080/13504622.2012.749980

Martin, S.C. (2003). The influence of outdoor schoolyard experiences on students' environmental knowledge, attitudes, behaviors, and comfort levels. *Journal of Elementary Science Education*, *15*(2), p. 51-56.

McKown, C., & Weinstein, R. S. (2002). Modeling the role of child ethnicity and gender in children's differential response to teacher expectations. *Journal of Applied Social Psychology, 32*, 159−184. https://doi.org/10.1111/j.1559-1816.2002.tb01425.x

McLeod, S. (2014). Lev Vygotsky. *Simply psychology*, 1-13. https://pdfs.semanticscholar.org/96ba/471b2677b7ca0b62e778dee40fa1078d5c03.pdf

McMillan, E., & Vasseur, L. (2010). Environmental education: Interdisciplinarity in action. *International Journal of Interdisciplinary Social Sciences*, *5*(3), 435-445 doi:10.18848/1833-1882/CGP/v05i03/51624.

Moalosi, W. T. S. (2013). Assessing Vygotsky's model for students learning. *Educational Research International*, *2*(3), 39-42. http://erint.savap.org.pk/PDF/Vol.2(3)/ERInt.2013(2.3-06).pdf

Nadin, S., & Cassell, C. (2006). The use of a research diary as a tool for reflexive practice: Some reflections from management research. *Qualitative Research in Accounting & Management.*

National Academies of Sciences, Engineering, and Medicine. (2018). *How people learn II: Learners, contexts, and cultures*. The National Academies Press. https://doi.org/10.17226/24783.

National Research Council. (2000). *How people learn.* National Academy Press.

National Research Council. (2004). *How students learn: History, mathematics, and science in the classroom.* National Academy Press.

National Research Council. (2007). *Taking science to school.* National Academies Press.

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* National Academies Press

National Science Teachers' Association (2020). *NSTA/ASTE Standards for Science Teacher Preparation*. https://static.nsta.org/pdfs/2020NSTAStandards.pdf

NGSS Lead States. 2013. *Next Generation Science Standards: For States, By State*s. Washington, DC: The National Academies Press.

Palavan, O., Cicek, V., & Atabay, M. (2016). Perspectives of elementary school teachers on outdoor education. *Universal Journal of Educational Research*, *4*(8), 1885-1893. https://doi.org/10.13189/ujer.2016.040819

Park, S., & Oliver, J. S. (2008). Revisiting the conceptualisation of pedagogical content knowledge (PCK): PCK as a conceptual tool to understand teachers as professionals. *Research in Science Education*, *38*(3), 261-284. https://doi.org/ 10.1007/s11165-007-9049-6

Park, S., & Suh, J. K. (2019). The PCK map approach to capturing the complexity of enacted PCK (ePCK) and pedagogical reasoning in science teaching. In *Repositioning pedagogical content knowledge in teachers' knowledge for teaching science* (pp. 187-199). Springer, Singapore.

Plumley, C. L. (2019). *2018 NSSME+: Status of elementary school science.* Horizon Research, Inc. http://www.horizon-research.com/horizonresearchwp/wp-content/uploads/2020/02/2018-NSSME-Status-of-Elementary-Science.pdf

Posnanski, T. J. (2002). Professional development programs for elementary science teachers: An analysis of teacher self-efficacy beliefs and a professional development model. *Journal of Science Teacher Education*, *13*(3), 189-220. https://doi.org/10.1023/a:1016517100186

Retelsdorf, J., Schwartz, K., & Asbrock, F. (2015). "Michael can't read!" Teachers' gender stereotypes and boys' reading self-concept. *Journal of Educational Psychology*, *107*(1), 186. https://doi.org/10.1037/a0037107

Rickinson, M., Dillon, J., Teamey, K., Morris, M., Choi, M. Y., Sanders, D., & Benefield, P. (2004). *A review of research on outdoor learning.* National Foundation for Educational Research and King's College London.

Rios, J. M., & Brewer, J. (2014). Outdoor education and science achievement. *Applied Environmental Education & Communication*, *13*(4), 234-240. https://doi.org/10.1080/1533015x.2015.975084

Rong, X. L. (1996). Effects of race and gender on teachers' perception of the social behavior of elementary students. *Urban Education*, *31*(3), 261-290. https://doi.org/10.1177/0042085996031003003

Rubie-Davies, C. M. (2010). Teacher expectations and perceptions of student attributes: Is there a relationship?. *British Journal of Educational Psychology*, *80*(1), 121-135. https://doi.org/10.1348/000709909x466334

Rubie-Davies, C. M. (2006). Teacher expectations and student self-perceptions: Exploring relationships. *Psychology in the Schools*, *43*(5), 537-552.

Shulman, L. S. (1986). Those who understand: Knowledge growth in teaching. *Educational Researcher, 15*(2), 4–14. https://doi.org/10.3102/0013189x015002004

Sjöblom, P., & Svens, M. (2018). Learning in the Finnish outdoor classroom: Pupils' views. *Journal of Adventure Education and Outdoor Learning*, 1-14. https://doi.org/10.1080/14729679.2018.1531042

Skinner, E. A., Chi, U., & The Learning-Gardens Educational Assessment Group 1. (2012). Intrinsic motivation and engagement as "active ingredients" in garden-based education: Examining models and measures derived from self-determination theory. *The Journal of Environmental Education*, *43*(1), 16-36.

Smith, P. S. (2020). *2108 NSSME+: Trends in U.S. science education from 2012 to 2018*. Horizon Research, Inc. https://files.eric.ed.gov/fulltext/ED611301.pdf

Sorhagen, N. S. (2013). Early teacher expectations disproportionately affect poor children's high school performance. *Journal of Educational Psychology*, *105*(2), 465. https://doi.org/10.1037/a0031754

Stevenson, K.T., Szczytko, R., Carrier, S.J., & Peterson, M.N. (2021). How outdoor science education can help girls stay engaged with science. *International Journal of Science Education*, *43*(7), 1090-1111.

Szczytko, R., Carrier, S., & Stevenson, K.T. (2018). Impacts of outdoor environmental education on attention, behavior, and learning outcomes for students with emotional, cognitive, and behavioral disabilities. *Frontiers.* DOI: 10.3389/feduc.2018.00046.

Thompson, J. J., & Windschitl, M. (2005). "Failing girls": Understanding connections among identity negotiation, personal relevance, and engagement in science learning from underachieving girls. *Journal of Women and Minorities in Science and Engineering, 11*(1), 1-26. https://doi.org/10.1615/jwomenminorscieneng.v11.i1.10

Timmermans, A. C., Kuyper, H., & van der Werf, G. (2015). Accurate, inaccurate, or biased teacher expectations: Do Dutch teachers differ in their expectations at the end of primary education?. *British Journal of Educational Psychology*, *85*(4), 459-478. https://doi.org/10.1111/bjep.12087

Tran, L. U., & King, H. (2011). Teaching science in informal environments: Pedagogical knowledge for informal educators. In *The professional knowledge base of science teaching* (pp. 279-293). Springer, Dordrecht. https://doi.org/10.1007/978-90-481-3927-9_16

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes.* (M.Cole, V. J. Steiner, S. Scribner, & E. Sauberman, Eds.). Harvard University Press. (Original work published in 1934)

Wang, J. R., Wang, Y. C., Tai, H. J., & Chen, W. J. (2010). Investigating the effectiveness of inquiry-based instruction on students with different prior knowledge and reading abilities. *International Journal of Science and Mathematics Education*, *8*(5), 801-820. https://doi.org/10.1007/s10763-009-9186-7

Wang, V. C. (2018). Theory and practice of adult and higher education. *International Journal of Information and Communication Technology Education*, *14*(2), 84-87.

Wells, N. M., & Lekies, K. S. (2006). Nature and the life course: Pathways from childhood nature experiences to adult environmentalism. *Children Youth and Environments*, *16*(1), 1-24.

Weinstein, R. S. (2002). *Reaching higher: The power of expectations in schooling.* Harvard University Press. https://doi.org/10.5860/choice.41-0458

Weinstein, R. S., Gregory, A., & Strambler, M. J. (2004). Intractable self-fulfilling prophecies fifty years after Brown v. Board of Education. *American Psychologist, 59*(6), 511−520. https://doi.org/10.1037/0003-066x.59.6.511

Wilson, R. (2018). *Nature and young children: Encouraging creative play and learning in natural environments.* New York, NY: Routledge. https://doi.org/10.4324/9781315148533

Wistoft, K. (2013). The desire to learn as a kind of love: Gardening, cooking, and passion in outdoor education. *Journal of Adventure Education & Outdoor Learning, 13*(2), 125-141. https://doi.org/10.1080/14729679.2012.738011

Woodcock, S., & Vialle, W. (2011). Are we exacerbating students' learning disabilities? An investigation of preservice teachers' attributions of the educational outcomes of students with learning disabilities. *Annals of Dyslexia, 61*(2), 223-241. https://doi.org/10.1007/s11881-011-0058-9

Woodcock, S., & Vialle, W. (2010) Attributional beliefs of students with learning disabilities. *The International Journal of Learning Annual Review, 17(7*), 177-192.

Zagona, A. L., Kurth, J. A., & MacFarland, S. Z. (2017). Teachers' views of their preparation for inclusive education and collaboration. *Teacher Education and Special Education*, *40*(3), 163-178.

Zoldosova, K., & Prokop, P. (2006). Education in the field influences children's ideas and interest toward science. *Journal of Science Education and Technology*, *15*(3-4), 304-313. https://doi.org/10.1007/s10956-006-9017-3

Appendix A
Interview Questions

1. How many years have you participated in the OSE program including this year? How many expeditions did your students participate in? How many of these were schoolyard vs. field expeditions? Which topics did you choose?

2. Do you attend OSE expeditions? If so, which ones and what was your role during the expeditions? Would you change your role in the future and if so, how?

3. Prior to your participation with OSE, how would you describe your comfort with science instruction?

4. Did this change during or following the OSE experience? If so, how and if not, why not?

5. Prior to your participation with OSE, how would you describe your comfort taking students outdoors for learning?

6. Did this change during or following the OSE experience? If so, how; and if not, why not?

7. What do you feel are the strongest features of the OSE program?

8. What part of the OSE programs did your students most appreciate?

9. What do you feel are the features of the OSE program that need improvement? How?

10. What part of the OSE program did your students least appreciate or fail to connect with? Do you have suggestions for modification?

11. Did you notice OSE' activities connect with some students more than others? Do you have ideas/groupings of students who benefit most from the OSE experience?

12. What did you notice about your students' behaviors in the outdoors compared to the classroom? Did any students surprise you in their reactions to the outdoor instruction?

13. Do you feel that the OSE experience supported your students' understanding of science concepts specific to 5th grade standards/objectives?

14. Do you feel the OSE experience exposed your students to science concepts beyond the standards?

15. Do you feel the OSE experience impacted students' attitudes about science?

16. Do you feel the OSE experience provided opportunities for you and your students to expand beyond science to other disciplines such as mathematics, social studies, language arts, art, or physical education? If so how and if not, can you describe how to expand it to be more interdisciplinary?

17. Do you have any other comments about the experience you would like to share?

Appendix B
Survey Questions

1. How often do you participate in OSE expeditions with your students?

Always Most of the Time About half the time Sometimes Never

2. How comfortable do you feel teaching science to your students?

Extremely comfortable, Moderately comfortable, Comfortable, Neither comfortable nor uncomfortable, Moderately uncomfortably, Extremely uncomfortable

3. How comfortable do you feel teaching science to your students through outdoor experiences?

Extremely comfortable, Moderately comfortable, Comfortable, Neither comfortable nor uncomfortable, Moderately uncomfortably, Extremely uncomfortable

4. To what degree has the OSE supported you in:
    a. Your science learning

A great deal, much, moderately, little, none

    b. Support your students' learning

A great deal, much, moderately, little, none

    a. Comfort level in teaching science

A great deal, much, moderately, little, none

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 2, 25-56

Coordinated and Intersecting: How Preservice Secondary Science Teachers Understand Science and Engineering Practices and Instructional Principles for Diverse Students

Stacey L. Carpenter
University of California, Santa Barbara

Alexandria K. Hansen
California State University, Fresno

Meghan Macias
University of California, Santa Barbara

Elisa M. Stone
University of California, Berkeley

Erik Arevalo
University of California, Santa Barbara

Julie A. Bianchini
University of California, Santa Barbara

ABSTRACT

Teacher education programs need to prepare their preservice teachers to both implement current science education reforms and teach in culturally and linguistically diverse classrooms. In this qualitative study, we used a framework of four instructional principles to investigate 31 preservice secondary science teachers' understanding of instruction aligned with current reforms and responsive to culturally and linguistically diverse students. We analyzed interview data to examine preservice teachers' evolving understanding of the principle of engaging students in disciplinary practices, specifically the eight science and engineering practices (SEPs) highlighted in current U.S. standards. We found that, over time, participants more often discussed using multiple SEPs together in coordinated ways. However, they consistently reported struggling with the SEPs of *developing and using models* and *using mathematics and computational thinking*. Further, we examined how participants discussed engaging students in disciplinary practices in intersection with the three other principles in our framework: providing students with language production opportunities, attending to and supporting disciplinary language demands, and using student funds of knowledge and other resources. We found that participants focused more often on language production and language support than on student funds of knowledge in intersection with SEPs. Further, participants most frequently discussed the SEP of *engaging in argument from evidence* in intersection with the other three principles. This study is useful for informing teacher educators on how to better support preservice secondary science teachers in developing their understanding of engaging students in disciplinary practices and other instructional principles for diverse students.

Keywords: science and engineering practices, preservice science teachers, secondary science, diverse students

Introduction

The current science education reform movement in the U.S. emphasizes the importance of providing opportunities for ambitious and equitable science learning for all students. *A Framework for K-12 Science Education* [Framework] (National Research Council [NRC], 2012) and the *Next Generation Science Standards* [NGSS] (NGSS Lead States, 2013), in particular, expect teachers to implement

© 2023 International Consortium for Research in Science & Mathematics Education (ICRSME)

rigorous learning goals and instructional approaches that engage all students in science and engineering practices to make sense of and use core ideas and crosscutting concepts. These documents also expect science teachers to "acquire effective strategies to include all students regardless of racial, ethnic, cultural, linguistic, socioeconomic, and gender backgrounds" (NGSS Lead States, 2013, Appendix D, p. 38). Thus, to adequately prepare beginning science teachers to teach science in ways aligned with current reforms and standards, teacher education programs must attend both to rigorous disciplinary instruction and to the diversity of students in U.S. classrooms (National Academy of Sciences et al., 2011).

Our study responds to recent shifts in teacher education that emphasize the complementary nature of effective methods for teaching science and effective methods for teaching culturally and linguistically diverse students (Bravo et al., 2014; Brown, 2017; Lee & Buxton, 2013; Tolbert et al., 2014). These new models of science teacher education provide beginning teachers with principle-based approaches for teaching science content in ways that are responsive to students' cultural and linguistic backgrounds (Bravo et al., 2014; Lyon et al., 2016; Nava et al., 2018; Roberts et al., 2017; Rutt & Mumba, 2020). At the heart of these principle-based approaches is instruction aligned with current reforms and standards, where teachers engage students in science and engineering practices in coordinated ways to make sense of complex problems or phenomena. These principle-based approaches also call for teachers to capitalize on opportunities for students to produce and use language while engaging in these practices; to correspondingly provide necessary language supports; and to use students' diverse cultures, languages, and experiences as resources for disciplinary learning. In this study, we used a framework of four interrelated instructional principles to examine preservice teachers' understanding of science instruction that is aligned with current reforms and responsive to culturally and linguistically diverse students. The principles of our framework include (1) engaging students in disciplinary practices, (2) providing students with rich language production opportunities, (3) attending to disciplinary language demands and providing language supports, and (4) building on student funds of knowledge and other resources.

We consider the principle of engaging students in disciplinary practices as the center of our framework. This principle is rooted in sociocultural perspectives that view learning as increased participation in a community's practices along with the development and use of knowledge from participating in those practices (Lave & Wenger, 1991). It connects to current reforms and standards, which encourage teachers to engage students in eight science and engineering practices (SEPs) that are both reflective of the work of scientists and engineers and central to student learning (NGSS Lead States, 2013; NRC, 2012). Teachers need to provide all students, regardless of linguistic or cultural background, with opportunities to engage in this kind of disciplinary work (Lee et al., 2013; Windschitl & Calabrese Barton, 2016).

For this study, we focused on how 31 preservice secondary science teachers from three teacher education programs described engaging students in disciplinary practices. We also examined how their understanding of instruction in these disciplinary practices intersected with the other three principles in our framework. We asked two sets of research questions:

1. How did preservice science teachers describe engaging students in SEPs? More specifically, how did they describe using multiple SEPs in coordinated ways during instruction? What challenges with engaging students in SEPs did they identify?
2. How did preservice science teachers discuss SEPs in intersection with other principles of effective instruction for culturally and linguistically diverse learners? More specifically, what types of language production opportunities, language demands and supports, and student funds of knowledge related to SEPs did preservice teachers identify?

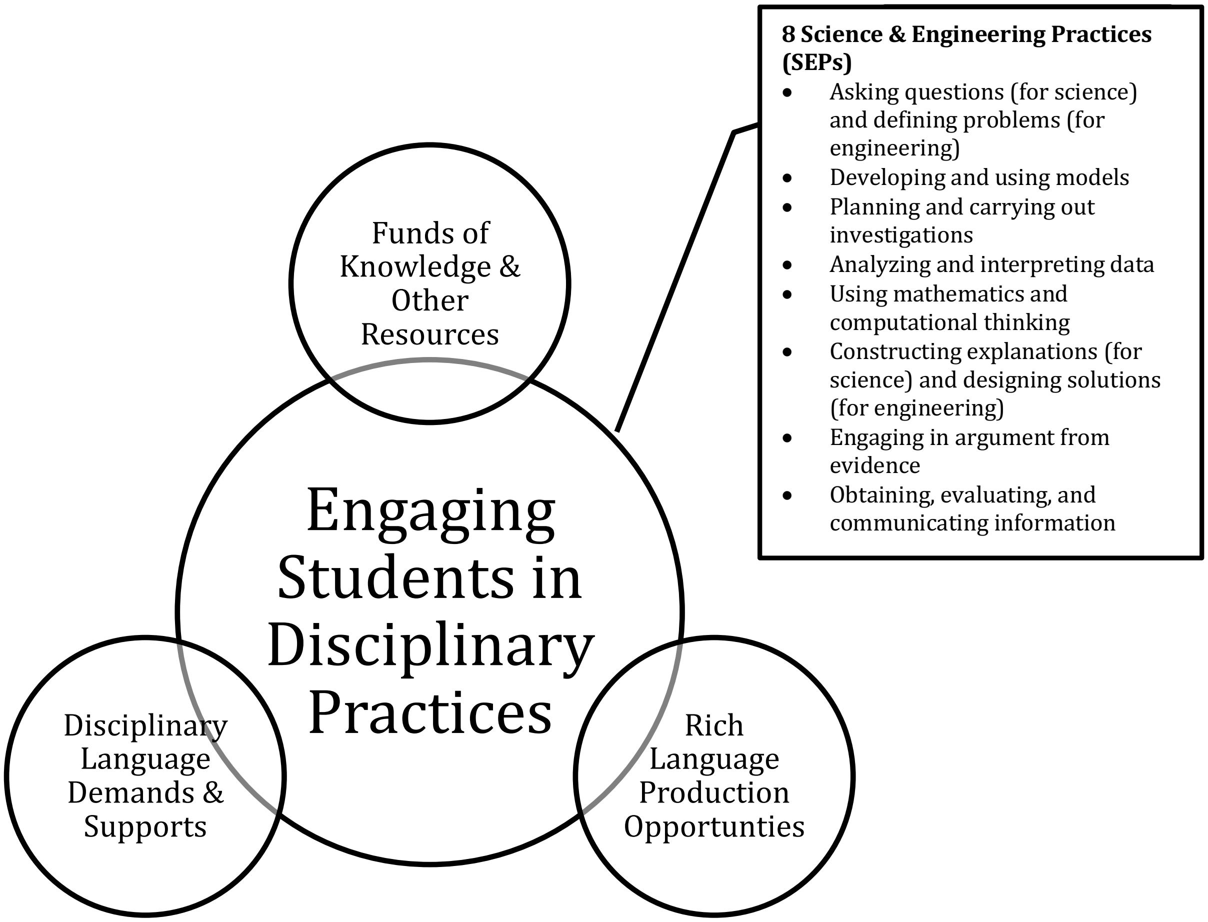
For both sets, we looked for patterns overall, over time, and across programs.

Conceptual Framework

As introduced above, our study was informed by a framework of four interrelated principles of effective science instruction for culturally and linguistically diverse students (see Figure 1). We designed this framework with colleagues (see also Moon et al., 2021; Roberts et al., 2017) based on literature on science instruction that is aligned with current reforms and responsive to culturally and linguistically diverse students. We used this framework to inform both our data collection and analysis.

Figure 1

Conceptual Framework



Disciplinary Practices

The central principle in our framework is engaging students in disciplinary practices. As students participate in the disciplinary work of science, they better understand scientific concepts as well as how scientific knowledge is developed (NRC, 2007, 2012). As described above, the Framework and NGSS specify eight science and engineering practices (SEPs) as important for K-12 science learning (see again Figure 1). The focus on these eight SEPs is rooted in sensemaking: Scientists, engineers, and students use these SEPs to build, refine, and use knowledge to make sense of the world and solve problems (Schwarz et al., 2017). As a means to sensemaking, the SEPs inherently build on and relate to each other; they are not a list of isolated steps. Thus, teachers should engage students in the SEPs in *coordinated* ways to make sense of phenomena.

Rich Language Production Opportunities

Another principle in our framework is providing students with opportunities for rich language production. Learning science involves learning the language unique to science and how to use that language to express ideas and build understanding (Lemke, 1990). The SEPs highlighted in the NGSS are language intensive; participating in these SEPs provides authentic contexts for students to produce and use language (Lee et al., 2013). As students use language to engage in these practices, they make sense of science ideas, which enhances their understanding of scientific concepts and the nature of science (Tolbert et al., 2014). Further, language production is both central to culturally relevant instruction, which positions students and teachers as social makers of knowledge (Ladson-Billings, 1995), and important for multilingual learners, as they develop both language and scientific understanding in science classes (Lyon et al., 2016).

Disciplinary Language Demands and Supports

Another principle in our framework is attending to disciplinary language demands and providing needed language supports. As teachers engage students in authentic disciplinary practices, they also need to attend to the associated language functions that drive the disciplinary practices (e.g., asking questions, constructing explanations) and understand the receptive and productive language demands required of students to engage in these practices (Lyon et al., 2016). Thus, beyond providing opportunities for students to produce language, teachers need to pay attention to those aspects of language that might prove challenging and to provide adequate scaffolding for students to interpret and produce language. Indeed, the language and literacy demands of the SEPs can be challenging for all students, including multilingual learners (Bunch, 2013). In short, as teachers engage students in the language intensive work of the SEPs, they need to attend to the corresponding language demands and provide appropriate supports as well.

Student Funds of Knowledge and Other Resources

The final principle in our framework is building on and using student funds of knowledge and other resources. Student funds of knowledge, the "historically accumulated and culturally developed bodies of knowledge and skills essential for household or individual functioning and well-being," are powerful resources teachers can use to inform instruction (Moll et al., 1992, p. 133). In addition to home-related funds of knowledge, students bring other resources that are valuable for science learning, such as their prior content knowledge, personal interests, and concerns about socioscientific issues (Basu & Calabrese Barton, 2007; Calabrese Barton & Tan, 2009; Campbell et al., 2016). Teachers can draw on these funds and resources in moment-by-moment instruction to engage students in SEPs (Razfar & Nasir, 2019). Contextualizing content knowledge and student engagement in SEPs by building on students' ideas and everyday experiences makes the science more meaningful to all students and can improve learning and participation for culturally and linguistically diverse students, in particular (Tolbert et al., 2019).

Literature Review

There has been limited work specifically examining *preservice secondary* science teachers' understanding of the SEPs articulated in the Framework and NGSS. In one study, Brownstein and Horvath (2016) analyzed preservice secondary science teachers' written responses to the edTPA teacher performance assessment to understand their use of SEPs in their lessons. Researchers found that preservice science teachers consistently described implementing three SEPs in their edTPA lessons: *analyzing and interpreting data*; *constructing explanations and designing solutions*; and *obtaining, evaluating, and communicating information*. The preservice teachers, however, showed limited use of the remaining SEPs. In a second study, French and Burrows (2018) administered a questionnaire to preservice secondary science teachers to gauge their knowledge of and confidence with implementing the SEPs. Researchers found that teachers generally had a working knowledge of the SEPs with respect to curriculum, student understanding, and instructional strategies. However, they also noted that the preservice teachers struggled with enacting certain SEPs, including *developing and using models*, *constructing explanations and designing solutions*, and *engaging in argument from evidence*.

In addition to the few studies that have focused on preservice secondary science teachers, studies of practicing secondary teachers as well as preservice and practicing elementary teachers provide insight into teachers' successes and struggles with SEPs. Prior research with practicing secondary science teachers has examined teachers' understanding of the epistemic nature of science practices as well as how teachers' goals align or misalign with the NGSS (Kawasaki & Sandoval, 2020; Kite et al., 2020). These prior studies found that even practicing secondary science teachers lacked a sophisticated understanding of the SEPs and underestimated how different their teaching practice needed to be to engage students in the SEPs. For example, Kite et al. found that teachers had a limited understanding of the role of models and computational thinking. Prior research with elementary teachers has investigated teachers' ideas about the SEPs as well as how and which SEPs teachers implement in their instruction (Berland et al., 2020; Dalvi et al., 2020; Kang et al., 2018; Kang et al., 2019; Smith & Nadelson, 2017). In previous studies, elementary teachers were found to be successful at integrating the SEPs with each other: Teachers treated the SEPs as comprising a single sensemaking system rather than eight distinct practices (Berland et al., 2020; Kang et al., 2019). However, teachers still struggled with implementing certain SEPs, such as *engaging in argument from evidence* (Kang et al., 2019).

Overall, prior studies have found that preservice and practicing teachers alike need further support in understanding what the SEPs are and how to implement them in their instruction. Our study seeks to extend this body of scholarship by focusing explicitly on preservice secondary science teachers' descriptions of engaging students in the SEPs in ways that intersect with principles of effective instruction for culturally and linguistically diverse students. This work uniquely informs researchers and teacher educators in how to better support preservice teachers so that they deepen their understanding of the SEPs and of how to implement them in relation to each other and to the three interrelated principles discussed above.

Method

Participants and Context

A total of 31 preservice secondary science teachers from three teacher education programs participated in this qualitative study. They represented 69% of the preservice science teachers enrolled in the three programs during the 2016-2017 academic year. Participant demographics are shown in Table 1. All three teacher education programs were housed at universities from the same state university system in the western U.S. The programs were similar in that they included science methods courses that focused on reform-based instruction, additional courses that specifically addressed teaching linguistically and culturally diverse students, and intensive classroom-based field experiences. The program at University 1 was an integrated undergraduate credential program in which preservice teachers earned a bachelor's degree and teaching credential simultaneously, whereas programs at

Universities 2 and 3 were postbaccalaureate credential programs. A summary of characteristics and the number of participants for each teacher education program is shown in Table 2.

Table 1

Preservice Teacher Participant Demographics

Gender	***n***
Female	21
Male	10
Race/Ethnicity	
White/European American	20
Asian/Asian American	7
Multiracial	2
Latinx	1
Other	1
First Language	
English	25
Language(s) other than or in addition to English	6

Note. All demographic data were self-reported.

Table 2

Teacher Education Program Information

University	**Type of Program**	**Length**	**Coursework & Field Placements**	**Credential & Degrees**	***n***
1	Undergraduate	Possible to complete in four years	University-based courses with short-term practicum placements throughout the four years, plus final semester of "student teaching" with associated university seminar	Minor in STEM education, bachelor's degree (in content area), and credential	8
2	Post-baccalaureate	11 months	University-based courses concurrent with field experiences	Potential to earn MEd the following year	11
3	Post-baccalaureate	13 months	University-based courses concurrent with field experiences	Potential to earn MEd concurrently with credential	12

Data Collection

Participants were individually interviewed twice during their teacher education programs (initial and follow-up interviews) using semi-structured interview protocols (Brenner, 2006; see

Appendices A and B for the protocols). Both the initial and follow-up protocols included the same 24 questions to elicit participants' ideas about science teaching and learning. The initial protocol contained an additional five questions about participants' interests in teaching; the follow-up protocol contained an additional nine questions about participants' teacher performance assessment portfolios (i.e., edTPA portfolios); both protocols included similar questions about participants' field placements. The set of 24 common questions included questions that, to varying degrees, addressed all four principles of our conceptual framework. For example, the question, "How would you engage students in discussions?" addresses the principle of language production opportunities, and the question, "What kinds of connections would you make between school science and students' lives outside of school?" addresses the principle of funds of knowledge. The set of 24 common questions also included specific questions about SEPs, including what participants learned about the SEPs, the two SEPs they most frequently implemented in their field experiences, the one SEP they considered as most important to teach students and the one or two SEPs they needed more help to understand or implement. There were also questions about participants' ideas about effective instruction for multilingual learners. We note that we used the term English Language Learner (ELL) in the interview protocols but have since revised our language to multilingual learner. We included all questions from both protocols in our analysis.

Participants at University 1 were interviewed toward the beginning and end of their student teaching semester; participants at Universities 2 and 3 were interviewed toward the beginning and end of their yearlong programs. Initial interviews occurred after participants had started their coursework and had some experience in field placement classrooms. All 62 initial and follow-up interviews lasted approximately one hour. All were audio recorded, professionally transcribed, and checked by researchers for accuracy. Transcripts were anonymized, and all participant names used in the Findings section are pseudonyms.

Data Analysis

We analyzed transcript data using three cycles of coding (Saldaña, 2016). For all coding, the unit of analysis was a main interview question from the interview protocol, including any probes or clarifications that were under the main question. For the first cycle of coding, we analyzed each transcript and coded for instances where participants implicitly or explicitly addressed one or more of the eight SEPs. We applied subcodes for each SEP that was implicitly or explicitly discussed. For the second cycle of coding, we further examined those responses identified in the first cycle where participants addressed one or more SEPs. To do this, we applied codes to mark responses where participants discussed using multiple SEPs in a coordinated way, where participants discussed challenges with engaging students in specific SEPs, and where participants described a clear intersection between SEPs and one or more of the remaining three principles of our Conceptual Framework (i.e., providing rich language production opportunities, attending to disciplinary language demands and supports, and using student funds of knowledge). For the third cycle of coding, we explored the intersections of SEPs with these three principles in more detail. To do this, we narrowed in on the SEP of *engaging in argument from evidence*, because we found that this was the SEP with the most intersections overall. We inductively created a coding scheme to characterize how preservice teachers described the language production opportunities, disciplinary language demands and supports, and student funds of knowledge associated with the SEP of *engaging in argument from evidence*. Descriptions of codes for cycles 1 through 3 are shown in Tables 3, 4, and 5, respectively.

Table 3

Codes for the First Cycle of Analysis

Code	Subcodes	Description
SEPs		Participant talks implicitly or explicitly about engaging students in one or more of the Science and Engineering Practices (SEPs) articulated in the NGSS. Includes examples of engaging students in SEP(s) during actual instruction, hypothetical descriptions of engaging students in SEP(s), or descriptions of why SEP(s) is/are important.
	Asking questions (for science) and defining problems (for engineering)	Includes asking questions about texts, phenomena observed, and conclusions drawn from models or investigations. For engineering, questions should define the problem to be solved and elicit ideas that lead to its solution.
	Developing and using models	Includes diagrams, physical replicas, mathematical representations, analogies, and computer simulations. Although models do not correspond exactly to the real world, they bring certain features into focus and obscure others.
	Planning and carrying out investigations	Includes opportunities to plan and carry out different kinds of investigations, spanning those structured by the teacher and those that emerge from students' own questions.
	Analyzing and interpreting data	Includes use of a range of tools for tabulation, graphical representation, visualization, and statistical analysis of data.
	Using mathematics and computational thinking	Includes using algebraic thinking and analysis, a range of linear and nonlinear functions, exponentials and logarithms, and computational tools for statistical analysis to analyze, represent, and model data. Simple computational simulations are created and used as well.
	Constructing explanations (for science) and designing solutions (for engineering)	Includes explanations and designs that are supported by multiple and independent student-generated sources of evidence consistent with scientific ideas, principles, and theories.
	Engaging in argument from evidence	Includes using appropriate and sufficient evidence and scientific reasoning to defend and critique claims and explanations about the natural and designed world(s).
	Obtaining, evaluating, and communicating information	Includes reading, producing, and evaluating genres of texts that are intrinsic to science and engineering.

Note. Descriptions taken from Appendix F of the *NGSS* (NGSS Lead States, 2013).

Table 4

Codes for the Second Cycle of Analysis

Code	Description
Coordinated	Participant describes coordinating or connecting multiple SEPs with one another so that SEPs are being used together or building from each other in one activity or lesson or over a series of activities/lessons.
Challenges	Participant describes challenges with engaging students in specific SEPs.
Intersection-language production opportunities	Participant describes students' production of language (talk or writing) related to specific SEPs and/or opportunities for students to produce language related to specific SEPs.
Intersection-disciplinary language demands and supports	Participant describes disciplinary language demands and/or language supports related to specific SEPs.
Intersection-funds of knowledge	Participant describes students' backgrounds, prior knowledge, experiences, everyday life, language, cultural strengths, or other resources that students bring in relation to specific SEPs.

Table 5

Codes for Third Cycle of Analysis

Category	Code	Description
Language production opportunities		
Modality	Talking	Participant refers to students engaging in argument from evidence specifically through talk (can include group work, working with a peer, debates).
	Writing	Participant refers to students engaging in argument from evidence specifically through writing.
	Both talking and writing	Participant refers to students engaging in argument from evidence specifically through talk AND writing.
	Unspecified	Participant refers to students engaging in argument from evidence, but type of language use is not explicit and could be interpreted as either writing or talking.
Context	Activity	Participant describes engaging in argument from evidence as part of or resulting from a lab, investigation, project, or larger activity.
	Assessment	Participant refers to engaging in argument from evidence as part of a formative or summative assessment.
	Classroom discussion	Participant refers to engaging in argument from evidence as part of whole-class or small-group discussions.
	Reading	Participant refers to engaging in argument from evidence as part of or related to a reading assignment (e.g., using evidence from a reading to support an argument).
Disciplinary language demands and supports	Writing assignment	Participant refers to engaging in argument from evidence as part of a writing assignment.
	Language demands	Participant discusses language demands specific to the practice of engaging in argument from evidence.
	Claim-Evidence-Reasoning framework (CER)	Participant refers to the Claim-Evidence-Reasoning (CER) framework to scaffold students engaging in argument from evidence.
	Graphic organizer	Participant refers to providing students with a graphic organizer or other visual chunking method to help students write or construct arguments (e.g., providing separate boxes for evidence, reasoning, etc.).
	Group work or peer collaboration	Participant refers to having students work in groups or with peers for engaging in argument from evidence.
	Guided or scaffolded questioning	Participant refers to using specific questions (oral or written) to guide or scaffold students' arguments or evidence; not necessarily assignment questions, but questions to scaffold students' arguments or evidence; can include asking students probing questions to explain or expand their reasoning/argument/evidence.
	Rubric or checklist	Participant refers to providing students with a rubric, checklist of required elements, specifications of requirements related to engaging in argument from evidence (e.g., include 3 pieces of evidence).
	Sentence frames or starters	Participant refers to using sentence starters or sentence frames for engaging in argument from evidence.
Funds of knowledge and other resources	Everyday science experiences	Participant refers to using science topics or examples because they are related to students' everyday life experiences in the context of engaging in argument from evidence.
	Linguistic resources	Participant refers to using students' home language(s) or everyday language(s) in the context of engaging in argument from evidence.
	Prior content knowledge	Participant refers to using students' prior content knowledge or skills in the context of engaging in argument from evidence. Can include content knowledge from previous courses, earlier in the same course, or other STEM learning experiences.
	Socioscientific issues	Participant refers to using social or global issues in the context of engaging in argument from evidence.

We established inter-coder reliability for each of our three cycles of coding in a stepwise fashion following the process of MacPhail et al. (2015). First, members of the research team coded the same interview excerpts independently, and a kappa coefficient was calculated (Fleiss, 1971). The researchers then discussed and resolved areas of disagreement. The process was repeated with additional interview data until a kappa coefficient of 0.8 was achieved and maintained. The researchers then coded remaining data independently or in pairs. To ensure that inter-coder reliability remained acceptable, the researchers coded additional data at designated times throughout the independent/paired coding process and a kappa coefficient of at least 0.8 was maintained. For each cycle, approximately 20% of the transcripts were coded collectively, and 80% were coded independently or in pairs. In addition to establishing inter-coder reliability, we tracked analytic decisions using a detailed audit trail (Guest et al., 2012) to increase the trustworthiness of our analysis (Brenner, 2006).

After coding was completed, we calculated percentages of codes to look for patterns overall, over time, and across universities. We note that because this is a qualitative study, any comparisons of percentages over time and across universities are not for statistical tests of differences but rather to understand general patterns in what participants discussed.

Findings

Finding Set 1: Engaging Students in Disciplinary Practices

Disciplinary Practices as Coordinated

For our first research question, to probe preservice teacher participants' understanding of engaging students in disciplinary practices, we examined responses where participants described using SEPs in a coordinated way. In these responses, participants either discussed coordinating multiple SEPs in a particular activity, lesson, or series of activities/lessons, or they discussed how certain SEPs connect with one another more generally. Overall, we found that 43% of all responses coded for SEPs were also coded as coordinated. In other words, in 43% of responses where participants addressed SEPs, they described engaging students in two or more SEPs in coordinated ways. We also found that the proportion of responses coded as coordinated increased from the initial to follow-up interviews. In initial interviews, 32% of responses addressing SEPs were considered coordinated. In comparison, in the follow-up interviews, 50% of responses addressing SEPs were considered coordinated. This pattern of an increase in the proportion of coordinated responses over time was consistent across universities. Again, we reiterate that this comparison of percentages over time indicates the general pattern of what participants discussed; the statistical significance of this difference was not determined.

In references where participants discussed coordinating SEPs, they most frequently connected the practice of *analyzing and interpreting data* with other SEPs. Over half of all responses coded as "coordinated" included this SEP (see Table 6). For example, in her follow-up interview, Harper described several SEPs, including *analyzing and interpreting data*, as part of the lesson series for her edTPA teacher performance assessment. In this lesson series, Harper guided students in completing a simulated DNA extraction activity to investigate mutations by comparing three species of a fictional organism called "gorks". As she described:

> They [students] definitely went through the series of the science and engineering practices in that they planned and carried out their own investigations as far as what the mutations were actually doing to the gorks—they figured out what would be the best way to actually test that, and what kind of data they would need to collect. And they also got practice in analyzing that data in comparing the different kinds [of DNA sequences] that they got, and then saying how

> that would affect the gorks. And then, in that final paper, [they] had a chance to be communicating the information they got as well as arguing from evidence. Their claim ended up being which of the three mutations was harmful, which was helpful, and which was neutral, and then based on the evidence that they had collected, here's why we're saying that.

Harper coordinated the practice of *analyzing and interpreting data* with *planning and carrying out investigations*; *obtaining, evaluating, and communicating information*; and *engaging in argument from evidence* to support her students in learning about genetic mutations.

Table 6

Percentage of Coordinated Responses That Included Each SEP

SEP	%
Asking questions and defining problems	28
Developing and using models	33
Planning and carrying out investigations	43
Analyzing and interpreting data	56
Using mathematics and computational thinking	22
Constructing explanations and designing solutions	40
Engaging in argument from evidence	49
Obtaining, evaluating, and communicating information	28

Note. Percentages were calculated as the number of responses that included a specific SEP out of the total number of responses coded as coordinated. Percentages do not sum to 100% because each response included multiple SEPs.

Challenges With Implementing Disciplinary Practices

To extend our investigation of participants' understanding of disciplinary practices, we identified responses where participants discussed challenges with implementing SEPs. We found that participants most frequently discussed the two SEPs of *developing and using models* and *using mathematics and computational thinking* as being challenging to implement. *Developing and using models* accounted for 22% of codes for SEPs identified as challenging; *using mathematics and computational thinking* accounted for 21%. This pattern remained consistent across initial and follow-up interviews and was similar across campuses.

For the SEP of *developing and using models*, several participants noted that they lacked an understanding of the practice; they did not understand what a model is. Participants also discussed struggles with implementing the practice, including issues with incorporating modeling into their instruction and a lack of clarity in how to support students' engagement in the practice. For example, in her follow-up interview, Mia described:

> I want to be confident that I'm implementing modeling or models in my class correctly or appropriately, but I struggle with letting students develop their own model. It's like, "What do you think?" But, I'm less competent at helping them revise their models or provide them with the perfect, "Well, this is what it actually is." I'm not sure that needs to happen, but I do know that I can't let students who develop grossly incorrect models to just stop at that point, but

> I'm not sure of the best ways to help them revise their models without invalidating what they did.

Other participants noted their own lack of experience with this practice, either as a K-12 student learning science or as a student teacher in their field placements. As Kayleigh noted in her initial interview, "With the models, I never experienced that growing up. So, it's hard for me to understand all this NGSS stuff because it's totally new."

For the SEP of *using mathematics and computational thinking*, several participants noted challenges with implementing the practice because of students' prior experiences with mathematics or perceived lack of mathematical skills. For example, in her follow-up interview, Sue described this SEP as challenging because of students' math anxiety:

> Because a lot of my students come into my classroom really terrified of math. Whether that's because they had bad math experiences in the past, teachers that they didn't necessarily agree with or like, or that they feel like they just don't know how to do it and aren't good at it…between 75% and 90% of my students say they are not good at math. It becomes very difficult to implement and go for any math things because they tend to just shut down as soon as they see it.

Similar to the challenges discussed with *developing and using models*, participants also noted a lack of clarity about what it means to engage students in *mathematics and computational thinking* or noted challenges with incorporating it into their curricula. In particular, several noted challenges with integrating mathematics into their biology curricula. For example, in her initial interview, Luna commented, "Especially [in] 9th-grade biology, we don't touch on any math equations. I don't know how if I can integrate types of math into what I'm teaching for biology… I'm not really sure where to start for that one."

Finding Set 2: Disciplinary Practices in Intersection With Other Principles

For our second research question, we analyzed where and how preservice teachers talked about disciplinary practices in intersection with the other three principles of effective instruction for culturally and linguistically diverse students described in our Conceptual Framework: providing rich opportunities for student language production, attending to disciplinary language demands and providing language supports, and using student funds of knowledge. Overall, we found that participants addressed one or more of these three principles in approximately 55% of their responses that addressed SEPs. As shown in Table 7, participants most often discussed providing language production opportunities in intersection with SEPs, followed by disciplinary language demands and supports. They least often discussed using student funds of knowledge or other resources in intersection with SEPs. This pattern remained consistent across initial and follow-up interviews and across universities.

Table 7

Percentage of Responses with One or More Intersections that Included Each Principle

Principle	%
Rich language production opportunities	84
Disciplinary language demands and supports	64
Funds of knowledge and other resources	24

Note. Percentages were calculated as the number of responses that included each specific principle out of the total number of responses that included an intersection with one or more principles and SEPs. Percentages do not sum to 100% because one to multiple principles could be included in each response.

To elaborate, we found that participants most frequently discussed the SEP of *engaging in argument from evidence* in intersection with one or more of the other three principles; nearly half of the responses addressing an intersection of SEPs with one or more of the three principles included *engaging in argument from evidence* (see Table 8). To clarify, this differed from the SEP, most often discussed in coordination with other SEPs, *analyzing and interpreting data*. This pattern was consistent across initial and follow-up interviews and across universities. The proportions of principles intersecting with this SEP also followed the same patterns as discussed above, with the principle of language production opportunities intersecting the most and the principle of funds of knowledge intersecting the least. Thus, we focused on *engaging in argument from evidence* to further explore the types of language opportunities, disciplinary language demands and supports, and student funds of knowledge discussed by the participants in relation to this SEP. We characterized how preservice teachers described (1) the context and modality of student language production associated with this SEP, (2) the types of disciplinary language demands and/or supports for this SEP, and (3) the types of student funds of knowledge or other resources used as students engaged in this SEP.

Table 8

Percentage of Responses with One or More Intersections that Included Each SEP

SEP	%
Asking questions and defining problems	17
Developing and using models	26
Planning and carrying out investigations	31
Analyzing and interpreting data	32
Using mathematics and computational thinking	14
Constructing explanations and designing solutions	30
Engaging in argument from evidence	48
Obtaining, evaluating, and communicating information	23

Note. Percentages were calculated as the number of responses that included a specific SEP out of the total number of responses that included an intersection with one or more principles and SEPs. Percentages do not sum to 100% because one to multiple SEPs could be included in each response.

Rich Language Production Opportunities

To characterize the ways participants described student language production associated with the SEP of *engaging in argument from evidence*, we examined both the language modality (i.e., writing, talking) and the context of student participation in the SEP (e.g., as part of an investigation or larger activity, as a class discussion). We note that participants' descriptions of *engaging in argument from evidence* did not always describe participation in the complete practice: Participants often described the practice as students making and supporting claims without addressing the evaluation, critique, and reconciliation components of argumentation (Berland et al., 2017). We found that in approximately 30% of responses, participants described students engaging in argument from evidence through writing, and in another, approximately 30% of responses, participants described students engaging in argument through talking. Further, we found that in approximately 30% of responses, the language modality was unspecified—the type of language use was not explicit and could be interpreted as either writing or talking. We found few instances (less than 10% of responses) where participants described students engaging in both written and oral arguments.

For the modality of written argumentation, participants regularly described students *engaging in argument from evidence* in the contexts of writing assignments connected to investigations or larger activities. For example, in her initial interview, Lanh described how students in her placement typically completed writing assignments using the claim, evidence, and reasoning framework to "wrap up what they've learned" after an activity or lesson. She further explained that students engaged in argument from evidence "through the claim, evidence, reasoning [writing assignments] that they have to do…which are pretty big because we have an activity or lesson that goes before they can actually create a claim, evidence, and provide reasoning." Less often, participants described students engaging in written argumentation in the contexts of general writing assignments, reading assignments, or assessments.

For the modality of oral argumentation, participants most often described students engaging in argument in the context of classroom discussions. For example, when describing how she facilitated discussions, Madelyn said, "I just really like asking questions and asking students to draw out what they're saying, or expand on what they're saying, or maybe provide an argument against what someone else is saying." Less often, participants described students engaging in oral argumentation in the context of an investigation or larger activity.

When the language modality was unspecified, participants typically described students engaging in argument in the context of an investigation or larger activity. For example, while describing how she implemented the SEP of *engaging in argument from evidence*, Kathryn stated that students "were asked to use that evidence, use their data from their lab, to explain whether or not their original prediction was correct, and why they thought that." Here, we clarify, it was unclear whether students' arguments about their predictions were written or oral.

For the few instances where participants described students engaging in argument through both talking and writing, the contexts were typically both specific writing assignments and classroom discussions connected to an investigation. For example, in the follow-up interview, Eric described:

> Essentially after the labs, they [students] have to write claims, and those claims have to be backed up with evidence they collected in the lab. And so, it allows them to engage in more argument from evidence because some students, their evidence might point to different things, and then we talk about it, and argue about it in a productive way so that the concepts are more explicitly laid out for them.

Here, Eric pointed out how students engaged in both written and oral argumentation stemming from laboratory activities. We add that, in this example, as did Madelyn above, Eric did include students arguing with one another about differing claims and evidence.

Disciplinary Language Demands and Supports

Regarding the principle of attending to disciplinary language demands and providing disciplinary language supports, we found that discussions of disciplinary language demands associated with the SEP of *engaging in argument from evidence* accounted for 10% of codes for this principle. The remaining 90% of codes for this principle comprised the various types of language supports that participants discussed related to this SEP. In their discussions of disciplinary language demands, participants often spoke of the demands of students' struggles with identifying and using evidence to support arguments. For example, in the follow-up interview, when describing an assignment about evolution, Kayla said:

> I wanted them [students] to use all of the evidence we had gathered to engage in an argument, which some students just defined each category [of evidence] instead of using the category as an argumentative tool. For instance, they knew that DNA was evidence for evolution, but they didn't know how to talk about, how to argue about DNA.

Here, Kayla acknowledged the challenges her students had with using the evidence they had gathered to support their arguments.

In their discussions of language supports for *engaging in argument from evidence*, participants most frequently mentioned using a claim, evidence, reasoning (CER) framework (McNeill & Krajcik, 2012). Discussions of the CER framework accounted for 30% of the codes for the principle of disciplinary language demands and supports. For example, in her initial interview, Luna noted, "I usually do a claim, evidence, reasoning to promote the writing." When asked about SEPs implemented most often in her placement, she replied:

> Let's say engaging in an argument from evidence. Like I said, I have them doing a claim, evidence, reasoning. I usually have a driving question that I introduce three to five lessons beforehand….Then students are then able to use those lessons and what they took from those lessons as evidence. They'll make their claim and they'll have evidence from all the activities we did. Then they'll do the reasoning part where they'll restate their claim and pull specific things from their evidence to support that claim.

Luna used the CER framework to support students engaging in argument as part of summative writing assignments following a series of lessons. The next most frequently discussed language support for *engaging in argument* was group work or peer collaboration (17% of codes for this principle). For example, in his follow-up interview, Timothy commented that argumentation "doesn't necessarily need to be a whole-class discussion. It could be in lab groups, anywhere that you can maximize opportunities for students to speak in that specific academic register where you have arguing from evidence, citing or stating your claims, reasoning your arguments, that sort of thing." Finally, participants less frequently discussed other language supports, such as rubrics or checklists, graphic organizers, and sentence frames.

Funds of Knowledge and Other Resources

Overall, we found that the principle of using student funds of knowledge and other resources was seldom discussed in intersection with the SEP of *engaging in argument from evidence*. Of the few instances, participants mainly discussed the resource of students' prior content knowledge to support argumentation. Sadie noted in her follow-up interview that students engaged in arguments about the rock cycle using conventions of molecular diagramming, which they had covered in a previous unit. As she described, "We connected it to what we had been talking about when we talked about convection and making pictures of those molecules then. They used those molecular diagrams to support their arguments." Other types of resources, including students' every day science experiences, linguistic resources, and awareness of socioscientific issues, were discussed once each. As an example of using a socioscientific issue to connect to student funds of knowledge, Kayla described an activity where students participated in a mock city council meeting and engaged in a debate about genetically modified organisms (GMOs). As she explained, "I gave them a fake case study about these GMO papayas in Hawaii. So, I had each student take on the role of a different person in the debate – so a farmer, the GMO person, someone who's an organic farmer." She used a current and familiar social issue as a context for engaging in argument.

Discussion and Implications

Our findings provide insight into the successes and struggles that preservice secondary science teachers experience with engaging students in the disciplinary practices of science and engineering and in the principles of effective instruction for diverse students. We examined how preservice teachers from three teacher education programs discussed the eight SEPs highlighted in recent U.S. reform documents. We also examined how they described rich language production opportunities, language demands and supports, and student funds of knowledge associated with these SEPs. We found that, over time, preservice teachers more often described the SEPs as coordinated. However, we also found that preservice teachers consistently identified struggles to understand and implement two SEPs: *using mathematics and computational thinking* and *developing and using models*. Further, we found that preservice teachers readily and consistently described opportunities for students to produce language with the SEPs and, to a lesser extent, language demands and supports associated with the SEPs. However, we found that preservice teachers struggled with the principle of using funds of knowledge or other resources to engage their students in the SEPs.

Strengthening Preservice Teachers' Understanding of Disciplinary Practices

Looking more closely at the findings for our first research question, we found that preservice teachers' descriptions of using the SEPs in coordinated ways increased over time: A higher percentage of their discussions included descriptions of using two or more SEPs in coordinated ways at the end of their teacher education programs compared to the beginning. This growth over time is promising and suggests that preservice teachers developed a better sense of the coordinated nature of SEPs through their teacher education experiences. This finding is consistent with other research that has found increases in teachers' use and understanding of SEPs as coordinated after learning opportunities focused on the SEPs (Berland et al., 2020; Kang et al., 2019). It contrasts with other studies, such as that by Kite et al. (2020), where relatively few practicing secondary science teachers were found to exhibit sophisticated understandings of scientific practices that extended beyond the rigid, linear scientific method. Our finding remains important because understanding the SEPs as coordinated is aligned with the larger goals of the NGSS, which emphasize that the SEPs should be conceived not simply as a list of practices to check off but rather as a coordinated way to build, use, and make sense of scientific knowledge (NGSS Lead States, 2013; Schwarz et al., 2017). However, we did not examine how the three teacher education programs in our study facilitated preservice teachers' growth in their understanding of the coordinated nature of SEPs—that was beyond of the scope of this study. This is a fruitful avenue for further research. Future studies should examine the opportunities for learning that teacher education programs provide to facilitate such growth.

We also examined which SEPs preservice teachers reported as especially challenging to implement. Identifying the SEPs that preservice teachers described as challenging is important to inform how teacher educators can better support them. We found one of the most common SEPs that preservice teachers deemed challenging was *using mathematics and computational thinking*. They reported struggling to accommodate students' varying mathematical backgrounds and skills, a lack of clarity on what the SEP entails, and uncertainty on how to incorporate it into their instruction, particularly biology instruction. Other studies have similarly identified *using mathematics and computational thinking* as a challenging SEP for teachers to understand and implement (Brownstein & Horvath, 2016; Kite et al., 2020). We also found preservice teachers reported the SEP of *developing and using models* as challenging. This resonates with other, more specific studies on modeling, which have found that teachers struggle with understanding and using models as tools for scientific inquiry, having students construct and evaluate models, and seeing models as more than illustrations of phenomena or patterns (Khan, 2011; Schwarz & Gwekwerere, 2007; Windschitl & Thompson, 2006). With such clear struggles surrounding these two SEPs, teacher education programs need to better support preservice teachers in understanding and implementing them.

We recommend that teacher educators leverage the SEPs that preservice teachers feel more comfortable with as a focal point for exploring the SEPs that are more challenging. Given the coordinated and overlapping nature of the SEPs, teacher educators can point out explicit connections and overlaps between SEPs that preservice teachers find challenging and SEPs they readily use and understand. As one example, in our study, *analyzing and interpreting data* was the SEP preservice teachers most often described as coordinated with other SEPs—perhaps indicating a high level of familiarity with and understanding of this SEP. Interestingly, although preservice teachers seemed to have facility with implementing *analyzing and interpreting data*, they still encountered difficulties with implementing *mathematics and computational thinking*—an SEP with clear connections to analyzing and interpreting data. Indeed, mathematics and computation are tools for analyzing and interpreting data that facilitate the analytic process along with making sense of and reasoning with data (Rivet & Ingber, 2017). Thus, highlighting the coordination between these two SEPs could support preservice teachers in better recognizing and understanding the use of mathematics and computational thinking in their curricula. As a second example, other researchers have suggested *developing and using models* as an anchor for engaging students in other SEPs (Passmore et al., 2013; Passmore et al., 2009). However, we found that *developing and using models* was a challenging SEP for preservice teachers to understand and implement. Our study provides evidence that the SEP of *analyzing and interpreting data* could serve as an alternative anchor practice, as a strong entry point to develop their understanding of the other SEPs, because preservice teachers are comfortable with it.

Strengthening Preservice Teachers' Understanding of Intersecting Principles

For our second research question, we examined how preservice teachers discussed SEPs in intersection with the instructional principles of providing language production opportunities, attending to disciplinary language demands and supports, and using student funds of knowledge and other resources. We found that in their discussions of SEPs, preservice teachers most often touched on the principle of providing language production opportunities, followed by the principle of attending to disciplinary language demands and supports. The SEPs are noted as being language intensive (Lee et al., 2013), and current reforms emphasize the need for students to engage in the language of science (NRC, 2007). Thus, it is promising that preservice teachers in our study recognized opportunities to use language through SEPs and provided support for this language use.

We also found that preservice teachers tended to focus on the language aspects of certain SEPs over others: They most often discussed the language principles in intersection with the SEP of *engaging in argument from evidence*. A closer examination of the types of language opportunities, demands, and supports related to *engaging in argument from evidence* showed that preservice teachers used supports like the claim, evidence, reasoning (CER) framework (McNeill & Krajcik, 2012) and peer collaboration to facilitate students engaging in argument through writing assignments connected to larger activities and through classroom discussions. The CER framework was a frequent support mentioned by preservice teachers from each of the three universities. Since the preservice teachers had a clear tool to support the disciplinary language demands of *engaging in argument from evidence*, perhaps they were better able to describe the language opportunities of this SEP.

Teacher educators could do more to help teachers recognize the language opportunities and demands associated with each of the SEPs. A focus on tools, like the CER framework, that support students' language use with each SEP could help teachers recognize and implement opportunities for language production through student engagement with all the SEPs. For example, Windschitl et al. (2018) developed a suite of tools to support student language production with *planning and carrying out investigations*, *constructing explanations*, and *engaging in argument from evidence*.

Although preservice teachers recognized opportunities and supports for student language production through SEPs, they struggled to recognize how student funds of knowledge or other resources could be used to engage students in SEPs. This finding resonates with prior studies that have documented similar struggles among beginning teachers to contextualize classroom science activity in students' lives outside of school (Bravo et al., 2014; Tolbert et al., 2019). In our examination of *engaging in argument from evidence*, we found few instances of preservice teachers acknowledging student funds of knowledge and resources in relation to this SEP; most consisted of discussing students' prior content knowledge. As Razfar and Nasir (2019) pointed out, student funds of knowledge about scientific practices can come from in-school, out-of-school, and in-between experiences, and teachers can draw on these various funds of knowledge in dynamic ways by considering how student funds come into play beyond curricular topics—for example, by connecting to students' values, beliefs, and contested ideologies as students engage in argument from evidence.

Building on the example given by Kayla, we recommend that teachers connect to their student funds of knowledge while *engaging in argument from evidence* by contextualizing the SEP in socioscientific issues (SSIs). SSIs are controversial social issues that have conceptual or procedural links to science and readily connect to student funds of knowledge (Sadler, 2004; Zeidler et al., 2009; Zeidler & Sadler, 2011). Further, grounding argumentation in SSIs can help teachers facilitate deeper argumentation beyond claims, evidence, and reasoning—which we also found as a struggle for preservice teachers. Indeed, argumentation is a central focus of SSI instructional frameworks (Aikenhead, 1985; Driver et al., 2000), where the SSIs examined and argued have personal meaning to students. As a result, students can construct more substantive arguments because of their interest in and connection with an SSI. In sum, including SSI frameworks in teacher education programs can help teachers connect their lessons to students' lives (Johnson et al., 2020) and improve students' ability to effectively argue in their classrooms.

Limitations

We recognize that our study has several limitations. First, although our interview questions asked preservice teachers about their practice, we did not examine their actual classroom instruction. As such, we were unable to determine how closely their reports resonated with their actual instruction. Second, our interviews asked a range of questions that addressed each of the four principles of our

Conceptual Framework. However, we did not ask questions that directly addressed the coordination of SEPs with one another or the intersection of SEPs with these principles. Had participants been asked about such intersections directly, they may have elaborated on their understandings. Finally, we did not specifically examine the opportunities for learning about SEPs and instruction for diverse students provided by the three teacher education programs in our study. A deeper examination of programmatic factors would generate additional recommendations for preparing preservice teachers for instruction that incorporates SEPs and instructional principles for diverse students.

Conclusion

Although a deeper examination of programmatic factors is needed, the inclusion of multiple programs in our study does contribute to a broader understanding of preparing preservice science teachers because studies of teacher education are often small in scale and consist of case studies of individual interventions (Sleeter, 2014). Indeed, our findings were generally consistent across the three teacher education programs included in our study, pointing to common successes and struggles that preservice teachers experienced. Said another way, this study offers insight into how to better support preservice teachers beyond improvements to single programs.

To conclude, we found that preservice teachers grew in their understanding of the coordinated nature of SEPs, and while they reported challenges with certain SEPs, they seemed successful with others. Thus, we recommend that teacher educators leverage these strengths to help preservice teachers better understand challenging SEPs. Further, we found that preservice teachers more readily recognized the intersections of SEPs with language opportunities and supports than with student funds of knowledge. More specifically, preservice teachers described oral and written language production opportunities for the SEP of *engaging in argument from evidence* along with specific supports, like the CER framework. Thus, we suggest that teacher educators consider other tools and instructional supports that can help preservice teachers draw on the language opportunities and student funds of knowledge related to all SEPs – to fully engage culturally and linguistically diverse students in reform-based science education.

This work was supported by the National Science Foundation DUE-1557283. We thank Trish Stoddart, Gretchen Andreasen, Leslie Bushong, Cheryl Forbes, Susann Pinter, and Sherry Seethaler for their guidance and assistance with this project.

Stacey L. Carpenter (staceycarpenterphd@gmail.com) Stacey L. Carpenter is currently a lecturer at Cal Poly Pomona. She was a postdoctoral scholar at the University of California, Santa Barbara for the Science and Mathematics Teacher Research Initiative.

Meghan Macias (mmacias2@wested.org) Meghan Macias is a Research Associate at WestEd. She received her Ph.D. in Education from the University of California, Santa Barbara.

Erik Arevalo (erarevalo@ucsd.edu) is a postdoctoral scholar at the University of California, San Diego in the biological sciences department. He received his PhD and MA from the University of California, Santa Barbara in Science Education.

Alexandria K. Hansen (khansen@mail.fresnostae.edu) is an Assistant Professor in the Department of Biology at California State University, Fresno. She also coordinates the Single Subject Science Teaching Credential Program. Previously, she earned her PhD and MA in STEM Education from the University of California, Santa Barbara, and her Secondary Science Teaching Credential from Loyola Marymount University in Los Angeles.

Elisa M. Stone (emstone@berkeley.edu) is the Director for CalTeach, an undergraduate STEM teacher education program at University of California (UC), Berkeley. She holds a PhD in Genetics from the State University of New York (SUNY) at Stony Brook, an MA in Education from UC, Berkeley and a California Single Subject Teaching Credential in Life Sciences.

Julie A. Bianchini (jbianchi@education.ucsb.edu) is a Professor of Science Education in the Department of Education at the University of California, Santa Barbara (UCSB). She also serves as Faculty Director of UCSB's CalTeach, part of a UC-wide initiative to recruit and better prepare science and mathematics teachers. Dr. Bianchini investigates ways to make science accessible, understandable, and relevant to culturally and linguistically diverse students, particularly at the secondary level.

References

Aikenhead, G. S. (1985). Collective decision making in the social context of science. *Science Education*, *69*(4), 453-575.

Basu, S. J. & Calabrese Barton, A. (2007). Developing a sustained interest in science among urban minority youth. *Journal of Research in Science Teaching*, *44*(3), 466-489. https://doi.org/10.1002/tea.20143

Berland, L. K., McNeill, K. L., Pelletier, P., & Krajcik, J. (2017). Engaging in argument from evidence. In C. V. Schwarz, C. Passmore, & B. J. Reiser (Eds.), *Helping students make sense of the world using next generation science and engineering practices* (pp. 229-257). NSTA Press.

Berland, L. K., Russ, R. S., & West, C. P. (2020). Supporting the Scientific Practices through Epistemologically Responsive Science Teaching. *Journal of Science Teacher Education*, *31*(3), 264-290. https://doi.org/10.1080/1046560x.2019.1692507

Bravo, M. A., Mosqueda, E., Solís, J. L., & Stoddart, T. (2014). Possibilities and limits of integrating science and diversity education in preservice elementary teacher preparation. *Journal of Science Teacher Education*, *25*(5), 601-619. https://doi.org/10.1007/s10972-013-9374-8

Brenner, M. E. (2006). Interviewing in educational research. In J. L. Green, G. Camilli, & P. B. Elmore (Eds.), *Handbook of complementary methods in education research* (pp. 357-370). Lawrence Erlbaum Associates.

Brown, J. C. (2017). A metasynthesis of the complementarity of culturally responsive and inquiry-based science education in K-12 settings: Implications for advancing equitable science teaching and learning. *Journal of Research in Science Teaching*, *54*(9), 1143-1173. https://doi.org/10.1002/tea.21401

Brownstein, E. M., & Horvath, L. (2016). Next Generation Science Standards and edTPA: Evidence of science and engineering practices. *Electronic Journal of Science Education*, *20*(4). https://ejrsme.icrsme.com/article/view/16090

Bunch, G. C. (2013). Pedagogical language knowledge: Preparing mainstream teachers for English learners in the new standards era. *Review of Research in Education*, *37*, 298-341. https://doi.org/10.3102/0091732X12461772

Calabrese Barton, A., & Tan, E. (2009). Funds of knowledge and discourses and hybrid space. *Journal of Research in Science Teaching*, 46(1), 50-73. https://doi.org/10.1002/tea.20269

Campbell, T., Schwarz, C. V., & Windschitl, M. (2016). What we call misconceptions may be necessary stepping-stones toward making sense of the world. *Science and Children*, *53*(7). https://doi.org/10.2505/4/sc16_053_07_28

Dalvi, T., Silva Mangiante, E., & Wendell, K. (2020). Identifying pre-service teachers' conceptions about the NGSS practices using a curriculum critique and revision (CCR) task. *Journal of Science Teacher Education*, 1-25. https://doi.org/10.1080/1046560x.2020.1791465

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*(3), 287-312. https://doi.org/10.1002/(SICI)1098-237X(200005)84:3%3C287::AID-SCE1%3E3.0.CO;2-A

Fleiss, J. L. (1971). Measuring nominal scale agreement among many raters. *Psychological Bulletin*, *76*(5), 378-382. https://doi.org/10.1037/h0031619

French, D. A., & Burrows, A. C. (2018). Evidence of science and engineering practices in preservice secondary science teachers' instructional planning. *Journal of Science Education and Technology*, *27*(6), 536-549. https://doi.org/10.1007/s10956-018-9742-4

Guest, C., MacQueen, K. M., & Namey, E. E. (2012). *Applied thematic analysis*. Sage.

Johnson, J., Macalalag, A. Z., & Dunphy, J. (2020). Incorporating socioscientific issues into a STEM education course: Exploring teacher use of argumentation in SSI and plans for classroom implementation. *Disciplinary and Interdisciplinary Science Education Research*, *2*(1). https://doi.org/10.1186/s43031-020-00026-3

Kang, E. J. S., Donovan, C., & McCarthy, M. J. (2018). Exploring elementary teachers' pedagogical content knowledge and confidence in implementing the NGSS science and engineering practices. *Journal of Science Teacher Education*, *29*(1), 9-29. https://doi.org/10.1080/1046560x.2017.1415616

Kang, E. J. S., McCarthy, M. J., & Donovan, C. (2019). Elementary teachers' enactment of the NGSS science and engineering practices. *Journal of Science Teacher Education*, *30*(7), 788-814. https://doi.org/10.1080/1046560x.2019.1630794

Kawasaki, J., & Sandoval, W. A. (2020). Examining teachers' classroom strategies to understand their goals for student learning around the science practices in the Next Generation Science Standards. *Journal of Science Teacher Education*, *31*(4), 384-400. https://doi.org/10.1080/1046560x.2019.1709726

Khan, S. (2011). What's missing in model-based teaching. *Journal of Science Teacher Education*, *22*(6), 535-560. https://doi.org/10.1007/s

Kite, V., Park, S., McCance, K., & Seung, E. (2020). Secondary science teachers' understandings of the epistemic nature of science practices. *Journal of Science Teacher Education*, 1-22. https://doi.org/10.1080/1046560x.2020.1808757

Lee, O., & Buxton, C. A. (2013). Integrating science and English proficiency for English language learners. *Theory Into Practice*, *52*, 36-42. https://doi.org/10.1080/07351690.2013.743772

Lee, O., Quinn, H., & Valdés, G. (2013). Science and language for English language learners in relation to Next Generation Science Standards and with implications for Common Core State Standards for English Language Arts and Mathematics. *Educational Researcher*, *42*(4), 223-233. https://doi.org/10.3102/0013189X13480524

Lemke, J. L. (1990). *Talking science: Language, learning, and values*. Ablex.

Lyon, E. G., Tolbert, S., Stoddart, T., Solis, J., & Bunch, G. C. (2016). *Secondary science teaching for English learners: Developing supportive and responsive learning contexts for sense-making and language development*. Rowman & Littlefield.

MacPhail, C., Khoza, N., Abler, L., & Ranganathan, M. (2015). Process guidelines for establishing Intercoder Reliability in qualitative studies. *Qualitative Research*, *16*(2), 198-212. https://doi.org/10.1177/1468794115577012

McNeill, K. L., & Krajcik, J. (2012). *Supporting grade 5-8 students in constructing explanations in science: The claim, evidence and reasoning framework for talk and writing.* Pearson Allyn & Bacon.

Moll, L. C., Amanti, C., Neff, D., & Gonzalez, N. (1992). Funds of knowledge for teaching: Using a qualitative approach to connect homes and schools. *Theory Into Practice*, *31*(2), 132-141. https://doi.org/10.1080/00405849209543534.

Moon, S., Carpenter, S. L., Hansen, A. K., Bushong, L., & Bianchini, J. A. (2021). Examining the effects of undergraduate STEM teacher recruitment and teacher education programs on preservice secondary science and mathematics teacher readiness and teacher performance assessment (edTPA) scores. *School Science and Mathematics*, *121*(8), 452-465. https://doi.org/10.1111/ssm.12498

National Academy of Sciences, National Academy of Engineering, & Institute of Medicine. (2011). *Expanding underrepresented minority participation: America's science and technology talent at the crossroads.* The National Academies Press. https://doi.org/doi:10.17226/12984

National Research Council. (2007). *Taking science to school: Learning and teaching science in grades K-8*. The National Academies Press. https://doi.org/10.17226/11625

National Research Council. (2012). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas*. The National Academies Press. https://doi.org/10.17226/13165

Nava, I., Park, J., Dockterman, D., Kawasaki, J., Schweig, J., Quartz, K. H., & Martinez, J. F. (2018). Measuring teaching quality of secondary mathematics and science residents: A classroom observation framework. *Journal of Teacher Education*, *70*(2), 139-154. https://doi.org/10.1177/0022487118755699

NGSS Lead States. (2013). *Next generation science standards: For states, by states*. The National Academies Press. https://doi.org/10.17226/18290

Passmore, C., Coleman, E., Horton, J., & Parker, H. (2013). Making sense of natural selection: Developing and using the natural selection model as an anchor for practice and content. *The Science Teacher*, *80*(6), 43-49. https://www.jstor.org/stable/43557251

Passmore, C., Stewart, J., & Cartier, J. (2009). Model-based inquiry and school science: Creating connections. *School Science and Mathematics*, *109*(7), 394-402. https://doi.org/10.1111/j.1949-8594.2009.tb17870.x

Razfar, A., & Nasir, A. (2019). Repositioning English learners' funds of knowledge for scientific practices. *Theory Into Practice*, *58*(3), 226-235. https://doi.org/10.1080/00405841.2019.1599231

Rivet, A. E., & Ingber, J. (2017). Analyzing and interpreting data. In C. V. Schwarz, C. Passmore, & B. J. Reiser (Eds.), *Helping students make sense of the world using next generation science and engineering practices* (pp. 159-180). NSTA Press.

Roberts, S. A., Bianchini, J. A., Lee, J. S., Hough, S., & Carpenter, S. L. (2017). Developing an adaptive disposition for supporting English language learners in science: A capstone science methods course. In A. W. Oliveira & M. H. Weinburgh (Eds.), *Science Teacher Preparation in Content-Based Second Language Acquisition* (pp. 79-95). Springer. https://doi.org/10.1007/978-3-319-43516-9_5

Rutt, A. A., & Mumba, F. M. (2020). Developing secondary pre-service science teachers' instructional planning abilities for language- and literacy-integrated science instruction in linguistically diverse classrooms. *Journal of Science Teacher Education*, 1-28. https://doi.org/10.1080/1046560x.2020.1760431

Sadler, T. D. (2004). Informal reasoning regarding socioscientific Issues: A critical review ofthe research. *Journal of Research in Science Teaching*, 41(5), 513-536. https://doi.org/10.1002/tea.20009

Saldaña, J. (2016). *The coding manual for qualitative researchers* (3rd ed.). Sage.

Schwarz, C. V., & Gwekwerere, Y. N. (2007). Using a guided inquiry and modeling instructional framework (EIMA) to support preservice K-8 science teaching. *Science Education*, *91*(1), 158-186. https://doi.org/10.1002/sce.20177

Schwarz, C. V., Passmore, C., & Reiser, B. J. (2017). Moving beyond "knowing about" science to making sense of the world. In C. V. Schwarz, C. Passmore, & B. J. Reiser (Eds.), *Helping students make sense of the world using next generation science and engineering practices* (pp. 3-21). NSTA Press.

Sleeter, C. (2014). Toward teacher education research that informs policy. *Educational Researcher*, *43*(3), 146-153. https://doi.org/10.3102/0013189X14528752

Smith, J., & Nadelson, L. (2017). Finding alignment: The perceptions and integration of the Next Generation Science Standards Practices by elementary teachers. *School Science and Mathematics*, *117*(5), 194-203. https://doi.org/10.1111/ssm.12222

Tolbert, S., Knox, C., & Salinas, I. (2019). Framing, adapting, and applying: Learning to contextualize science activity in multilingual science classrooms. *Research in Science Education*, *49*(4), 1069-1085. https://doi.org/10.1007/s11165-019-9854-8

Tolbert, S., Stoddart, T., Lyon, E. G., & Solis, J. (2014). The Next Generation Science Standards, Common Core State Standards, and English learners: Using the SSTELLA framework to prepare secondary science teachers. *Issues in Teacher Education*, *23*(1), 65-90. https://www.itejournal.org/spring-2014-issue/

Windschitl, M., & Calabrese Barton, A. (2016). Rigor and equity by design: Locating a set of core teaching practices for the science education community. In D. H. Gitomer & C. A. Bell (Eds.), *Handbook of research on teaching* (5th ed., pp. 1099-1158). American Educational Research Association.

Windschitl, M., & Thompson, J. (2006). Transcending simple forms of school science investigation: The impact of preservice instruction on teachers' understandings of model-based inquiry. *American Educational Research Journal*, *43*(4), 783-835. https://doi.org/10.3102/00028312043004783

Windschitl, M., Thompson, J., & Braaten, M. (2018). *Ambitious science teaching*. Harvard Education Press.

Zeidler, D. L., Sadler, T. D., Applebaum, S., & Callahan, B. E. (2009). Advancing reflective judgment through socioscientific issues. *Journal of Research in Science Teaching*, 46(1), 74–101. https://doi.org/10.1002/tea.20281

Zeidler, D. L., & Sadler, T. D. (2011). An inclusive view of scientific literacy: Core issues and future directions of socioscientific reasoning. In C. Linder, L. Ostman, D. A. Roberts, P. Wickman, G. Erickson, & A. MacKinnon (Eds.), *Promoting scientific literacy: Science education research in transaction* (pp. 176–192). New York: Routledge.

Appendix A
Initial Interview Protocol

Thank you for agreeing to be interviewed today. The purpose of this interview is to learn about some of the successes and challenges you are experiencing as a teacher candidate. We are studying science and mathematics teacher education to better support beginning teacher learning. We ask that you try to be as candid and specific as possible.

The information from this interview will not affect your course grades, your teaching placements, or your standing in the Teacher Education Program. If there is a question you do not wish to answer, you can ask that it be skipped. If you later wish to revise an answer or to ask that an answer be deleted, you are free to do so as well.

The interview should last about 60 minutes. It is divided into several parts. Do I have your permission to begin recording the interview?

[Turn on recorder]

Background Information (Initial interview only)

First, I'd like to ask you a few questions about your interest in teaching.

1) What are some reasons why you decided to become a teacher and to teach ________ *[specific credential subject area (known from survey), e.g., biology, chemistry]* in particular?

2) Why did you decide to enroll in the _______ teacher education program?

3) *[if involved in undergrad recruitment program]* How did [undergrad recruitment program] help prepare you for teaching?

4) Where do you hope to teach after completing the program? In what kind of school, or what kinds of students, would you like to teach? Any particular grade levels and/or courses you would like to teach?

Conceptions of Science Teaching

These next few questions are about your ideas about effective science teaching.

5) Thinking back to middle school and high school, please describe a typical science lesson that you experienced as a student.

6) What do you think are the characteristics of an excellent science teacher?

7) What have you learned about effective science instruction from your teacher education program so far?

8) What more would you like to learn or feel you need to learn about effective science instruction?

For the next few questions, imagine that you are teaching a secondary science course, for example, in your student teaching placement.

9) If an observer walked into your classroom, what do you hope the observer would notice about what you are doing as a teacher?

10) What do you hope the observer would notice about the disciplinary core ideas, cross-cutting concepts, and/or science and engineering practices you are teaching?

11) What do you hope the observer would notice about what the students are doing?

12) How would you engage students in discussions?

13) How would you engage students in reading and writing?

14) What kinds of connections would you make between school science and students' lives outside of school?

Science Practices

These next few questions are about the *Next Generation Science Standards* science and engineering practices.

15) In general, what have you learned about the eight science and engineering practices from the *Next Generation Science Standards* in your teacher education program or from your prior experiences?

This is a list with the eight science and engineering practices from the *NGSS* [*at end of document*].

16) Which **two** have you implemented, or seen implemented, most often in your current student teaching placement? What are some examples of how these two practices have been implemented in your placement?

17) Out of all eight, which **one** do you think is most important to teach students? Why?

18) Which **one or two** practices do you think you need more help to understand or implement?

Conceptions of Learners

These next few questions are about students and student learning.

19) How do you think students learn science?

20) Why do you think some students succeed and other students struggle in school science courses?

21) Do you think students should be tracked according to ability in secondary science? What are the advantages and disadvantages of tracking?

Conceptions of Effective Practices for English Language Learners

These next few questions are about science instruction for diverse learners.

22) Classrooms are becoming increasingly culturally and linguistically diverse. How prepared do you feel to teach in a culturally and linguistically diverse classroom?

23) How do you define an English language learner (ELL)?

24) How do you think ELL students differ from one another?

25) What do you think ELL students bring as resources to increase the richness in class?

26) What knowledge and skills do you think it takes to be an effective secondary science teacher of English Language Learners?

For the next few questions, imagine that you are teaching a secondary science class with English language learners as well as native English speakers, for example, in your student teaching placement.

27) What supports for ELLs would you consider as you planned your instruction?

28) What factors would you consider when developing or selecting science texts for ELLs?

29) What would you consider when designing and using science assessment materials for ELLs?

Practicum Experience

These final questions are about your current practicum placement.

30) In what secondary school are you currently placed?

31) In what science class or classes are you currently placed?

32) What are the student demographics of the class or classes, in terms of gender, ethnicity, and ELLs?

33) What kinds of instructional responsibilities have you had so far?

34) How much autonomy do you have with your teaching, for example, in selecting topics and deciding what strategies to implement?

Science and Engineering Practices from NGSS

Asking questions (for science) and defining problems (for engineering)

Developing and using models

Planning and carrying out investigations

Analyzing and interpreting data

Using mathematics and computational thinking

Constructing explanations (for science) and designing solutions (for engineering)

Engaging in argument from evidence

Obtaining, evaluating, and communicating information

Appendix B
Follow-Up Interview Protocol

Thank you for agreeing to be interviewed today. This interview will be similar to the one you did previously.

The interview should last about an hour. It is divided into a few parts. Do I have your permission to audio record the interview?

[Turn on recorder]

Conceptions of Science Teaching

The first few questions are about your ideas about effective science teaching.

1) What do you think are the characteristics of an excellent science teacher?

2) What have you learned about effective science instruction from your teacher education program?

3) What more would you like to learn or feel you need to learn about effective science instruction?

For the next few questions, imagine that you are teaching a secondary science course, for example, in your student teaching placement or when you have your own classroom.

4) If an observer walked into your classroom, what do you hope the observer would notice about what you are doing as a teacher?

5) What do you hope the observer would notice about the disciplinary core ideas, cross-cutting concepts, and/or science and engineering practices you are teaching?

6) What do you hope the observer would notice about what the students are doing?

7) How would you engage students in discussions?

8) How would you engage students in reading and writing?

9) What kinds of connections would you make between school science and students' lives outside of school?

Science Practices

These next few questions are about the *Next Generation Science Standards* science and engineering practices.

10) In general, what have you learned about the eight science and engineering practices?

This is a list with the eight science and engineering practices from the *NGSS* [*at end of document*].

11) Which **two** have you implemented most often in your current student teaching placement? What are some examples of how these two practices have been implemented in your placement?

12) Out of all eight, which **one** do you think is most important to teach students? Why?

13) Which **one or two** practices do you think you need more help to understand or implement?

Conceptions of Learners
These next few questions are about students and student learning.

14) How do you think students learn science?

15) Why do you think some students succeed and other students struggle in school science courses?

16) Do you think students should be tracked according to ability in secondary science? What are the advantages and disadvantages of tracking?

Conceptions of Effective Practices for English Language Learners
These next few questions are about science instruction for diverse learners.

17) Classrooms are becoming increasingly culturally and linguistically diverse. How prepared do you feel to teach in a culturally and linguistically diverse classroom?

18) How do you define an English language learner (ELL)?

19) How do you think ELL students differ from one another?

20) What do you think ELL students bring as resources to increase the richness in class?

21) What knowledge and skills do you think it takes to be an effective secondary science teacher of English Language Learners?

For the next few questions, imagine that you are teaching a secondary science class with English language learners as well as native English speakers, for example, in your student teaching placement.

22) What supports for ELLs would you consider as you planned your instruction?

23) What factors would you consider when developing or selecting science texts for ELLs?

24) What would you consider when designing and using science assessment materials for ELLs?

Practicum/Student Teaching Experience
These questions are about your current practicum or student teaching placement.

25) In what secondary school are you currently placed?

26) In what science class or classes are you currently placed?

27) What are the student demographics of that class, in terms of gender, ethnicity, and ELLs?

28) In your placement, how aligned do you feel your teaching is to the *Next Generation Science Standards*?

29) How much support do you feel you get to teach in ways that are aligned with the *NGSS*?

EdTPA
This final set of questions is about your edTPA teaching event.

30) For your edTPA, what was the central focus of your lesson sequence?

31) How did you address the *NGSS*?

32) How did you support ELLs?

In your edTPA lesson sequence, in what ways did you…
33) Engage students in scientific sense-making?

34) Engage students in scientific discourse?

35) Support students' English language and literacy development?

36) Make connections between lesson activities and students' lives outside of school?

37) What kinds of support did you receive in completing your edTPA?

38) What additional support would you have liked?

Thank you!

Science and Engineering Practices from NGSS

Asking questions (for science) and defining problems (for engineering)

Developing and using models

Planning and carrying out investigations

Analyzing and interpreting data

Using mathematics and computational thinking

Constructing explanations (for science) and designing solutions (for engineering)

Engaging in argument from evidence

Obtaining, evaluating, and communicating information

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 2, 57-80

ICRSME
International Consortium for Research in
Science & Mathematics Education

Primary School Teachers' Noticing Skills Regarding Students' Thinking: The Case of Whole Number Subtraction

Reyhan Tekin Sitrava
Kırıkkale University

Mine Işıksal Bostan
Middle East Technical University

Seçil Yemen Karpuzcu
Kutahya Dumlupinar University

ABSTRACT

This qualitative case study investigates how, and to what extent, primary school teachers notice students' mathematical thinking in the context of whole number subtraction. A task involving a student's invented strategy was used to collect data. Three noticing questions connected to the task were asked to 45 teachers. Their written answers were analyzed to reveal teachers' noticing skills based on attending, interpreting, and deciding how to respond. The study results revealed that most participants provided limited evidence of attending and interpreting skills, and some responses showed a lack of evidence. More specifically, they could not identify the relationship between digits in ones and tens places. Further, their interpretations did not directly focus on students' solutions and consisted of general statements and misconceptions.

Keywords: teachers' noticing, professional noticing of children's mathematical thinking, students' invented strategies, whole number subtraction, primary school teachers

Introduction

The classroom is a complex environment where many situations are experienced simultaneously; however, teachers cannot focus on all situations in this environment (Sherin et al., 2011). Teachers' knowledge and beliefs about the students, the content of the lesson, the curriculum, and teaching and learning affect everything that needs to be considered during teaching (van Es, 2011). Among these constructs, teachers' knowledge and beliefs about students is a keystone while building instruction, and it has an essential role in students' learning (Darling-Hammond & Ducommun, 2010; Hoth et al., 2018). Based on the knowledge of students, the teachers plan the activities, the problems, and the teaching strategies before the lesson. However, they should be aware of all the situations in the classroom and emphasize if any situations will support students' learning during the lesson (Stockero et al.,2017). In order to be aware of the instant events that emerge during the lesson and to be able to make in-the-moment instructional decisions, the teachers need to have noticing skills which are the essential components of teaching expertise (Jacobs et al., 2011; Sherin et al., 2011; Sherin & Star, 2011).

A better understanding of noticing allows further development in mathematics teaching and learning (Amador et al., 2021a; Jacobs et al., 2011). Therefore, the need to understand teacher noticing

© 2023 International Consortium for Research in Science & Mathematics Education (ICRSME)

has emerged, and the interest of researchers in teacher noticing has increased abruptly in recent years (Stahnke et al., 2016). In order to meet this need partially, in this research study, we aimed to explore primary school teachers' noticing skills of students' mathematical thinking in the context of whole number subtraction.

Teacher Noticing

As a professional vision, teacher noticing is one of the critical components of teaching. However, it can be challenging because it requires one to perceive multiple aspects of the classroom environment (van Es & Sherin, 2021). Generally, teacher noticing includes attending and making sense of the specific events in a classroom and transforming the teachers' attention and interpretation in a teaching setting (Jacobs & Spangler, 2017). More specifically, teachers need to choose the events they will focus on and determine their focus duration to manage "a blooming, buzzing confusion of classroom events" (Sherin & Star, 2011, p. 73). Then, they need to interpret what they see and make connections between the observed events and related issues of the events (Sherin et al., 2011). Processing all these issues during mathematics teaching is not easy for teachers (Estapa et al., 2017).

Although researchers have different ideas about the components of teacher noticing and how to measure and develop it, they agree that identifying the noteworthy events and making sense of them are two essential features of teacher noticing (Sherin et al., 2011). For instance, van Es and Sherin (2008) claimed that mathematical thinking, pedagogy, classroom environment, and classroom management are the events that teachers should notice. Based on their claims, van Es and Sherin developed a framework, *Learning to Notice*, which has two main dimensions, each with four levels (baseline, mixed, focused, extended): what teachers notice and how teachers notice. The *what teachers notice* dimension of the framework includes noticing the classroom environment, students' behaviors and learning, teacher pedagogy, particular students' mathematical thinking, and the relationship between particular teaching strategies and resultant students' mathematical thinking. On the other hand, the *how teachers notice* dimension consists of providing comments and making connections between events and principles of teaching and learning.

In their subsequent work, van Es and Sherin (2021) introduced a new framework based on noticing as an active process. This considers the active interaction with the environment by enabling more observation and interpretation. These authors expanded the dimensions of attending and interpreting, which they put forward on their original Learning to Notice framework (van Es & Sherin, 2002) and included one dimension called shaping. Van Es and Sherin (2021) grounded shaping on the interaction between teacher and students instantly within a classroom environment. This interaction aims to get additional information related to students thinking, which serves for attending and interpreting, and curriculum materials.

Consequently, van Es and Sherin (2021) based their revised Learning to Notice framework on three dimensions, attending, interpreting, and shaping. Although another most cited framework for noticing, called Professional Noticing of Children's Mathematical Thinking, involves attending and interpreting dimensions (Jacobs et al., 2010), both frameworks discussed these dimensions from different points of view. More specifically, van Es and Sherin (2002; 2021) focused on attending and interpreting the noteworthy events that occurred in classroom settings, but Jacobs et al. took students' understanding into consideration while defining attending and interpreting dimensions. In addition, as a third dimension, Jacobs et al. (2010) presented deciding how to respond, meaning teachers' next instructional move to extend and support a student's thinking. To put it differently, in their framework called Professional Noticing of Children's Mathematical Thinking, Jacobs et al. focused more on students' thinking and defined teacher noticing as the ability to attend to the mathematical details in students' strategies, interpret students' mathematical understanding of the particular subject reflected in their strategies, and make decisions to support and improve student's learning based on their understandings. Since this study aims to investigate the extent to which primary school teachers notice students' mathematical thinking in the context of whole number subtraction, the Professional Noticing of Children's Mathematical Thinking framework served as a theoretical framework for the study.

Professional Noticing of Children's Mathematical Thinking

Within the scope of Professional Noticing of Children's Mathematical Thinking theory, Jacobs et al. (2010) identified a particular focus among the levels of the Learning to Notice framework, which is an extended level, and selected a particular slice of teaching which is teachers' in-the-moment decisions while they are responding to students' strategies. Jacobs et al. emphasized that the reasons for selecting a particular focus for noticing are attending more to how and to what extent teachers notice students' mathematical thinking rather than attending to the variety of what teachers notice. From this point of view, these authors attached particular importance to a group of teachers' expertise, with a specialized type of noticing, on which the theoretical framework of this study is based. It was built as a set of three interrelated skills: attending to children's strategies, interpreting children's understandings, and deciding how to respond based on children's understandings (Jacobs et al., 2010).

The first dimension describes the teacher's explanations related to how a student approaches the mathematical situation/activity/ problem, how he solves it, what materials and strategies he uses, and what the details of his strategies are (Jacobs et al., 2010). The second dimension is defined as a teacher's reasoning consistent with the mathematical details specific to a particular student's strategy and the research on students' mathematical development rather than revealing a holistic picture of students' mathematical understanding (Jacobs et al., 2010). Therefore, it is the ability of the teacher to interpret mathematically how a student understands the subject and how consistent this knowledge is with the knowledge of students. In addition, Jacobs et al. distinguished this interpretation from superficial evaluations, as many researchers did (van Es & Sherin, 2008). The last dimension deals with to what extent teachers use the knowledge they have learned from students' understanding in a given situation while deciding what to do with their next moves. Jacobs et al. (2010) stated that to decide how to respond based on students' understanding, the teacher has also attended to students' strategies and interpreted their understanding. This means that these three skills are interrelated with each other.

Students' Invented Strategies in Whole Number Subtraction

The National Council of Teachers of Mathematics (NCTM; 2014) emphasized that analyzing students' thinking is an important tool for teachers to make instructional decisions for improving students' learning. However, understanding students' thinking is challenging for teachers, and they need to spend substantial time and effort analyzing and interpreting students' thoughts (NCTM, 2014; Son, 2016). Regarding this, Franke et al. (2007) stated that one of the ways of understanding students' thinking and reaching their minds is to understand their invented strategies. Invented strategies are defined as different from standard algorithms and do not require using materials (Carpenter et al., 1998). Students start school with a great deal of knowledge about the concepts, and they can construct strategies for solving mathematics problems, especially for adding and subtracting (Carpenter & Fennema, 1992). These strategies play a prominent role in helping students develop number sense and learn multi-digit operations (Carpenter et al., 1994). Furthermore, it is claimed that inventing strategies requires making sense of mathematics (Carpenter et al., 1998) since these are flexible methods that change according to numbers and circumstances (Van de Walle et al., 2013). In order to connect students' strategies and standard algorithms, teachers need to attend to and interpret students' invented strategies (Carpenter et al., 1998).

Number-based invented strategies are categorized under three labels: decomposition, sequential, and varying strategies (Verschaffel et al., 2007). In decomposition strategies or partial differences strategies (Son, 2016) that involve decomposing the minuend and the subtrahend based on the place value of each number, the decomposed numbers (tens and ones) are subtracted separately. For instance, for the question, 57 – 28 = ?; the solution is 50 – 20 = 30, 7 – 8 = -1, so the answer is 30 + (-1) = 29.

The current study uses a subtraction operation involving a decomposition strategy to analyze primary school teachers' noticing skills of children's mathematical thinking. In the literature, this strategy is the most common strategy invented by students in subtraction problems (Carpenter et al., 1998). The decomposition strategy was also used in studies, which aim to understand pre-service teachers' interpretations (Son, 2016), or to compare pre-service teachers' and practicing teachers' content knowledge (Philipp et al., 2008). Furthermore, this strategy finds the answer by subtracting tens and ones separately without regrouping them. This requires students to conceptualize negative numbers or debt that will be subtracted later, which highlights conceptual understanding. Thus, in this study, by giving a solution involving the decomposition strategy to teachers, we had an opportunity to investigate the teachers' deeper mathematical understanding regarding subtraction with whole numbers through various connections within number relationships.

Rationale of the Study

In recent years, the construct of teacher noticing has gained significant importance in mathematics education research areas. Studies on noticing show variations in the dimensions of noticing (Amador et al., 2016; Blömeke et al., 2015; Jacobs et al., 2010; van Es & Sherin, 2021), the interventions used to measure or analyze teacher noticing skills (Choy, 2013; Sherin & van Es, 2009; Stockero, 2014), and the critical issues that are significant to notice (Jacobs et al., 2010; Star & Strickland, 2008; van Es & Sherin, 2008; 2021). Some of those studies analyzed teacher noticing using video clubs (Amador et al., 2021b; Girit-Yildiz et al., 2023; González &Vargas, 2020; Ivars et al., 2020; Ulusoy & Cakiroglu, 2021; Warshauer et al., 2021), others used students' written work or verbal responses (Dogan-Coskun et al., 2021; Jacobs et al.,2010; Roller, 2016; Sánchez-Matamoros et al., 2019; Tekin-Sitrava et al., 2021). In addition, some studies focused on teachers noticing the events in classroom environments (Stahnke & Blömeke, 2021; van Driel et al., 2021); however, others attended to teachers' noticing students' mathematical thinking (Teuscher et al., 2017).

When we turn our attention to participants, the studies conducted with pre-service teachers showed that pre-service teachers generally focused on the general aspects of the classroom, such as teacher actions (Santagata et al., 2007), management and student-teacher interaction (Star & Strickland, 2008), rather than focusing on students' thinking (Jacobs et al., 2010). This is not surprising since they do not have teaching experience, which is one of the variables that influence teacher noticing (Sherin et al., 2011) and provides support for attending to and interpreting children's understandings (Jacobs et al., 2010; Schoenfeld, 2011). Therefore, it would be hard to get an in-depth exploration of pre-service teachers' knowledge and interpretation regarding students' thinking. On the other hand, in-service teachers have more knowledge and skills to make students understand the concepts within the complex classroom environment in which student learning occurs (Miller, 2011). From this point of view, the participants, who can make instant decisions to reveal students' understanding by observing the complex and dynamic structure of the classroom environment, would provide more comprehensive and in-depth information about teachers' ability to notice. Additionally, working with experienced teachers may allow us to enhance the theoretical framework by articulating the extent to which in-service teachers attend to the details of students' strategies, how they interpret students' understanding as students reflected in their strategies, and how to respond to the basis of students' understanding. Therefore, it would be significant to study with in-service teachers to get more in-depth data on teacher noticing skills and a more categorized framework.

The focus of this study, whole number subtraction, is one of the important topics in primary school mathematics curriculum (Ministry of National Education [MoNE], 2018), and it has a crucial role in teaching and learning mathematics conceptually (Van de Walle et al., 2013). Therefore, teachers need to have the ability to attend to, interpret, and respond to students' understanding to empower them mathematically (Thanheiser, 2009). Although various studies aim to investigate in-service and pre-service teacher knowledge of whole number subtraction operation and students' strategies for whole number subtraction, there is a gap in the literature on teacher noticing of students' strategies for subtracting whole numbers. In other words, although teachers' and students' understanding of whole number subtraction was examined in the previous studies (e.g., Roy, 2014; Thanheiser, 2009), there are limited studies related to teachers' attending to students' strategies for whole number subtraction, their interpreting of students' understanding based on students' strategies, and the decisions on how to respond based on students' understandings (e.g., Son, 2016; Yeo & Webel, 2019). However, most of the studies focused on examining the noticing skills of prospective teachers who do not have any teaching experience. Therefore, it would be significant to examine the noticing skills of in-service teachers who have experience with students' invented strategies and in monitoring students' behaviors, learning, and understanding (Miller, 2011). Thus, we seek answers to the following research questions in the present study:

1. To what extent do primary school teachers attend to students' strategies in the context of whole number subtraction?
2. To what extent do primary school teachers interpret students' mathematical thinking in the context of whole number subtraction?
3. What is the nature of teachers' decisions to respond to students' mathematical thinking in the context of whole number subtraction?

Method

Design of the Study

Since this study aimed to investigate primary school teachers' professional noticing skills in the context of whole number subtraction, a qualitative case study method was used to reveal the findings and support the study's methodological perspective. The study focuses on one group of in-service, primary school teachers; thus, the study is a single case study. Moreover, since the aim is to investigate in-service primary school teachers' noticing skills in the context of only one mathematics subject, whole number subtraction, it includes only one unit of analysis. Thus, the study design is a single-case holistic design (Yin, 2009).

Context and Participants

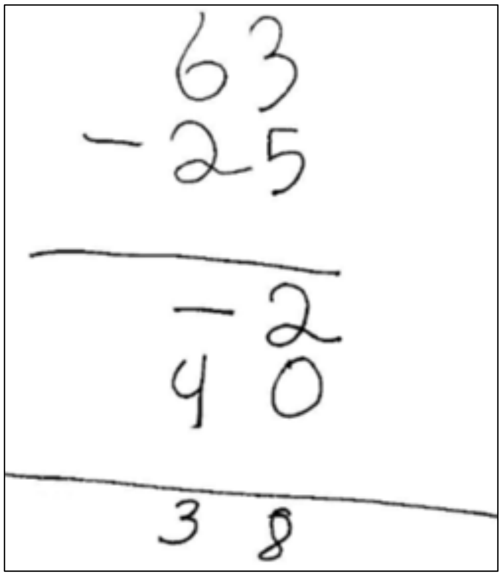
Turkey has a centralized national education system in which all public/private primary schools follow a primary school mathematics curriculum prepared by the MoNE (2018). In addition, primary school teachers graduated from four-year college programs from departments of primary education in Faculties of Education after the 1992-1993 academic year (Dursunoğlu, 2003). The participants in this study included 45 in-service primary school teachers. Three participants were male, and 42 were female, with teaching experiences ranging between five to 40 years. While most of them graduated from the faculty of education, participants became teachers by graduating from different sources, such as a bachelor's degree from any faculty, a two-year college, and a teacher high school. Of the 45 participants, ten worked in a public school, and the rest worked in private schools when the data was collected.

Student Invented Task

This study used a task involving a student-invented strategy adapted from the study of Philipp et al. (2008) to collect data. The scenario in this task also appeared in various sources (Campbell et al., 1998; Schifter et al., 1999; Son, 2016). The task involves a student's written response to a whole number subtraction problem and three further professional noticing questions in accordance with that response.

Figure 1

Mert's Solution to the Task



Note. In the scenario, a second-grade student, Mert, solves the problem of 63 – 25 = X correctly using an invented strategy, which is regarded as a decomposition strategy or partial differences strategy in whole number subtraction (Son, 2016).

For the study, three noticing questions connected to the scenario were developed from research on professional noticing of children's mathematical thinking (Jacobs et al., 2010). Each question corresponds to component skills of attending, interpreting, and deciding how to respond. The professional noticing questions are as follows:

1. Evaluate how Mert solved this problem and explain whether this method is appropriate for whole number subtraction in detail.
2. Explain what you learned from Mert's solution method about Mert's understanding on subtraction operation in detail.
3. Pretend that you are the teacher of Mert. What problem or problems would you pose next? Explain your rationale for posing that problem(s).

The first question was developed to identify the mathematically significant details of Mert's solution and determine whether it was correct. The aim of asking the second question was to assess teachers' professional skills in terms of interpreting children's understandings. Finally, in the third question, participants were asked to explain how they selected a further problem or problems to respond to the student.

Data Analysis

Data were analyzed qualitatively through thematic analysis regarding repeating coding and themes (Miles & Huberman, 1994). Additionally, the frequencies of categories were also described to detect patterns in themes, which are categories under the dimensions of the noticing. More specifically, to attain this study's goals, primary school teachers' written responses to the student's invented task questionnaire were analyzed according to the dimensions of the Professional Noticing of Children's Mathematical Thinking framework developed by Jacobs et al. (2010). While analyzing the first two dimensions, which are attending to children's strategies and interpreting children's understanding, teachers' explanations related to student approaches to the mathematical task and details of their strategy did not quite match the categories of Professional Noticing of Children's Mathematical Thinking framework. In other words, some responses were too detailed to be considered in the limited category, and too superficial to be considered in the robust category. Thus, more detailed categorization was needed to code the teachers' attending and interpreting skills. Accordingly, attending and interpreting skills were coded as Lack, Limited, Substantial, and Robust Evidence, which were presented by Tekin-Sitrava et al. (2021). However, besides these categories, it was suitable to code some teachers' interpretations as No Response, Wrong Interpretation, No Evidence of Interpretation, and Just Attention. About the deciding how to respond dimension, we did not categorize teachers' responses with respect to the level of evidence as indicated by Jacobs et al. (2010) and Tekin-Sitrava et al. (2021). Rather, we focused on the nature of the responses, so deciding how to respond skills of in-service teachers were investigated under the following categories: *Unrelated and General, Ignorance, Acknowledging,* and *Responding to child and incorporating***.** Details of each category are illustrated in the findings. After analyzing teachers' responses based on the Professional Noticing of Children's Mathematical Thinking framework dimensions, frequency analysis was performed to determine the number of teachers falling into categories in each dimension. Three mathematics educators (authors) analyzed data and discussed the inconsistencies until the coders reached 100% consensus.

Findings

Based on the research questions, the findings of the study are presented in three sections. The extent to which teachers attend and interpret is explained in the sections *Attending to Children's Strategies* and *Interpreting Children's Mathematical Understandings,* and what kind of decisions teachers make are explained in the section *Deciding How to Respond on the Basis of Children's Understandings.*

Attending to Children's Strategies

Attending to children's strategies is defined as teachers' explanations of students' approaches to the mathematical situation/activity/problem, students' usage of materials, students' strategies to solve the problem, and the details of these strategies (Jacobs et al., 2010). Based on the data analysis, the evidence of attention was categorized under four headings: *robust evidence of attention, substantial evidence of attention, limited evidence of attention,* and *lack of evidence of attention.* The details of each category and the frequency of responses for each category are given in Table 1. Then, evidence from teachers' responses is explained respectively.

Table 1

The Details of Attending to Children's Strategies Dimension and the Frequency of Each Category

Attending	Frequency
Lack of Evidence of Attention to Children's Strategies	
Identifying the solution/mathematical concepts correctly, but independent from tne students' answer	15 (33.33 %)
Identifying the solution correctly, but the mathematical concepts are missing, or the explanation includes a general statement	
Identifying the solution as incorrect	
Limited Evidence of Attention to Children's Strategies	
Correctly identifying the solution as true, but there is some naïve conceptions and misconceptions while describing the students' subtraction operation	13 (28.89%)
Correctly identifying the solution as true by using alternative numbers/alternative solution strategy without referring to the present situation	
Substantial Evidence of Attention to Children's Strategies	
Correctly identifying the solution as true and the subtraction between the numbers in the ones and tens place, but the connection between the numbers in the ones and tens place is missing	16 (35.56%)
Attention to solution is reasonable but not appropriate for that grade level (second grade)	
Robust Evidence of Attention to Children's Strategies	
Correct attention to solution through identifying the relationship between numbers in the ones and tens place.	1 (2.22%)

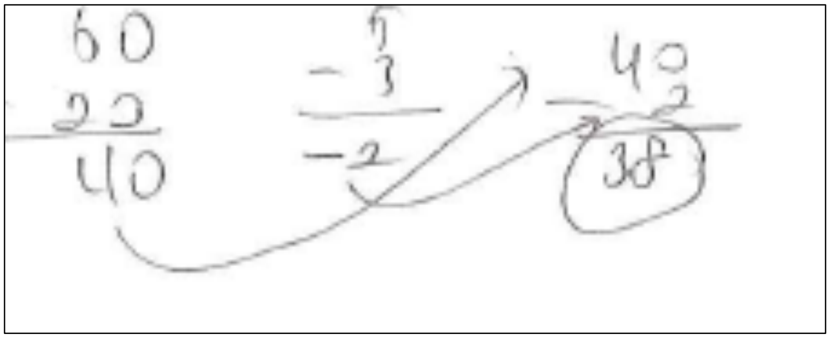
Robust Evidence of Attention to Children's Strategies

Data gathered from primary school teachers showed that only one of them could identify the solution as "true" by subtracting the ones and tens separately, and then establishing the relationship between results obtained from these subtractions. To accompany their work shown in Figure 2, Teacher 16 stated:

> The solution is correct. First, he found the difference between tens. Then, he calculated the difference between ones. Lastly, in order to find an answer, he subtracted this excessive amount (2) and found 38.

Figure 2

Teacher 16's Work



Teacher 16 identified the relationship between numbers in ones and tens place and explained the minus sign in front of 2 as an excessive amount which is subtracted from the difference between tens. Therefore, his attention to the children's strategy was regarded as *robust evidence of attention to children's strategies*.

Substantial Evidence of Attention to Children's Strategies

Most of the teachers (35.56%) correctly identified the subtraction operation by taking into account place value concepts and subtracted the ones from ones and tens from tens. However, the relationship between the numbers in the ones and tens place was missing. Also, the interpretation of '2' in the operation was neglected. Those responses were categorized as *substantial evidence of attention to children's strategies*. For instance, one of the teacher's expressions is as follows:

> He subtracted 3 (ones place of minuend) from the 5 (ones place of subtrahend) and wrote the result to the ones place. Then, he subtracted 2 (tens place of subtrahend) from the 6 (tens place of minuend) and wrote the result to the tens place. The answer is true. This is another way of thinking when you should subtract the small number from the larger number (Teacher 33).

As it could be realized from the above script, Teacher 33 did not relate the numbers obtained from the subtraction operation in the ones and tens place. In addition to this, some teachers tried to mention the role of the 2 in the subtraction operation; however, the connection between the numbers in the ones and tens place is still omitted. Why we subtract 2 from 40 or what -2 stands for is not obvious in the teachers' responses. For instance,

> He correctly solved the question, he subtracted ones from ones and found -2. Then, he subtracted tens from tens and found 40. Then, he calculated the difference of these two numbers (Teacher 19).

As could be deduced from the above response, the teacher correctly attended the subtraction operation by taking the difference between ones and tens separately. However, she did not interpret the (-2) and just explained the result by only taking the difference between ones and tens.

In this category, analysis of findings revealed another important issue regarding the student's solution. Some teachers' attention to the solution could be accepted as reasonable, but inappropriate for second grade. In other words, the concept of integer was not an appropriate explanation for the second-grade student. For instance,

> He subtracted 2 tens from 6 tens. Then, he thought about integers, and subtracted 5 from 3 and found -2. He subtracted the difference of ones from the difference of tens and found the answer (Teacher 1).

As can be seen from this example, Teacher 1 considered 2 as an integer. But, as a second grade student, Mert did not learn the subtraction operation of integers.

Limited Evidence of Attention to Children's Strategies

The analysis of in-service teachers' noticing skills revealed that 13 (28.89%) teachers' responses can be categorized as *limited evidence of attention to children's strategies*. Compared to the substantial evidence, the responses under this category consist of correctly identifying the student solution as true, but naïve conceptions regarding the subtraction operation. To exemplify,

> He conducted the operation mentally. He subtracted the tens and ones separately. Then, since he wrote the subtraction results in reverse order, he further subtracted two from 40 (Teacher 14).

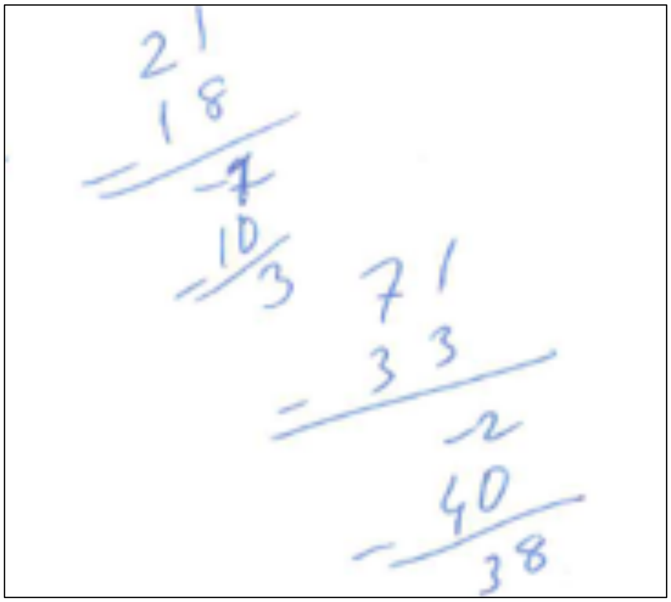
As could be deduced from the above response, the teacher could not interpret the difference between "-2" and "40". Furthermore, the teacher had a naïve conception that since the student wrote the operation in reverse order, he further conducted subtraction.

In some cases, the teachers used another number to check whether the solution was proper or not. In other words, the teachers correctly identified the solution as true by using alternative numbers without referring to the present situation. The teachers confirmed the operation by using alternative numbers. For instance, Teacher 12's explanation, and work shown in Figure 3, are as follows:

> He subtracted the ones and then tens from each other. Then he found the difference between these two results. The solution is correct when we tried the other number. Nice work! (Teacher 12).

Figure 3

Teacher 12's Work

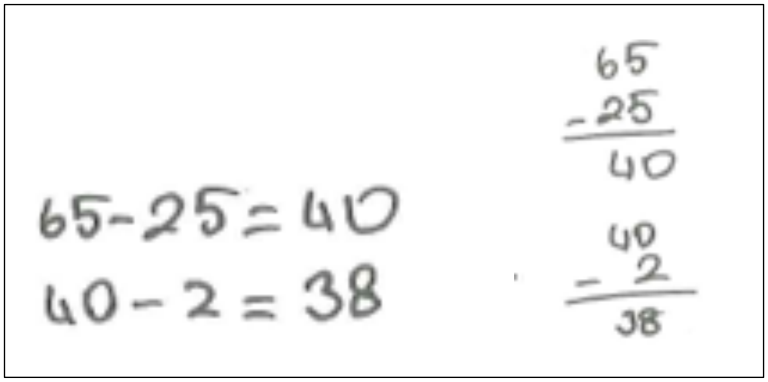


Identifying the solution without referring to the student's solution was popular among teachers' responses. In another example, another teacher identified the operation correctly, but he checked the correctness of the solution by using another solution strategy. Teacher 17's explanation and work in Figure 4 are as follows:

> I think the solution is true (Teacher 17).

Figure 4

Teacher 17's Work



65-25=40
40-2=38

$$\begin{array}{r} 65 \\ -25 \\ \hline 40 \end{array} \qquad \begin{array}{r} 40 \\ -2 \\ \hline 38 \end{array}$$

Some of these teachers applied Mert's solution strategy using alternative numbers without referring to Mert's solution, and some presented an alternative solution. Thus, those responses are categorized as *limited evidence of attention to children's strategies.*

Lack of Evidence of Attention to Children's Strategies

Fifteen teachers' (33.33%) responses did not provide strong evidence of attention to the student's strategies. Indeed, in this category of response, teachers identified the solution/mathematical concepts correctly, but independently from the student's answer, or they used general strategies and ignored the details in the student's solution. Examples from teachers' answers are as follows:

> He conducted the operation in reverse order. The answer is correct. He just used the logic of regrouping that yields the correct solution (Teacher 31).
>
> He just subtracted the ones and tens from each other. Based on the subtraction operation rule (the way of writing) the operation is wrong, but it is true conceptually (Teacher 13).

As could be deduced from the responses, the teachers used general ideas and did not refer to the student's solution. In addition to the categorization of attention, one of the teachers directly rejected the student's solution and evaluated it as wrong. Teacher 37 stated that

> He solved the problem by using the subtraction operation. He most probably learned this methodology from his family. The technique is not correct. It is even an inhibitor for the following years (Teacher 37).

To summarize, more than one-third of the primary school teachers provided *lack of evidence of attention to children's strategy*. Except for one teacher, the other teachers could not identify the relationship between numbers in the ones and tens place, even though it is a vital issue to the subtraction operation.

Interpreting Children's Mathematical Understandings

Similar to the attention dimension, the evidence of interpretation of children's mathematical understanding was also categorized under five headings: *robust evidence of interpretation of children's understandings, substantial evidence of interpretation of children's understandings, limited evidence of interpretation of* *children's understandings, lack of evidence of interpretation of children's understanding,* and *no evidence of interpretation of children's understanding.* The details of each category and the frequencies of responses for each category are given in Table 2. Then, evidence from teachers' verbatim is given.

Table 2

The Details of Interpreting Children's Mathematical Understanding Dimension and the Frequency of Each Category

Interpretation	Frequency
No response	2 (4.44%)
Wrong Interpretation	3 (6.67%)
No evidence of Interpretation, Just Attention	8 (17.78%)
Lack of Evidence of Interpretation of Children's Understanding	
Interpreting the solution/usage of mathematical concepts correctly, but independently from student answer	18 (40%)
General statement	
Limited Evidence of Interpretation of Children's Understandings	
L1: Consisting of general statement on ones and tens interpretation	7 (15.56%)
L2: Correctly identifying the solution as true, but there is a misconception while interpreting the subtraction operation	5 (11.11%)
Substantial Evidence of Interpretation of Children's Understandings	
Correctly identify the solution as true, correctly identifying the subtraction between the numbers in the ones and tens place, but the interpretation of the connection between the numbers in the ones and tens place is missing	2 (4.44%)
Interpretation of solution is reasonable but not appropriate for that grade level (second grade)	
Robust Evidence of Interpretation of Children's Understandings	
A correct interpretation of the solution through identifying the relationship between numbers in the ones and tens place	0

Substantial Evidence of Interpretation of Children's Understandings

Findings revealed that there was no response regarding the robust evidence of interpretation of children's understandings. Only two (4.44%) teachers' responses were categorized under the heading of *substantial evidence of interpretation of children's understandings.* In this category, teachers correctly identified the solution as true and correctly interpreted the subtraction between the numbers in the ones and tens place. However, the interpretation of the connection between the numbers in the ones and tens place is missing. For instance,

> Mert conceptualized the subtraction operation. While finding the difference between ones and tens, he used natural numbers and instead of trading tens into ones, he treated each number as an integer. Indeed, this is the written way of the mental strategies that we use while calculating (Teacher 1).

As could be understood from the above script, the teacher interpreted the ones and tens correctly, but the interpretation of the relationship between the numbers in the ones and tens place is missing.

Limited Evidence of Interpretation of Children's Understandings

Seven teachers' (15.56%) responses consisted of general statements regarding ones and tens interpretation without referring to the student solution (L1). Teacher 12's response could be categorized under this heading.

> He definitely understood the subtraction operation. He separated the ones and tens and performed subtraction separately. Thus, I believe that he understands the operation (Teacher 12).

In some cases (11.11%), teachers correctly identified the solution as true, but had some naïve conceptions while interpreting the subtraction operation (L2). For instance,

> For each digit, he subtracted the smaller number from the larger number. Thus, for the ones digit, he found his own way to subtract the bigger number from the smaller number (Teacher 35).

As could be understood from the above script, Teacher 35's interpretations included some naïve conceptions like smaller numbers should be subtracted from the bigger number while subtracting.

Lack of Evidence of Interpretation of Children's Strategies

In this category of responses, 18 teachers (40%) interpreted the student solution in a general manner without referring to the ones and tens place. For instance,

> I think this is a really creative solution. This solution is invented by a student, and it is really reasonable (Teacher 20).

As could be understood from the above script, the teacher made general interpretations about the subtraction operation but did not discuss in detail the student's solution. For this reason, similar responses were categorized under the heading of *lack of evidence of interpretation of children's understandings.*

No Evidence of Interpretation, Just Attention

Data analysis revealed some teachers' (17.78 %) responses were directly related to the solution and involved no interpretation. In other words, these responses are similar to those in the attention dimension and directly focus on the students' solution. The following quotation illustrates this approach:

> He did it correctly, he subtracted ones and tens separately. After the subtraction operation, he combined the result (Teacher 3).

To sum up, more than half of the primary school teachers interpreted Mert's solution as true although most interpretations were not directly related to Mert's solution and consisted of general statements and misconceptions.

Deciding How to Respond on the Basis of Children's Understandings

Deciding how to respond on the basis of children's understanding dimension was categorized under four headings: *responding to child and incorporating, acknowledging, ignorance,* and *unrelated and general.* The details of each category and the frequencies of responses for each category are given in Table 3.

Table 3

The Details of Deciding How to Respond Dimension and the Frequency of Each Category

Deciding	Frequency
No response	5 (11.11%)
Unrelated and General	
L0: Misconception about knowledge of teaching subtraction, unrelated response	1 (2.22%)
L1: General pedagogy, including some mathematical concepts, models, representation	17(37.78%)
Ignorance	
L0: Ignorance of students thinking and presentation of unrelated evidence from curriculum	3 (6.67%)
L1: Ignorance of students thinking and reference to traditional algorithm	6 (13.33%)
Acknowledging	
L0: Asking the student to explain her/his strategy, trying to understand student strategy	5 (11.11%)
L1: Performing the same operation by using different numbers with/without giving any rationale.	7 (15.56%)
Responding to child and incorporating	
L0: Incorporating further understanding unrelated to the students' strategy/thinking	0
L1: Incorporating further understanding (e.g., to make some generalization) without rationale	1 (2.22%)
L2: Incorporating further understanding (e.g., to make some generalization) with rationale	0

Responding to Child and Incorporating

Only one teacher (2.22%), Teacher 36, incorporated further understanding regarding student's solution. This teacher's work in Figure 5 and explanation is as follows. She stated that

> I will ask him the following questions. Since I want him to generate a solution strategy for this kind of questions (Teacher 36).

Figure 5

Teacher 36's Work

□ - 23 = 38 ve 63 - □ = 38

As can be seen in the script, she tried to make some generalizations about the subtraction operation by asking questions where the minuend and subtrahend were not given.

Acknowledging

The acknowledging category of deciding how to respond on the basis of children's understanding is divided into two categories. In the first category (L0), five teachers (11.11%) stated that they asked the student to explain his/her strategy because they wanted to understand the student's solution. For instance

> I asked him why he performed the operation like this. By this way, I tried to learn the justification of his solution (Teacher 21).

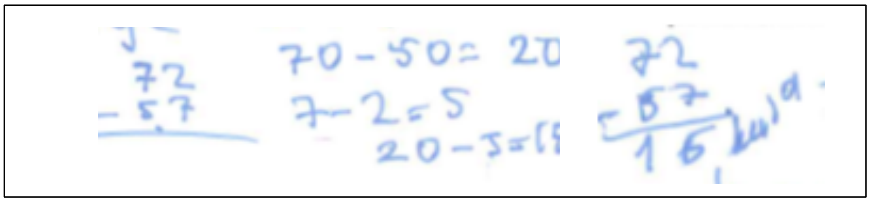
The teachers in this category did not aim to support or extend student's understanding. Instead, they wanted to understand the student's reasoning behind performing this kind of operation.

Teachers' responses under the second category (L1) were differentiated based on the provision of rationale. Five teachers (11.11%) stated that they performed the same operation using different numbers without giving any rationale. Teacher 10, who is in this group, stated the following and their work is shown in Figure 6:

> We have 72 cases of lemon, and we sold 57 of them. How many cases of lemon do we have at the end? (Teacher 10)

Figure 6

Teacher 10's Work

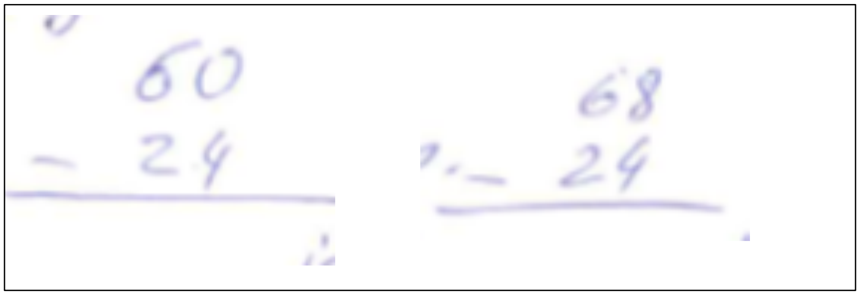


Apart from five teachers, two teachers (4.45%) stated that they performed the same operation using different numbers and giving further rationale. They reported as follows and their work shown in Figure 7:

> I asked him if we could operate more easily. Then, I asked the following questions because I tried to understand whether he had chosen the above method since he did not know the trading of tens into ones (Teacher 11).

Figure 7

Teacher 11's Work



Although Teacher 10 and four other teachers asked for the same operation with different numbers within word problems, they did not state any rationale related to changing the numbers and asking the operation in the form of a word problem. However, only two teachers explained the reasoning behind asking subtraction operations with different numbers.

Ignorance

Teachers' responses that are categorized under this heading are divided into two categories. More specifically, in the first category (L0), three teachers (6.67%) ignored student's thinking and presented evidence from the curriculum that was not directly related to the second-grade curriculum. For instance, one said:

> I asked him a similar question. Then, I asked if he can conduct the same operation by using three digit numbers(Teacher 30).

Six teachers (13.33%) ignored student's thinking and revisited the traditional algorithm (L1):

> Firstly, I asked Mert a subtraction operation involving one-digit numbers. Then, I asked a subtraction operation that did not call for trading tens into ones. Then, by using these operations to help them conceptualize the subtraction operation, I taught subtraction that requires trading of ten into ones (Teacher 19).

As it can be realized from these examples, Teacher 30 stated that he would ask a subtraction operation using three-digit numbers. However, second grade students perform subtraction operation with numbers up to 100 according to the mathematics curriculum (MoNE, 2015). Moreover, Teacher 19 did not attend to Mert's solution while deciding how to respond to him. Instead, she focused on teaching traditional algorithms. The teachers' explanations in this category were similar to those presented previously.

Unrelated and General

Data analysis revealed that 18 (40%) teachers responded to children without considering their solution strategy. Two categories emerged from these responses. Although the teachers did not consider the student's solution strategies in both categories, there is a discrepancy between them regarding the provision of mathematical concepts, terminology, and materials. In the first category (L0), one teacher (2.22 %) responded unrelatedly to Mert's solution. Her explanation is presented below.

> I asked the logic of the operation performed. I wanted him to share his solution with his friends. By this way, students can practically see that they can reach the solution through alternative methods (Teacher 3).

In the second category, 17 (37.78%) teachers expressed some mathematical ideas while responding to students, but they were too general and irrelevant to the student's solution. For instance,

> First of all, I asked Mert to explain his solution. I told him to write another problem. Then, I asked him to solve the written problem by using this method. As the teacher, I asked problems regarding the subtraction operation and asked him to solve the given problems by using different methods and compare the results (Teacher 8).

Of the 45 teachers, five (11.11%) could not provide any response to students regarding Mert's understandings.

To summarize, while responding to students, more than half of the primary school teachers disregarded Mert's solution. Instead, they came up with unrelated and general responses or revisited the traditional algorithm. On the other hand, a few teachers focused on Mert's solution strategy and aimed to get further understanding and make a generalization by performing the subtraction operation with different numbers.

Discussion and Conclusion

This study intended to investigate primary school teachers' noticing skills of students' mathematical thinking regarding whole number subtraction. The findings of the research questions will be discussed under three headings: attending to students' strategies, interpreting students' mathematical understanding, and deciding how to respond on the basis of children's understanding. Then, the descriptive findings will be discussed holistically, and implications will be made.

Attending to Children's Strategies

The findings of the study revealed that only one teacher provided robust remarkable evidence, while most of them showed a lack and limited evidence of attention. As teachers, they know how to solve the subtraction algorithm and can evaluate the correctness of the student's solution, but they do not understand how the students solve the problem. Thus, it can be concluded that solving the subtraction algorithm or evaluating whether the solution is correct was sufficient to identify relevant mathematical details in the student's solution. Consistent with these findings, previous studies resulted in teachers being able to solve the problems; however, they presented general statements about students' solutions rather than giving mathematical details in the solution (e.g., Sanchez-Matamoros et al., 2019). In a similar study conducted by Doğan-Coşkun et al. (2021) with pre-service teachers, more than half of the pre-service teachers demonstrated limited evidence of attention. Moreover, Fernandez et al. (2013) expressed that pre-service teachers had difficulty identifying the mathematically significant details involving proportional and non-proportional reasoning. However, Kılıç (2019) resulted that pre-service teachers could attend to students' thinking in the context of equations. Although the attending skill is regarded as the easiest skill (Jacobs et al., 2010), many studies concluded that in-service and pre-service teachers could not identify the details of students' solution strategies. Jacobs et al. stated that attending does not only require teachers' ability to determine noteworthy situations in a complex learning environment, but also requires having knowledge that enables teachers to determine mathematically significant details. Accordingly, LaRochelle et al. (2019) pointed out that to attend to students' strategies, teachers need to know different strategies that the students develop to solve the problems. From this point of view, the findings led us to conclude that the teachers providing lack of and partial evidence might have limited knowledge of students' strategies in the context of whole number subtraction.

Interpreting Children's Mathematical Understanding

Parallel to the attending expertise, approximately 65% of the teachers had interpreting skills under limited evidence, which included *lack of evidence*, *no evidence of interpretation, just attention*, and *wrong interpretation*. The teachers, who failed to identify the mathematical concepts, such as the ones and tens, did not interpret the relationship between these concepts. This important finding lets us conclude that attending to mathematical concepts in children's strategy plays a significant role in interpreting their mathematical understanding. This result is not surprising since it is necessary to give reasoning about students' strategies and understand how they perform the operation. However, even if a few teachers could explain the student's strategy, they could not interpret the student's understanding. This might be because teachers could focus on the procedural aspects of the operations rather than the conceptual aspects. Therefore, teachers might attach importance to explaining the steps of the strategy and disregard the students' understanding of their underlying reasoning. From this point of view, it can be concluded that although attending expertise is important for interpreting expertise, it does not guarantee to interpret students' understanding.

Deciding How to Respond on the Basis of Children's Understandings

The findings revealed that only one teacher (2.04%) incorporated further understanding to generalize the subtraction operation. Based on the suggested question, it can be realized that the teacher thought that if the unknowns are the minuend or subtrahend, then this question will make the students think deeply and generalize the subtraction operation. Although the teacher tried to enrich students' understanding by changing the unknown of the subtraction operation and to help the students to generalize, the question this teacher posed did not entice the students to think differently and invent new strategies. Because the question, which is generated by changing the unknown, does not necessitate conceptual knowledge of place value, knowledge of properties of operations, such as the associative, commutative, and distributive property, number relationships, connecting to subtraction operation, and other mathematical concepts. The students may only use the meaning and relationship between addition and subtraction to solve the subtraction operations whose minuend or subtrahend is unknown. In this case, the teacher had failed to suggest a question encouraging students to explore a new strategy.

On the other hand, the responses coded as acknowledging and ignorance did not include any questions provoking students to think about concepts deeply and extending their understanding. These teachers might be providing general responses because of difficulty attending to the students' strategies and interpreting students' understanding from their strategies. When a teacher does not understand students' strategies and interpret their understanding based on the important points of the strategies, they are likely to respond to the students in a general and superficial way (Amador et al., 2016). In order to support/ extend their current thinking, the teachers should notice how the students solve the problem and what knowledge/understanding allows them to solve the problem in this way. This result confirms the outcomes of the previous studies by concluding that responding expertise depends on both attending and interpreting expertise; thus, it can be regarded as the most complicated skill of teacher noticing (Crespo, 2002; Jacobs et al., 2010).

Another important reason for providing general responses might be their lack of content and pedagogical content knowledge. Indeed, Tyminski et al. (2014) emphasized that teachers should have coordinated and integrated knowledge to engage in deciding how to respond to students' thinking. As discussed earlier, to decide the best response to extend/support students' understanding, the teachers should attend to students' strategies, which require using their own knowledge related to the concepts. To this end, the teachers need to have rich specialized content knowledge (SCK) (Ball et al., 2008). Furthermore, to decide the best response, the teachers should interpret students' mathematical understanding, which necessitates knowledge of content and students (KCS). Therefore, providing no or limited evidence of attending and interpreting children's understanding might be attributed to teachers' deficiency in their mathematical content knowledge and knowledge of students (Son & Sinclair, 2010; Son, 2016). As attending, interpreting, and responding expertise are interrelated, SCK and KCS also play a foundational role in teachers' skills in responding expertise (Casey et al., 2018; Tyminski et al., 2014). Besides, knowledge of content and teaching (KCT) involves the knowledge needed to decide "which examples to use to take students deeper into the content" (Ball et al., 2008, p. 401), so responding expertise is closely related to teachers' KCT. As a result, the teachers' difficulty in attending, interpreting, and responding may arise from their lack of content and pedagogical content knowledge, and they necessitate having extensive teacher knowledge (Dreher & Kuntze, 2015; Jacobs et al., 2010).

Regarding the overall noticing skills, the descriptive findings suggest that the teaching experience may not positively impact teacher's noticing skills, which is different from what Jacobs et al. (2010) claim. The participants of the present study had teaching experience ranging from five to 40 years, but their noticing skills were not as robust as expected. Thus, it seems that teaching experience does not ensure having robust noticing skills. Instead of teaching experience, experience in students' invented strategies may have more effect on teacher noticing skills. The teachers, experienced in students' invented strategies, could analyze and interpret how the students make sense of the problem and solve it and what kind of knowledge they have. By doing so, they could respond to students using their invented strategies to support/extend their understanding (Crespo, 2002). From this point of view, it is significant to differentiate the terms of *teaching experience* and *experience in students' invented strategies*. Although teachers with at least five years of teaching experience are defined as experienced teachers (Berliner, 2001), they may not have any experience creating a classroom where students invent, share, and discuss their strategies. In such a case, the term experience has been misused, and it should be redefined since teacher noticing focuses more on children's understanding than the regular teaching that occurs during the mathematics lesson, which can be directly related to experience. Thus, in the literature of teacher noticing, the teaching experience might be regarded as being experienced in understanding/reasoning students' invented strategies rather than the number of years they have taught.

Last but not least is the study's contribution to the noticing literature in that it extended the categories of Professional Noticing of Children's Mathematical Thinking Framework developed by Jacobs and his colleagues (2010). Although the present study is grounded in Jacobs et al.'s study, our data, which was gathered by means of an invented strategy, has enabled us to add specific features to the framework. Since the invented strategies necessitate making sense of mathematics but do not require the usage of materials and are more flexible than the standard algorithm, we need to develop a more specific and analytic framework for teacher's noticing special to students' invented strategies

The main dimensions, *attending*, *interpreting,* and *responding*, were the same as those presented by Jacobs et al. (2010), and the names of the categories of attending and interpreting skills were the same as the categorization of Tekin-Sitrava et al.'s (2021). However, we have set the characteristics of each dimension, taking the domain specificity of whole number subtraction into consideration to ensure we contribute to the literature. Also, rather than presenting teachers' responding skills as robust, limited, and lack, we categorized them as *responding to child and incorporating*, *acknowledging*, *ignorance,* and *unrelated and general,* which gives greater insight into teachers' responses. With this categorization, we aimed to evaluate the teachers' responses in terms of whether the teachers consider the students' strategies while responding and whether they support/extend students' understanding. In conclusion, it could be emphasized that this framework enables one to analyze teacher noticing on the basis of students' invented strategies, which has a critical role in connecting students' strategies and the standard algorithm before introducing the standard algorithm.

Implications and Suggestions for Further Research

In light of the present study's findings, sharing some possible implications for teachers and teacher educators and recommendations for further research studies would be significant. Firstly, the findings revealed that most of the teachers had a lack of or limited evidence of noticing skills in the context of students' invented strategies. To enhance noticing skills, teachers might participate in professional development programs to deal with various students' invented strategies, interpret them, and decide how to respond to students based on their invented strategies. Those programs could be enriched by generating collaborative discussion environments among teachers on particular student solutions. This way, teachers find a chance to share their ideas with other teachers with various teaching experiences and could improve their noticing skills on student thinking.

This research was conducted with teachers; however, similar research can be conducted with prospective teachers, and valuable implications could be made for teacher education programs. Thus, further research is recommended to evaluate prospective teachers noticing skills through design courses where prospective teachers could discuss student-invented strategies through written scenarios or video clips. In this regard, teacher education programs might give theoretical and practical opportunities to pre-service teachers so that they will work through invented strategies and their noticing skills will be fostered. Studies with different content areas, including measurement, geometry, and statistics, could be conducted in order to depict teachers' noticing skills in alternative content areas in mathematics. In addition, from a research perspective, our framework is more specific than the current noticing frameworks since it focuses on students' invented strategies. Thus, it could be applied and tested in different contexts with participants from different contexts and backgrounds.

The authors received no financial support for the research, authorship, and/or publication of this manuscript.

Reyhan Tekin Sitrava (reyhantekin@kku.edu.tr) is an associate professor in mathematics education at the Department of Science and Mathematics Education of Kırıkkale University's Education Faculty. Her research interest includes mathematics teacher education. More specifically, her current interests are mathematics teacher knowledge, teacher noticing skills, and children's mathematical understanding.

Mine Işıksal Bostan (misiksal@metu.edu.tr) is a professor of mathematics education at the Department of Science and Mathematics Education at the Middle East Technical University. Her research interest includes the development of prospective teachers' mathematical knowledge and pedagogical content knowledge. She is also interested in prospective and in-service teachers' noticing skills of student thinking. Her current interests further include the development of HLT to enhance students' understanding of mathematics in various content areas.

Seçil Yemen Karpuzcu (secil.karpuzcu@dpu.edu.tr) is an assistant professor in mathematics education at the Department of Science and Mathematics Education of Kutahya Dumlupinar University's Education Faculty. Her research interest includes the use of technology in mathematics education and mathematics teacher education. Her current interests are using technology to learn mathematics and in the professional development of mathematics 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.

Amador, J. M., Bragelman, J., & Superfine, A. C. (2021a). Prospective teachers' noticing: A literature review of methodological approaches to support and analyze noticing. *Teaching and Teacher Education*, *99*, 103256. https://doi.org/10.1016/j.tate.2020.103256

Amador, J. M., Estapa, A., Kosko, K., & Weston, T. (2021b). Prospective teachers' noticing and mathematical decisions to respond: Using technology to approximate practice. *International Journal of Mathematical Education in Science and Technology*, *52*(1), 3–22.

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

Berliner, D. C. (2001). Learning about and learning from expert teachers. *International Journal of Educational Research, 35,* 463–482.

Blömeke, S., Gustafsson, J. E., & Shavelson, R. J. (2015). Beyond dichotomies: Competence viewed as a continuum. *Zeitschrift für Psychologie*, *223*(1), 3.

Campbell, P. F., Rowan, T. E., & Suarez, A. R. (1998). What criteria for student-invented algorithms? In L. J. Morrow & M. J. Kenney (Eds.), *The teaching and learning of algorithms in school mathematics: 1998 yearbook of the National Council of Teachers of Mathematics*. NCTM.

Carpenter, T. P., & Fennema, E. (1992). Cognitively guided instruction: Building on the knowledge of students and teachers. *International Journal of Educational Research*, *17*(5), 457–470.

Carpenter, T.P., Fennema, E., Fuson, K., Hiebert, J., Human, P., Murray, H., Olivier, A., & Wearne, D. (1994). *Teaching mathematics for learning with understanding in the primary grades.* Paper Presented at the annual meeting of the American Educational Research Association, New Orleans, LA, 1994.

Carpenter, T. P., Franke, M. L., Jacobs, V. R., Fennema, E., & Empson, S. B. (1998). A longitudinal study of invention and understanding in children's multidigit addition and subtraction. *Journal for Research in Mathematics Education, 29*(1), 3–20.

Casey, S., Lesseig, K., Monson, D., & Krupa, E. E. (2018). Examining preservice secondary mathematics teachers' responses to student work to solve linear equations. *Mathematics Teacher Education and Development*, *20*(1), 132–153.

Choy, B. H. (2013). Productive mathematical noticing: What it is and why it matters. In V. Steinle, L. Balland, & C. Bardini (Eds.), *Proceedings of 36th annual conference of Mathematics Education Research Group of Australasia* (pp. 186–193). MERGA.

Crespo, S. (2002). Praising and correcting: Prospective teachers investigate their teacherly talk. *Teaching and Teacher Education*, *18*(6), 739–758.

Darling-Hammond, L., & Ducommun, C. E. (2010). Recognizing and developing effective teaching: What policy makers should know and do. *National Education Association (NEA) and American Association of Colleges for Teacher Education (AACTE).*

Dogan-Coskun, S., Tekin-Sitrava, R., & Isiksal-Bostan, M. (2021). Pre-service elementary teachers' noticing expertise of students' mathematical thinking: The case of fractions. *International Journal of Mathematical Education in Science and Technology*, 54(6), 1–18.

Dreher, A., & Kuntze, S. (2015). Teachers' professional knowledge and noticing: The case of multiple representations in the mathematics classroom. *Educational Studies in Mathematics*, *88*(1), 89–114.

Dursunoğlu, H. (2003). Cumhuriyet döneminde ilköğretime öğretmen yetiştirmenin tarihi gelişimi [Historical development of teacher training in primary education in the Republican Period]. *Milli Eğitim Dergisi*, 160.

Franke, M. L., Kazemi, E., & Battey, D. (2007). Mathematics teaching and classroom practice. In F. K. Lester, Jr., (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 225–256) Information Age Publishing.

Girit-Yildiz, D., Osmanoglu, A., & Gundogdu Alayli, F. (2023). Providing a video-case-based professional development environment for prospective mathematics teachers to notice students' misconceptions in measurement. *Journal of Mathematics Teacher Education*, *26*, 179-209.

González, G., & Vargas, G. E. (2020). Teacher noticing and reasoning about student thinking in classrooms as a result of participating in a combined professional development intervention. *Mathematics Teacher Education and Development*, *22*(1), 5–32.

Hoth, J., Kaiser, G., Döhrmann, M., König, J., & Blömeke, S. (2018). A situated approach to assess teachers' professional competencies using classroom videos. In O. Buchbinder & S. Kuntze (Eds.), *Mathematics teachers engaging with representations of practice* (pp. 23–45). Springer.

Ivars, P., Fernández, C., & Llinares, S. (2020). A learning trajectory as a scaffold for pre-service teachers' noticing of students' mathematical understanding. *International Journal of Science and Mathematics Education*, *18*(3), 529–548. https://doi.org/10.1007/s10763-019-09973-4

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. C., Philipp, R. A., & Schappelle, B. P. (2011). Deciding how to respond on the basis of children's understandings. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 97–116). Routledge.

Jacobs, V. R., Philipp, R. A., & Sherin, M. G. (2011). Preface. In M. G. Sherin, V. R. Jacobs & R. A. Philipp (Eds.), *Mathematics teacher noticing: seeing through teachers' eyes* (pp. xxv– xxvii). Routledge.

Jacobs, V. R., & Spangler, D. A. (2017). Research on core practices in K-12 mathematics teaching. In J. Cai (Ed.), *Compendium for research in mathematics education* (pp. 766–792). NCTM.

Kilic, H. (2019). Pre-service mathematics teachers' noticing skills and scaffolding practices. *International Journal of Science and Mathematics Education*, *16*, 377–400.

LaRochelle, R., Nickerson, S. D., Lamb, L. C., Hawthorne, C., Philipp, R. A., & Ross, D. L. (2019). Secondary practicing teachers' professional noticing of students' thinking about pattern generalisation. *Mathematics Teacher Education and Development*, *21*(1), 4–27.

Miles, M. B., & Huberman, M. A. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Sage

Miller, K. F. (2011). Situation awareness in teaching: What educators can learn from video-based research in other fields. In M. G. Sherin, V. R. Jacobs & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 51–65). Taylor and Francis.

Ministry of National Education (2018). *Matematik dersi öğretim programı (ilkokul ve ortaokul 1, 2, 3, 4, 5, 6, 7 ve 8. Sınıflar)* [Mathematics curricula for primary and middle school 1,2,3,4, 5, 6, 7, and 8 grades]. Ankara, Turkey: MEB.

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

Philipp, R., Schappelle, B., Siegfried, J., Jacobs, V., & Lamb, L. (2008). *The effects of professional development on the mathematical content knowledge of K-3 teachers*. Paper presented at the American Educational Research Association, New York, NY.

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*(5), 477–498.

Roy, G. J. (2014). Developing prospective teachers' understanding of addition and subtraction with whole numbers. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, *2*, 1-15.

Sánchez-Matamoros, G., Fernández, C., & Llinares, S. (2019). Relationships among prospective secondary mathematics teachers' skills of attending, interpreting, and responding to students' understanding. *Educational Studies in Mathematics*, *100*(1), 83–99.

Santagata, R., Zannoni, C., & Stigler, J. W. (2007). The role of lesson analysis in pre-service teacher education: An empirical investigation of teacher learning from a virtual video-based field experience. *Journal of Mathematics Teacher Education*, *10*(2), 123–140.

Schifter, D., Bastable, V., & Russell, S. J. (1999). *Making meaning for operations: Casebook*. Dale Seymour Publications.

Schoenfeld, A. H. (2011). Noticing matters. A lot. Now what? In M. G. Sherin, V. R. Jacobs & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 223–238). Routledge.

Sherin, B., & Star, J. R. (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.

Sherin, M., Jacobs, V., & Philipp, R. (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. 3–13). Routledge.

Sherin, M. G, Russ, R., & Colestock, A. (2011). Accessing mathematics teachers' in-the-moment noticing. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 79–94). Routledge.

Sherin, M. G., & van Es, E. A. (2009). Effects of video club participation on teachers' professional vision. *Journal of Teacher Education, 60*(1), 20–37.

Son, J. W., & Sinclair, N. (2010). How preservice teachers interpret and respond to student geometric errors. *School Science and Mathematics*, *110*(1), 31–46.

Son, J. W. (2016). Moving beyond a traditional algorithm in whole number subtraction: Preservice teachers' responses to a student's invented strategy. *Educational Studies in Mathematics*, *93*(1), 105–129. https://doi.org/10.1007/s10649-016-9693-8

Stahnke, R., & Blömeke, S. (2021). Novice and expert teachers' situation specific skills regarding classroom management: What do they perceive, interpret, and suggest? *Teacher and Teacher Education*, *98*, 1-14. https://doi.org/10.1016/j.tate.2020.103243

Stahnke, R., Schueler, S., & Roesken-Winter, B. (2016). Teachers' perception, interpretation, and decision-making: A systematic review of empirical mathematics education research. *ZDM–Mathematics Education*, *48*(1), 1–27. https://doi.org/10.1007/s11858-016-0775-y

Star, J. R., & Strickland, S. K. (2008). Learning to observe: Using video to improve preservice mathematics teachers' ability to notice. *Journal of Mathematics Teacher Education, 11*(2), 107–125. https://doi.org/10.1007/s10857-007-9063-7

Stockero, S. L. (2014). Transitions in prospective mathematics teacher noticing. In J. J. Lo, K. R. Leathamand, & L.R. van Zoest (Eds), *Research trends in mathematics teacher education* (pp.239–259). Springer.

Stockero, S. L., Rupnow, R. L., & Pascoe, A. E. (2017). Learning to notice important student mathematical thinking in complex classroom interactions. *Teaching and Teacher Education*, *63*, 384–395.

Tekin-Sitrava, R., Kaiser, G., & Isiksal-Bostan, M. (2021). Development of prospective teachers' noticing skills within initial teacher education. *International Journal of Science and Mathematics Education*, *20*, 1611-1634. https://doi.org/10.1007/s10763-021-10211-z.

Teuscher, D., Leatham, K. R., & Peterson, B. E. (2017). From a framework to a lens: Learning to notice student mathematical thinking. In E. Schack, M. Fisher, & J. Wilhelm (Eds)., *Teacher noticing: Bridging and broadening perspectives, contexts, and frameworks* (pp. 31–48). Springer.

Thanheiser, E. (2009). Preservice elementary school teachers' conceptions of multidigit whole numbers. *Journal for Research in Mathematics Education, 40*(3), 251–281.

Tyminski, A. M., Zambak, V. S., Drake, C., & Land, T. J. (2014). Using representations, decomposition, and approximations of practices to support prospective elementary mathematics teachers' practice of organizing discussions. *Journal of Mathematics Teacher Education*, *17*(5), 463–487. https://doi.org/10.1007/s10857-013-9261-4

Ulusoy, F., & Çakıroğlu, E. (2021). Exploring prospective teachers' noticing of students' understanding through micro-case videos. *Journal of Mathematics Teacher Education*, *24*(3), 253–282. https://doi.org/10.1007/s10857-020-09457-1

Van de Walle, J. A., Karp, K. S., & Bay-Williams J. M. (2013). *Elementary and middle school mathematics: Teaching developmentally* (8th Ed.). Pearson Education.

van Driel, S., Crasborn, F., Wolff, C. E., Brand-Gruwel, S., & Jarodzka, H. (2021). Exploring preservice, beginning and experienced teachers' noticing of classroom management situations from an actor's perspective. *Teaching and Teacher Education*, *106*, 103435.

van Es, E. A. (2009). Participants' roles in the context of a video club. *Journal of the Learning Sciences, 18*(1), 100–137. https://doi.org/10.1080/10508400802581668 van Es, E. A. (2011). A framework for learning to notice student thinking. In M. G. Sherin, V. R. Jacobs, & R. A. Philipp (Eds.), *Mathematics teacher noticing: Seeing through teachers' eyes* (pp. 134–151). Routledge.

van Es, E., & 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., & Sherin, M. G. (2008). Mathematics teachers' "learning to notice" in the context of a video club. *Teaching and Teacher Education, 24*(2), 244–276.

van Es, E. A., & Sherin, M. G. (2021). Expanding on prior conceptualizations of teacher noticing. *ZDM–Mathematics Education*, *53*(1), 17–27.

Verschaffel, L., Greer, B., & De Corte, E. (2007). Whole number concepts and operations. In F. Lester (Ed.), *Handbook of research in mathematics teaching and learning* (pp. 557–628). (2nd edition). MacMillan.

Warshauer, H. K., Starkey, C., Herrera, C. A., & Smith, S. (2021). Developing prospective teachers' noticing and notions of productive struggle with video analysis in a mathematics content course. *Journal of Mathematics Teacher Education*, *24*(1), 89–121. https://doi.org/10.1007/s10857-019-09451-2

Yeo, S., & Webel, C. (2019). Preservice teachers' use of noticing practices to evaluate technological resources. Paper presented at the *North American Chapter of the International Group for the Psychology of Mathematics Education*.

Yin, R. K. (2009). *Case study research: design and methods* (4th ed.). Sage.

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 2, 81-98

ICRSME
International Consortium for Research in
Science & Mathematics Education

Change in Emergent Multilingual Learners' Mathematical Communication: Attending to Language Use and Needs

Kathy Horak Smith
Tarleton State University

Natalie Smith Jones
Tarleton State University

Cecilia Silva
Texas Christian University

Beth Riggs
Tarleton State University

Molly H. Weinburgh
Texas Christian University

ABSTRACT

Part of learning a new discipline is learning the language used in the discipline. For mathematics, emergent multilingual learners (EML) must learn English and mathematical symbols in order to make meaning and communicate. The mathematics community's understanding of communication is complex and includes the use of natural language, incorporation of representations (mathematical symbols and visuals), and manipulation of tools and technology. In our research, we use this notion of communication as we examine the way students think about their abilities to communicate in and about mathematics. We specifically ask: (1) How do fifth- and eighth-grade EMLs change in their understanding of mathematics communication with intentional instruction as captured on the Mathematics Communication Inventory (MCI) composite scores? (2) If there is change, how do fifth- and eighth-grade EMLs' scores compare? (3) How does the use of academic language to communicate in mathematics change over time for EMLs with intentional instruction? Two groups of students (15 fifth and 17 eighth graders) enrolled in a newcomers' program informed this research. Data were collected using an open-ended pre- and post-writing assessment. The results strongly suggest that students began to recognize the extent to which they used mathematics for communication, after explicit instruction, to reveal modes of communication in mathematics that are easily and constantly used by students. The change over time was different for the two age groups for total words/symbols and unique words.

Keywords: sociocultural constructivism; mathematics communication; mathematics communication inventory; emergent multilingual learners

Introduction

At the national level in the United States (U.S.), mathematics educators stress that students should communicate *about* mathematics and *in* mathematics using a variety of methods (National Governors Association Center for Best Practices & Council of Chief State School Officers [NGA Center & CCSO], 2010; National Council of Teachers of Mathematics [NCTM], 2000). Effective mathematics communication is important for all students regardless of heritage language (Chen & Li, 2008). Students entering U.S. schools with no English must learn mathematics concepts, develop communication skills, and learn English. With a growing population of non-English speaking students,

© 2023 International Consortium for Research in Science & Mathematics Education (ICRSME)

mathematics teachers need to scaffold with emerging multilingual students (EMLs) in terms of both mathematics and language skills. This scaffolding calls for incorporating best practices from two disciplines: mathematics and English Language Arts. Our research focuses on the development of mathematical understanding and communication abilities of EML students in a program that intentionally addressed both.

Theoretical Framework

This study is grounded in a theoretical framework based on semiotic mediation as it is described in two areas: socio-cultural constructivism as posited by Vygotsky (1978; 1986), and academic literacies as articulated by Gee (2004a; 2004b; 2008) and Lemke (2002; 2004). Language is portrayed as a tool for meaning-making (Vygotsky), used within a community (Gee), and having hybrid discursive practices (Lemke). All three portrayals focus on meaning as found in symbols, artifacts, and language.

Socio-cultural Constructivism

Socio-cultural constructivism (Luria, 1976; Vygotsky, 1978; 1986; Vygotsky & Luria, 1994) posits the epistemological stance that knowledge is constructed by individuals as they interact within the social context of community. Vygotsky and his students particularly stress language as a tool for developing conceptual understandings. This theory of learning incorporates forms of scaffolding in which a more knowledgeable other supports the learning when the learner cannot continue alone. The support allows the learner to move forward rather than remaining at his/her current level. The social tool of semiotic items (language, symbols) is important in this scaffolding process.

Literacies

The ability to communicate is further illuminated by Gee (2004a). He outlined a position whereby members of a specialized community use a social language recognizable by that specific community. He points out that the use of the language also makes members recognizable by those outside that particular community. This distinction is seen with mathematicians as there is a specific way of thinking, believing, and viewing the discourse of mathematics. Ubiquitous words, such as mean, whole, median, and sum, are recognized for specific mathematical meaning. The contexts in which the words are spoken or written situate their meaning. For example, in oral language the following use of the homophones whole and hole are understood only by knowing something about the context: "Create the whole using manipulatives″ versus ″Create the hole using a shovel."

The Discourse of mathematics uses four modes as theorized by Lemke (2002). He posits that natural language, visual representation, mathematical symbolism, and manual technical operations work together to form “a single unified system of meaning making” (p. 1). Natural language refers to the written and oral communication as defined by linguists (Lyons, 1991). Lemke (2002) proposes that natural language is not precise enough to represent phenomena mathematically. Historically, as humans recognized the need for more precision in their communication, it became necessary to extend natural language from typological (qualitative meaning) to include topological (quantitative meaning). To illustrate, “The Statue of Liberty is tall” is an example of typological semiotics, where “The Statue of Liberty is 93 meters tall” exemplifies topological semiotics. Mathematical symbols provide information that communicates a variety of information. For example, $T(x,y) = k(x\ 2 + y\ 2)$ or $f(x) = a_0 + \sum_{n=1}^{\infty}\left(a_n \cos\frac{n\pi x}{L} + b_n \sin\frac{n\pi x}{L}\right)$. These symbols and expressions (sentences) are recognizable by mathematicians, but often are meaningless to non-mathematicians.

Lemke (2004) further expands the notion of communication to include visual representations such as drawings, charts, graphs, and tables. Although not unique to mathematics, these representations help organize ideas and communicate mathematical thinking. The 4th mode, manual technical operations, includes meaningful actions and practices. These actions utilize tools within a specific environment to further communicate and make meaning.

Literature Review

Linguistic Challenges of Mathematics

An often expressed misconception is that mathematics is universal and free of influence (Brown et al., 2009; Jourdain & Sharma, 2016). However, the language of mathematics can be confusing (Jourdain & Sharma, 2016) and part of learning a new discipline is learning the language used in the discipline. For mathematics, an EML must learn English and mathematical symbols as well as develop an understanding of how the different systems interact for making meaning (Schleppegrell, 2007; 2011). Teachers often point out vocabulary as a challenge in teaching mathematics and fail to recognize the grammatical patterning that is involved. This is particularly true when word problems are included in the instruction (Fatmanissa & Novianti, 2021).

Truxaw and Rojas (2014) described how difficult it can be to use a developing language while learning new mathematics concepts. These authors pointed to mathematics language as being more abstract, specific, and culturally determined than conversational language. Moschkovich (2015) stressed that conceptual understanding in mathematics requires words, vocabulary, and definitions. She suggested that using multimodal communication can be beneficial for overcoming some of the challenges. Jourdain and Sharma (2016), in a review of the literature, stressed the need for pedagogical practices to pay attention to how language is used in mathematics.

Mathematics Communication

Beal et al. (2010) pointed out that little attention has been paid to low levels of mathematics achievement by EMLs. Their research showed an achievement gap between EMLs and non-EMLs. These authors suggested that there is a minimum reading proficiency associated with mathematics achievement. In fact, two studies provided evidence that mathematical difficulties may reflect deficient language skills rather than quantitative processes (LeFevre et al, 2010; Vukovic, 2012). In order for EMLs to communicate in mathematics, they must use different skills from those in everyday communication. Mathematics communication requires using abstractions and symbols (Olivares, 1997). Vukovic and Lesaux (2013) stated that "general verbal ability appears to impact children's performance by influencing the mathematical thinking that involves the symbolic number system" (p. 89-90). In addition, the specificity of the elements of mathematics 'sentences' often means that the order cannot be rearranged. This is not true of everyday and informal speech.

NCTM

NCTM (2000) recognized that "(c)ommunication can support students' learning of new mathematical concepts as they act out a situation, draw, use objects, give verbal accounts and explanations, use diagrams, write, and use mathematical symbols" (p. 61). The process standard of communication generally refers to the skills of writing, speaking, listening, and reading about mathematics. The language of mathematics is complicated and dense (Schleppegrell, 2007; 2011). Without opportunities to use it, students struggle to communicate thoughts and ideas. By middle school, students should be able to describe, clarify, and extend their mathematics thinking in written and oral forms. Their communication skills are further enhanced through use of multiple representations, including drawings, organized visuals (graphs, tables, and charts), mathematical symbols, manipulatives, or technologies (NCTM, 2000).

Common Core State Standards

Additionally, communication was addressed in two of the six Guiding Principles for School Mathematics found in the Common Core State Standards for Mathematics (CCSSM) (NGA Center & CCSO, 2010). The two are: (a) Teaching and Learning, and (b) Tools and Technology. The Teaching and Learning Guiding Principle discussed in *Principles to Actions* (NCTM, 2014) suggests that students engage "in meaningful learning through individual and collaborative experiences that promote their ability to make sense of mathematical ideas and reason mathematically" (p. 5). The Tools and Technology Principle states that math tools and technology are "essential resources to help students learn and make sense of mathematical ideas, reason mathematically, and communicate their mathematical thinking" (p. 5).

Methods

The mathematics community's understanding of communication is complex and includes the use of natural language, incorporation of representations (mathematical symbols and visuals), and manipulation of tools and technology. In our research we use this notion of communication (Weinburgh et al., 2014) as we examine the way students think about their abilities to communicate in and about mathematics. We specifically ask: (1) How do fifth- and eighth-grade EMLs change in their understanding of mathematics communication with intentional instruction as captured on the Mathematics Communication Inventory (MCI) (Smith et al., 2015) composite scores? (2) If there is change, how do fifth- and eighth-grade EMLs' scores compare? (3) How does the use of academic language to communicate in mathematics change over time for EMLs with intentional instruction?

Participants

Two groups of students (15 fifth and 17 eighth graders) enrolled in a newcomers' program (Silva et al., 2008) within a local urban school system informed this research. Students had been in the U.S. less than three years and were expected to exit the newcomer program, entering mainstream classrooms in the fall. They were classified as "advanced high language proficient" on the Texas English Language Proficiency Assessment System. Twenty languages from 10 different countries were represented.

Context/Instruction

Students attended a three-week (80 hours) enrichment experience focused on a Crime Scene Investigation (CSI) theme in which science, mathematics, and English instruction were integrated (Silva et al., 2012; Weinburgh & Silva, 2011; Weinburgh et al., 2014). Students voluntarily attended the program for six hours each day. The program was developed and taught by four of the authors: two mathematics educators, a science educator, and a bilingual educator. Grade-appropriate investigations helped students solve a mystery using forensic practices to eliminate suspects. Students examined footprints, fingerprints, blood, ink, and DNA samples. To support students in collecting and analyzing scientific data, mathematics content areas of (1) measurement; (2) data collection and analysis; (3) proportional reasoning; (4) patterns, relationships, and algebraic thinking; and (5) numerical reasoning were emphasized. Furthermore, the process standards of connection, proof and reasoning, communication, problem solving, and multiple representations were highlighted.

Week 1

First, students photographed and measured the crime scene to create sketches. The variation in the sketches produced by the students allowed for a discussion regarding the need for precision using scale and proportions. Responding to this discussion, the students used accurate measurements to create a two-dimensional scaled drawing of the crime scene. Later, students made a chart of possible suspects and used the patterns found in fingerprints to eliminate the first suspects. They engaged in multiple iterations of blood typing. As an extension and connection to their lives, the students confirmed their own blood types with their parents to calculate the percentage of each blood type within the class and compared the findings to national statistics. Students used their problem-solving skills and new knowledge of the percentage of blood types in the population to help eliminate other possible suspects. Through discussions and multiple representations, EMLs from each grade level shared their particular findings. These findings resulted in a full set of collaborative data for use by both groups.

Week 2

Fifth graders used chromatography to compare the ink on the packing slip to ink in pens taken from the list of suspects generated by the students, allowing for the elimination of other suspects. Eighth graders used their own foot-print measurements to discover the ratio between foot length and height. Based on this ratio, they calculated the offender's approximate height, thereby eliminating more suspects. In addition, the eighth-grade students used patterns and relationships to read electrophoreses results from their DNA samples to identify the perpetrator from the remaining suspects.

Week 3

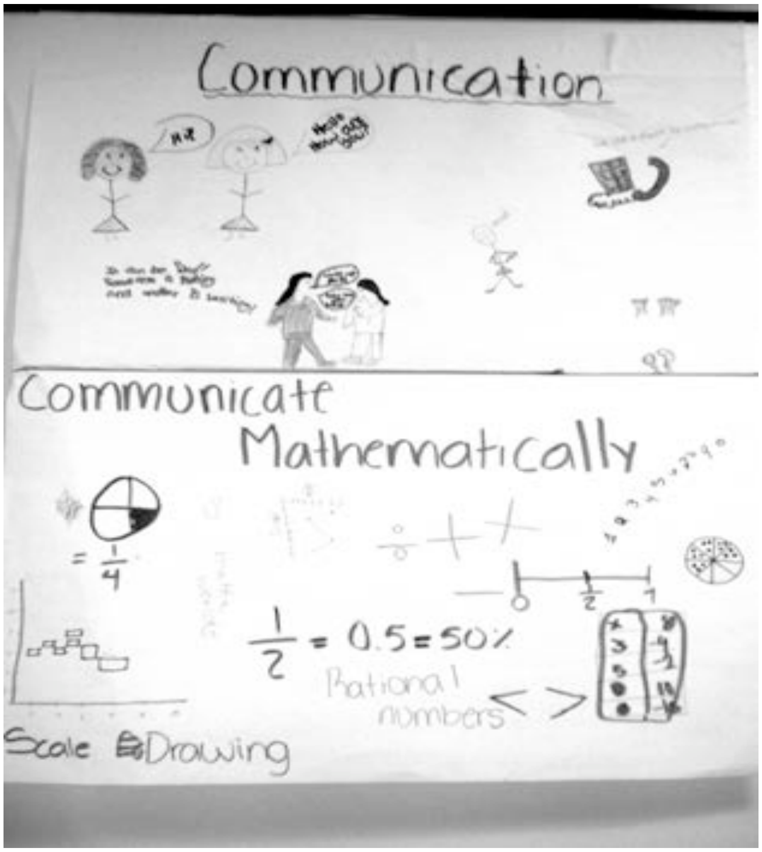
Students engaged in discussions about the various ways of communicating in mathematics. Class discussions included how to communicate within the classroom with peers and teachers through the use of words, pictures, organized visuals, mathematical symbols, and through the use of tools. The students made a list of different ways to communicate. Based on prior experiences, other discussions included ways in which people outside the classroom use mathematics, including their parents.

Students were asked to reflect on their understandings of how they used mathematics to communicate their findings. Various forms of communication were recorded and posted on a classroom chart. Students initially recognized that they were using numbers, charts, and graphs to communicate their findings. Students began to add ideas like written words, drawing pictures, cooking, using manipulatives, and hand gestures. The chart became a growing reference that was used throughout the summer program.

In another activity, students worked in groups to identify ways in which mathematics communication is used outside of the classroom. Figure 1 shows a group display of how students began to think about using mathematics to communicate.

Figure 1

Students Produce a Poster to Represent How They Now Think About Math Communication



Data Collection

Data was collected using an open-ended pre- and post-writing assessment. Students were asked to respond to the prompt "How do I, as a mathematician, communicate information?" This prompt allowed students maximum leeway in expressing understanding of how to communicate in and about mathematics, as well as time needed for responding. Thus, student responses varied in length, complexity, and types of communication systems.

Data Analysis

Data were analyzed in two phases. First, we scored the pre- and post-writing using the MCI. The MCI is an analytical framework grounded in Lemke's notion of hybrid language (2004), NCTM's *Principles and Standards for School Mathematics* (2000), NGA Center and CCSO initiative (2010), and Silverstein's theoretical work in multifunctional communicative semiotic (1995; 2004). The MCI has five coded categories: (1) Mode I, (2) Mode II, (3) Content, (4) Process, and (5) Placement/Context. The MCI was presented to a panel of mathematicians for feedback as to its appropriateness for capturing change in communication. In addition, word and mathematical symbols were counted. The second phase of data analysis compared the data from phase one using the Statistical Package for the Social Science (SPSS).

Phase One Analysis

Writing samples (Appendix) were scored using the MCI. Each sub-category could receive a score of 0 or 1. Mode I and Mode II parallel Lemke's hybrid language. Content and Process were taken from the NCTM standards, CCSSM, and current state standards. Placement/Context was added by the researchers to reflect Gee's (2004b) notion of situated meaning. In addition to the five categories, words and symbols were counted for total number of words/symbols, total mathematics words/symbols, and unique mathematics words/symbols. The total number of words was included because the students were EML learning English beyond mathematics.

Mode I reflected how the students presented their information. For example, the students could use natural language, symbols, visuals, or organized visuals as a part of their explanations. All students received credit for natural language because they wrote something; some students received credit for the use of symbols, visuals, and organized visuals. We determined that manual technical operations would never be used in Mode I.

Mode II reflects what the student described as methods of communication. For the sample to receive a one for natural language, the student had to write about talking, writing, or speaking about mathematics. To receive a one for mathematical symbols, a student had to show how s/he used mathematical symbols. To receive a one for visuals, a student had to describe using pictures or drawings. For a student to receive a one for organized visuals, s/he had to write about using tables, charts, graphs, or other organized visuals. A student might write that a person could use a graph to show the household bills (see Appendix). Lastly, to receive a one for manual technical, the student needed to discuss how s/he used gestures or tools. An example that most students described was using cups/spoons to measure ingredients.

For Content, the coders examined the way the student expressed an understanding of the six content strands based on NCTM standards, CCSSM, and current state standards. Process has five areas as found in the NCTM (2000) standards. Placement/Context included formal, informal, and academic. The sample received credit for informal if the student described a situation such as cooking or driving, and academic if it took place in the school setting. Word and symbol counts were conducted.

Each student's pre- and post-writing was transcribed with all representations (symbolic and visual) included. We practiced using the instrument on data collected from students who were missing a pre- or post- writing sample. Once interrater reliability was established at 91%, coding for the study began.

Phase Two Analysis

In order to answer Research Question 1, a comparison of the scores from the MCI on students' pre-writing and post-writing assessments were conducted using a one-tailed *t*-test. In order to answer Research Question 2, the five sub-scores on the MCI were examined using a one-tailed *t*-test for each of the grade levels.

Both groups improved significantly in understanding of mathematics communication, as shown in Table 1. For the fifth-grade EMLs, MCI mean scores increased from 3.87 to 11.8. This increase was significant at the .05 level, $t(14) = -6.65$, $p < 0.001$. Means for the eighth-graders increased from 2.94 to 15.65, which was significant at the .05 level, $t(16) = -12.62$, $p < 0.001$.

Table 1

Analysis of Pre-Writing and Post-Writing Scores on the MCI

		PreWriting		Post-Writing			
Grade	*n*	*M*	*SD*	*M*	*SD*	*t*	*p*
5th	15	3.87	1.81	11.8	5.09	-6.65	<.001
8th	17	2.94	1.14	15.65	3.89	-12.62	<.001

Research Question 2

Both fifth- and eighth-grade EMLs showed improvement on all five sub-scores which warranted Research Question 3. Analysis revealed differences between the groups from the pre- to the post-writing.

Fifth-grade Change

A comparison of pre- to post- scores by the five categories is shown in Table 2 for the fifth-grade students.

Table 2

Analysis of Pre-Writing and Post-Writing Sub-scores on the MCI

Fifth-Grade		
Sub-score Category	*t*	*p*
Mode I	-2.78	0.007
Mode II	-4.66	<.001
Content Standards	-6.46	<.001
Process Standards	-4.62	<.001
Context	-4.58	<.001

Eight-Grade		
Sub-score Category	*t*	*p*
Mode I	-7.38	<.001
Mode II	-6.17	<.001
Content Standards	-10.87	<.001
Process Standards	-8.72	<.001
Context	-13.96	<.001

During the pre-writing task, students predominately used natural language as their form of communication. Only one student used an organized visual in the writing sample. In the post-writing task, only four students exclusively used natural language. The other 11 students used two or more forms (natural language, mathematical symbols, visuals and/or organized visuals), providing evidence that the students were thinking about a variety of communication strategies.

On Mode II, the fifth-grade students' pre-writing task scores ranged from 0-2. Three students scored 0, nine students scored 1, and three students scored 2. The students only discussed using two forms of communication: natural language and mathematical symbols. Students mostly wrote about being able to talk generically using mathematics. Five students wrote about using mathematics symbols. In contrast, the post-writing task scores ranged from 0-5. The majority of the students' scores ranged from 2-5 on the post- test. Only two of the students scored 0 on the post-test and 13 students scored 2 or higher.

For mathematics content standards, the majority of the students (14) scored either 0 or 1 on the pre-writing task with number and operations being the most common content standard discussed. Only one student scored 2, and s/he touched on number and operations and financial literacy. The post-writing task still had one-fifth of the students scoring 0, the other four-fifths scored between 1 and 4, with most students scoring 3. One student discussed using graphs, one student wrote about using algebra, and two students wrote about using geometry.

Process standards had seven different measures: 1) communication, 2) problem solving, 3) representations, 4) proof and reasoning 5) connections with other content areas, 6) connections within math, and 7) connections in other contexts. On the pre-writing task, seven students scored 0, seven students scored 1, and one student scored 2. On the post-writing task, one student scored 0 and two scored 1. Again, those scoring 1 only discussed using the process standard of communication. The other 12 students scored between 2 and 5, with four scoring 2, four scoring 3, three scoring 4, and one scoring 5. No student hit all seven areas.

Fifth-grade students also showed change in Placement/Context. In the pre-writing task, only four students placed their writing within a specific context: academic setting and business context. On the post-writing samples, four students contextualized mathematics in only one area: informal context or academic context. Five of six students scoring 2 on the MCI placed the mathematics in informal and academic settings. There were four students who were able to write about communicating mathematics in all three contexts.

Sub-scores in each category on the MCI from pre-writing to post-writing were analyzed to determine if the differences were significant at the .05 level. For the fifth-graders, all categories except Mode I (natural language) showed significant improvement (see Table 2).

Eighth-grade Change

A comparison of pre- to post-scores by category are shown in Table 3.

Table 3

Total Word/Symbol Count & Unique Math Word/Symbol Count

Total Word/Symbol Count

		Pre-Writing		Post-Writing			
Grade	*n*	*M*	*SD*	*M*	*SD*	*t*	*p*
5	15	24.93	14.14	245.87	126.34	-6.98	<.001
8	17	24.59	14.99	352.82	151.31	-9.31	<.001

Unique Mathematical Word/Symbol Count

		Pre-Writing		Post-Writing			
Grade	*n*	*M*	*SD*	*M*	*SD*	*t*	*p*
5	15	2.53	2.39	17.93	15.47	-3.85	<.001
8	17	1.12	1.76	23.94	10.87	-8.42	<.001

In the pre-writing task, the only form of communication used by students for Mode I was natural language. In the analysis of the post writing task only two students used natural language exclusively. The majority of the students used three (7 students) and four (5 students) forms of communication.

Looking at Mode II pre-scores on the MCI, the students predominately scored 1. These 12 students wrote about communicating with natural language. Of the three who scored 2 on the instrument, two described using symbols and one wrote about using manual technical operations (rulers, calculators, manipulatives, money, etc.). On the post-writing task, only two students scored 1. One student discussed using natural language and the other discussed using manual technical operations. Of the 15 remaining students, all wrote about using mathematical symbols, four discussed using visuals (pictures and drawings), 11 wrote about using organized visuals (tables, charts, graphs, etc.), and eight described using manual technical tools. These data indicated that students were able to apply the modes to describe how they would communicate mathematically with others.

On the pre-test for Content, only three students identified a content area (number and operations) that they communicate about mathematics. In the post-writing, students expanded the content to include measurement (13), number and operations (17), and financial literacy (16). These post-writing findings show a tremendous increase on which content areas within mathematics can be used to communicate mathematically.

In analyzing the students' pre-writing task for Process, only ten of 17 students wrote about using one of seven process standard components. All ten students discussed using at least one of the four hybrid processes to communicate mathematically. No other process standard was identified in the pre-writing task. On the post-writing task, only one student scored 1. The remaining 16 students scored between 2 and 6. The four process standards included the most were communication (10), problem solving (16), representations (10) and connections within other contexts (15).

When analyzing the pre-writing task about the context, only three students placed their writing within a specific context. One discussed communicating about mathematics in an informal setting, while two students wrote about communicating within an academic setting. For the post-writing example, students were more diverse in the situations where they communicated mathematically. Communicating in both informal and business situations was discussed by 16 students. In contrast, only 12 students discussed using mathematics to communicate in an academic setting.

Sub-scores in each category of the MCI from pre-writing to post-writing were analyzed to determine if the differences were significant at the .05 level. For the eighth-graders, all categories of the MCI showed a significant improvement (see Table 2).

Research Question 3

In order to investigate the ways in which academic language used to communicate in mathematics changes over time for EMLs with intentional instruction, several pre/post word and mathematical symbol counts were conducted. For this study, two counts were reported: total words/symbols used, and unique math words/symbols used by the students.

Total Words/Symbols

The total word/symbol count enabled the authors to explore the increase in students' general writing, including their comfort level in using English to write about mathematics. As shown in Table 3, a statistical increase in total word/symbol count after intentional instruction was noted. The total word/symbol count for fifth-graders was significant at the .05 level, $t(14) = -6.98, p < 0.001$. The fifth-grade students' writing samples average increased from approximately 25 words/symbols to 245 words/symbols. The total word/symbol count for eighth-grade increased from 25 words/symbols to

350 words/symbols. The eighth-grade data were significant at the .05 level, $t(16) = -9.31, p < 0.001$.

Unique Words

The unique mathematical word/symbol count looked solely at mathematics vocabulary and symbols present in the student writing samples. Words/symbols were counted only once even if the word/symbol was used several times. For example, when a student wrote "We communicate with a graph, and there are different types of graphs. Line graph, bar graph, pie graph, double bar graph," the word "graph" was only counted once. In this instance, line graph, bar graph, and pie graph were considered to be unique math terms. A similar procedure was used for math symbols.

Table 3 indicates that the fifth- and eighth-grade students' use of unique mathematics vocabulary/ symbols also increased. The use of unique mathematics vocabulary/symbol count for the fifth-graders was significant at the .05 level, $t(14) = -3.85, p < 0.001$. They increased their use of unique mathematics terminology from 2.5 words per pre-writing sample to 18 unique words per post-writing sample. Whereas the eighth-graders were significant at the .05 level, $t(16) = -8.42, p < 0.001$, their increase was greater. The eighth-graders went from using one unique mathematical term per pre-writing sample to using 24 unique terms in the post-writing sample.

Discussion

As teachers of a growing number of EMLs, we were interested in seeing if intentional instruction that included metacognitive activities would change students' understanding and use of mathematics communication. Although the two grade levels' gains were different, there were similarities in how communication skills developed for both. Therefore, we chose to frame this discussion by sub-categories on the MCI rather than by grade level, followed by a discussion of the changes observed in word/symbol use.

Mode I

Mode I shows the students' understanding of how to present information in a mathematical format. The NCTM Standards call for the students to "(c)reate and use representations to organize, record, and communicate mathematical ideas" (NCTM, 2000, p. 67). However, neither the fifth- nor eighth-grade students came into the summer program showing such an understanding. The data show that given meta-linguistic and meta-mathematical instruction, the students were able to expand their repertoire of communication modes. During the pre-writing task, the students predominately used natural language as their form of communication but moved to using other modes of the hybrid language by the end of the intervention. As indicated by Lemke (2002), natural language is necessary, but not sufficient to communicate fully the precision and complexity of mathematical thought. In addition, Cai et al. (1996), Johnson (2013), and Jourdain and Sharma (2016) discuss the importance of using natural language as well as mathematical expressions, and visual representations. All three posit that these are imperative to students' mathematical understandings.

Mode II

Mode II reflects what the student described as methods of communication. Initially, they predominately used natural language to describe using language to communicate (i.e., I talk, read, and write about math to my friends.). In the pre-writing task, the students discussed using only two forms of communication: natural language and mathematical symbols. After intentional instruction, students were able to expand their repertoire to include mathematical symbols, visuals, organized visuals, and manual technical operations (mathematical tools, gestures, etc.) as a way to communicate mathematics to others. This expansion is an important finding because students who have an understanding of the importance of using different modes of communicating mathematically fulfill the NCTM (2000) call for the ability to "(s)elect, apply, and translate among mathematical representations to solve problems (p. 67) and to "(u)se representations to model and interpret physical, social, and mathematical phenomena" (p. 67).

Content Standards

The scores on the pre-writing task provide evidence that students were only aware of communicating about number/operations and financial literacy. Much of the instruction during the 3-week period required the students to communicate using measurement, data analysis, and number/operations. In the post-writing, the majority of students discussed using measurement, numbers/operations, and financial literacy to communicate mathematically. Limited movement toward mathematical thinking of geometry and algebra was also noted.

We feel the reason students discussed financial literacy (using money) is because of its practicality in their lives. It was not surprising that geometry and algebra were not mentioned by many of the children, since these were not the topics of instruction. We were somewhat surprised that we did not observe more movement regarding data analysis since we used tables, charts, and graphs almost daily to communicate findings in both the mathematics and science classroom.

Process Standards

The process standards are important for students because they cut across all mathematics content areas. Our state guidelines posit that "[t]he process standards weave the other knowledge and skills together so that students may be successful problem solvers and use mathematics efficiently and effectively in daily life" (Texas Education Agency, 2012, p. 1). With knowledge of process skills, both the fifth- and eighth-grade students are more likely to be successful learning the content. As noted in the results section, the students moved beyond only talking or writing about mathematics, to using each of the process standards to show how someone would do mathematics. The biggest gains were found in the process standard of problem solving and connections. This gain can be explained by the emphasis put on problem solving within the integration and overlap of science, language, and mathematics, and with the discussions about real world applications of mathematics. Kosko and Norton (2012) discuss the importance of connecting mathematics not only to other content areas and other contexts outside of school, but also about helping students make connections within mathematics.

The process standard of reasoning and proof increased the least, which did not surprise us because the complexity involved in this process is well documented. Even though we asked the children as a regular part of instruction to explain their answers, we provided less scaffolding for this skill than for others. However, we believe with more time, such as a full semester or academic year, students' understanding of the processes of reasoning and proof would increase.

Placement/Context

Placement and/or context denote the informal, formal, or academic setting in which the students believe mathematics communication is used. The identification of the placement or contexts is very telling about student awareness of using mathematics to communicate in daily life. In looking at the results, if the students mentioned a context during their pre-writing sample, they generally placed the communication within the classroom (academic). However, this occurred in only four of the students' writings with only one student noting that businesses (formal) use mathematics to communicate about sales. Even more astounding was the fact that the eighth-graders did not describe mathematics communication within any context. None of the students wrote about communicating mathematics in an informal setting, yet that is probably the most prevalent context in which to communicate math on a daily basis.

During the post–writing sample, the students demonstrated drastic changes in their views on placement/context. For both age groups, students gave examples within all three areas (informal, formal, and academic). It is obvious that the students became more aware of using mathematics to communicate within informal, everyday situations such as going to the grocery store, driving a car, giving directions, paying bills, cooking, etc. They also identified a multitude of business examples as seen in an example of one student (see Appendix).

Word/Symbol Use

Word and symbol use was divided into total and unique mathematical words/symbols used by EMLs. One might expect that after a 3-week summer program focusing on language, total word/symbol count would increase. This expectation was supported by the data. However, a striking finding was the increase in mathematics word count and unique mathematics terminology. Not only did the students increase the number of words/symbols used, but they placed them in mathematical sentences and appropriate language contexts.

Implications

The results strongly suggest that students began to recognize the extent to which they used math for communication after explicit instruction to reveal modes of communication in mathematics that are easily and constantly used by students. This recognition displayed itself as an increase in the use of the modes of the hybrid language in the post-writing. Of particular note for classroom teachers is that all but one of the students increased in the actual use of natural language, mathematical expressions, and visual representations as they described ways in which they could communicate in mathematics.

Classroom mathematics teachers teach the students the mechanics of communication, but often do not explain that what they are demonstrating is communication. For example, teachers show students how to read graphs. This instruction is necessary, but not sufficient. Students, especially EMLs, benefit from intentional instruction in metacognitive strategies in order to understand the underlying modes of communication. Thus, teachers should also point out the communicative function of the graph. As students become aware that they are doing some form of mathematics all the time, they realize that much of their daily communication includes mathematics. It is possible that the pre-writing reflects that students do not recognize that visual representations and symbolic representations are forms of communication. They do not immediately move from natural language to using visual and mathematical expressions in their explanation. Our study shows that with intentional instruction about modes of communication within the mathematics classroom, students begin to move beyond natural language. These findings are consistent with what we know about natural language in the sense that learners that develop metalinguistic awareness are better language users.

Even though our research does not attempt to show improvement in mathematics achievement, others have done so. Cohen et al. (2015) found that having second grade students write about communicating math improved their mathematics content scores over those who did not write. Their definition of writing, like the writing examples our study, includes natural language, mathematical symbols, and visuals. Their writing samples were not limited to reasoning and proof examples either, but included many types of writing such as poetry. Even though Cohen et.al. (2015) state that their research cannot be applied to other grade levels, we believe that when children communicate about mathematics using the four modes of communication, their new breadth of understanding will move them to thinking about mathematics in many contexts.

The authors received no financial support for the research, authorship, and/or publication of this manuscript.

Kathy Horak Smith (ksmith@tarleton.edu) is currently a Professor and Head of the Department of Mathematics at Tarleton State University in Stephenville, TX. She has over 30 years' experience working (19 years at Tarleton) with pre-service and in-service teachers on integrating mathematics, science, and language. Kathy's teaching at Tarleton focuses on mathematics education for pre-service teachers. Emerging from this work are 17 peer-reviewed journal publications, four book chapters, two books, two state-wide mathematics professional development modules for in-service teachers, and over 100 professional and teacher practice presentations.

Cecilia Silva (c.silva@tcu.edu) is a professor emerita at Texas Christian University. Her research focuses on bilingual education and is interested on the integration of multimodal and content instruction for Active Bilingual Learners/Users of English (ABLE).

Molly H. Weinburgh (m.weinburgh@tcu.edu) is the Andrew Chair for Research in Mathematics & Science Education and Director of the Andrews Institute of Mathematics & Science Education at Texas Christian University. Her honors include the Chancellor's Distinguished Achievement as a Creative Teacher and Scholar, Piper Professorship, Fellow of the American Association for the Advancement of Science, and the 2020 Award for Lifetime Outstanding Service to The Association for Science Teacher Education. Her scholarship focuses on academic language acquisition and conceptual understanding in science by emerging multilingual students.

Natalie Smith Jones (nmsmith9@gmail.com) began working with the MSL Summer Program beginning her junior year while at Tarleton State University and continued for six years. She was a graduate student at Texas Christian University at the time this research was conducted, collected, and analyzed. Natalie has five years of experience teaching middle school mathematics.

Beth Riggs (eriggs@tarleton.edu) is a Professor of Mathematics at Tarleton State University. She teaches a variety of mathematics and mathematics education courses at the undergraduate and graduate level. Her research interests include effective pedagogy and self-efficacy in the mathematics classroom.

References

Beal, C. R., Adams, N. M. & Cohen, P. R. (2010). Reading proficiency and mathematics problem solving by high school English language learners. *Urban Education, 45*(1), 58-74. https://doi.org/10.1177/0042085909352143

Brown, C. L., Cady, J. A., & Taylor, P. M. (2009). Problem solving and the English language learner. *Mathematics Teaching in the Middle School*, *14*(9), 532–539.

Cai, J., Jakabscin, M. S., & Lane, S. (1996). Assessing students' mathematical communication. *School Science and Mathematics*, *96*(5), 238-246.

Chen, X. & Li, Y. (2008). Language proficiency and mathematics learning. *School Science and Mathematics*, *108*(3), 90-93.

Cohen, J. A., Miller, H. C., Casa, T. M., & Firmender, J. M. (2015). Characteristics of second graders' mathematical writing. *School Science and Mathematics*, *115*(7), 344-355.

Fatmanissa, N. & Novianti, M. N. R. (2021). Linguistic challenges in solving mathematics word problems: A case of EFL university students. *Advances in Social Science, Education and Humanities Research, 627*, 16-23.

Gee, J.P. (2004a). Language in the science classroom: Academic social languages as the heart of school-based literacy. In E. W. Saul (Ed.), *Crossing borders in literacy and science instruction* (pp. 13-32). National Science Teachers Association Press.

Gee, J. P. (2004b). *Situated language and learning: A critique of traditional schooling.* Routledge.

Gee, J. P. (2008). What is academic language? In A. S. Rosebery & B. Warren (Eds.), *Teaching science to English language learners: Building on students' strengths* (pp. 57- 70). National Science Teachers Association Press.

Johnson, H. (2013). What's literacy got to do with science and math? *School Science and Mathematics*, *113*(3), 107-108.

Jourdain, L. & Sharma, S. (2016). Language challenges in mathematics education: A literature review. *Waikato Journal of Education, 21*(2), 43-56.

Kosko, K. W. & Norton, A. (2012). Relationships between the process standards: Process elicited through letter writing between preservice teachers and high school mathematics students. *School Science and Mathematics*, *112*(6), 340-348.

Lemke, J. L. (2002). Mathematics in the middle: Measure, picture, gesture, sign, and word. In M. Anderson, A. Saenz-Ludlow, S. Zellweger, & V. Cifarelli (Eds.), *Educational perspectives on mathematics as semiosis: From thinking to interpreting to knowing* (pp. 215-234). Legas Publishing.

Lemke, J. L. (2004). The literacies of science. In E.W. Saul (Ed.), *Crossing borders in literacy and science instruction* (pp. 33-47). National Science Teachers Association Press.

LeFevre, J. L., Fast, L., Skwarchuk, S., Smith-Chant, B. L., Bisanz, J., Kamawar, D., et al. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development*, *81*(6), 1753-1767. https://doi.org/10.1111/j.1467-8624.2010.01508.x.

Luria, A. R. (1976). *Cognitive development: Its cultural and social foundations.* (M. Lopez- Morillas & L. Solotaroff, Trans.). Harvard University Press.

Lyons. J. (1991). *Natural language and universal grammar: Essays in linguistic theory.* Cambridge University Press.

Moschkovich, J. N. (2015). Academic literacy in mathematics for English learners. *The Journal of Mathematical Behavior, 40*, 43-62. https://doi.org/10.1016/j.jmthb.2015.01.005

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

Olivares, R.H. (1997). Communication in mathematics for students with limited English proficiency. In P. Elliott & M. Kenney (Eds.), *Communication in mathematics, K-12 and beyond.* (pp. 219-230). National Council of Teachers of Mathematics.

Schleppegrell, M. J. (2007). Linguistic challenges of mathematics teaching and learning: A research review. *Reading & Writing Quarterly, 23*, 139-159.

Schleppegrell, M. J. (2011). Language in mathematics teaching and learning: A research review. In J. Moschkovich (Ed.), *Language and mathematics education: Multiple perspectives and directions for research* (pp. 73-112). Information Age.

Silva, C., Weinburgh, M. H., Smith, K. H., Barreto, G., & Gabel, J. (2008). Partnering to develop academic language for English language learners in mathematics and science. *Childhood Education, 85*(2), 107-112.

Silva, C., Weinburgh, M. H., Smith, K., Malloy, R. & Marshall, J. (2012). Toward integration: A model of science and literacy. *Childhood Education*, *88*(2), 91-95.

Silverstein, M. (1995). Shifters, linguistic categories and cultural description. In B. G. Blount (Ed.), *Language, culture and society: A book of readings* (pp. 187-221). Waveland Press.

Silverstein, M. (2004). “Cultural” concepts and the language-culture nexus. *Current Anthropology, 45*(5), 621-652.

Smith, K. H., Silva, C., Weinburgh, M. H., & Smith, N. (2015). *Emerging bilinguals: Developing a Rubric for Communicating in Mathematics*. Paper presented at the annual meeting of the Research Council on Mathematics Learning. Las Vegas, NV, USA.

Texas Education Agency. (2012). *Chapter 111: Texas essential knowledge and skills for mathematics: Subchapter B. Middle school.* http://ritter.tea.state.tx.us/rules/tac/chapter111/ch111b.html

Truxaw, M. P., & Rojas, E. D., (2014). Challenges and affordances of learning mathematics in a second language. *Journal of Urban Mathematics Education, 7*(2), 21-30.

Vukovic, R. K. (2012). Mathematics difficulty with and without reading difficulty: Findings and implications from a four-year longitudinal study. *Exceptional Children, 78*(3), 230-300.

Vukovic, R. K. & Lesaux, N. K. (2013). The relationship between linguistic skills and arithmetic knowledge. *Learning and Individual Differences, 23*, 87-91.

Vygotsky, L. (1978). *Mind in society: The development of higher psychological processes.* Harvard University Press.

Vygotsky, L. (1986). *Thought and language.* MIT Press.

Vygotsky, L. & Luria, A. (1994). Too and symbol in child development. In R. Van Der Veer & J. Valsiner (Eds.), *The Vygotsky reader* (pp. 99-172). Blackwell.

Weinburgh, M. H. & Silva, C. (2011). Math, science, and models. *Science & Children*, *48*(10), 38-42.

Weinburgh, M. H., Silva, C., & Smith, K. H. (2014). Is this a science, mathematics, or language arts lesson? Practical advice for teachers of students learning English. In D. F. Berlin & A. L. White (Eds.), *Initiatives in mathematics and science education with global implications* (pp. 93-106). International Consortium for Research in Science and Mathematics Education.

Appendix

Pre – Instruction Journal Entry

How do I as a mathematician communicate information?

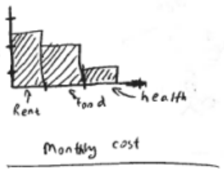
-I don't know.

Word Count: 11 Without question: 3

Post – Instruction Journal Entry

How do I as a mathematician communicate information?

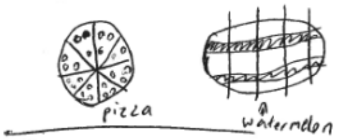
Use a graph to show the house and family bills.



We can use it in steps, or procedures

Go 21st street: Granbury: Walmart
(to) (to)

Maybe cut pizza or apple in equal pieces or any other round objects



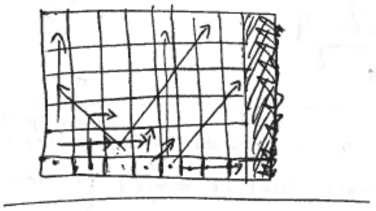
maybe we can use code writing with our friends

"You know, yesterday I proved my dad that I am > tommy
(better) (bigger)

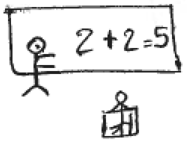
Use it in time table? (graph) (table)

Time	Subject
1:00	Math
2:00	Science
3:15	English

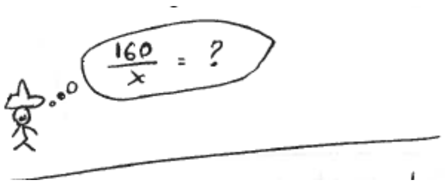
Use in chess (game) to show how each unit can be moved.



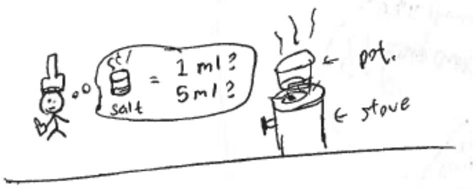
Use it as a teacher to teach younger ones.



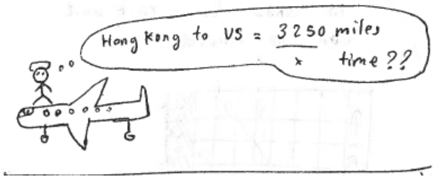
Use it as a fisherman to see how long it would take to comeback if he went 160 feet away from the port.



Use it as a chef to see how ~~me~~ much salt or sugar are needed to add.

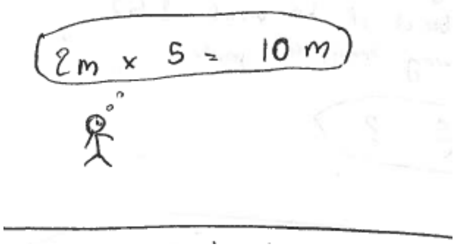


Use it as a pilot to know the approm. time
to reach to destination.

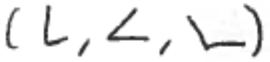


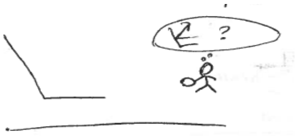
Use it as a blender or baranter (someone Who mizes drink at lounge) to know what to and how much amount to mix.

Use it as an angineer or architector to know the measurements of building on the sketch and to scale it up



Use pie chart to show your favorite activities

use angles  to predict where ~~an~~ will the object bounce back when thrown on a slope.



Word Count: 210 Without question: 202

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 2, 99-119

Women Have Lower Physics Self-efficacy and Identity Even in Courses in Which They Outnumber Men: A Sign of Systemic Inequity?

Sonja Cwik
University of Pittsburgh

Chandralekha Singh
University of Pittsburgh

ABSTRACT

The motivational beliefs of students, who were mainly bioscience majors interested in careers in health professions, in mandatory large introductory level algebra-based physics courses were surveyed. Although female students outnumbered male students in these courses, they had lower physics motivational beliefs including self-efficacy and identity at the beginning of the physics course, and this gender gap increased by the end of the course. Moreover, the present study used a slightly modified version of the physics identity framework by Hazari et al. (2010) to investigate whether the relation between gender and physics identity was mediated by other motivational beliefs, including perceived recognition by others, self-efficacy, and interest. The model shows that perceived recognition by others, self-efficacy, and interest mediated students' physics identity and there was no direct path from gender to identity. The increased gender gap in these beliefs measured at the end of the physics courses may signify inequity and the non-inclusive nature of the physics learning environment. These findings related to the gender gap in physics motivational beliefs are valuable because they may signify that classroom representation alone will not change the pernicious effects of systemic gender inequities in physics perpetuated by society and bolstered further by the physics learning environments.

Keywords: equity, inclusion, gender, motivational beliefs, physics, undergraduate

Introduction

While prior research has investigated issues pertaining to women's representation in the science, technology, engineering, and mathematics (STEM) fields (American Institute of Physics, 2020; Blue et al., 2018; Bowling et al., 2017; Brainard & Carlin, 1998; Center, 2015; Harackiewicz et al., 2016; Kost-Smith et al., 2010; Lorenzo et al., 2006; Miyake et al., 2010; Seymour, 2002; Seymour et al., 1997; Tsui, 2007; Vincent-Ruz et al., 2018; Walton et al., 2015; Whitten et al., 2003), most studies in college physics have focused on courses in which women are outnumbered by men. Moreover, in explaining participation and learning in STEM disciplines, student motivational beliefs have been shown to play an important role (Sari et al., 2018; Smith et al., 2013; Stets et al., 2017; Swarat, 2008; Yang, 2016). However women have been shown to have lower motivational beliefs than men (Hanson et al., 2020; Louis & Mistele, 2012; E. M. Marshman et al., 2018; Stewart et al., 2020). In particular, students' identity has been shown to play an important role, not only in students' in-class participation and performance, but also in their choices of future courses and careers (Carlone & Johnson, 2007; Gee, 2000; Hazari et al., 2010; Stets et al., 2017; Tonso, 2006; Vincent-Ruz & Schunn, 2018). Identity

© 2023 International Consortium for Research in Science & Mathematics Education (ICRSME)

in this context (e.g., physics identity) refers to students' views about whether they see themselves as a *physics person* or a person who can excel in physics (Hazari et al., 2017; Hazari et al., 2013; Hazari et al., 2010; Monsalve et al., 2016).

In the present study, a version of the physics identity framework developed by Hazari et al. (2010) was used, which adapted the science identity framework by Johnson et al. (2017). Johnson et al.'s (2007) science identity framework includes three dimensions: competence (I think I can), performance (I am able to do), and recognition (I am recognized by others). Hazari et al. (2010) modified the framework specifically for physics. Competence and performance were defined as students' beliefs in their ability to understand the subject and students' belief in their ability to perform physics tasks. Additionally, recognition was framed as recognition by others as being a good physics student. Lastly a fourth dimension, interest, was added to the framework since students have highly varying levels of interest in physics (Hazari et al., 2017; Hazari & Cass, 2018). In future studies by Hazari for introductory students, performance and competence are combined into one variable (Hazari et al., 2020). In a slightly reframed version of Hazari's physics identity framework by Kalender et al. (2019b), the framework used in the study, performance/competence was framed as self-efficacy (closely related to competency belief). Additionally, recognition was framed explicitly as perceived recognition by students for clarity.

Self-efficacy refers to students' belief in their ability to accomplish tasks or solve problems (Bandura, 1977; 1994). It has been shown to influence students' engagement, learning, and persistence in science courses, in addition to contributing to students' science identity (Britner, 2008; Cheryan et al., 2017; Felder et al., 1995; Lindstrøm & Sharma, 2011; Sawtelle et al., 2012; Schunk & Pajares, 2002; Zimmerman, 2000). For example, when tackling difficult problems, students with high self-efficacy tend to view the problems as challenges that can be overcome, whereas those with low self-efficacy tend to view them as personal threats to be avoided (Bandura, 1994). However, in introductory physics courses in which women are underrepresented, studies have found a gender gap in self-efficacy that widens by the end of the course, even in interactive engagement courses (E. Marshman et al., 2018; Nissen & Shemwell, 2016). Additionally, self-efficacy has been shown to predict students' engagement, learning, and persistence in science courses (Bandura, 1994; Bouffard-Bouchard et al., 1991; Cavallo et al., 2004; Correll, 2004; Fencl & Scheel, 2005; Vincent-Ruz & Schunn, 2017).

Another motivational belief that influences physics identity is interest. Interest in a particular discipline may affect students' perseverance, persistence, and achievement (Bailey et al., 2017; Harackiewicz et al., 2002; Hidi, 2006; Lichtenberger & George-Jackson, 2013; Strenta et al., 1994; Tims et al., 2014; Wang & Degol, 2013). One study showed that changing the curriculum to stimulate the interest of female students helped improve all of the students' understanding at the end of the year (Häussler & Hoffmann, 2002). Within expectancy-value theory, interest and self-efficacy are constructs that predict students' academic outcomes and career expectations (Wigfield & Eccles, 1992). In this study the focus is on intrinsic interest, or an individual's personal interest, and enjoyment in engaging with physics concepts in this study.

The third belief, perceived recognition by others, has been shown to play an important role in a student's identity (Vincent-Ruz & Schunn, 2018) and motivation to excel (Goodenow, 1993). In a study on students' perception of support, teacher support was more strongly linked to the motivation and engagement of girls than boys (Goodenow, 1993). However, studies have shown that female students are not recognized appropriately in many STEM disciplines, even before they enter college (Archer et al., 2017; Bian et al., 2017; Kalender et al., 2019a). One study found that science faculty members in biological and physical sciences exhibit biases against female students by rating male students significantly more competent (Moss-Racusin et al., 2012). Moreover, prior research suggests that students' perceived recognition by instructors and teaching assistants (TAs) can impact their self-efficacy and interest in physics and students' physics self-efficacy can impact their interest in physics (e.g., Doucette et al., 2020; Doucette & Singh, 2020).

Prior studies have shown that women have lower physics motivational beliefs, including identity, than men (Archer et al., 2017; Lock et al., 2013; Monsalve et al., 2016). However, most of the studies in the college context concerning physics motivational beliefs are conducted in classes in which women are underrepresented (Hazari et al., 2010; Kalender et al., 2019a). Nevertheless, pervasive societal stereotypes and biases about who can excel in physics that bombard women from a young age could impact their physics motivational beliefs, including their identity, even in physics courses in which they outnumber men, (e.g., introductory physics courses for bioscience majors in which women are not underrepresented). The students' physics motivational beliefs in these courses could influence not only their participation and performance in the physics courses, which is very important in itself, but also in other STEM disciplines in which they are intending to major or may be considering a career. For example, past studies have shown that the factors that influence students' physics identity influence their engineering career choices (Godwin et al., 2016) and predict students' engineering identity (Patrick et al., 2018), when the survey was administered in college English composition courses (where the majority of students were first year students with intentions to major in a variety of disciplines).

Similarly, these physics identity factors could influence students in a physics course for bioscience and pre-health majors. In particular, the focus here is on students who are bioscience majors primarily interested in health-related professions, for whom two physics courses are mandatory. The reason these physics courses are mandatory for students is not only because the foundational knowledge of physics is central to developing deep understanding of bioscience, but also because the kind of reasoning skills students develop in physics can help them in their majors and future careers. In fact, a majority of these bioscience majors are on the pre-health track and want to become health professionals. Higher physics motivational beliefs including an identity as a person who can excel in physics can help them engage and perform better (e.g., on the Medical College Admission Test (MCAT) in which 5-8% of the material consists of physics concepts) to realize their career goals.

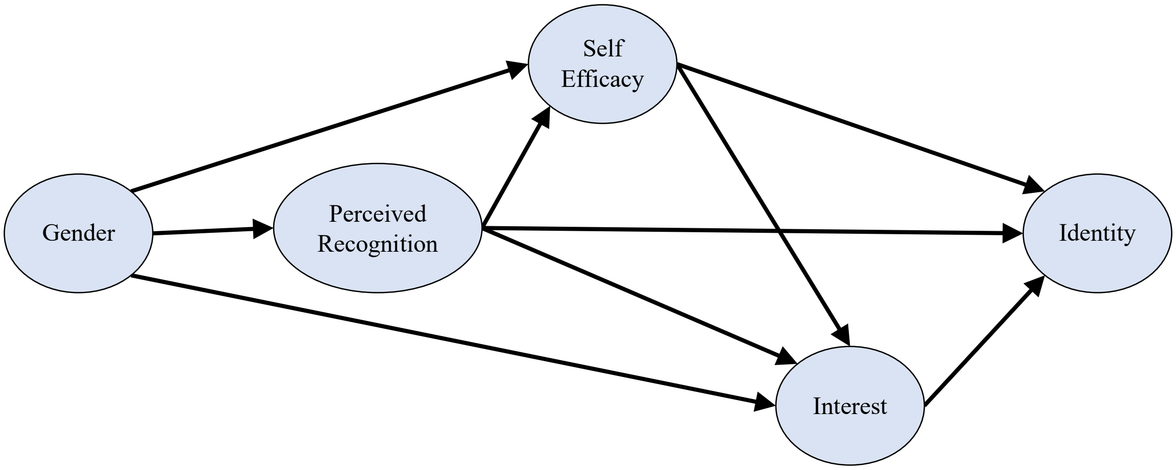
Explicit and implicit societal stereotypes and biases that women internalize about physics could influence their motivational beliefs, including their physics identity. One common societal stereotype is that genius and brilliance are important factors to succeed in physics (Leslie et al., 2015). However, genius is often associated with boys (Upson & Friedman, 2012), and girls from a young age tend to shy away from fields associated with innate brilliance or genius (Bian et al., 2017). Studies have found that by the age of six, girls are less likely than boys to believe they are "really really smart" and less likely to choose activities that are made for "brilliant people" (Bian et al., 2017). These types of stereotypes may continue to negatively impact women from childhood through college physics courses. In formal and informal situations, women are often made to feel that they cannot excel in physics-related fields. Formal situations include K-12 education, in which teachers and counselors often treat men and women differently and give them differential advice about courses to take, and informal situations, which include interacting with family, friends, media and others. Additionally, negative societal stereotypes and biases may lead to a chilly climate for women in science classes, and lack of attention to making science curricula relevant to the interests of many female students (Blickenstaff, 2005). These issues emanating from structural societal inequities pertaining to physics can lower the physics motivational beliefs of women compared to men.

Moreover, in prior investigations, researchers have often hinted at the fact that underrepresentation of women in those courses is likely related to their lower physics motivational beliefs and the gender gap increasing from the beginning to the end of those courses (E. M. Marshman et al., 2018). However, it is important to investigate whether a similar gender gap in physics motivational beliefs exist in physics courses in which women are not numerically underrepresented, and whether these gender gaps get worse from the beginning to the end of the courses. In particular, it is important to investigate whether numerical overrepresentation of women (e.g., in physics courses for bioscience majors) would eliminate or significantly reduce the gender gap in physics motivational beliefs at the beginning and end of these courses. If this is not the case, it may signify that classroom representation alone will not change the pernicious effects of systemic gender inequities in physics perpetuated by society and bolstered further by the physics learning environments. It may point to the need to view these gender-gaps as more deep-rooted than simply due to women's underrepresentation in a physics course and speaks to the need to address inequities in physics learning environments that exacerbate these gender gaps. Since physics motivational beliefs can impact student engagement and performance in physics courses and have the potential to impact their career choices, the deterioration of a gender gap in these courses would be markers of inequity that must be addressed.

Thus, this study focuses on the hypothesis that male and female students' physics motivational beliefs, including their identity, may be different even in courses in which women are not underrepresented. The physics learning environment in this situation may also exacerbate the situation and make the gender gap worse. Figure 1 shows the schematic representation of the identity framework for this investigation (identical to the one used for calculus-based introductory physics courses in which women are outnumbered by men (Kalender et al., 2019b)). In the framework shown in Figure 1, the relation between gender and physics identity is mediated by students' perceived recognition, physics self-efficacy, and interest (Flowers III & Banda, 2016; Godwin et al., 2016; Lock et al., 2013; Potvin & Hazari, 2013; Sawtelle et al., 2012). Models which also have a direct path from gender to physics identity were tested and they were not statistically significant. Therefore, that direct path from gender to physics identity is not shown for clarity.

Figure 1

Schematic Representation of the Physics Identity Model



Note. Schematic representation of the physics identity model and how perceived recognition, self-efficacy, and interest mediate the relation between gender and identity. From left to right, all possible paths were considered (including the one from gender to identity, although it is not shown here since it was not statistically significant).

In accordance with the framework, male and female students' self-efficacy, interest, and perceived recognition, and how these motivational beliefs predict identity in introductory physics courses for bioscience majors in which women outnumber men were investigated. Structural equation modeling (SEM) was used to investigate how the motivational beliefs that comprise physics identity relate to each other in this model. As noted earlier, it is not clear in the context of an introductory physics course in which the majority of students are women whether there are gender differences in these motivational beliefs and how perceived recognition, self-efficacy, and interest mediate male and female students' physics identity. From a statistical point of view, SEM can establish correlations among factors but cannot establish causal effects (Tomarken & Waller, 2005). While past models have theorized the directionality between the constructs in Figure 1 in multiple ways (Godwin et al., 2016; Hazari et al., 2020; Kalender et al., 2019a), this model was chosen because it has the ability to empower instructors so that they adopt effective practices, understand their role in empowering students, and recognize and affirm their work. Additionally, physics interest is not fixed and can increase or decrease depending on students' perceived recognition and self-efficacy, which in turn are dependent on how inclusive and equitable the physics learning environment is. The research questions are delineated based upon the framework, then the methodology is described, then results and discussion are provided, and finally, instructional implications and future directions are discussed. The following research questions were answered by analyzing data from a validated survey administered to students in large algebra-based physics courses at a research university in the United States in which women outnumber men and used mediation analysis via SEM:

Research Questions

RQ1 Are there gender differences in the physics motivational beliefs (self-efficacy, interest, perceived recognition, and identity) and how do they change from the pre-test (at the beginning of the course) to the post-test (at the end of the course)?

RQ2 Can gender differences in students' physics identity be explained with gender differences in physics perceived recognition, self-efficacy, and interest at the end of an introductory algebra-based physics sequence?

Methodology

Participants

The participants were students at a large, public research university in the United States The students were administered a validated written survey at the beginning (pre) and end (post) of the first semester of a two-semester sequence in a traditionally taught introductory algebra-based physics course in which women outnumbered men. Data were used from 501 students who completed the survey on paper in the first week and the last two weeks of recitation class. The students completed the posttest in the week preceding the final exam for the course. These courses are typically taken by students primarily majoring in bioscience in their junior or senior year of undergraduate studies, with approximately 50 to 70% of the students expressing a desire to pursue future careers in health professions. The university provided demographic information such as age, gender, and ethnic/racial information using an honest broker process by which the research team received the information without knowledge of the identities of the participants. The gender data provided by the university included only binary options for male and female, although gender is a socio-cultural and a nonbinary construct (less than 1% of the students did not provide this information and thus were not included in this study). Based on the university data from the participants, 35% identified as male and 65% identified as female students. Thus, female students outnumber male students significantly in this physics class.

Instrument Validity

The survey items were constructed from items validated by others (Adams et al., 2006; Glynn et al., 2011; *PERTS Academic Mindsets Assessment*, 2020) and re-validated in our own context using one-on-one student interviews (E. M. Marshman et al., 2018), exploratory factor analysis (EFA), confirmatory factor analysis (CFA) (Cohen, 2013), analyzing the Pearson correlation between different constructs (Cohen, 2013), and using Cronbach alpha (Cronbach, 1951) At the beginning of the survey, students were instructed to answer the questions on the survey with regard to the physics course they were in. The survey items asked about different motivational beliefs at the beginning and end of the course. These motivational constructs included students' physics identity (1 item), self-efficacy (4 items), interest (4 items), and perceived recognition (3 items). The *physics identity* question focuses on whether students see themselves as a physics person (Hazari et al., 2013; Hazari et al., 2010; Shanahan & Nieswandt, 2009). *The physics self-efficacy* questions measure students' confidence in their ability to understand and answer physics problems (Adams et al., 2006; Glynn et al., 2011; Godwin et al., 2016; Hazari et al., 2013; Learning Activation Lab, 2017; Schell & Lukoff, 2010). The *interest in physics* questions measure students' enthusiasm and curiosity to learn physics and ideas related to physics (Learning Activation Lab, 2017). The *perceived recognition* questions measure the extent to which a student believes that other people see them as a physics person (Hazari et al., 2013). The interest questions were adapted from the Activation Lab Survey on science fascination/interest (Learning Activation Lab, 2017). The first self-efficacy question was taken from the Peer Instruction Self-efficacy instrument (Schell & Lukoff, 2010). Three other self-efficacy questions were adapted from Godwin et al.'s (2016) performance/competence questions with minor changes based upon individual interviews during the validation of the survey instrument at our institution. The physics identity and perceived recognition questions were adapted from Godwin et al. with no fine-tuning required based upon interviews as students interpreted the questions as intended (Godwin et al., 2016).

After performing EFA to ensure that the items factored according to different constructs as envisioned, a CFA was conducted to establish a measurement model for the constructs and used in SEM. The square of CFA factor loadings (lambda) indicates the fraction of variance explained by the factor. The model fit indices were good and all of the factor loadings (lambda) were above 0.50, which indicate good loadings (Cohen, 2013). The results of the CFA model are shown in Table 1. The Cronbach alpha was used to measure the internal consistency of the items. The Cronbach alpha is 0.79 for the self-efficacy questions, 0.76 for interest questions, and 0.89 for perceived recognition questions which are considered reasonable (Cronbach, 1951).

Table 1

Survey Questions and Factor Loadings (Lambda) from the Confirmatory Factor Analysis (CFA)

Construct and Item	Lambda
Physics Identity	
I see myself as a physics person	1.000
Physics Self-efficacy	
I am able to help my classmates with physics in the laboratory or recitation.	0.587
I understand concepts I have studied in physics.	0.750
If I study, I will do well on a physics test.	0.759
If I encounter a setback in a physics exam, I can overcome it.	0.723
Physics Interest	
I wonder about how physics works.	0.567
In general, I find physics. †	0.751
I want to know everything I can about physics.	0.741
I am curious about recent discoveries in physics.	0.637
Physics Perceived Recognition	
My family sees me as a physics person.	0.889
My friends see me as a physics person.	0.921
My physics instructor and/or TA sees me as a physics person.	0.702

Note. Survey questions corresponding to each of the motivational constructs, along with factor loadings from the Confirmatory Factor Analysis (CFA) for all students (N = 501). The questions in the study were designed on a Likert scale of 1 (low endorsement) to 4 (high endorsement) (Likert, 1932). The rating scale for some of the self-efficacy and interest questions was *NO!, no, yes, YES!* while the rating scale for the physics identity and perceived recognition questions was *strongly disagree, disagree, agree, strongly agree.* All *p*-values (of the significance test of each item loading) are $p < 0.001$.
† the rating scale for this question was very boring, boring, interesting, very interesting.

Zero-order pair-wise Pearson correlations r are shown in Table 2. These Pearson r values signify the strength of the relationship between constructs. The inter-correlations vary in the strength of their correlation, but none of the correlations are so high that the constructs cannot be considered separate, consistent with prior studies (Kalender et al., 2019b).

Table 2

Pearson Inter-Correlations Between Factors

Pearson Correlation Coefficient				
Observed Variable	1	2	3	4
1. Perceived Recognition	--	--	--	--
2. Self-Efficacy	0.61	--	--	--
3. Interest	0.60	0.60	--	--
4. Physics Identity	0.78	0.59	0.62	--

Note. Pearson inter-correlations are given between all the predictors and outcomes for the post-test in Physics 1. All *p*-values < 0.001.

The highest inter-correlation was between physics identity and perceived recognition (0.78), which is consistent with prior studies in calculus-based courses. Perceived recognition questions ask about external identity (perception of whether other people recognize an individual as a physics person), whereas the physics identity question asks about internal identity (whether an individual sees oneself as a physics person) so there tends to be a high correlation between these constructs. However, the correlation is low enough that they can be considered separate constructs.

In addition, preformed item response theory (IRT) was completed in order to check the response option distances for the survey constructs (Embretson & Reise, 2000). The parametric grades response model using STATA was used to test the measurement precision of the response scale. The two response scales were *NO!, no, yes, YES!* and *strongly disagree, disagree, agree, strongly agree.* The parametric grades item response model calculates the location parameter for each response and calculates the difference between the locations. For the first group (NO!, no, yes, YES!), the response scale discrimination values were 1.54 and 2.08. For the second set of answers (strongly disagree, disagree, agree, strongly agree), the response scale discrimination values were 1.51 and 1.26. The numerical values for the location differences need to be similar, as they are for both of these scales, so the means for these values can be used. In addition, IRT-based domain scores were estimated (Embretson & Reise, 2000). The correlation coefficient between the Likert mean scale values for each observed variable and the IRT-based domain scores were calculated to be 0.99 between interest questions, 0.98 between self-efficacy questions and 0.99 between perceived recognition and identity questions (Embretson & Reise, 2000; Samejima, 1969). Since the factor scores derived from the IRT are so highly correlated with the mean score, using mean values for the scores is acceptable when analyzing the data (De Winter & Dodou, 2010; Embretson & Reise, 2000; Norman, 2010). In particular, because the psychological distance between adjacent response items and across items was approximately similar and complex factor scores derived from IRT or CFA are so highly correlated with the mean scores, it is reasonable to use mean scores (Embretson & Reise, 2000; Samejima, 1969). Moreover, the frequencies of student responses for each question are included in Appendix A.

Analysis

Initially, data was analyzed using descriptive statistics and compared female and male students' mean scores on various constructs for statistical significance using *t*-tests and computed the effect sizes using Cohen's *d* (Cohen, 2013). In general, $d = 0.20$ indicated a small effect size, $d = 0.50$ indicates a medium effect size, and $d = 0.80$ indicates a large effect size (Cohen, 2013). For predictive relationships between different motivational constructs, Structural Equation Modeling (SEM) was used as a statistical tool in R (lavaan package) with a maximum likelihood estimation method (Team, 2013). The SEM is an extension of multiple regression. It conducts several multiple regressions simultaneously between different latent variables (or factors or constructs) in one estimation model and can have multiple outcome variables. This is an improvement over multiple regression since it can calculate overall goodness of fit and it allows for all estimates to be standardized simultaneously. This enables a direct comparison among different structural components, along with calculations of factor loadings for all factors (or constructs or latent variables). The SEM also has an option to handle missing data using the full estimation maximum likelihood, or ML estimation feature which improves both power and generalizability since it imputes missing data so that students only missing some data are not dropped. The model fit for SEM was reported by using the Comparative Fit Index (CFI), Tucker-Lewis Index (TLI), Root Mean Square Error of Approximation (RMSEA), and Standardized Root Mean Square Residuals (SRMS). Commonly used thresholds for goodness of fit are as follows: CFI and TLI > 0.90, and SRMR and RMSEA < 0.08 (MacCallum et al., 1996).

The model estimates were performed using gender moderation analysis to check whether any of the relations between variables show differences across gender by using lavaan to conduct multi-group SEM. Initially, different levels of measurement invariance in the multi-group SEM model were tested with gender moderation. In each step, different elements of the model were fixed to equality across gender and compared the results to the previous step using the Likelihood Ratio Test. Since there was not statistically significant moderation by gender, the theoretical model in a gender mediation analysis was tested, using gender as a variable directly predicting all latent variables to examine the resulting structural paths between constructs.

Results and Discussion

Pertaining to **RQ1**, Table 3 and Table 4 show that women had statistically significantly lower mean values than men for all motivational beliefs in the model. For example, both women's and men's scores were low in physics identity with women scoring below the negative value (2). Additionally, when the scores in the pre-test and post-test as shown in Table 5 are compared, the scores significantly change for women's self-efficacy, interest, and physics identity as well as for men's self-efficacy from when they enter the class until the end of the course. These differences increase throughout the semester and the gender gap widens by the end as evident from the larger Cohen's d values for women in Table 5. This occurs despite the fact that women outnumber men in this course. In addition, in Appendix A, the percentage of men and women who selected each response to the questions are provided. This provides a sense of how students shifted their answers from the pre-test to the post-test. From Tables 6 and 7 in Appendix A, it is found that in general both men and women shift from higher values to lower values for each motivational factor. For example, in the perceived recognition and identity questions, most students selected response 2 in the pretest while most students selected response 2 in the posttest. Additionally from Table 6 and Table 7, for the self-efficacy and interest questions, the overall trend is that the percentage of students who answered 3 or 4 decreased from the pretest to the posttest and the percentage of students who answered 1 or 2 increased from the pretest to the posttest.

Table 3

Mean Pre Predictor and Outcome Values by Gender and Effect Sizes (Cohen's d)

Predictors and Outcomes	**Mean**		**Cohen's *d***
	Men	**Women**	
Perceived Recognition	2.19	2.00	0.32
Self-Efficacy	3.07	2.83	0.58
Interest	2.74	2.45	0.58
Physics Identity	2.16	1.87	0.46

Note. All *p*-values < 0.001.

Table 4

Mean Post Predictor and Outcome Values by Gender and Effect Sizes (Cohen's D).

Predictors and Outcomes	**Mean**		**Cohen's *d***
	Men	Women	
Perceived Recognition	2.21	1.92	0.43
Self-Efficacy	2.90	2.55	0.66
Interest	2.66	2.27	0.69
Physics Identity	2.03	1.64	0.59

Note. All *p*-values < 0.001.

Table 5

Cohen's d (d) for Men and Women's Outcomes from the Pre-Test to the Post-Test

Predictors and Outcomes	**Men**		**Women**	
	d	*p-value*	*d*	*p-value*
Perceived Recognition	0.05	0.686	0.13	0.128
Self-Efficacy	0.33	0.005	0.54	< 0.001
Interest	0.15	0.203	0.34	< 0.001
Physics Identity	0.18	0.137	0.34	< 0.001

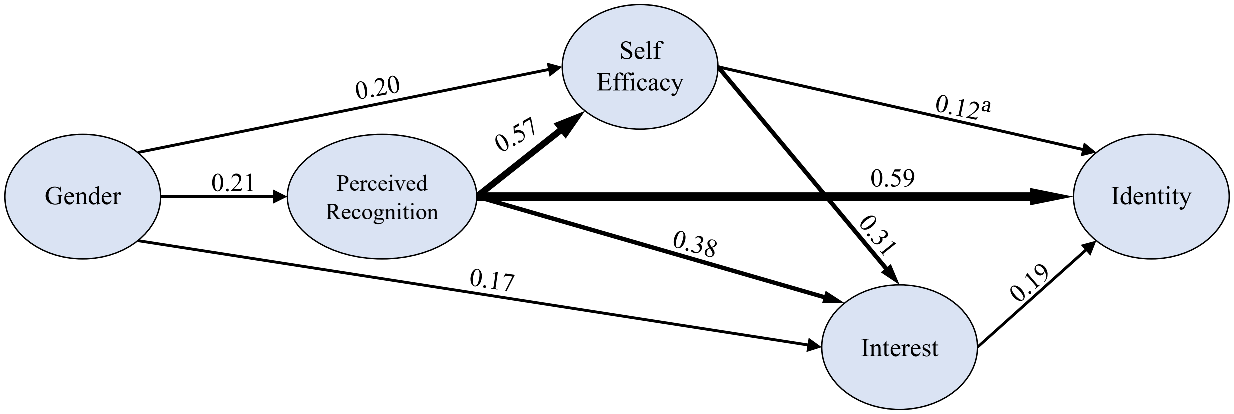
Note. Statistical significance refers to *p*-values.

Pertaining to **RQ2**, SEM was used to investigate the relationships between the constructs and to unpack each construct's contribution to explaining the physics identity of women and men. Initially, gender moderation was tested between different constructs using multi-group SEM (between male and female students) and investigated whether the relationships between the different motivational constructs were different across gender. There were no group differences at the level of weak and strong measurement invariance, including no difference at the level of regression coefficients. Therefore, gender mediation analysis was used to understand how gender mediates physics identity at the end of the semester in the introductory physics course.

The results of the SEM are presented visually in Figure 2. The model fit indices indicate a good fit to the data (acceptable fit thresholds in parentheses): CFI = 0.981 (> 0.90), TLI = 0.974 (> 0.90), RMSEA = 0.043 (<0.08), and SRMR = 0.030 (< 0.08).

Figure 2

Result of the Path Analysis Part of the SEM



Note. Result of the path analysis part of the SEM showing mediation between gender and physics identity through perceived recognition, self-efficacy, and interest. The line thickness qualitatively denotes the relative magnitude of the standardized regression coefficients β shown. All p-values for β are indicated by no superscript for $p < 0.001$ and "a" for $p = 0.013$. Gender does not directly predict physics identity.

All three of the mediating constructs (perceived recognition, self-efficacy, and interest) predict physics identity at the end of the physics course similar to past results (Hazari et al., 2010; Kalender et al., 2019b) in other contexts. Perceived recognition had the largest direct effect (β = 0.59) with smaller effects from self-efficacy (β = 0.12) and interest (β = 0.19).

Additionally, gender predicts perceived recognition, self-efficacy, and interest. However, the relation between gender and physics identity is mediated only by the mediating constructs, and after accounting for these indirect paths, there is no direct path from gender to identity. In other words, women appear to have a lower physics identity because they have lower perceived recognition, self-efficacy, and interest. These results are very similar to those for calculus-based introductory physics courses (Kalender et al., 2019b) in which women are severely underrepresented and the majority of students are engineering majors. In addition, in Appendix B, data are presented for how the factors that predict students' physics identity also predict students' science identity. Although it is not one of the main research questions, Figure 3 in Appendix B shows that physics self-efficacy also predicted students' science idenity.

Summary, Implications, and Future Directions

In this research involving both descriptive and inferential quantitative analyses, gaps in physics motivational beliefs are found that disadvantage women in mandatory introductory physics courses for bioscience majors in which women are not outnumbered by men. This result is similar to what has been found earlier in introductory calculus-based courses in which women are severely underrepresented (Hazari et al., 2010; Kalender et al., 2019b). In particular, prior research in calculus-based courses in which women are underrepresented shows that, compared to men, women have a greater decrease in their motivational beliefs (e.g., about whether instructors and TAs see them as people who can excel in physics) (Kalender et al., 2019b). These findings are true even in algebra-based physics courses in which women are not underrepresented (e.g., both men and women have a mean recognition below the positive lower threshold, i.e., score of 3, and women score significantly lower than men). Tables 3 and 4 show that women had a greater decrease in physics perceived recognition, self-efficacy, interest, and identity compared to men. The inferential analysis using SEM in Figure 2 shows how perceived recognition, self-efficacy and interest mediate students' physics identity. The gender differences in students' perceived recognition, self-efficacy, and interest predict their identity as a person who can excel in physics and the correlations between these motivational beliefs predicting students' physics identity are similar to those in calculus-based introductory courses in which women are underrepresented (Hazari et al., 2010; Kalender et al., 2019b). Moreover, there is no direct path from gender to physics identity, similar to the calculus-based courses (Kalender et al., 2019b).

The model provides support for the physics identity framework in a new context and helps us understand the role of different motivational factors in predicting physics identity of women and men in algebra-based physics courses where women are not underrepresented. This is important since identity in a particular discipline is context-dependent and factors that influence physics identity can relate to each other differently in different contexts. Additionally, the model shows that TAs and instructors could play a critical role to increase students' self-efficacy, interest, and identity in physics. The findings suggest that women feel less recognized by their instructors and TAs than men which influence women's self-efficacy, interest, and identity in physics.

The existence of a gender gap at the beginning of the physics course may be due to a variety of reasons including societal stereotypes, experiences in previous science courses, and lack of female physics role models in media. The exacerbation of the gender gaps in motivational beliefs such as physics perceived recognition, self-efficacy, interest, and identity from the beginning to the end of the physics course for bioscience majors may be a sign of inequity and can disadvantage women in terms of their participation, engagement, and outcomes in the physics courses. These persistent gaps can also impact their choice of majors and future careers in STEM. The increased gender gap in these beliefs at the end of the physics courses may signify inequity and the non-inclusive nature of the learning environment. The trends in the gender gap in physics motivational beliefs are highly troubling. Further investigation of their cause is required because this may signify that classroom representation alone will not change the pernicious effects of systemic gender inequities in physics perpetuated by society and bolstered further by the physics learning environments. For example, the increased motivational belief gaps may at least partly be due to physics instructors and TAs unwittingly reinforcing gender stereotypes about physics and communicating lower expectations for women. If female students are not given the same type of positive feedback as male students, this could have a negative effect on perceived recognition and self-efficacy of women. Changing the narrative, increasing students' sense of belonging (Binning et al., 2020), and making the physics learning environment equitable and inclusive (e.g., by not letting men dominate the conversations in class and affirming students more when they make progress in the course) has the potential to decrease the gender gap in student's perceived recognition and other motivation beliefs.

A limitation of this quantitative study focusing on descriptive and inferential quantitative analysis is that insight can only be found in the relative values of the motivational beliefs of men and women at the beginning and end of the course, and how the relation between gender and physics identity is mediated by physics perceived recognition, self-efficacy, and interest. However, causal effects cannot be established. Therefore, future studies should investigate factors in the physics learning environment that exacerbate gender gaps in motivational beliefs even in these courses in which women outnumber men. In addition, future work can investigate approaches to improving student motivational beliefs in these types of physics courses in which women are not underrepresented and investigate whether approaches that are effective in these courses would also be successful in courses in which women are underrepresented. One potential approach for improving students' motivational beliefs is through brief social-psychological classroom interventions (e.g., mindset/sense of belonging), that have been shown to be effective in boosting women's grades in some science courses (Binning et al., 2020; Harackiewicz et al., 2016; Walton et al., 2015; Yeager & Walton, 2011).

This work was supported by the National Science Foundation award DUE-1524575.

Sonja Cwik (Sonja.cwik@pitt.edu) is an Assistant Professor in the Physics Department at Chapman University. She received her bachelors degree from Wellesley College and her masters and Ph.D. from the University of Pittsburgh. Her research interests include understanding women's motivational beliefs in introductory physics courses.

Chandralekha Singh (clsingh@pitt.edu) is a Distinguished Professor in the Department of Physics and Astronomy and the Founding Director of the Discipline-based Science Education Research Center (dB-SERC) at the University of Pittsburgh. She is a Past President of the American Association of Physics Teachers. She obtained her bachelors and masters degrees from the Indian Institute of Technology Kharagpur and her Ph.D. in theoretical condensed matter physics from the University of California Santa Barbara. She was a postdoctoral fellow at the University of Illinois Urbana Champaign, before joining the University of Pittsburgh. She has been conducting research in physics education for more than two decades. She co-led the US team to the International Conference on Women in Physics in Birmingham, UK in 2017. She is a Fellow of the American Physical Society, American Association for the Advancement of Science, and the American Association of Physics Teachers.

References

Adams, W. K., Perkins, K. K., Podolefsky, N. S., Dubson, M., Finkelstein, N. D., & Wieman, C. E. (2006). New instrument for measuring student beliefs about physics and learning physics: The Colorado Learning Attitudes about Science Survey. *Physical Review Special Topics - Physics Education Research, 2*(1), 010101. https://doi.org/10.1103/PhysRevSTPER.2.010101

American Institute of Physics. (2020). *Percent of physics bachelors and PhDs earned by women, classes of 1975 through 2016*. https://www.aip.org/statistics/data-graphics/percent-physics-bachelors-and-phds-earned-women-classes-1975-through-2016

Archer, L., Moote, J., Francis, B., DeWitt, J., & Yeomans, L. (2017). The "exceptional" physics girl: A sociological analysis of multimethod data from young women aged 10–16 to explore gendered patterns of post-16 participation. *American Educational Research Journal, 54*(1), 88-126. https://doi.org/10.3102/0002831216678379

Bailey, J. M., Lombardi, D., Cordova, J. R., & Sinatra, G. M. (2017). Meeting students halfway: Increasing self-efficacy and promoting knowledge change in astronomy. *Physical Review Physics Education Research, 13*(2), 020140. https://doi.org/10.1103/PhysRevPhysEducRes.13.020140

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological review, 84*(2), 191.

Bandura, A. (1994). Self-efficacy. In R. J. Corsini (Ed.), *Encyclopedia of Psychology* (2nd ed., Vol. 3, pp. 368-369). Wiley.

Bian, L., Leslie, S.-J., & Cimpian, A. (2017). Gender stereotypes about intellectual ability emerge early and influence children's interests. *Science, 355*(6323), 389-391.

Binning, K., Kaufmann, N., McGreevy, E., Fotuhi, O., Chen, S., Marshman, E., Kalender, Z. Y., Limeri, L., Betancur, L., & Singh, C. (2020). Changing social norms to foster the benefits of collaboration in diverse workgroups. *Psychological Science, 31*(9), 1059-1070. https://doi.org/10.1177/0956797620929984

Blickenstaff, J. C. (2005). Women and science careers: Leaky pipeline or gender filter? *Gender and Education, 17*(4), 369-386. https://doi.org/10.1080/09540250500145072

Blue, J., Traxler, A. L., & Cid, X. C. (2018). Gender matters. *Physics Today, 71*(3), 40. https://doi.org/10.1063/pt.3.3870

Bouffard-Bouchard, T., Parent, S., & Larivee, S. (1991). Influence of self-efficacy on self-regulation and performance among junior and senior high-school age students. *International Journal of Behavioral Development, 14*(2), 153-164.

Bowling, K. G., Miller, L. M., Patel, R., & Bass, K. M. (2017). Impacts of virtual clinical trials simulations on science knowledge and attitudes. *The Electronic Journal for Research in Science & Mathematics Education, 21*(7). https://ejrsme.icrsme.com/article/view/17592

Brainard, S. G., & Carlin, L. (1998). A six-year longitudinal study of undergraduate women in engineering and science. *Journal of Engineering Education, 87*(4), 369-375.

Britner, S. L. (2008). Motivation in high school science students: A comparison of gender differences in life, physical, and earth science classes. *Journal of Research in Science Teaching, 45*(8), 955-970.

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. https://doi.org/10.1002/tea.20237

Cavallo, A. M. L., Potter, W. H., & Rozman, M. (2004). Gender differences in learning constructs, shifts in learning constructs, and their relationship to course achievement in a structured inquiry, yearlong college physics course for life science majors. *School Science and Mathematics, 104*(6), 288-300. https://doi.org/10.1111/j.1949-8594.2004.tb18000.x

Center, N. R. (2015). *Science & engineering degree attainment: 2004-2014*. National Student Clearinghouse Research Center. https://nscresearchcenter.org/snapshotreport-degreeattainment15/

Cheryan, S., Ziegler, S. A., Montoya, A. K., & Jiang, L. (2017). Why are some STEM fields more gender balanced than others? *Psychological Bulletin, 143*(1), 1. https://doi.org/https://doi-org.pitt.idm.oclc.org/10.1037/bul0000052

Cohen, J. (2013). *Statistical power analysis for the behavioral sciences*. Routledge.

Correll, S. J. (2004). Constraints into preferences: Gender, status, and emerging career aspirations. *American Sociological Review, 69*(1), 93-113.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*(3), 297-334.

De Winter, J., & Dodou, D. (2010). Five-point likert items: *t* test versus Mann-Whitney-Wilcoxon (Addendum added October 2012). *Practical Assessment, Research, and Evaluation, 15*(1), 11.

Doucette, D., Clark, R., & Singh, C. (2020). Hermione and the secretary: How gendered task division in introductory physics labs can disrupt equitable learning. *European Journal of Physics, 41*(3), 035702. https://doi.org/10.1088/1361-6404/ab7831

Doucette, D., & Singh, C. (2020). Why are there so few women in physics? Reflections on the experiences of two women. *The Physics Teacher, 58*(5), 297-300.

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Lawrence Erlbaum Associates Publishers.

Felder, R. M., Felder, G. N., Mauney, M., Hamrin Jr, C. E., & Dietz, E. J. (1995). A longitudinal study of engineering student performance and retention: Gender differences in student performance and attitudes. *Journal of Engineering Education, 84*(2), 151-163. https://doi.org/https://doi.org/10.1002/j.2168-9830.1995.tb00162.x

Fencl, H., & Scheel, K. (2005). Research and teaching: Engaging students - An examination of the effects of teaching strategies on self-efficacy and course climate in a nonmajors physics course. *Journal of College Science Teaching, 35*(1), 20.

Flowers III, A. M., & Banda, R. (2016). Cultivating science identity through sources of self-efficacy. *Journal for Multicultural Education, 10*(3), 405-417.

Gee, J. P. (2000). Chapter 3: Identity as an analytic lens for research in education. *Review of Research in Education, 25*(1), 99-125.

Glynn, S. M., Brickman, P., Armstrong, N., & Taasoobshirazi, G. (2011). Science motivation questionnaire II: Validation with science majors and nonscience majors. *Journal of Research in Science Teaching, 48*(10), 1159-1176. https://doi.org/10.1002/tea.20442

Godwin, A., Potvin, G., Hazari, Z., & Lock, R. (2016). Identity, critical agency, and engineering: An affective model for predicting engineering as a career choice. *Journal of Engineering Education, 105*(2), 312-340. https://doi.org/10.1002/jee.20118

Goodenow, C. (1993). Classroom belonging among early adolescent students: Relationships to motivation and achievement. *The Journal of Early Adolescence, 13*(1), 21-43. https://doi.org/10.1177/0272431693013001002

Hanson, L. M., Westerlund, J. F., & Vaughn, P. W. (2020). Gender differences in student attitudes towards science in secondary school classrooms with and without NSF GK-12 resident scientists. *The Electronic Journal for Research in Science & Mathematics Education, 24*(4), 50-68. https://ejrsme.icrsme.com/article/view/20033

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., & Elliot, A. J. (2002). Predicting success in college: A longitudinal study of achievement goals and ability measures as predictors of interest and performance from freshman year through graduation. *Journal of Educational Psychology, 94*(3), 562-575. https://doi.org/10.1037/0022-0663.94.3.562

Harackiewicz, J. M., Canning, E. A., Tibbetts, Y., Priniski, S. J., & Hyde, J. S. (2016). Closing achievement gaps with a utility-value intervention: Disentangling race and social class. *Journal of Personality and Social Psychology, 111*(5), 745.

Häussler, P., & Hoffmann, L. (2002). An intervention study to enhance girls' interest, self-concept, and achievement in physics classes. *Journal of Research in Science Teaching, 39*(9), 870-888. https://doi.org/10.1002/tea.10048

Hazari, Z., Brewe, E., Goertzen, R. M., & Hodapp, T. (2017). The importance of high school physics teachers for female students' physics identity and persistence. *The Physics Teacher, 55*(2), 96-99. https://doi.org/10.1119/1.4974122

Hazari, Z., & Cass, C. (2018). Towards meaningful physics recognition: What does this recognition actually look like? *The Physics Teacher, 56*(7), 442-446. https://doi.org/10.1119/1.5055325

Hazari, Z., Chari, D., Potvin, G., & Brewe, E. (2020). The context dependence of physics identity: Examining the role of performance/competence, recognition, interest, and sense of belonging for lower and upper female physics undergraduates. *Journal of Research in Science Teaching, 57*(10), 1583-1607.

Hazari, Z., Potvin, G., Lock, R. M., Lung, F., Sonnert, G., & Sadler, P. M. (2013). Factors that affect the physical science career interest of female students: Testing five common hypotheses. *Physical Review Special Topics-Physics Education Research, 9*(2), 020115.

Hazari, Z., Sonnert, G., Sadler, P. M., & Shanahan, M.-C. (2010). Connecting high school physics experiences, outcome expectations, physics identity, and physics career choice: A gender study. *Journal of Research in Science Teaching, 47*(8), 978-1003. https://doi.org/10.1002/tea.20363

Hidi, S. (2006). Interest: A unique motivational variable. *Educational Research Review, 1*(2), 69-82. https://doi.org/10.1016/j.edurev.2006.09.001

Johnson, A., Ong, M., Ko, L. T., Smith, J., & Hodari, A. (2017). Common challenges faced by women of color in physics, and actions faculty can take to minimize those challenges. *The Physics Teacher, 55*(6), 356-360. https://doi.org/https://doi.org/10.1119/1.4999731

Kalender, Z. Y., Marshman, E., Schunn, C. D., Nokes-Malach, T. J., & Singh, C. (2019a). Gendered patterns in the construction of physics identity from motivational factors. *Physical Review Physics Education Research, 15*(2), 020119. https://doi.org/10.1103/PhysRevPhysEducRes.15.020119

Kalender, Z. Y., Marshman, E., Schunn, C. D., Nokes-Malach, T. J., & Singh, C. (2019b). Why female science, technology, engineering, and mathematics majors do not identify with physics: They do not think others see them that way. *Physical Review Physics Education Research, 15*(2), 020148. https://doi.org/10.1103/PhysRevPhysEducRes.15.020148

Kost-Smith, L. E., Pollock, S. J., & Finkelstein, N. D. (2010). Gender disparities in second-semester college physics: The incremental effects of a "smog of bias". *Physical Review Special Topics - Physics Education Research, 6*(2), 020112. https://doi.org/10.1103/PhysRevSTPER.6.020112

Learning Activation Lab. (2017). *Activation lab tools: Measures and data collection instruments.* http://www.activationlab.org/tools/.

Leslie, S.-J., Cimpian, A., Meyer, M., & Freeland, E. (2015). Expectations of brilliance underlie gender distributions across academic disciplines. *Science, 347*(6219), 262-265.

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.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology, 22*(140), 55. https://legacy.voteview.com/pdf/Likert_1932.pdf

Lindstrøm, C., & Sharma, M. D. (2011). Self-efficacy of first year university physics students: Do gender and prior formal instruction in physics matter? *International Journal of Innovation in Science and Mathematics Education (formerly CAL-laborate International), 19*(2), 1-19.

Lock, R. M., Hazari, Z., & Potvin, G. (2013). Physics career intentions: The effect of physics identity, math identity, and gender. *AIP Conference Proceedings, 1513*(1), 262-265. https://doi.org/10.1063/1.4789702

Lorenzo, M., Crouch, C. H., & Mazur, E. (2006). Reducing the gender gap in the physics classroom. *American Journal of Physics, 74*(2), 118-122.

Louis, R. A., & Mistele, J. M. (2012). The differences in scores and self-efficacy by student gender in mathematics and science. *International Journal of Science and Mathematics Education, 10*(5), 1163-1190.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130.

Marshman, E., Kalender, Z. Y., Nokes-Malach, T., Schunn, C., & Singh, C. (2018). Female students with A's have similar physics self-efficacy as male students with C's in introductory courses:

A cause for alarm? *Physical Review Physics Education Research, 14*(2), 020123. https://doi.org/10.1103/PhysRevPhysEducRes.14.020123

Marshman, E. M., Kalender, Z. Y., Schunn, C., Nokes-Malach, T., & Singh, C. (2018). A longitudinal analysis of students' motivational characteristics in introductory physics courses: Gender differences. *Canadian Journal of Physics, 96*(4), 391-405.

Miyake, A., Kost-Smith, L. E., Finkelstein, N. D., Pollock, S. J., Cohen, G. L., & Ito, T. A. (2010). Reducing the gender achievement gap in college science: A classroom study of values affirmation. *Science, 330*(6008), 1234-1237.

Monsalve, C., Hazari, Z., McPadden, D., Sonnert, G., & Sadler, P. M. (2016). Examining the relationship between career outcome expectations and physics identity. *Proceedings of the Physics Education Research Conference*, Sacramento, CA.

Moss-Racusin, C. A., Dovidio, J. F., Brescoll, V. L., Graham, M. J., & Handelsman, J. (2012). Science faculty's subtle gender biases favor male students. *Proceedings of the National Academy of Sciences of the United States of America, 109*(41), 16474. https://doi.org/10.1073/pnas.1211286109

Nissen, J. M., & Shemwell, J. T. (2016). Gender, experience, and self-efficacy in introductory physics. *Physical Review Physics Education Research, 12*(2), 020105. https://doi.org/10.1103/PhysRevPhysEducRes.12.020105

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Sciences Education, 15*(5), 625-632.

Patrick, A., Seepersad, C., & Borrego, M. (2018). A combined model for predicting engineering identity in undergraduate students. *2018 American Society of Engineering Education Annual Conference & Exposition*, Salt Lake City, UT.

PERTS Academic Mindsets Assessment. (2020). https://survey.perts.net/share/dlmooc

Potvin, G., & Hazari, Z. (2013). The development and measurement of identity across the physical sciences. *Proceedings of the 2013 Physics Education Research Conference*, Portland, OR.

Samejima, F. (1969). Estimation of latent ability using a response pattern of graded scores. *Psychometrika monograph supplement, 1968*(1), i-169. https://doi.org/10.1002/j.2333-8504.1968.tb00153.x

Sari, U. S., Alıcı, M., & Faruk, Ö. (2018). The effect of STEM instruction on attitude, career perception and career interest in a problem-based learning environment and student opinions. *The Electronic Journal for Research in Science & Mathematics Education, 22*(1). https://ejrsme.icrsme.com/article/view/17861

Sawtelle, V., Brewe, E., & Kramer, L. H. (2012). Exploring the relationship between self-efficacy and retention in introductory physics. *Journal of Research in Science Teaching, 49*(9), 1096-1121.

Schell, J., & Lukoff, B. (2010). Peer instruction self-efficacy instrument [Developed at Harvard University] (unpublished).

Schunk, D. H., & Pajares, F. (2002). The development of academic self-efficacy. In A. Wigfield & J. S. Eccles (Eds.), *Development of achievement motivation* (pp. 15-31). Elsevier.

Seymour, E. (2002). Tracking the processes of change in US undergraduate education in science, mathematics, engineering, and technology. *Science Education, 86*(1), 79-105. https://doi.org/10.1002/sce.1044

Seymour, E., Hewitt, N. M., & Friend, C. M. (1997). *Talking About Leaving: Why Undergraduates Leave the Sciences* (Vol. 12). Westview Press.

Shanahan, M. C., & Nieswandt, M. (2009). Creative activities and their influence on identification in science: Three case studies. *Journal of Elementary Science Education, 21*(3), 63-79. www.jstor.org/stable/43156177

Smith, J. L., Lewis, K. L., Hawthorne, L., & Hodges, S. D. (2013). When trying hard isn't natural: Women's belonging with and motivation for male-dominated STEM fields as a function of effort expenditure concerns. *Personality and Social Psychology Bulletin, 39*(2), 131-143.

Stets, J. E., Brenner, P. S., Burke, P. J., & Serpe, R. T. (2017). The science identity and entering a science occupation. *Social Science Research, 64*, 1-14.

Stewart, J., Henderson, R., Michaluk, L., Deshler, J., Fuller, E., & Rambo-Hernandez, K. (2020). Using the social cognitive theory framework to chart gender differences in the developmental trajectory of STEM self-sfficacy in science and engineering students. *Journal of Science Education and Technology, 29*(6), 758-773.

Strenta, A. C., Elliott, R., Adair, R., Matier, M., & Scott, J. (1994). Choosing and leaving science in highly selective institutions. *Research in Higher Education, 35*(5), 513-547. www.jstor.org/stable/40196139

Swarat, S. (2008). What makes a topic interesting? A conceptual and methodological exploration of the underlying dimensions of topic interest. *The Electronic Journal for Research in Science & Mathematics Education*, *12*(2), 1-26. https://ejrsme.icrsme.com/article/view/7773

Team, R. C. (2013). *R: A language and environment for statistical computing.* Scientific Research Publishing

Tims, H. E., Turner, G. E., Corbett, K., Deemer, E. D., & Mhire, J. (2014). Cyber value and interest development: Assessment of a STEM career intervention for high school students. *The Electronic Journal for Research in Science & Mathematics Education, 18*(1), 1-15. https://ejrsme.icrsme.com/article/view/11946

Tomarken, A. J., & Waller, N. G. (2005). Structural equation modeling: Strengths, limitations, and misconceptions. *Annual Review of Clinical Psychology, 1*(1), 31-65. https://doi.org/10.1146/annurev.clinpsy.1.102803.144239

Tonso, K. L. (2006). Student engineers and engineer identity: Campus engineer identities as figured world. *Cultural Studies of Science Education, 1*(2), 273-307. https://doi.org/10.1007/s11422-005-9009-2

Tsui, L. (2007). Effective strategies to increase diversity in STEM Fields: A review of the research literature. *The Journal of Negro Education, 76*(4), 555-581. www.jstor.org/stable/40037228

Upson, S., & Friedman, L. F. (2012). Where are all the female geniuses? *Scientific American Mind, 23*(5), 63-65.

Vincent-Ruz, P., Binning, K., Schunn, C. D., & Grabowski, J. (2018). The effect of math SAT on women's chemistry competency beliefs. *Chemistry Education Research and Practice, 19*(1), 342-351.

Vincent-Ruz, P., & Schunn, C. D. (2017). The increasingly important role of science competency beliefs for science learning in girls. *Journal of Research in Science Teaching, 54*(6), 790-822. https://doi.org/10.1002/tea.21387

Vincent-Ruz, P., & Schunn, C. D. (2018). The nature of science identity and its role as the driver of student choices. *International Journal of STEM Education, 5*(1), 48.

Walton, G. M., Logel, C., Peach, J. M., Spencer, S. J., & Zanna, M. P. (2015). Two brief interventions to mitigate a "chilly climate" transform women's experience, relationships, and achievement in engineering. *Journal of Educational Psychology, 107*(2), 468-485.

Wang, M.-T., & Degol, J. (2013). Motivational pathways to STEM career choices: Using expectancy–value perspective to understand individual and gender differences in STEM fields. *Developmental Review, 33*(4), 304-340.

Whitten, B. L., Foster, S., & Duncombe, M. (2003). What works for women in undergraduate physics? *Physics Today, 56*(9), 46. https://doi.org/10.1063/1.1620834

Wigfield, A., & Eccles, J. S. (1992). The development of achievement task values: A theoretical analysis. *Developmental Review, 12*(3), 265-310. https://doi.org/https://doi.org/10.1016/0273-2297(92)90011-P

Yang, L. H. (2016). Interpersonal relationships and the development of student interest in science. *The Electronic Journal for Research in Science & Mathematics Education, 20*(1), 18-38. https://ejrsme.icrsme.com/article/view/15757

Yeager, D. S., & Walton, G. M. (2011). Social-psychological interventions in education: They're not magic. *Review of Educational Research, 81*(2), 267-301.

Zimmerman, B. J. (2000). Self-efficacy: An essential motive to learn. *Contemporary Educational Psychology, 25*(1), 82-91.

Appendix A

Percentages of men and women who selected each answer choice for each question are shown. This provides a sense of how students shifted their answers from the pre-test to the post-test.

Table 6

Percentages of Women Who Answered Each Question

	Women							
	Pre-Test				Post-Test			
Question	1	2	3	4	1	2	3	4
Physics Identity								
1	27%	59%	13%	1%	45%	45%	9%	1%
Physics Perceived Recognition								
2	31%	53%	15%	1%	40%	46%	13%	1%
3	30%	55%	13%	2%	38%	46%	13%	3%
4	16%	52%	29%	3%	26%	38%	32%	4%
Physics Self-efficacy								
5	14%	45%	39%	2%	11%	29%	54%	6%
6	4%	19%	68%	9%	6%	29%	59%	6%
7	1%	6%	66%	27%	11%	39%	41%	9%
8	1%	14%	66%	19%	7%	39%	48%	6%
Physics Interest								
9	24%	53%	20%	3%	23%	35%	33%	9%
10	5%	32%	58%	5%	13%	38%	46%	3%
11	5%	44%	45%	6%	18%	54%	26%	2%
12	5%	33%	54%	8%	15%	40%	41%	4%

Note. Percentages of women who answered each question by the options they selected, with 1 being the low value (NO! and strongly disagree) and 4 being the high value (YES! and strongly agree). The rating scale for the self-efficacy and interest questions was *NO!, no, yes, YES!,* while the rating scale for the physics identity and perceived recognition questions was *strongly disagree, disagree, agree, strongly agree.*

Table 7

Percentages of Men Who Answered Each Question

	Men							
	Pre-Test				Post-Test			
Question	1	2	3	4	1	2	3	4
Physics Identity								
1	12%	62%	24%	2%	21%	56%	21%	2%
Physics Perceived Recognition								
2	17%	64%	18%	1%	25%	49%	23%	3%
3	15%	63%	21%	1%	21%	48%	28%	3%
4	8%	52%	35%	5%	12%	41%	39%	8%
Physics Self-efficacy								
5	3%	38%	53%	6%	5%	21%	63%	11%
6	1%	9%	80%	10%	2%	19%	65%	14%
7	0%	2%	54%	44%	5%	14%	57%	24%
8	0%	5%	65%	31%	3%	22%	60%	15%
Physics Interest								
9	13%	42%	37%	8%	10%	26%	51%	13%
10	2%	17%	72%	10%	3%	28%	59%	10%
11	1%	30%	58%	11%	3%	49%	41%	7%
12	3%	17%	68%	12%	5%	31%	56%	8%

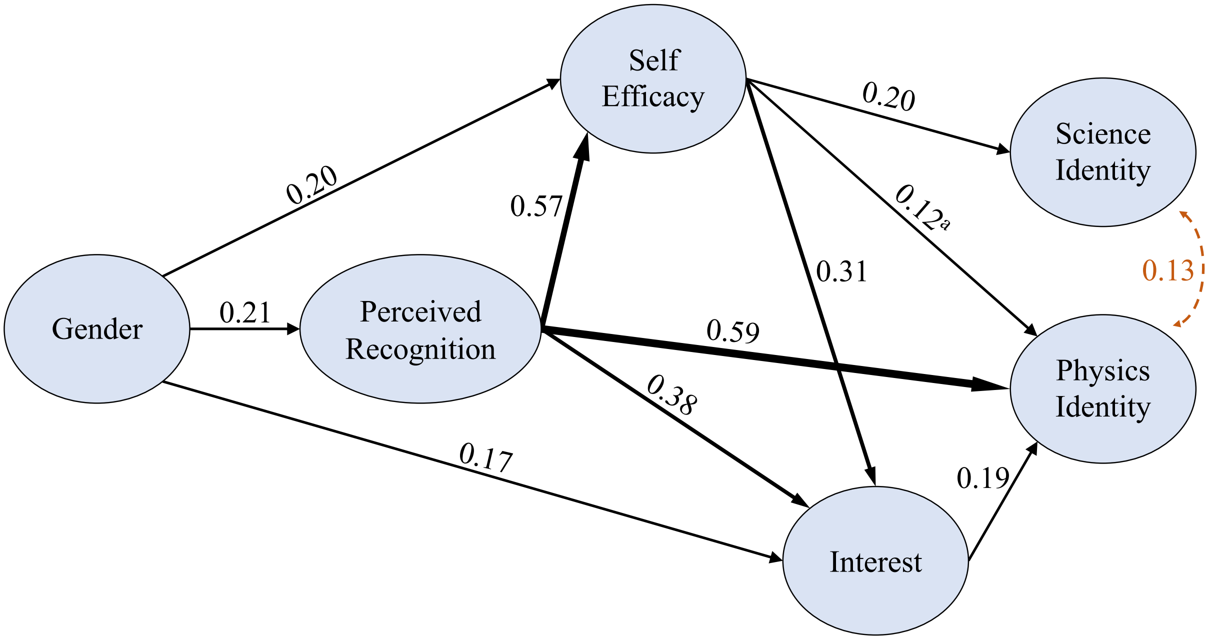
Note. Percentages of men who answered each question by the options they selected, with 1 being the low value (NO! and strongly disagree) and 4 being the high value (YES! and strongly agree). The rating scale for the self-efficacy and interest questions was *NO!, no, yes, YES!,* while the rating scale for the physics identity and perceived recognition questions was *strongly disagree, disagree, agree, strongly agree.*

Appendix B

Below, the path analysis part of the SEM for how the factors that predict physics identity also predict students' science identity is provided. The science identity question on the survey asked students whether they "strongly disagree, disagree, agree, strongly agree" with "I see myself as a scientist". The survey item was adapted from the survey by Godwin et al. (2016) and was re-validated in our own context. The results of the SEM are presented visually in Figure 3. The model fit indices indicate a good fit to the data (acceptable fit thresholds in parentheses): CFI = 0.978 (> 0.90), TLI = 0.971 (> 0.90), RMSEA = 0.043 (<0.08), and SRMR = 0.033 (< 0.08). All three of the mediating constructs (physics perceived recognition, self-efficacy, and interest) predict physics identity at the end of the physics course similar to Figure 2. The mediating constructs of physics self-efficacy also predict science identity at the end of the physics course.

Figure 3

Result of the Path Analysis Part of the SEM with Science Identity



Note. Result of the path analysis part of the SEM with mediation between gender and science/physics identity through physics perceived recognition, self-efficacy, and interest. Each line thickness qualitatively denotes the relative magnitude of the standardized regression coefficients β shown. The dashed lines indicate covariance. All *p*-values for β are indicated by no superscript for $p < 0.001$, "a" for $p = 0.015$, "b".

ELECTRONIC JOURNAL FOR RESEARCH
IN SCIENCE & MATHEMATICS EDUCATION
VOL. 27, NO. 2, 120-139

The Examination of the Relationship Between Scientific Literacy and Some Cognitively Based Individual Differences in Terms of Gender

Feride Sahin
Manisa Celal Bayar University

Salih Ates
Gazi University

ABSTRACT

The aim of this study is to examine whether the structural model that Sahin and Ates (2020) put forward about the relationship among students' scientific literacy levels, logical thinking ability, cognitive styles, mental capacity, and mental rotation ability differ in terms of gender. A causal-comparative model approach was used in this study. The sample of the research was 823 seventh-grade students in the central districts of Ankara. This sample was created using the random stratified sampling method. In the study, it is seen that 64% of the variance in the scientific literacy scores of the females and 48% of the variance in the scientific literacy scores of the males are explained by the cognitively based variables. Comparing male and female students in the data analyses, the predictive effect of the mental rotation ability of female students on the logical thinking ability was not significant. In male students, when only the direct effects were examined, the effect of cognitive styles on scientific literacy was not significant. Based on the results of the study, it is thought that examining the interactions among the activities related to learning, teaching, and assessment process used in science classes, and cognitive styles will play an important role in the process of removing inequalities of opportunity arising from possible gender-cognitive style interaction in the process of raising scientifically literate individuals.

Keywords: cognitive styles, gender, logical thinking ability, mental capacity, mental rotation ability, scientific literacy

Introduction

The vision of the Science Curriculum in Turkey is defined as educating all students as scientifically literate regardless of their individual differences (Turkish Ministry of Education [TME], 2017). Raising scientifically literate individuals is not only Turkey's aim but also many countries' (Pruitt, 2014). By determining the effect of various individual differences in the process of raising scientifically literate people, the activities to be carried out to reach the scientific literacy vision could be built on a more solid basis.

Many studies have been conducted on individual differences that are thought to affect science achievement from past to present (Schmeck & Grove, 1979; Schlatter et al., 2021). When these studies are examined, it is seen that the individual differences discussed may be in cognitive, affective, and psychomotor dimensions (Jonassen & Grabowski, 1993). Regarding the individual differences in cognitive characteristics, it is suggested that the variables of logical thinking ability, cognitive styles, spatial ability, mental capacity, locus of control, creativity, and fluent/crystallized intelligence are

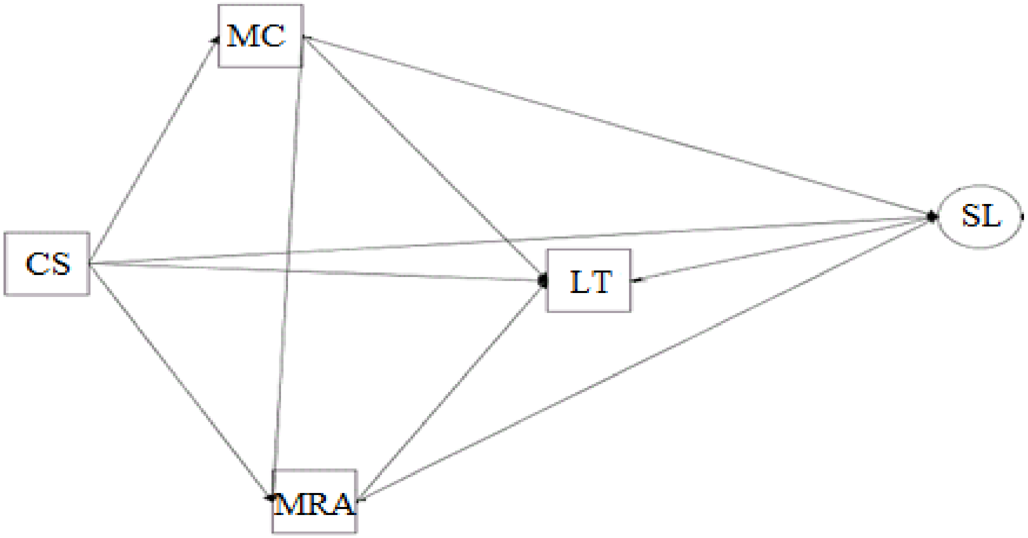
© 2023 International Consortium for Research in Science & Mathematics Education (ICRSME)

frequently mentioned in the literature. When the studies conducted between 1975 and 2017 in Web of Science and ERIC databases were examined using the keywords "individual difference" and "science achievement" in the field of science education, it was observed that cognitive-based individual differences, which are discussed in over 5% of the articles reviewed, include logical thinking ability, spatial thinking skills, cognitive styles and mental capacity (Sahin, 2018).

When the studies investigating the relationship between the above-mentioned cognitive-based individual differences and students' science achievement are addressed, it is seen that science achievement is mostly defined as the ability to solve end-of-units or chapter problems in science textbooks (Ates & Cataloglu, 2007). Although the main purpose of science education today is to raise scientifically literate individuals, the number of studies that define science achievement in the dimension of scientific literacy and reveal the interaction between these variables from a holistic perspective is quite limited. In their study in which they defined science achievement in the dimension of scientific literacy, Sahin and Ates (2020) revealed the relationship among those cognitively based individual differences and the scientific literacy level of students with a holistic model and tested the model seen in Figure 1.

Figure 1

The Theoretical Model of the Relationships Between Scientific Literacy Levels (SL), Logical Thinking Ability (LT), Cognitive Styles (CS), Mental Capacity (MC), and Mental Rotation Ability (MRA).



According to the model, the variable that directly predicts the scientific literacy levels of seventh-grade students is their logical thinking abilities. All other variables predict scientific literacy both directly and indirectly. When it comes to total effects in the study, it was found that field-dependent/field-independent cognitive styles had the highest predictive effect on scientific literacy, logical thinking abilities, mental capacities, and mental rotation abilities of seventh-grade students.

It is regarded that the proposed structural model is vital in terms of examining the interaction between various cognitive-based individual differences that may affect students' scientific literacy levels in a multifaceted manner. Besides, it is thought that it is also crucial to examine whether this model and findings differ in terms of gender, which is one of the other notable individual differences widely researched (Dimitrov, 1999; Scantlebury, 2011).

Conceptual Framework

In studies examining the relationship between science achievement and individual differences, it is seen that science achievement is mostly defined as the measurement of end of unit or chapter problem-solving ability related to subjects. Considering the 21st-century skills, the one-way approach to science achievement is quite limited and has some basic problems (Ates & Cataloglu, 2007; Pruitt, 2014). Evaluating students' science achievement in the context of scientific literacy, which emphasizes science education reforms in many countries, is necessary for realizing the visions of the curriculum.

In this study, science achievement was discussed in the dimension of scientific literacy. The historical development of the concept of scientific literacy suggests that there are many different definitions. According to the Project 2061 and Science for All Americans report, a scientifically literate individual is aware of the strengths and weaknesses of science, mathematics, and technology as well as interconnected human initiatives, comprehends basic science concepts and principles, knows the natural world, and recognizes both its diversity and unity. It is also a person who uses scientific knowledge in ways of scientific thinking for individual and social purposes. One of the most up-to-date definitions of scientific literacy that fits this definition is the one made by Fives et al. (2014). They define scientific literacy as the ability of individuals to have knowledge about the nature and processes of any field of science and thus to use science in daily life in a pragmatic and meaningful way.

There are other tools in the literature for the measurement of students' scientific literacy levels. One of these measurement tools is the Test of Basic Scientific Literacy developed by Laugksch and Spargo (1996). This scale includes questions about content knowledge, the nature of science, science-technology-society, and environment sub-dimensions. In addition to this scale, the Test of Scientific Literacy Skills developed by Gormally et al. (2012) is another scale used to measure the level of scientific literacy. This scale measures the skills of understanding research methods that provide access to scientific information, and the ability to organize, analyze, and interpret quantitative data, and scientific information. Another scale frequently used in the literature is Fives et al.'s (2014) Scientific Literacy scale. Fives et al. (2014) evaluated scientific literacy in terms of the role of science, scientific thinking, and acting, science and society, science media literacy, mathematics in science, motivation, and beliefs towards science.

In addition to these tools used in the measurement of scientific literacy are the Trends in International Mathematics and Science Study (TIMSS) and Program for International Student Assessment (PISA) exams conducted by the International Association for the Evaluation of Educational Achievement (IEA) and Organisation for Economic Co-operation and Development (OECD). PISA is international research conducted by the OECD in three-year cycles that evaluates the knowledge, and skills of 15-year-old students in certain fields. PISA research is carried out to measure the mathematical literacy, science literacy, and reading skills of students in the 15-age group who continue their formal education. In each cycle of the research, an area is selected as a weighted area, and in-depth analyzes are carried out in that area (TME, 2016). In the PISA 2015 research, the field of scientific literacy was chosen as the predominant field. The difference between gender groups among students' scientific literacy levels determined in PISA exams varies from country to country. When the scientific literacy performance of the PISA 2015 application is analyzed according to the gender variable, the average score of female students in all participating countries is higher than that of male students. But this difference is not statistically significant.

The same results are valid for the sample in Turkey; the difference between the mean scores of male, and female students in favor of females is not statistically significant (Tas et al., 2016). Similar results for Turkey were also seen in PISA scientific literacy measurements made in 2009 and 2012. In the PISA scientific literacy measurements conducted in 2015, the average of female students is higher than that of male students, but when the effect of other variables (economic, social, and cultural status, grade level, grade repetition) is controlled, it is seen that the average of boys is higher than that of females (Üstün et al., 2020).

Although individuals need to be scientifically literate both individually and socially, this is a challenging process. This situation may be caused by the functional definition of scientific literacy, teachers, teaching materials used, assessment and evaluation approaches, characteristics of the teaching environment, and the effects of individual differences, as well as the interaction of these variables with each other (Lipuma, 2008). Therefore, it is of paramount importance to investigate the effect of various variables such as individual differences on the development of scientific literacy.

Logical thinking ability is one of the variables frequently examined among individual differences in cognitive characteristics. (Ersanlı et al., 2018). According to Lawson (2004), logical thinking ability is the mental plans, strategies, or rules used to process information, and draw conclusions beyond directly observable experiences. Kuhn and Dean (2004) define logical thinking ability as the strategies, and rules that individuals use to ensure coordination between evidence, and theoretical hypotheses, and to create causal inferences. In both definitions, control of variables, proportional reasoning, correlation reasoning, and probabilistic reasoning, and hypothetical deductive reasoning are used as logical mathematical skills that assist scientific reasoning (Lawson, 2004; Zimmerman, 2000). Past studies in this field indicate that this ability begins to develop during adolescence and is a crucial factor affecting learning scientific concepts (Acar et al., 2015, Cheng et al., 2018, Lawson, 1995; Shayer & Adey 1992, Stender et al., 2018). When the change in the logical thinking ability of students in terms of gender groups was examined, it was found that in some studies there was a difference in favor of females (Demirtaş, 2011), in some studies in favor of boys (Valanides, 1996; Yenilmez et al., 2005), and in some studies, no difference was observed between the groups (Al-Zoubi et al., 2009; Piraksa et al., 2014; Talib et al., 2018; Yüzak, 2012).

Logical thinking ability is a concept related to the cognitive development of individuals, and one of the most accepted theories of cognitive development is Piaget's Theory of Cognitive Development (Lawson, 1995). The propositions of this theory have been prevalently used in many fields (Flavell, 1996). However, as a result of some studies to test Piaget's ideas, criticisms of the theory were put forward. One of the biggest criticisms of the theory is that it ignores the effect of individual differences on cognitive development. The theory cannot explain why some individuals progress faster during the cognitive development periods than others (Greenberg, 1995). Due to these findings, some studies preferred to examine and reinterpret Piaget's theory with new experimental findings. These researchers are called Neo-Piagetians (Knight & Sutton, 2004). Pascual-Leone's Constructivist Theory of Operators, the first Neo-Piagetian theorist, attributes the inability to perform activities with similar logical structures by students at the same developmental stage to the differences in students' mental capacities. Mental capacity refers to a limited resource that increases with age from childhood and can show variability within the age, enabling the activation of the schemas in mind related to the activity (Pascual-Leone & Johnson, 2005; 2011).

The science education literature review suggests that mental capacity is a cognitive variable that affects students' science achievement (Tsaparlis, 2005; Stamovlasis, 2010). In studies examining the difference between the mental capacities of different gender groups, it is seen that there is no statistically significant difference between the mental capacity scores of the gender groups (Hindal, 2007; Hindal et al., 2013). Pascual-Leone expresses that individuals' structural mental capacity cannot be used entirely during cognitive tasks, and various factors can affect their use of it. One of these factors is the cognitive styles of individuals (Pascual-Leone, 1970). According to Morgan (1997), cognitive styles are specific features that an individual uses in acquiring or learning information and can affect all individual activities. Although there are different classifications for cognitive styles in the literature, the most researched cognitive style in educational research is field-dependent/field-independent cognitive style classification (McKeachie & Svinicki, 2013). Field-independent learners are less affected by external stimuli in analyzing the complex structure of the field they are in and in the process of finding and extracting a particular element from a complex whole, compared to dependent learners who receive it from external stimuli. While field-dependent individuals give importance to external stimuli that affect their perceptions, field-independent individuals are more affected by internal stimuli rather than external stimuli (Jonassen & Grabowski, 1993).

When the literature on field-dependent/field-independent cognitive styles is examined, it is concluded that this variable is one of the strongest predictors of academic achievement (Terrell, 2002). Additionally, it is observed that field-independent students studying in some disciplines are more successful (Catalоğlu & Ates, 2014, Morris et al., 2019; Özarslan & Bilgin, 2016). In some studies, the tendency to have field-dependent/field-independent cognitive styles in terms of gender were examined, and no statistically significant difference was observed based on the scores obtained (Horzum & Alper, 2006; Maghsudi, 2007; Witkin et al., 1977). While in a few studies, it was found that male learners tended to be field-independent (Kirk, 2000; Onyekuru, 2015).

Mental rotation ability, another individual difference discussed in this research, is one of the vital components of spatial ability and is used extensively both in daily life and in courses such as biology, physics, chemistry, and geometry (Ganley et al., 2014). Meta-analysis and review papers also infer that mental rotation is an important factor for student success not only in science, but also in technology, engineering and mathematics (Castro-Alonso & Uttal, 2019; Maeda & Yoon, 2013; Langlois et al., 2019).In terms of the gender variable, it is stated that the difference between the scores obtained by gender groups is higher in favor of male learners (Hirnstein et al., 2009; Maeda & Yoon, 2013; Pietsch & Jansen, 2012; Titze et al., 2010). However, in all the studies examining the effect of individual differences mentioned above, it is seen that science achievement mainly focuses on specific dimensions of scientific literacy, such as knowledge or skill. In this paper, science achievement was defined in the dimension of scientific literacy, and its relationship with the specified variables was reexamined in terms of gender.

The Present Study

Gender has been one of the most widely researched variables for many years, considering that it affects achievement in science education research (Kahle & Meece, 1994; Scantlebury & Baker, 2007; Scantlebury, 2011). To this end, in order for countries to progress socially and economically, it is a fundamental requirement to prevent inequalities of opportunity in terms of gender in the education and training process (National Research Council [NRC], 1996; Next Generation Science Standards [NGSS] Lead States, 2013). One of the main purposes of research on gender is expressed as to reveal the performance differences/similarities between male and female students, and their reasons, and prevent inequalities of opportunity among students by attempting to understand the learning ways of students (OECD, 2009). By examining whether the model, as mentioned earlier, differs in terms of gender groups, this study aims to serve the main purpose of preventing possible inequalities of opportunity that may occur between groups. In this context, the research question and sub-problems of this study are stated below:

> Are the paths specified in the proposed theoretical model statistically significant in the samples formed by students of different genders?

In the context of this research question, the sub-problems related to the model are as follows:

a. Are the paths specified in the proposed theoretical model statistically significant in the sample of female students?
b. Are the paths specified in the proposed theoretical model statistically significant in the sample of male students?

Method

Research Model

A causal-comparative design was used in this study. This model is used to describe the causes or consequences of pre-existing differences between or among groups of individuals (Fraenkel et al., 2012).

Sample

The research population includes seventh-grade students studying in central districts of Ankara. Using the stratified random sampling method, 823 seventh-grade students studying in Ankara constitute the research sample. In this method, the population is divided into sub-units to represent itself. Elements are then sampled from each of these subunits. The inclusion of the elements in the subunits into the sample is made in proportion to the ratio of the subunits to the total population (Fraenkel et al., 2012). In the application of this sampling method, Ankara central districts were determined as sub-units. Considering the ratio of the number of students in the central districts of Ankara to the total number of students in the population, the number of students in each central district was determined. The students participating in the study consisted of 446 (52%) females and 377 (46%) males.

Data Collection Tools

Demonstrated Scientific Literacy (SL)

In this study, the Demonstrated Scientific Literacy Test, developed by Fives et al. (2014) for secondary school students, was used to determine the students' scientific literacy levels. The test was translated and adapted to Turkish by Sahin and Ates (2018). The test is an assessment tool consisting of multiple-choice questions, including items about the role of science, scientific thinking and doing, science and society, science media literacy, and mathematics in science. In the relevant study, the chi-square value (χ^2= 178.41, *N*=823, *df*=135, *p*=0.00) as a result of the CFA analysis performed to test the construct validity of the test, was statistically significant. The values of (χ^2/df) =1.32; RMSEA= 0.02; CFI= 0.97; TLI= 0.96; WRMR= 0.94 were obtained. These values show that the data fit the model well (Kline, 2005). In this study, the KR-20 reliability coefficient of the test was found to be 0.66. These results show that the measurement is valid and reliable (Alpar, 2013; Kline, 2005).

Test of Logical Thinking (TOLT)

The Test of Logical Thinking developed by Tobin and Capie (1981) was used to measure the logical thinking abilities of secondary school students in this study. The test consists of 10 items. The test was translated and adapted to Turkish by Geban et al. (1992) and the Cronbach α reliability coefficient was found to be 0.77. As a result of the CFA analysis performed to test the construct validity of the test in the relevant study, it is seen that the chi-square value (χ^2=43.55, *N*=790, *df*=31, *p*=0.00) was statistically significant. The values of (χ^2/df) =1.40; RMSEA= 0.02; CFI= 0.99; TLI= 0.98; WRMR= 0.77 were obtained. These values indicate that the data fit the model well. In this study, the KR-20 reliability coefficient of the test was found to be 0.63. These results show that the measurement is valid and reliable (Alpar, 2013; Kline, 2005).

The Group Embedded Figures Test (GEFT)

In this study, the Group Embedded Figures Test was used to determine students' field-dependent/field-independent cognitive styles (Witkin et al., 1971). The test consists of 18 items. The test was adapted to Turkish by Çakan (2003). In the relevant study, the chi-square value (χ^2= 396.83, *N*=804, *df*=135, *p*=0.00) was statistically significant in the CFA analysis performed to test the construct validity of the test. In addition to that, the values of (χ^2/df)=2.93; RMSEA= 0.05; CFI= 0.98; TLI= 0.97; WRMR= 1.28 were obtained. These values suggest that the data fit the model well. In this study, the KR-20 reliability coefficient of the test was found to be 0.89. These results show that the measurement is valid and reliable (Alpar, 2013; Kline, 2005).

Figural Intersection Test (FIT)

In this study, the Figural Intersection Test was used to determine the functional mental capacity of the students (Pascual-Leone & Johnson, 2011). The test consists of 36 items. The test was adapted to Turkish by Sahin (2018). As a result of the CFA analysis performed to test the construct validity of the test in this study, the chi-square value (χ^2=1099.38, *N*=785, *df*=594, *p*=0.00) was found to be statistically significant. Additionally, the values of (χ^2/df)=1.85; RMSEA= 0.03; CFI= 0.94; TLI= 0.94; WRMR= 1.34 were obtained. These values infer that the data fit the model well. In this study, the KR-20 reliability coefficient of the test was found to be 0.89. These results show that the measurement is valid and reliable (Alpar, 2013; Kline, 2005).

Mental Rotation Test (MRT)

The basic format of the Mental Rotation Test developed by Peters et al. (1995) was used to determine mental rotation abilities, a dimension of students' spatial abilities. The test consists of 24 items. The test was translated and adapted to Turkish by Yildiz and Tüzün (2011). In the relevant study, the chi-square value (χ^2= 482.69, *N*=759, *df*=252, *p*=0.00) was statistically significant in the CFA analysis performed to test the construct validity of the test. The values of (χ^2/df)=1.91; RMSEA= 0.04; CFI= 0.91; TLI= 0.90; WRMR= 1.15 were obtained. These values show that the data fit the model well. In this study, the KR-20 reliability coefficient of the test was found to be 0.77. These results show that the measurement is valid and reliable (Alpar 2013; Kline, 2005).

Application

The data collection process for this study was carried out by the researchers in the spring term of the 2015-2016 academic year. Before the implementation, official permissions regarding the implementation had been obtained from institutions affiliated with the TME and detailed planning was made. In the planning, the issues that may reduce student motivation regarding the implementation were taken into consideration. For example, no implementation was made during the exam weeks of the students or in their favorite sports lessons. For the same purpose, the data collection process was planned to last for 2.5 months, and the process was completed before the air temperatures increased. In addition, five measuring instruments were not applied one after the other to avoid test fatigue. Completion of implementation in a school was spread over two weeks. After the planning, the implementation was made in the schools at the specified times. Before the application, all participants were informed of the purpose and importance of the study. During the application, the participants were not forced, and only voluntary participants were included in the application. The scales of the participants who did not want to continue the application were deemed invalid. The tests used in the research were applied in the same order in all schools. The researcher took care to have the same attitude in all classes where data was collected. After the data collection process, the scoring of the tests was done by the researcher using pre-prepared answer keys. In this way, the threat of instrumentation, scoring, and participant attitude has been tried to be prevented.

Data Analysis

In this study, the SPSS 23 package program was used for descriptive analysis, while the Mplus 7.0 program was used to test the structural model in both gender groups.

Findings

Descriptive Analysis

Table 1 shows the descriptive statistics of the total scores obtained from the tests used in this study. When the skewness and kurtosis coefficients in the table are examined, it is seen that these coefficients are within the limits of ±1, indicating that the total scores show a normal distribution (Mertler & Vannatta, 2005).

Table 1

Descriptive Analysis of Tests

Variable	Group	*n*	Min	Max	Mean	Std.Dev	Skewness	Kurtosis
SL	Total	823	1	17	8.23	3.31	0.16	-0.60
	Female	446	2	17	9.01	3.21	0.13	0.23
	Male	377	1	17	8.05	3.36	0.27	-0.61
LT	Total	823	0	10	5.43	2.23	0.13	-0.59
	Female	446	1	10	5.50	2.21	0.11	-0.51
	Male	377	0	10	5.33	2.25	0.16	0.25
CS	Total	823	0	18	7.78	4.93	0.22	-0.98
	Female	446	0	18	8.07	4.91	0.14	-0.97
	Male	377	0	18	7.41	4.93	0.34	0.25
MC	Total	823	0	34	19.77	6.17	-0.34	-0.09
	Female	446	0	34	19.49	6.43	-0.37	0.01
	Male	377	1	32	20.06	5.81	-0.29	-0.38
MRA	Total	823	14	44	27.21	5.15	0.89	0.85
	Female	446	14	44	26.85	4.67	0.95	1.45
	Male	377	15	44	27.67	5.61	0.81	0.25

When Table 1 is examined, while female students have a higher mean score than male students in terms of scientific literacy, logical thinking ability, and cognitive styles, it is seen that the mean scores of male students are higher than female students in terms of mental capacity and mental rotation ability variables. Some of these differences between mean scores are statically significant, whereas the others are statistically insignificant.

Table 2

ANOVA Results of Test Means by Gender

Variable	Source	Sum of Squares	*df*	Mean Square	*F*	**p*
SL	Inter-group	190.69	1	190.69	17.73	.000*
	Intra-group	8831.18	821	10.76		
	Total	9021.87	822			
LT	Inter-group	6.66	1	6.66	1.34	.248
	Intra-group	4093.71	821	4.99		
	Total	4100.37	822			
CS	Inter-group	95.78	1	95.78	3.95	.047*
	Intra-group	19900.30	821	24.24		
	Total	19996.08	822			
MC	Inter-group	60.25	1	60.25	1.58	.209
	Intra-group	31253.12	821	38.07		
	Total	31313.36	822			
MRA	Inter-group	147.57	1	147.57	5.59	.018*
	Intra-group	21688.34	821	26.42		
	Total	21835.91	822			

As seen in Table 2, there was no significant difference between the mean scores of male and female students in terms of mental capacity scores (*F* (1, 821) =1.58; $p \geq 0.05$); there is a significant difference in favor of male students in terms of mental rotation ability (*F* (1, 821) =5.59; $p<0.05$). While the mean score difference between male and female students is significant in favor of female students in terms of scientific literacy (*F* (1, 821) =17.73; $p< 0.05$) and field-dependent/field-independent cognitive styles scores (*F* (1,821) =3.95; $p< 0.05$), there is no statistically significant difference between the mean scores in terms of logical thinking ability (*F* (1,821) =1.34; $p \geq 0.05$).

Testing the Structural Model on Male and Female Students

The structural model was tested using the Mplus software. In the analysis, it was seen that the chi-square value (χ^2= 212.46, *N*=377, *df*=203, *p*=0.31) was not significant in the sample of male students. Additionally, the values of (χ^2/df) =1.05; RMSEA= 0.01; CFI= 0.99; TLI= 0.99; WRMR= 0.79 were obtained. These values suggest that the data fit the model well. The chi-square value (χ^2= 262.79, *N*=445, *df*=203, *p*=0.00) was found to be significant in the sample consisting of female students. The values of (χ^2/df) =1.29; RMSEA= 0.03; CFI= 0.96; TLI= 0.95; WRMR= 0.91 were obtained and conclude that the data fit the model well. The direct, indirect, and total effects obtained as a result of the analyzes for both samples, and the percentages of variance explained by other variables for dependent and mediator variables in the model, are given in Table 3.

Table 3

Direct, Indirect, and Total Effects in the Model and Effect Sizes (Male and Female Students)

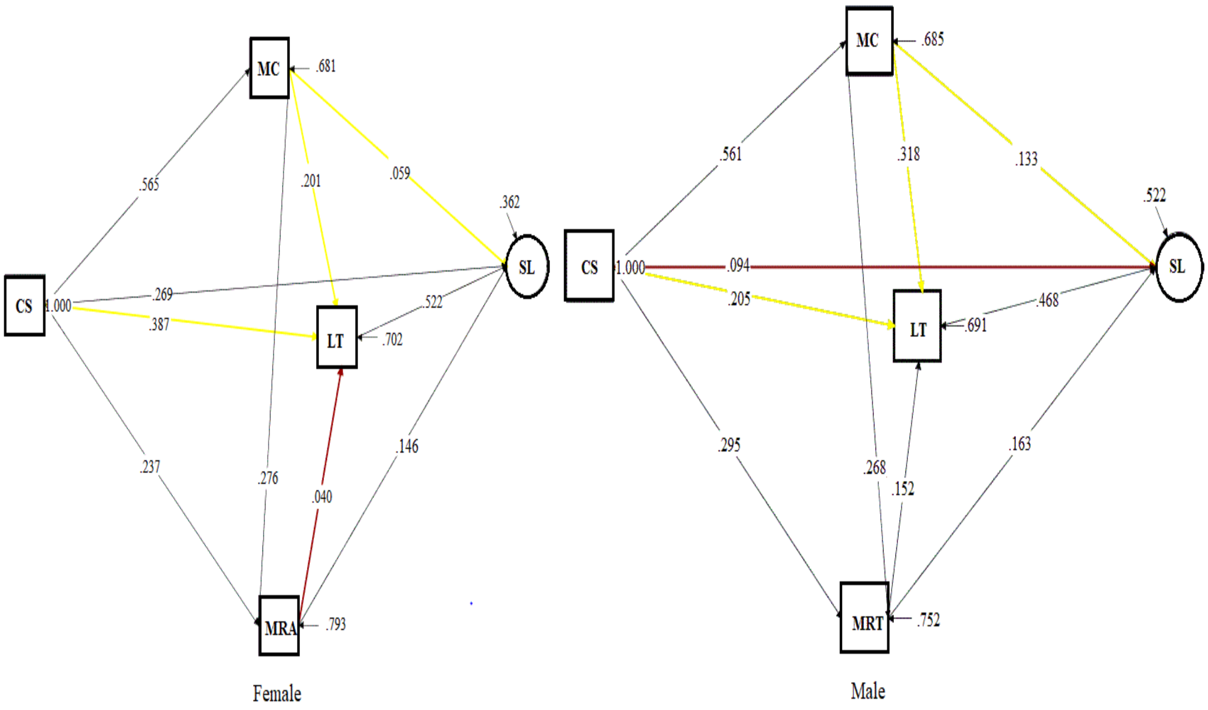
Dependent Variable	Independent Variable	Standardized Regression Coefficients					
		Direct		Indirect		Total	
		Male	Female	Male	Female	Male/ Effect size	Female/ Effect size
SL	LT	0.47*	0.52*	-	-	0.47*/Big	0.52*/Big
R^2=0.48	CS	0.09	0.27*	0.36*	0.36*	0.45*/Big	0.63*/Big
(Male)	MRA	0.16*	0.15*	0.07*	0.02	0.23*/Medium	0.17*/Medium
R^2=0.64	MC	0.13*	0.06	0.21*	0.15*	0.34*/Medium	0.21*/Medium
(Female)							
LT	CS	0.20	0.39 *	0.25*	0.13*	0.45*/ Big	0.52*/ Big
R^2=0.31	MC	*	0.20*	0.04*	0.01	0.36*/Medium	0.21*/Medium
(Male)	MRA	0.32*	0.04	-	-	0.15*/Small	0.04/ Small
R^2=0.30		0.15*					
(Female)							
MRA	CS	0.30*	0.24*	0.15*	0.15*	0.45*/ Big	0.39*/Medium
R^2=0.25	MC	0.27*	0.28*	-	-	0.27*/Medium	0.28*/Medium
(Male)							
R^2=0.21							
(Female)							
MC	CS	0.56*	0.57*	-	-	0.56*/ Big	0.57*/ Big
R^2=0.32							
(Male)							
R^2=0.32							
(Female)							

When the variables in Table 3 are considered in terms of the total predictive effect, it can be inferred that only the mental rotation skills of female students have no statistically significant effect on logical thinking ability (β=0.04, p>0.05). Apart from this, the predictive effects of scientific literacy, logical thinking ability, mental rotation ability, and mental capacity variables on total prediction by other variables in both samples are statistically significant. When Table 3 is reviewed in terms of its direct predictive effect, the effect of cognitive styles on the direct prediction of scientific literacy in the sample formed by male students (β= 0.09, p>0.05), the effect of mental capacity to directly predict scientific literacy (β= 0.06, p>0.05) and the effect of mental rotation ability to directly predict logical thinking ability (β= 0.04, p>0.05) in the sample of female students were not statistically significant. When Table 3 is addressed in terms of indirect prediction effect, in the sample formed by female students, the effect of mental rotation ability to indirectly predict scientific literacy (β= 0.02, p>0.05) and the effect of mental capacity to indirectly predict logical thinking ability (β= 0.01, p>0.05) was not found to be statistically significant. Apart from these, the predictive effects of scientific literacy, logical thinking ability, mental rotation ability, and mental capacity variables on indirect prediction by other variables in both samples are statistically significant.

As seen in Table 3, a total of 4 variables that predict scientific literacy in male students can explain 48% of the variance in scientific literacy. This value is 64% for female students. Three variables that predict logical thinking ability in male students can explain 31% of the variance in logical thinking ability. This value is 30% for female students. Two variables predicting mental rotation ability in male students can explain 25% of the variance in mental rotation ability. This value is 21% for female students. One variable predicting mental capacity in male students can explain 32% of the variance in mental capacity. This value is 32% for female students. When Table 3 is analyzed in terms of effect size created by the effects of the variables on the total prediction made, it is seen that the effect sizes are in the same category for all variables in both samples (Kline, 2005). The values in Table 3 are summarized in Figure 2.

Figure 2

Structural Model of the Relationships Between Scientific Literacy (SL), Logical Thinking Ability (LT), Cognitive Styles (CS), Mental Capacity (MC), and Mental Rotation Ability (MRA) in Gender Groups



Note: → Path coefficients with no significant predictive effect, → Path coefficients differing in size in both samples

One of the main differences seen in the model, which was tested separately for male and female students, is related to the direct predictive effect of the cognitive style variable on scientific literacy. In the sample of female students, the effect on this prediction was found to be statistically significant, while in the sample formed by male students, the effect on this prediction is not statistically significant. Another significance is related to the direct predictive effect of mental rotation ability on logical thinking ability. While the effect on this prediction was statistically significant in the sample formed by male students, the effect on this prediction is not statistically significant in the sample formed by female students. The direct predictive effect of mental capacity on cognitive literacy in the model also differs in both groups. While the effect on this prediction was statistically significant in the sample formed by male students, the effect on this prediction is not statistically significant in the opposite gender.

Discussion

In this study, the structural model that Sahin and Ates (2020) proposed with regards to the interaction between scientific literacy level and logical thinking ability, field-dependent/field-independent cognitive styles, mental capacities, and mental rotation abilities of seventh-grade students was examined in terms of male and female students. As a result of the research, 64% of the variance in the scientific literacy scores of the females and 48% of the variance in the scientific literacy scores of the male are explained by logical thinking ability, cognitive style, mental capacity, and mental rotation ability. As stated by Sahin and Ates (2020), this value was found to be 55% in the study in which male and female students were assessed together and all the predictive effects in the theoretical model were found to be statistically significant.

When both models were examined to explain the difference in the variance values that explain scientific literacy in gender groups in the research, it is noteworthy that there is a significant difference between the groups in terms of the effects of students' cognitive styles and mental capacities to directly or indirectly predict scientific literacy levels. This remarkable difference is also observed in terms of the predictive effect of mental rotation abilities on logical thinking abilities. When the direct effects were analyzed in the study, it was found that the cognitive styles of male students did not have a statistically significant effect on predicting scientific literacy. In contrast, in female students, it is seen that the effect of mental capacity to predict scientific literacy is not statistically significant. At the same time, the effect of mental rotation ability to predict logical thinking ability is not statistically significant. Besides these, there is no significant difference between gender groups.

The most notable result of those observed among gender groups in the study is that the effect of cognitive styles on the direct predictor of scientific literacy is not statistically significant in males but significant in females. The difference between field-dependent/field-independent cognitive styles of females and males is frequently mentioned in the relevant literature. In some of these studies, the difference between the scores of males and females in cognitive styles measurements was quite minor and not significant (Horzum & Alper, 2006; Idika, 2017; Maghsudi, 2007; Witkin et al., 1977), while in others this difference was stated to be higher in favor of males (Kirk, 2000; Onyekuru, 2015).

In the analysis of variance in the study, although the difference observed between field-dependent/field-independent cognitive styles scores between females and males was found to be significant in favor of females, this result, obtained through testing the theoretical model, provides information beyond whether there is a significant difference between the cognitive style scores of both genders. The result obtained by testing the structural model in this study proposes that the predictive effect of the increase in field-independent cognitive style tendencies of female students on the level of scientific literacy is higher and more significant than that of males. It is a key point to interpret this situation in terms of scientific literacy vision. To this end, the result of the research shows that the increase in cognitive styles scores of female students affects scientific literacy achievement more, thanks to the fact that the teaching environment and materials are designed according to the characteristics of different cognitive styles by being aware of the cognitive styles structure of learners.

In this respect, there are studies implying that planning strategies to be used in lessons mediate the effect of cognitive style on achievement (Tinajero et al., 2012). Hence, science teachers need to be aware of the barriers to learning arising from field-dependent/field-independent cognitive styles and consider this issue when presenting teaching materials to learners. Organizing the teaching materials to be used during teaching, and simplifying the irrelevant or complex contexts in the activities in order of importance, can facilitate the learning of dependent students. It is tough for field-dependent students to distinguish the important ones among the many pieces of information presented in the lectures in which the lecture method is applied. Using teaching methods with much social interaction, such as in-class discussion and collaborative teaching methods, could help prevent this disadvantage of field-dependent students (Jonassen & Grabowski, 1993). Although it is hard for teachers to apply the appropriate teaching method to each student in the classroom, it is thought that the disadvantages arising from the field-dependent/field-independent cognitive style could decrease with the diversification of the teaching and assessment techniques used.

Another remarkable finding in the study is related to the direct predictive effect of mental rotation ability on logical thinking ability. The results obtained by testing the structural model in this study show that the predictive effect of mental rotation ability in the theoretical model on logical thinking ability is not statistically significant for female students. Although this effect is significant for male students, it is categorized as a minor effect. In the analysis of variance in the study, it was found that the difference in performance between genders was significant in favor of males. These findings are consistent with the relevant literature (Hirnstein et al., 2009; Maeda & Yoon, 2013; Pietsch & Jansen, 2012; Priest, 2019; Titze et al., 2010). The result obtained by testing the structural model in the research provides information beyond whether there is a significant difference between the groups. The result obtained by testing the structural model in this study suggests that the predictive effect of the increase in male students' mental rotation ability performance on logical thinking ability is greater and more significant than that of the female students. It is reported in the literature that the difference observed between gender groups in mental rotation ability may be due to the learning opportunities that individuals had in the past and biological, environmental, and sociocultural factors (Schoning et al., 2007; Sundberg, 1994). Sorby (1999) asserts that playing with toys, playing with three-dimensional computer games, and doing some sports activities in childhood contribute to developing spatial ability. Although a difference was observed between gender groups in terms of mental rotation ability, students can develop this skill through appropriate activities (Miller & Halpern, 2013; Sanchez, 2012). Studies conclude that participation in spatial activities, especially at an early age, can affect students' performance in spatial activities. As reported in two meta-analysis studies (Baenninger & Newcombe, 1989; Uttal et al., 2013) conducted on the ability to improve mental rotation ability with teaching, while education focused on spatial activities and materials allow the development of mental rotation ability, this development also positively affects other components of spatial ability. Researchers also indicate that it is vital to start activities to develop spatial ability at an early age before the gender gap in spatial ability widens (Newcombe & Frick, 2010).

Another finding obtained among gender groups is related to the direct predictive effect of mental capacity on scientific literacy. While this effect is statistically significant for male students, it is insignificant for female students. Meanwhile, in the analysis of variance for mental capacity, it was found that the mean score difference in favor of male students between both genders was not significant. This result shows parallelism with studies in this field (Hindal, 2007; Hindal et al., 2013). One of the main results obtained by testing the structural model in this study is related to the predictive effect of mental capacity on scientific literacy rather than the differentiation found between gender groups. This result obtained from the research shows that the increase in functional mental capacity of students affects the scientific literacy performance of male students more than female students, thanks to the strategies that increase the functional mental capacity by removing unnecessary information from the activities organized for students and increasing the motivation of the students. An individual's mental capacity is explained as a cognitive variable that defines the ability to process many phenomena or concepts simultaneously (Pascual-Leone, 1970). According to the Constructivist Operators Theory, the more complex the activity, the greater the demand for mental capacity. As a matter of fact, studies tried to reduce the field effect by removing unnecessary information from the activities, and it was observed that there is an increase in student performance in this way (Boujaoude et al., 2004; Danili & Reid, 2004; Niaz, 1988a, b; Niaz & Robinson, 1992; Tsaparlis, 1998; Tsaparlıs et al., 1998). It is thought that the spatial effect could be reduced by alleviating unnecessary or extra information through the context or life-based activities in the lessons. In addition, by increasing the social context, there could be an increase in student motivation. In this way, it is thought that there may be an increase in student performance. As can be seen in the research results, although the structural mental capacity cannot be interfered, a functional mental capacity area can be created with the above-mentioned strategies. For this reason, it is thought that the concept of functional mental capacity can be known by the teachers and the implementation of the aforementioned strategies can contribute to the development of scientific literacy levels of all students.

Conclusions and Suggestions

This research concluded that cognitive styles and mental capacity variables differ according to gender groups in terms of their predictive effect on scientific literacy. Therefore, it is expressed that the preparation of the teaching materials (written, visual, and interactive) be used in the lessons by considering how these two variables play a key role in preventing inequalities of opportunity in terms of gender. For this purpose, field experts should examine the effects of textbooks and other teaching materials in terms of these variables and cross-check their efficiency.

Since this research was conducted in the central districts of Ankara, to generalize the research results to Turkey, it can be tested in model gender groups with students from different regions and at different grade levels. In the study, a model was created by considering only certain cognitive variables. To examine the scientific literacy levels of students in-depth, it is regarded as important to test new models in which only affective variables take place, or that cognitive and affective variables take place together in different gender groups.

The authors received no financial support for the research, authorship, and/or publication of this manuscript. This article is derived from Feride Sahin's Ph.D. dissertation entitled "An investigation of the relationship between seventh-grade students' scientific literacy and some cognitive variables" conducted under the supervision of Salih Ates.

Feride Sahin (feridecelik84@gmail.com) graduated from Selçuk University, Ahmet Keleşoğlu Education Faculty, Science Education Department. She received his master's degree in science education from Selçuk University. She received his doctorate degree in science education from Gazi University. Dr. Şahin is currently working at Manisa Celal Bayar University, Faculty of Education, Department of Science Education.

Salih Ates (s.ates@gazi.edu.tr) graduated from Gazi University, Gazi Education Faculty, Physics Teaching Department. He received his master's and doctorate degrees in Science Education from the University of Kentucky in the United States. Dr. Ates is currently working at Gazi University, Gazi Faculty of Education, Department of Science Education.

References

Acar, Ö., Türkmen, L., & Bilgin, A. (2015). Examination of gender differences on cognitive and motivational factors that ınfluence 8th graders' science achievement in Turkey. *Eurasia Journal of Mathematics, Science & Technology Education*, *11*(5) 1027-1040. https://doi.org/10.12973/eurasia.2015.1372a

Alpar, R. (2013). *Uygulamalı çok değişkenli istatistiksel yöntemler*. Detay.

Al-Zoubi, T., El-shar'a, I., & Al-Salam, M.K. (2009). The scientific reasoning level of students' in the faculty of science in Al-Hussein Bin Talal University and its affection of gender, teaching level, and specialization. *An-Najah University Journal for Research Humanities*, *23*(2), 401-437. https://journals.najah.edu/article/357/

Ates, S., & Cataloglu, E. (2007). The effects of students' reasoning abilities on conceptual understandings and problem-solving skills in introductory mechanics. *European Journal of Physics*, *28*(6), 1161. http://dx.doi.org/10.1080/02635140701250618.

Baenninger, M., & Newcombe, N. (1989). The role of experience in spatial test performance: A meta-analysis. *Sex Roles*, *20*(5), 327–344. https://doi.org/10.1007/BF00287729

BouJaoude, S., Salloum, S., & Abd-El-Khalick, F. (2004). Research report: Relationships between selective cognitive variables and students' ability to solve chemistry problems. *International Journal of Science Education*, *26*(1), 63-84. https://doi.org/10.1080/0950069032000070315

Cakan, M. (2003). Cross-cultural aspect of the group embedded figures test: Norms for Turkish eighth graders. *Perceptual and Motor Skills*, *97*(2), 499-509. https://doi.org/10.2466/pms.2003.97.2.499

Castro-Alonso, J. C., & Uttal, D. H. (2019). Science education and visuospatial processing. In C. A. Juan (Ed.), *Visuospatial processing for education in health and natural sciences* (pp. 53-79). Springer.

Cataloglu, E., & Ates, S. (2014). The effects of cognitive styles on naive impetus theory application degrees of pre-service science teachers. *International Journal of Science and Mathematics Education*, *12*(4), 699-719. https://doi.org/10.1007/s10763-013-9430-z

Cheng, S. C., She, H. C., & Huang, L. Y. (2018). The impact of problem-solving instruction on middle school students' physical science learning: Interplays of knowledge, reasoning, and problem solving. *Eurasia Journal of Mathematics, Science and Technology Education*, *14*(3), 731-743. https://doi.org/10.12973/ejmste/80902

Danili, E., & Reid, N. (2004). Some strategies to improve performance in school chemistry based on two cognitive factors. *Research in Science and Technological Education,* 22, 203–23. https://doi.org/10.1080/0263514042000290903

Dimitrov, D. M. (1999). Gender differences in science achievement: Differential effect of ability, response format, and strands of learning outcomes. *School Science and Mathematics, 99*(8), 445-450. https://doi.org/10.1111/j.1949-8594.1999.tb17507.x

Ersanlı, E., Gencoglu, C., & Duran, V. (2018). A meta-synthesis study on investigations of dissertations on education regarding reasoning skills in Turkey. *International Online Journal of Educational Sciences*, *10*(5). http://dx.doi.org/10.15345/iojes.2018.05.015

Fives, H., Huebner, W., Birnbaum, A. S., & Nicolich, M. (2014). Developing a measure of scientific literacy for middle school students. *Science Education*, *98*(4), 549-580. https://doi.org/10.1002/sce.21115

Flavell, J.H. (1996) Piaget's legacy, *Psychological Science*, *7*, 200–203. https://www.jstor.org/stable/pdf/40062945.pdf.

Fraenkel, J. R., Wallen, N. E., & Hyun, H. H. (2012). *How to design and evaluate research in education*. McGraw-Hill.

Ganley, C. M., Vasilyeva, M., & Dulaney, A. (2014). Spatial ability mediates the gender difference in middle school students' science performance. *Child development*, *85*(4), 1419-1432. https://doi.org/10.1111/cdev.12230

Geban, Ö., Askar, P. & Özkan, I. (1992). Effects of computer simulated experiment and problem solving approaches on students' learning outcomes at the high school level. *Journal of Educational Research*, *86*(1), 5-10. https://doi.org/10.1080/00220671.1992.9941821

Gormally, C., Brickman, P., & Lutz, M. (2012). Developing a test of scientific literacy skills (TOSLS): Measuring undergraduates' evaluation of scientific information and arguments. *CBE—Life Sciences Education*, *11*, 364-377. https://www.researchgate.net/publication/233880797

Greenberg, D. (1995). *Free at last: The Sudbury Valley School*. Sudbury Valley School.

Hindal, H. S. (2007). *Cognitive characteristics of students in middle schools in State of Kuwait, with emphasis on high achievement* [Doctoral dissertation, Glasgow University]. Glasgow University Theses and Dissertations Archive. https://theses.gla.ac.uk/5717/1/2007hindalphd.pdf

Hindal, H., Reid, N., & Whitehead, R. (2013). Gender and learner characteristics. *European Journal of Educational Research*, *2*(2), 83-96. https://doi.org/10.12973/eu-jer.2.2.83

Hirnstein, M., Bayer, U. & Hausmann, M. (2009). Sex-specific response strategies in mental rotation. *Learning and Individual Differences*, *19*, 225-228. https://doi.org/10.1016/j.lindif.2008.11.006.

Horzum, M. B., & Alper, A. (2006). Fen Bilgisi dersinde olaya dayalı öğrenme yöntemi, bilişsel stiller ve cinsiyetin öğrenci başarısına etkisi. *Ankara Üniversitesi Eğitim Bilimleri Fakültesi Dergisi*, *39*(2), 151-175. https://app.trdizin.gov.tr/makale/TnpRNU16TXo/fen-bilgisi-dersinde-olaya-dayali-ogrenme-yontemi-bilissel-stilin-ve-cinsiyetin-ogrenci-basarisina-etkisi

Idika, M. I. (2017). Influence of cognitive style and gender on secondary school students' achievement in and attitude to chemistry. *Advances in Social Sciences Research Journal*, *4*(1), 129-139. https://doi.org/10.14738/assrj.41.2585

Jonassen, D. H., & Grabowski, B. (1993). *Handbook of ındividual difference learning, and ınstruction*. Routledge.

Kahle, J.B & Meece, J. (1994). Research on gender issue in the classroom. In D. L. Gabel (Eds). *Handbook of Research on Science Teaching and Learning*, (pp.542-558). Macmillan.

Kirk, G. R. (2000). *The relationship of attitudes toward science, cognitive style and self-concept to achievement in chemistry at the secondary school level* [Doctoral dissertation, Florida International University]. Florida International University ProQuest Dissertations Publishing. https://www.proquest.com/openview/ac0e6532411650cdeed81f4b17185c83/1?pq-origsite=gscholar&cbl=18750&diss=y

Kline, R.B. (2005). *Principles and practice of structural equation modeling*. The Guilford.

Kuhn, D. & Dean, D. (2004). Connecting scientific reasoning and causal inferences. *Journal of Cognition and Development, 5*(2), 261–288. https://doi.org/10.1207/s15327647jcd0502_5

Langlois, J., Bellemare, C., Toulouse, J., & Wells, G. A. (2019). Spatial abilities training in anatomy education: A systematic review. *Anatomical Sciences Education*. *13*(1), 71-79. https://doi.org/10.1002/ase.1873

Laugksch, R. C., & Spargo, P. E. (1996). Development of a pool of scientific literacy test-items based on selected AAAS literacy goals. *Science Education*, *80*(2), 121-143. https://doi.org/10.1088/0963-6625/5/4/003

Lawson, A. E. (1995). *Science teaching and the development of thinking: Stages in the development of procedural knowledge*. Wadsworth.

Lawson, A. E. (2004). The nature and development of scientific reasoning. *International Journal of Science and Mathematics Education, 2*(3), 307–338. https://link.springer.com/content/pdf/10.1007/s10763-004-3224-2.pdf

Lipuma, M., (2008, May 27). *Obstacles to science literacy.* https://webcache.googleusercontent.com/search?q=cache: TTARK WcvAJ:https://web.njit.edu/~lipuma/LipumaJMObstacles to Scientific Literacy.doc+&cd=1&hl=tr&ct=clnk&gl=tr

Maeda, Y., & Yoon, S.Y. (2013). A meta-analysis on gender differences in mental rotation ability measured by the Purdue spatial visualization tests: Visualization of rotations (PSVT:R). *Educational Psychology Review, 25*, 69-94. https://doi.org/10.1007/s10648-012-9215-x

Maghsudi, M. (2007). The interaction between field dependent/independent learning styles and learners' linguality in third language acquisition. *Interactive Multimedia Electronic Journal of Computer-Enhanced Learning*, *7*(5), 1-23. http://ut.pr/biblioteca/Glossa2/Journal/dec2007/Linguality%20in%20Third%20Language%20Acquisition.pdf

McKeachie, W., & Svinicki, M. (2013). *Mckeachie's teaching tips: strategies, research, and theory for college and university teachers*. Wadsworth Cengage Learning.

Miller, D. I., & Halpern, D. F. (2013). Can spatial training improve long-term outcomes for gifted STEM undergraduates? *Learning and Individual Differences, 26, 141*–152. https://doi.org/10.1016/j.lindif.2012.03.012.

Morgan, H. (1997). *Cognitive styles and classroom learning*. Praeger.

Morris, S., Farran, E. K., & Dumontheil, I. (2019). Field-independence associates with mathematics and science performance in 5-to 10-year-olds after accounting for domain-general factors. *Mind, Brain, and Education*. *13*(4), 268-278. https://doi.org/10.1111/mbe.12214

National Research Council. (1996). *The national science education standards*. National Academy.

Newcombe, N. S., & Frick, A. (2010). Early education for spatial intelligence: Why, what, and how. *Mind, Brain, and Education, 4* (3), 102–111. https://doi.org/10.1111/j.1751-228X.2010.01089.x

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

Niaz, M. (1988a). The information processing demand of chemistry problems and its relation to Pascual-Leone's functional M-capacity. *International Journal of Science Education, 10*, 231–38. http://dx.doi.org/10.1080/0950069880100211.

Niaz, M. (1988b). Manipulation of M-demand of chemistry problems and its effects on students performance: A neo-Piagetian study. *Journal of Research in Science Teaching, 25*, 643–57. http://dx.doi.org/10.1002/tea.3660250804.

Niaz, M., & Robinson, W. R. (1992). Manipulation of logical structure of chemistry problems and its effect on student performance. *Journal of Research in Science Teaching, 29,* 211–226. https://doi.org/10.1002/tea.3660290303

OECD. (2009). *Equally prepared for life? How 15-year-old boys and girls perform in school.* http://www.pisa.oecd.org/dataoecd/59/50/42843625.pdf

Onyekuru, B. U. (2015). Field dependence-field ındependence cognitive style, gender, career choice and academic achievement of secondary school students in emohua local government area of rivers state. *Journal of Education and Practice*, *6*(10), 76-85. http://files.eric.ed.gov/fulltext/EJ1081668.pdf

Özarslan, M., & Bilgin, İ. (2016). Öğrencilerin alan bağımlı/bağımsız bilişsel stillerinin ve bilimsel düşünme yeteneklerinin maddenin doğası kavramlarını anlamalarına ve fen dersine yönelik tutumlarına etkisi. *Mustafa Kemal Üniversitesi Sosyal Bilimler Enstitüsü Dergisi*, *13*(33), 94-110. https://dergipark.org.tr/tr/pub/mkusbed/issue/19579/208938

Pascual-Leone, J. (1970). A mathematical model for the transition rule in Piaget's developmental stages. *Acta Psvchologica*, *32*, 301-345. https://doi.org/10.1016/0001-6918(70)90108-3.

Pascual-Leone, J., & Johnson, J. (2005). A dialectical constructivist view of developmental intelligence. In O. Wilhelm & R. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 177-201). Sage.

Pascual-Leone, J., & Johnson, J. (2011). A developmental theory of mental attention: Its application to measurement and task analysis. In P. Barrouillet & V. Gaillard (Eds.), *Cognitive development and working memory: A dialogue between neoPiagetian theories and cognitive approaches,* (pp.13-46). Psychology.

Pietsch, S., & Jansen, P. (2012). Different mental rotation performance in students of music, sport and education. *Learning and Individual Differences, 22*(1), 159-163. https://doi.org/10.1016/j.lindif.2011.11.012

Piraksa, C., Srisawasdi, N., & Koul, R. (2014). Effect of gender on student's scientific reasoning ability: A case study in Thailand. *Procedia-Social and Behavioral Sciences*, *116,* 486-491. https://doi.org/10.1016/j.sbspro.2014.01.245

Priest, M. J. (2019). *Taking gender out of mental rotation: reexamining gender differences in mental rotation using virtual reality and 3d models* [Unpublished doctoral dissertation]. University of Wyoming.

Pruitt, S. L. (2014). The next generation science standards: The features and challenges. *Journal of Science Teacher Education*, *25*(2), 145-156. https://doi.org/10.1007/s10972-014-9385-0

Sahin, F. (2018). Yedinci sınıf öğrencilerinin bilimsel okuryazarlık düzeyi ile bazı bilişsel degişkenler arasındaki ilişkinin incelenmesi. [Doctoral dissertation, Gazi University]. Turkey Higher Education Council Thesis Center. https://tez.yok.gov.tr/UlusalTezMerkezi/tezSorguSonucYeni.jsp

Sahin, F., & Ates, S. (2018). Ortaokul öğrencilerine yönelik bilimsel okuryazarlık ölçeği adaptasyon çalışması. *Gazi Eğitim Fakültesi Dergisi*, *38*(3), 1173-1205. https://doi.org/10.17152/gefad.406601

Sahin, F., & Ates, S. (2020). Examination of the relationship between seventh-grade students' scientific literacy among certain cognitive variables. *Education And Science*, *45*(203), 63-89. http://dx.doi.org/10.15390/EB.2020.8552

Sanchez, C. A. (2012). Enhancing visuospatial performance through video game training to increase learning in visuospatial science domains. *Psychonomic Bulletin & Review, 19* (1),58–65. https://doi.org/10.3758/s13423-011-0177-7

Scantlebury K. & Baker D. (2007). Gender issues in science education research: Remembering where the difference lies. In S. K. Abel, & N. G. Lederman (Eds). *Handbook of Research on Science Education,* (pp.257-286). Lawrence.

Scantlebury, K. (2011). *Still part of the conversation: gender ıssues in science education*. In B. J. Fraser, K. Tobin, & C. J. McRobbie (Eds). *Second International Handbook of Science Education*, (pp.499-512). Springer.

Schlatter, E., Lazonder, A. W., Molenaar, I., & Janssen, N. (2021). Individual differences in children's scientific reasoning. *Education Sciences*, *11*(9), 471. https://doi.org/10.3390/educsci11090471

Schmeck, R. R., & Grove, E. (1979). Academic achievement and individual differences in learning processes. *Applied Psychological Measurement*, *3*(1), 43-49. https://doi.org/10.1177/014662167900300106

Schöning, S., Engelien, A., Kugel, H., Schäfer, S., Schiffbauer, H., Zwitserlood, P., & ... Konrad, C. (2007). Functional anatomy of visuo-spatial working memory during mental rotation is influenced by sex, menstrual cycle, and sex steroid hormones. *Neuropsychologia*, *45*(14), 3203-3214. 10.1016/j.neuropsychologia.2007.06.011

Shayer, M., & Adey, P. S. (1992). Accelerating the development of formal thinking in middle and high school students III: Testing the permanency of effects. *Journal of Research in Science Teaching*, *29*(10), 1101-1115. https://doi.org/10.1002/tea.3660291007

Sorby S. (1999). Developing 3-D Spatial visualization skills. *Engineering Desing Graphics Journal*, *63*(2), 21-32. http://diggingdeeper.pbworks.com/.

Stamovlasis, D. (2010). Methodological and epistemological issues on linear regression applied to psychometric variables in problem solving: Rethinking variance. *Chemistry Education Research and Practice*, *11*(1), 59-68. https://doi.org/10.1039/C001048K

Stender, A., Schwichow, M., Zimmerman, C., & Härtig, H. (2018). Making inquiry-based science learning visible: The influence of CVS and cognitive skills on content knowledge learning in guided inquiry. *International Journal of Science Education*, *40*(15), 1812-1831. https://doi.org/10.1080/09500693.2018.1504346

Sundberg, S. E. (1994). *Effect of spatial training on spatial ability and mathematical achievement as compared to traditional geometry instruction* [Doctoral dissertation, University of Missouri]. University of Missouri ProQuest Dissertations Publishing. https://www.proquest.com/openview/e180e40a661b32cf3d509c8003755af3/1?pq-origsite=gscholar&cbl=18750&diss=y

Talib, C.A., Rajan, S.T., Hakim, N.W.A., Malik, A.M.A., Siang, K.H., & Ali, M. (2018, November 8-9). *Gender difference as a factor in fostering scientific reasoning skills among students*. [Paper presentation]. IEEE 10th International Conference on Engineering Education, Kuala Lumpur, Malaysia.

Taş, U. E., Arıcı, Ö., Ozarkan, H. B., & Özgürlük, B. (2016*). Pısa 2015 Ulusal Raporu*. MEB.

Terrell S. T., (2002), *The use of cognitive style as a predictor of membership in middle and High School Programs for the academically gifted.* [Paper presentation]. The Annual Meeting of the American Educational Research Association, New Orleans, Louisiana.

Tinajero, C., Lemos, S. M., Araujo, M., Ferraces, M. J., & Paramo, M. F. (2012). Cognitive style and learning strategies as factors which affect academic achievement of Brazilian University Students. *Psicologia-Reflexao E Critica*, *25*(1), 105–113. https://doi.org/10.1590/S0102-79722012000100013

Titze, C., Jansen, P. & Heil, M. (2010). Mental rotation performance in fourth graders: No effects of gender beliefs (yet?). *Learning and Individual Differences*, *20*, 459-463. https://doi.org/10.1016/j.lindif.2010.04.003.

Turkish Ministry of Education. (2016). *Uluslararası öğrenci değerlendirme programı PISA 2015 ulusal raporu,* MEB.

Turkish Ministry of Education. (2017). *Science lesson curriculum (primary and secondary school 3, 4, 5, 6, 7 and 8th grades.* Talim ve Terbiye Kurulu Başkanlığı.

Tobin, K. G., & Capie, W. (1981). The development and validation of a group test of logical thinking. *Educational and Psychological Measurement*, *41*(2), 413-423. https://doi.org/10.1177/001316448104100220

Tsaparlis, G. (1998). Dimensional analysis and predictive models in problem solving. *International Journal of Science Education, 20*, 335–50. http://dx.doi.org/10.1080/0950069980200306.

Tsaparlis, G. (2005). Non-algorithmic quantitative problem solving in university physical chemistry: a correlation study of the role of selective cognitive factors. *Research in Science & Technological Education*, *23*(2), 125-148. http://dx.doi.org/10.1080/02635140500266369.

Tsaparlis, G., M. Kousathana, and M. Niaz. (1998). Molecular-equilibrium problems: Manipulation of logical structure and of M-demand, and their effect on student performance. *Science Education, 82*, 437–54.

Uttal, D. H., Meadow, N. G., Tipton, E., Hand, L. L., Alden, A. R., Warren, C., & Newcombe, N. S. (2013). The malleability of spatial skills: A meta-analysis of training studies. *Psychological Bulletin*, *139* (2), 352–402. https://doi.org/10.1037/a0028446

Üstün, U., Özdemir, E., Cansız, M., & Cansız, N. (2020). What are the factors affecting Turkish students' science literacy? A hierarchical linear modelling study using PISA 2015 data.

Hacettepe University Journal of Education, 35(3), 720-732. https://doi.org/10.16986/HUJE.2019050786

Valanides, N.C. (1996). Formal reasoning and science teaching. *School Science and Mathematics*, *96*(2), 99-107. https://doi.org/10.1111/j.1949-8594.1996.tb15818.x

Witkin, H. A., Moore, C. A., Goodenough, D. R., & Cox, P. W. (1977). Field-dependent and field-independent cognitive styles and their educational implications. R*eview* of *Educational Research, 7*(1), 1-64. https://doi.org/10.2307/1169967

Witkin, H. A., Oltman, P. K., Raskin, E., & Karp, S. A. (1971). *A manual for the embedded figures test*. Consulting Psychologist.

Yenilmez, A., Sungur, S., & Tekkaya, C. (2005). Investigating students' logical thinking abilities: The effects of gender and grade level. *Hacettepe University Journal of Education*, *28*(28), 219-225. https://eric.ed.gov/?id=EJ808059

Yıldız, B., & Tüzün, H. (2011). Üç-boyutlu sanal ortam ve somut materyal kullanımının uzamsal yeteneğe etkileri. *Hacettepe Üniversitesi Eğitim Fakültesi Dergisi*, *41*, 498-508. http://dergipark.ulakbim.gov.tr.

Zimmerman, C. (2000). The development of scientific reasoning skills. *Developmental Review, 20*(1), 99–149. https://doi.org/10.1006/drev.1999.0497