THE CASE FOR AN INTERDISCIPLINARY MODEL OF INQUIRY STEM INSTRUCTION FOR 21ST CENTURY STUDENT SUCCESS

A Dissertation Presented to the Faculty of the

University of Virginia School of

Education and Human Development

In Partial Fulfillment of Requirements for the Degree

Doctor of Philosophy

by

Ashley Elizabeth Hunt, M.A.

May 2021

Abstract

Improving the quality of science, technology, engineering, and mathematics (STEM) education has been a consistent research and policy focus of the 21st century. Of particular importance is the need to create equitable STEM learning experiences that accommodate the growing cultural, linguistic, and ethnic diversity in classrooms across the county. This three-manuscript dissertation makes the case for an instructional framework for elementary STEM that allows students to develop and practice social and emotional skills, acquire conceptual knowledge, and understand how what they learn applies to their own lives and communities. The first study used discourse analysis to characterize small group discussions in a sample of fourth grade science classrooms. When students effectively used social gestures, they were better equipped to engage in productive science discussions. The second study explored the association between student perception of classroom climate and their engagement in science. Multilevel regression analyses found a positive association between social support and engagement that was stronger in more linguistically diverse classrooms. The final study examined whether fifth grade mathematics teachers were observed using reform instruction and emotionally supportive interactions. Results highlighted that educators were most likely to use equity-promoting instruction in schools that served more of the economically privileged families in the district. Taken together, the three manuscripts provide an empirical foundation for an interdisciplinary model of elementary-level instruction that advances scholarly understanding of the inextricable link between social and emotional development and authentic STEM learning experiences.

Education, Leadership, Foundations, and Policy
School of Education and Human Development
University of Virginia
Charlottesville, VA

APPROVAL OF THE DISSERTATION

This dissertation, *The Case for an Interdisciplinary Model of Inquiry STEM Instruction for 21st Century Student Success*, has been approved by the Graduate Faculty of the University of Virginia School of Education and Human Development in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

____
Dr. Sara E. Rimm-Kaufman (Chair)

____
Dr. Jennifer L. Chiu

____
Dr. Natalia Palacios

____
Dr. Erik A. Ruzek

____
Date

DEDICATION

This dissertation is dedicated to the memory of more than 250,000 Americans who died of COVID-19 in the time between my proposal and defense, the professionals in STEM fields devoted to understanding, containing, and treating the virus, and the educators who have worked tirelessly to teach and care for their students through a global crisis.

ACKNOWLEDGEMENTS

I have so many people to thank for getting me this far.

My advisor Dr. Sara Rimm-Kaufman has been my #1 champion and partner from day one. I will never be able to fully express my gratitude for your time, care, and mentorship. Anything I accomplish in my career will be a reflection of the support and wisdom you have wrapped around me through each step.

In addition to their time and expertise as members of my committee, I have to thank: Dr. Erik Ruzek, for his unwavering encouragement through so many methods; Dr. Natalia Palacios, for teaching me to speak with purpose and confidence (and to breathe occasionally); and Dr. Jennie Chiu, for her presence throughout this journey as an instructor, as the second reader of my comprehensive exams, and finally as part of my dissertation committee.

My mom, who missed my defense to cover for a fellow teacher in quarantine, has taught me to live with intention and integrity through her enormous capacity for believing in the good of others. I am endlessly grateful for her and the rest of my huge, mixed, family: my grandparents, Robert and Mary Hunt and Lawrence House; my aunts, Katie Whitfield, Lauren Richardson, Marianne Spence, Jeanie House, and Nya House, who have loved me like their own. Thank you to my younger siblings – Sam, Hannah, Joe, Sean, and Brian – for being my lifelong case studies on child development. I am so proud of each of you.

I am thrilled to acknowledge Dr. Tatiana Hill and Dr. Ann Partee as my colleagues, friends, and cohort of champions that have been by my side since our first visit to grounds as prospective students. It has been a privilege to live and learn alongside you and the rest of the brilliant EP-ADS community and VEST fellows both current and past.

An enormous thanks to the Connect Science development team: to Dr. Eileen Merritt, who showed me how to approach research with curiosity and wonder; to Dr. Candy Lapan, for 40 hours in a rental van packed with school supplies; and to Rebecca McGregor for being one of the most kind and hardworking people I've ever known.

Thank you to my friends, colleagues, and teachers – to Eileen, Kathleen, and John Waters, and the rest of the Waters family; to Kelsey Kern, who I will never be able to thank enough so I might as well stop trying; to the Sal's gals; to Ben Throckmorton for making sure my dogs and I were walked regularly; to Eliana Saltares, Lorena Ortiz, Cheri Fancsali; to my partners at the Discovery School, Meredith Evans and Erin Fahey; and to Brandon Daughtry Slocum and Joe Leticia.

Thank you to the Institute for Education Sciences for funding this work through Grants R305B140026 and R305A150272. The opinions expressed throughout are those of the author(s) and do not represent views of the U.S. Department of Education.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Linking Document

Improving learning experiences and opportunities in the fields of science, technology, engineering, and mathematics (STEM) has been a consistent focus for K-12 educators and scholars alike throughout the 21st century. Federal education policies like the Next Generation Science Standards (NGSS, NRC, 2012a) and recommendations from organizations like the National Council for the Teaching of Mathematics (NCTM, 2012; NCTM, 2014) share several common themes. Recommendations are built around practices instead of individual content strands and procedural approaches (Merritt, Palacios, Banse, Rimm-Kaufman, & Leis, 2017; Schoenfeld, 2020). Instruction is designed to be student centered rather than teacher centered, creating an environment where learning is driven by peer collaboration and discourse. Perhaps most importantly, educators are urged to situate STEM learning in the real world and to link academic content to students' lived experiences (Colley & Windschitl, 2016). Together, these instructional themes comprise a concept-driven approach to active learning often described as *inquiry* teaching and learning. Transforming STEM education through a focus on inquiry has the potential to generate equitable learning experiences to better accommodate the racial, cultural, and linguistic diversity represented in U.S. schools (Tang, el Turkey, Cilli-Turner, Savic, Karakok, & Plaxco, 2017).

Inquiry learning necessitates communication and self-regulatory skills that have not been traditionally considered targets of STEM instruction. Even though social and emotional functioning is widely accepted as predictive of student success in school and beyond, (Greenberg, Weissberg, & O'Brien, 2003) limited research has explicitly examined the overlap between social and emotional learning instruction (SEL) and STEM. Most frameworks for STEM learning implicate or even explicitly describe social and emotional skills that underlie engaged participation in learning. For example, the NRC describes *21st century skills* as the cognitive, interpersonal, and intrapersonal domains to promote deep learning (NRC, 2012b). Behaviors and attitudes like respectful communication, perseverance, and social engagement help students to make meaning of STEM content (Hesse, Care, Buder, Sassenberg, & Griffi, 2017; Hewlett Foundation, 2013). As developmental and learning sciences have increasingly recognized the importance and cognitive demand of social and emotional behaviors, terms like *noncognitive skills* and *soft skills* have become less common. Nevertheless, they still appear (and are of increasing interest) in the study of economics and workforce development.

Prioritizing inquiry changes the role of the classroom social environment in the context of learning. For example, positive teacher-student relationships are instructional assets that give educators information about student ideas, experiences, and identities that can be leveraged towards making content personally relevant. Furthermore, students in inquiry learning classrooms must be able to work together, talk about their ideas, and consider different approaches and perspectives (NCTM 2014; NRC, 2012a). The variety of social competencies that underlie inquiry STEM instruction present an opportunity for students to develop and practice new skills beyond traditional models of academic achievement and performance.

Mirroring their distinct lines of research, SEL and STEM are generally viewed as distinct (and often competing) aspects of teacher preparation and professional learning. As a result, STEM instructional methods rarely embed support for cultivating social skills and a culture that allows for actively engaged participation. This dissertation builds a theoretical model of inquiry instruction based on empirical evidence of the interplay between STEM and SEL observed in classroom-based research. Taken together, the three manuscripts present a rich, methodologically diverse investigation of the intersection between development, inquiry STEM learning, and student success. The findings build on a growing body of interdisciplinary research establishing student-centered, equitable STEM learning opportunities as a possible driver of student academic and developmental progress.

Conditions for Effective Inquiry Instruction

This work highlights three conditions shared at the intersection between STEM and SEL: classrooms must be relational, equitable, and participatory. The *relational* dimension of the classroom refers to necessary social conditions like positive a climate and productive peer interactions. The *equitable* dimension ensures equal access to challenging and engaging STEM learning experiences that communicate respect for students' academic, linguistic, and cultural backgrounds. Finally, the *participatory* dimension reflects a culture of curiosity that prompts active involvement and learning experiences that hinge on full participation. Participatory STEM extends beyond typical definitions of engagement by designing lessons that give students experiences of collective action and civic engagement (Stitzlein, 2017). The Relational-Equitable-

Participatory instructional framework seeks to characterize classroom environments where students develop social-emotional skills and competencies, acquire meaningful conceptual knowledge, and learn how STEM practices and approaches apply to problems in their own lives and communities.

Relational

Challenging inquiry learning tasks require students to work together, support their claims with evidence, and take action towards solving problems (Duschl & Osborne, 2002). Unsurprisingly, collaboration among students more effectively meets learning goals when children experience positive relationships and interactions with each other and their teacher (Pianta et al, 2003). The social climate and quality of teacher-student relationships shape the relational conditions of a STEM learning environment (Hoisington & Winokur, 2019; Kochenderfer-Ladd & Ladd, 2016).

One way the relational dimension of the classrooms shapes inquiry STEM instruction is through discourse. Classic work describes the presence of two types of discourse in classrooms: academic and social (Cummins, 1979). Academic discourse refers to the vocabulary and discipline-specific phrases associated with content, whereas social discourse refers to the relational aspects of communication necessary to sustain a successful dialog. For example, a science activity where students engage in argument from evidence requires the combination of both types of discourse.

Imagine that students are discussing the pros and cons of different energy sources. Students need to understand both the meaning of “renewable” and “non-renewable” and how the words relate to energy sources to participate in the discussion. Academic discourse could emerge through student use of vocabulary related to energy science and production of electricity. At the same time, social discourse plays a role. As students are learning and using new vocabulary, they might make a mistake like conflating natural gas and gasoline. The way students choose to correct a peer informs what happens next: the remaining conversation unfolds much differently if a student accepts the feedback from their peer than if they become defensive or upset about being corrected or questioned further.

Equitable

Positioning equity as a central goal of STEM instruction is essential. Importantly, this means designing accessible materials and activities for a given classroom or group, not over-simplifying concepts or holding lower expectations of students with unique learning needs (NCTM, 2014). Eliciting and incorporating students' prior knowledge and experiences represents one inquiry STEM practice with clear connections to equity (NCTM, 2014; NRC, 2012a). Returning to the student who mistook gasoline for natural gas, this time the teacher intervenes to determine what caused the misconception. Most elementary students are probably with the idea of using gasoline to fuel cars but might not have ever seen or heard of natural gas. Both include the word "gas," so it is conceivable they might be similar or the same. Teachers who are skilled at leveraging students' prior knowledge and past experiences use errors and disagreements among students to make their own thinking visible and explicit.

Teachers can also promote equitable STEM learning experiences by using a variety of grouping strategies to give students opportunities to learn with and from their peers. For example, a teacher might sometimes pair an Englisher Learner (EL) student with a native English speaker, allowing each child to practice the skill of working with someone different from themselves. On other occasions, the teacher might pair EL students and allow them to work in their shared native language. This grouping choice honors students' linguistic identities and allows them to communicate concepts that they might not be ready to express (or in some cases, do not even exist) in English.

Participatory

Situating STEM learning in socially or culturally relevant problems is one approach to promoting a participatory learning environment. Inquiry learning experiences that help students understand real problems like climate change and the overuse of non-renewable resources can empower them to think about the actions they could take to implement solutions to the problem in the future (Tang et al., 2017). For example, in one fourth-grade inquiry science curriculum, students meet content learning standards about energy and resources by investigating energy use in their own homes and communities (Connect Science, 2018). The focus on relatable, personally relevant STEM problems in inquiry instruction is also related to promoting civic engagement (Condon & Wichowsky, 2018).

Teachers can promote participatory STEM learning through the thoughtful design of learning activities. Think again of the students using evidence to build arguments about energy sources. When students engage directly with others about a topic in small groups or pairs, the expectation is that each group member contributes to complete the assignment. However, if the entire class watches two students debate energy sources in from of their classmates, the remaining students act as spectators rather than active participants. To build participatory attitudes towards STEM learning, a teacher might use both methods: students build their arguments and identify evidence in small groups that they then share in class-wide discussion. Since each group contributes to reaching consensus about a set of statements about energy sources, individual students can consider the perspectives of their peers and consolidate their own thinking.

Equity and Emergent Bilingualism

Identifying meaningful engagement as a central component of inquiry instruction raises questions about the role of individual differences that may contribute to students' willingness to fully engage in STEM learning. One important student characteristic that plays a significant role in ethnic and racial disparities in education is the EL status of emergent bilingual children. Nearly half of districts and schools in the U.S. enroll students that receive EL services, increasingly in classrooms alongside their English-fluent peers (Larkin, 2020; Zehler, Fleischman, Hopstock, Stephenson, Pendick, & Sapru, 2003;). Almost *every* elementary school teacher is a teacher of students eligible for EL services, and the vast majority do not have the cross-cultural knowledge or language ability to provide bilingual instruction. This is especially the case in under-resourced communities where families with financial insecurity, limited education, or recent immigration status are most likely to live and work (Lewis, Ream, Bocian, Cardullo, Hammond, & Fast, 2012). Given the rapid growth of linguistic diversity across the country, more generalized instructional guidance is needed to support mainstream teachers who struggle to differentiate their STEM instruction for emergent bilingual children in their classrooms.

Some elementary STEM interventions integrate explicit EL supports (e.g., the Promoting Science among English language learners project, Maerten-Rivera, Ahn, Lanier, Diaz, & Lee, 2016). This dissertation calls attention to how inquiry practices have the potential to create more equitable STEM learning environments for language diverse classrooms that deliberately acknowledge the diversity of cultural and ethnic identities that students bring to the classroom that accompany their EL status (Lewis et al., 2012).

An Interdisciplinary Approach to 21st Century Student Success

As with SEL, limited work has explicitly considered how to support emergent bilingual children in challenging STEM learning environments where teachers use inquiry practices. Nevertheless, the overlap between the two as they operate in elementary STEM classrooms is undeniable. Taking an interdisciplinary approach to investigating these phenomena provides two key advantages when conceptualizing pathways to 21st century success. First, the work can draw on existing research from a spectrum of perspectives and methods (Ottmar, 2019). Considering the results of designs from randomized-controlled trials (RCTs) to case studies creates a comprehensive picture of the instruction and outcomes valued in each field. Second, this approach allows for more seamless translation between disciplines that describe similar practices and outcomes but fail to reach practitioners. Integrating instructional recommendations to align with the goals of a relational, equitable, and participatory inquiry STEM instruction can refine recommendations for best practices at the elementary level. The field needs a greater understanding of the classroom conditions under which children develop academic, social, and emotional competencies. Furthermore, it is important to identify features of ambitious STEM instruction that contribute to engagement and positive academic and social outcomes for all students.

Description of Manuscripts

This dissertation synthesizes results from three manuscripts about inquiry STEM learning in upper elementary classrooms. The first study explored how science discourse, cognitive demand, and problem solving emerged during science discussions among fourth grade students. The second study measured the association between social support, collaborative learning, and academic engagement and the implications of learning in a positive social climate for EL students. The final study examined whether observed teacher use of reform mathematics practices and emotionally supportive interactions in fifth grade classrooms related to improved outcomes for students. Taken together, the manuscripts comprise an empirical foundation for an interdisciplinary model of elementary-level inquiry STEM instruction that is relational, equitable, and participatory.

Manuscript 1: Mixed Methods Analysis of Academic and Social Discourse in Science

The first study used discourse analysis to characterize how small groups of students used scientific argumentation practices and social gestures during science discussions. Conversation topics included energy sources, the relationship between objects in space, the pros and cons of technology, and structural and behavioral adaptations in animals. Qualitative analyses of conversations in 14 fourth-grade classrooms revealed the interplay of social and academic science discourse practices.

Results indicated that students in intervention classrooms had more productive, respectful conversations about their ideas. In contrast, content-related discourse was observed less frequently in comparison classrooms, where students were highly focused on negotiating elements of the task, (e.g., deciding who would draw each piece of a diagram). Descriptive analyses reflecting the frequency and content of student use of argumentation practices and social gestures provided quantitative evidence of the themes and patterns observed during qualitative analyses.

Results of these sequential mixed methods analyses suggested that when students effectively used social gestures, they were better equipped to engage in productive science discussions. Furthermore, significant differences between the intervention and comparison group indicated that giving teachers access to NGSS-aligned professional learning and materials was associated with more effective teaching and learning through scientific argumentation. Students were more likely to have productive science discussions when they were focused on a single, clear question, had access to reference materials, and did not have to produce a final product beyond the group conversation.

Manuscript 2: Multilevel Analysis of the Association Between Classroom Climate and Academic Engagement in Science

Theory and research suggest that a variety of teaching practices and individual attributes contribute to student engagement (Saeki & Quirk, 2015). The second study investigated the association between classroom climate and concurrent student academic engagement in science. This study used two student-report surveys to examine how much of the variance in social and behavioral engagement were explained by classroom- versus student-level factors. Multilevel regression analyses described the association between student-reported classroom climate and academic engagement in science. Moderation analyses investigated how the role of positive social support varied between classrooms that enrolled EL students versus those that were 100% English-fluent.

Results indicated the importance of positive, proactive social supports in the classroom to promote behavioral and social engagement in science. Model results indicated that, while holding other covariates constant, a one-point increase in positive support related to a 0.19-point increase in behavioral engagement and a 0.42-point increase in social engagement, respectively (*SE*s < .05; unstandardized betas). Conversely, negative social experiences were not significant. Findings showed that variance in engagement was explained almost entirely at the student level. Nevertheless, climate at the classroom-level was more predictive of social, (ICC = .03) than behavioral (ICC = .00) engagement in science. Finally, the association between positive support and engagement in science was stronger in linguistically diverse classrooms. In 100% English-fluent classrooms, a one-point increase in positive support was, on average, associated with a 0.17-point increase in behavioral engagement ($p < .05$). However, in classrooms where more than half of students received EL services, the same increase in social support was associated with a 0.35-point increase in behavioral engagement ($p < .001$). A pattern of similar magnitude and positive direction was observed for social engagement.

On one hand, results indicate that a positive climate might help students overall, and ELs in particular, engage more fully in science learning. On the other, it also indicates that students in less supportive environments might do little to engage in science learning. This decreased participation has serious potential consequences for learning and developmental outcomes, especially for students from historically marginalized backgrounds.

Manuscript 3: Examination of the Effect and Accessibility of Reform Mathematics Instruction in Emotionally Supportive Classrooms

Reform mathematics instruction (NCTM, 2012; NCTM, 2014) places unique demands on students – they need to take intellectual risks and drive their own learning. It is important that students experience certain social supports for authentic inquiry learning to happen. To investigate how reform mathematics instruction and emotional support emerged in the real world, the study used observation data from fifth grade mathematics classrooms that participated in an RCT of an SEL intervention.

Fifth grade teachers (N = 59) and their students (N = 387) were sampled from a large suburban school district. Classroom observations were coded using the M-Scan to quantify use of reform mathematics practices and the Classroom Assessment Scoring System (CLASS) to assess emotionally supportive interactions. Factor analyses fit a two-factor solution to the observation data, which was used as a measure of equity-promoting mathematics practice in analyses.

Data collected using the M-Scan and CLASS observation protocols yielded a psychometrically sound two-factor model of equity-promoting practices that consisted of reform mathematics practices and emotionally supportive interactions. Observed use of equity-promoting mathematics practices was only mildly associated with student outcomes and did not explain a significant portion of the variance in student achievement or social skills at the end of the school year. Comparison of school, family, and student characteristics among classrooms above and below average for observed use of equity-promoting practices found that the schools where teachers were least likely to use the observed practices were also served more of the low-income families in the district. Findings from the study highlight opportunities for measurement of equity-promoting STEM instructional approaches in future research to better understand what conditions prompt educators to use ambitious mathematics instruction in diverse classrooms.

Summary

Taken together, the manuscripts provide an empirical foundation for an interdisciplinary model of elementary-level inquiry STEM instruction that is relational, equitable, and participatory. Through a variety of methods, data sources, and analytic approaches, this work establishes foundational support for instruction that integrates STEM content with SEL instruction in an upper elementary context.

References

Colley, C., & Windschitl, M. (2016). Rigor in Elementary Science Students' Discourse: The Role of Responsiveness and Supportive Conditions for Talk. *Science Education*, *100*(6), 1009–1038. https://doi.org/10.1002/sce.21243

Condon, M. (2018). *Developing citizen-scientists*. *119*(2).

Cummins, J. (1979). Linguistic interdependence and the educational development of bilingual children. *Review of Educational Research, 49,* 222-251.

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, *38*(1), 39–72. https://doi.org/10.1080/03057260208560187

Greenberg, M. T., Weissberg, R. P., O'Brien, M. U., Zins, J. E., Fredericks, L., Resnik, H., & Elias, M. J. (2003). Enhancing School-Based Prevention and Youth Development Through Coordinated Social, Emotional, and Academic Learning. *American Psychologist*, *58*(6–7), 466–474. https://doi.org/10.1037/0003-066X.58.6-7.466

Harkins, T., Merritt, E., Rimm-Kaufman, S.E., Hunt, A. & Bowers, N. (2019). *Connect Science*. Curriculum Manual, University of Virginia, Arizona State University, & Harkins Consulting, LLC, Charlottesville, Virginia.

Hesse, F.W., Care, E., Buder, J., Sassenberg, K, & Griffi, P. (2017). A framework for teachable collaborative problem solving skills. *Assessment and Teaching of 21st Century Skills* (37-56). Dordrecht: Springer.

Abry, T., Rimm-Kaufman, S. E., Larsen, R. A., & Brewer, A. J. (2013). The influence of fidelity of implementation on teacher-student interaction quality in the context of a randomized controlled trial of the Responsive Classroom approach. *Journal of School Psychology*, *51*(4), 437–453. https://doi.org/10.1016/j.jsp.2013.03.001

Caraballo, L. (2014). The student as an assemblage of success: Constructing multiple identities amidst classes, gendered, and raced discourses of achievement. *Curriculum and Teaching Dialogue*, *16*(1), 103–123.

Colley, C., & Windschitl, M. (2016). Rigor in Elementary Science Students' Discourse: The Role of Responsiveness and Supportive Conditions for Talk. *Science Education*, *100*(6), 1009–1038. https://doi.org/10.1002/sce.21243

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, *38*(1), 39–72. https://doi.org/10.1080/03057260208560187

Henderson Pinter, H., Merritt, E. G., Berry, R. Q., & Rimm-Kaufman, S. E. (2018). The importance of structure, clarity, representation, and language in elementary mathematics instruction. *Investigations in Mathematics Learning*, *10*(2), 106–127. https://doi.org/10.1080/19477503.2017.1375354

Jacob, R., Hill, H., & Corey, D. (2017). The Impact of a Professional Development Program on Teachers' Mathematical Knowledge for Teaching, Instruction, and Student Achievement. *Journal of Research on Educational Effectiveness*, *10*(2), 379–407. https://doi.org/10.1080/19345747.2016.1273411

Jone, S. M., Barnes, S. P., Bailey, R., & Doolittle, E. J. (2017). Promoting social and emotional competencies in elementary school. *Future of Children*, *27*(1), 49–72.

https://doi.org/10.1353/foc.2017.0003

Kochenderfer-Ladd, B., & Ladd, G. W. (2016). Integrating academic and social-emotional learning in classroom interactions. In K. R. Wentzel & G. B. Ramani (Eds.), *Handbook of Social Influences in School Contexts: Social-Emotional, Motivation, and Cognitive Outcomes* (pp. 349–366). Routledge. https://doi.org/10.4324/9781315769929.ch22

Martin, D. P., & Rimm-Kaufman, S. E. (2015). Do student self-efficacy and teacher-student interaction quality contribute to emotional and social engagement in fifth grade math? *Journal of School Psychology*, *53*(5), 359–373. https://doi.org/10.1016/j.jsp.2015.07.001

Merritt, J., Lee, M. L., Rillero, P., & Kinach, B. M. (2017). Problem-based learning in K-8 mathematics and science education: A literature review. *Interdisciplinary Journal of Problem-Based Learning*, *11*(2), 5–17.

Ottmar, E. (2019). The effects of deeper learning opportunities on student achievement: Examining differential pathways. *Psychology in the Schools*, *56*(5), 840–855. https://doi.org/10.1002/pits.22237

Saeki, E., & Quirk, M. (2015). Getting students engaged might not be enough: the importance of psychological needs satisfaction on social-emotional and behavioral functioning among early adolescents. *Social Psychology of Education*, *18*(2), 355–371. https://doi.org/10.1007/s11218-014-9283-5

Schoenfeld, A. H. (2020). Mathematical practices, in theory and practice. *ZDM Mathematics Education*, *52*(6), 1163–1175. https://doi.org/10.1007/s11858-020-

01162-w

Tang, G., El Turkey, H., Cilli-Turner, E., Savic, M., Karakok, G., & Plaxco, D. (2017). Inquiry as an entry point to equity in the classroom. *International Journal of Mathematical Education in Science and Technology*, *48*, S4–S15. https://doi.org/10.1080/0020739X.2017.1352045

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting Positive Youth Development Through School-Based Social and Emotional Learning Interventions: A Meta-Analysis of Follow-Up Effects. *Child Development*, *88*(4), 1156–1171. https://doi.org/10.1111/cdev.12864

Kochenderfer-Ladd, B., & Ladd, G. W. (2016). Integrating academic and social-emotional learning in classroom interactions. In K. R. Wentzel & G. B. Ramani (Eds.), *Handbook of Social Influences in School Contexts: Social-Emotional, Motivation, and Cognitive Outcomes* (pp. 349–366). Routledge. https://doi.org/10.4324/9781315769929.ch22

Larkin, D.B. (2020). *Teaching Science in Diverse Classrooms: Real Science for Real Students* (1st ed.). Routledge. https://doi.org/10.4324/9780429199721

Lewis, J. L., Ream, R. K., Bocian, T. M., Cardullo, R. A., Hammond, K. A., & Fast, L. A. (2012). Con Cariño: Teacher caring, math self-efficacy, and math achievement among hispanic english learners. *Teachers College Record*, *114*(7).

Maerten-Rivera, J., Ahn, S., Lanier, K., Diaz, J., & Lee, O. (2016). Effect of a Multiyear Intervention. *The Elementary School Journal*, *116*(4), 600–624.

Merritt, E. G., Palacios, N., Banse, H., Rimm-Kaufman, S. E., & Leis, M. (2017).

Teaching practices in Grade 5 mathematics classrooms with high-achieving English learner students. *Journal of Educational Research*, *110*(1), 17–31. https://doi.org/10.1080/00220671.2015.1034352

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

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

National Research Council. (2012b). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: The National Academies Press.

O'Conner, R., De Feyter, J., Carr, A., Luo, J. L., & Romm, H. (2017). A review of the literature on social and emotional learning for students ages 3–8: Characteristics of effective social and emotional learning programs (part 1 of 4). *Regional Educational Laboratory Mid-Atlantic*, *(REL 2017–245)*, 1–31. https://ies.ed.gov/ncee/edlabs/projects/project.asp?projectID=443

Ottmar, E. (2019). The effects of deeper learning opportunities on student achievement: Examining differential pathways. *Psychology in the Schools*, *56*(5), 840–855. https://doi.org/10.1002/pits.22237

Pianta, R. C., & La Paro, K. (2003). Improving early school success. *Educational Leadership*, 60(*7*), 24–29.

Saeki, E., & Quirk, M. (2015). Getting students engaged might not be enough: the importance of psychological needs satisfaction on social-emotional and behavioral functioning among early adolescents. *Social Psychology of Education*, *18*(2), 355–371. https://doi.org/10.1007/s11218-014-9283-5

Sandilos, L. E., Baroody, A. E., Rimm-Kaufman, S. E., & Merritt, E. G. (2020). English learners' achievement in mathematics and science: Examining the role of self-efficacy. *Journal of School Psychology*, *79*(June 2018), 1–15. https://doi.org/10.1016/j.jsp.2020.02.002

Schoenfeld, A. H. (2020). Mathematical practices, in theory and practice. *ZDM Mathematics Education*, *52*(6), 1163–1175. https://doi.org/10.1007/s11858-020-01162-w

Stitzlein, S.M. (2017). *American public education and the responsibility of its citizens.* Oxford University Press.

Tang, G., El Turkey, H., Cilli-Turner, E., Savic, M., Karakok, G., & Plaxco, D. (2017). Inquiry as an entry point to equity in the classroom. *International Journal of Mathematical Education in Science and Technology*, *48*, S4–S15. https://doi.org/10.1080/0020739X.2017.1352045

Zehler, A. M., Fleischman, H. L., Hopstock, P. J., Stephenson, T. G., Pendick, M. L., & Sapru, S. (2003). *Policy report: Summary offindings related to LEP and Sp-Ed LEP students*. Washington, DC: U.S. Department of Education.

“Because the sun is really not that big”: An exploration of fourth graders tasked with arguing from evidence

Ashley Hunt, University of Virginia

Sara E. Rimm-Kaufman, University of Virginia

Eileen G. Merritt, Arizona State University

Nicole Bowers, Arizona State University

Abstract

Little is known about integration of social and emotional learning (SEL) with science instruction. We used a sequential mixed methods design to examine: 1) How do fourth graders use argumentation practices and social gestures in science class? and, 2) How do argumentation practices and social gestures differ between intervention and comparison classrooms? Intervention classrooms implemented Connect Science. Fourteen student conversations in seven classrooms were coded for argumentation practices (i.e., claims, evidence, questions) and social gestures (i.e., agreement, disagreement, assertive speech, prosocial speech). Across all classrooms, science conversations were most productive when students used social gestures to support use of argumentation practices. Without social gestures, conversations were disconnected or highly assertive. Proportionally, Connect Science students discussed science content more and spent less time discussing logistics than comparison students. Findings include recommendations for conditions (i.e., SEL instruction, science reference materials, and time) to enhance scientific discourse and argumentation in elementary school classrooms.

Manuscript 1: "Because the sun is really not that big": An exploration of fourth graders tasked with arguing from evidence

> "Learning the fine art of speaking with the possibility of being heard, and listening with the possibility of being changed, is a practical contribution to finding one's way in a wildly diverse democracy."
>
> -Bill Ayers, *I Shall Create!*

More and more, society expects teachers to guide students' social and emotional development in addition to their academic learning. Nevertheless, students often progress through the public education system without acquiring critical social skills that underlie personal and professional success (NRC, 2012b). The social-emotional learning (SEL) research literature and practice guidelines identify effective approaches to supporting students' awareness of and ability to manage their attention, thoughts, and behaviors in school and beyond (CASEL, 2012; Durlak, Weissberg, Dymnicki, Taylor, & Schellinger, 2011; Jones, Barnes, Bailey, & Doolittle, 2017; Taylor, Oberle, Durlak, & Weissberg, 2017). Several decades of research have established that students who develop SEL skills are also more likely to flourish academically (Durlak et al., 2011; Taylor et al., 2017).

A growing body of work shows the potential for SEL curricula to teach social and emotional skills that transfer to countless settings throughout students' lives. Programs with evidence of effectiveness like PATHS (Promoting Alternative Thinking Strategies) and Positive Action provide teachers a structured curriculum of SEL lessons and activities addressing explicit social and emotional skills, behaviors, and attitudes (Kam, Greenberg, & Walls, 2003; Lewis et al., 2016). Teachers lead activities like helping students recognize and label emotions or role-playing interpersonal situations (Jones et al., 2017; Rimm-Kaufman & Hulleman, 2015). However, the most clearly stated mission for schools remains focused on academic performance.

New work explores approaches that integrate SEL practices with academic instruction (Harris et al., 2015; NRC, 2012b). The links between SEL and science instruction in particular have noteworthy implications for practice. The Next Generation Science Standards (NGSS) position science as a discovery process through which students engage in scientific and engineering practices (SEPs) as they investigate phenomena (NRC, 2012a). While typical, teacher-centered science instruction has focused on memorization of vocabulary and following rigid procedures, (Banilower et al., 2018; Reiser, 2013; Trygstad, Smith, Banilower, & Nelson, 2013) NGSS describes a student-centered approach where students learn by engaging with scientific practices (NRC, 2012a).

For example, scientific argumentation requires students to build collective knowledge about a topic supported by evidence through verbal or written arguments, implicating the use of social gestures like voicing agreement and disagreement, (Hmelo-Silver, Duncan, & Chinn, 2007). The SEPs described by NGSS place explicit demands on students' ability to articulate their understanding (or lack thereof) through discourse. Despite the conceptual overlap and evidence that "dual-purpose" approaches to instruction that predict improved academic, social, and emotional outcomes (Kochenderfer-Ladd & Ladd, 2016), science educators rarely integrate the goals of SEL.

Study Purpose

Young students need opportunities to build and practice the complex skills that allow them to respectfully engage in authentic science practices. At the same time, teachers need support to meet the ambitious learning goals of NGSS. The current study uses an exploratory sequential mixed methods design to describe how small groups of fourth grade students used argumentation practices and social gestures to facilitate rich science discussions. We also explore how conversations differed when teachers used Connect Science, an NGSS-aligned curriculum that integrated SEL instruction, versus typical science instruction.

Review of Literature

Sociocultural learning theory asserts that a climate of openness, mutual respect, and normative communication helps students feel comfortable taking the social risks associated with making and evaluating scientific arguments (Colley & Windschitl, 2016; Vygotsky, 1978). Science discussions help students learn when group members verbally engage in a "shared thinking process" (Rogoff & Toma, 1997). Mortimer & Scott (2003) operationalized such a process as *dialogic-interactive* (D-I); engaging multiple voices in talk about multiple ideas generates knowledge, whereas one-sided interactions about a single viewpoint do not. Accordingly, D-I discourse serves as the foundation for productive discussions.

NGSS Implementation

Typical science instruction in elementary classrooms often consists of vocabulary, memorization of facts, and other activities with limited conceptual depth. The Next

Generation Science Standards (NGSS) call for the use of practices that promote inquiry, creative problem-solving, and collaborative engagement with science content (NRC, 2012a). The standards provide a framework for creating educational experiences that mirror authentic practices of scientists through three dimensions: science and engineering practices, (SEPs) disciplinary core ideas, and cross-cutting concepts. However, assessments and pacing guides have also been slow to adapt to NGSS-aligned instruction, (Trygstad et al., 2013). Standards have not been paired with necessary materials or professional support, an especially pressing problem in a time where the majority of the teaching force has not received training on inquiry approaches and three-dimensional instruction during their pre-service education (Pasley, Trygstad, & Banilower, 2016).

Despite many states adopting standards informed by the NGSS since their release, science education in practice has, on average, changed little. While implementation combines the three dimensions of NGSS, our study of the teaching and learning of argumentation focuses on a subset of SEPs that, taken together, distinguish *communication* in science as a necessary aspect of *learning* in science:

1. Asking questions and defining problems (SEP 1)
2. Obtaining, evaluating, and communicating information from scientific texts (SEP 8)
3. Constructing explanations and designing solutions to explain phenomena or solve problems (SEP 6)
4. Engaging in argument from evidence (SEP 7)

Science instruction based in these practices requires students to demonstrate their understanding through discussion, which necessitates a higher level of language and communication skills, (Lee, Miller, & Januszyk, 2014). Teachers report insufficient time for quality implementation and feeling "overwhelmed and intimidated," (p. 809) by NGSS (Hanuscin & Zangori, 2016; Penfield & Lee, 2010). Although resources to support implementation have become increasingly available, (e.g., NRC, 2015) they have focused on describing experiences that promote intended learning outcomes; they do not provide instructional guidance on how to boost students' SEL skills to make NGSS come alive in the classroom. Leveraging knowledge of effective SEL programs and practices might help teachers become comfortable with more time and language-intensive approaches to science teaching and learning.

Argumentation in Elementary Science

Scientists develop ideas through collaboration and critique in writing and through dialog. The ability to discuss and improve scholarly work, often referred to as "talking science," (Lemke, 1998), is itself a critical skill for scientists beyond their content expertise. *Argumentation* is a specific type of science discourse during which individuals demonstrate content understanding through the use of evidence, explanation, and reasoning (Erduran & Jimenez-Aleixandre, 2008). From a sociocultural perspective, scientific argumentation provides participants a conversational framework for establishing and evaluating ideas as a group (Tippett, 2009).

One way to describe effective scientific discourse uses a framework described by Mortimer & Scott (2003). Students' interactions in the classroom are classified across two dimensions: they can be *dialogic*, representing multiple ideas or perspectives, or *authoritative*, representing a single point of view; they are also *interactive*, engaging multiple voices, or *non-interactive*, characterized by a single voice. Accordingly, the framework describes four communicative approaches:

1) dialogic-interactive (D-I), or multiple ideas presented by multiple voices;

2) authoritative-interactive (A-I), or one idea presented by multiple voices;

3) dialogic-noninteractive (D-N), or multiple ideas presented by a single voice (for example, a student makes and evaluates a claim without response or interruption); and

4) authoritative-noninteractive (A-N), or one idea presented by a single voice (for example, a student makes a claim that receives no response).

Interactions categorized as A-I are the most common in traditional instruction and discourse in the classroom: one voice (usually the teacher's) imparts facts and knowledge to students (Tippett, 2009). For example, a specific type of A-I interaction known as the Inquire-Response-Evaluate (IRE) pattern has been observed commonly in classrooms, (Cazden, 2001). Conversely, student-centered science instruction prompts predominantly D-I conversations where students learn by and through conversations with each other and their teacher (Driver, Newton, & Osborne, 2000; Manz, 2015).

The educational value of conversation goes beyond sharing facts; it also reflects a group's ability to reach scientific understanding. Over time and with practice, students can begin to use argumentation practices and social gestures to communicate and refine their understanding of science content (Berland, 2011; Duschl & Osborne, 2002). While student-centered instruction might be more effective for engaging students in deep science learning than typical instructional approaches, guiding the development of skills that underlie productive discussions places unique demands on educators in the classroom, (Driver et al., 2000; Hayes & Trexler, 2016; NRC, 2012b). Since the ability to make and evaluate arguments from evidence implicates both communication skills and scientific knowledge, argumentation represents an opportunity to explore the integration of science and SEL instruction.

Social-Emotional Learning

Social-emotional learning (SEL) refers to a growing body of work that identifies effective ways of developing students' awareness of and ability to manage their attention, thoughts, and behaviors in school. The Collaborative for Academic, Social, and Emotional Learning (CASEL) developed one of the most widely-used frameworks to characterize social-emotional skills as comprised of five key competencies: self-awareness, self-management, social awareness, relationship skills, and responsible decision-making (CASEL, 2017). Importantly, the ability to learn and apply the skills, knowledge, and attitudes associated with SEL have intrinsic value as standalone competencies (Jones et al., 2017). In addition to developmental benefits, research has generated substantial evidence that implementation of programs and approaches to school-level SEL relate to a variety of positive academic and performance outcomes for students.

Summarizing decades of evaluation research, a meta-analysis of 213 studies of K-12 school-based SEL interventions found significant effects at follow up on academic and social and outcomes. On average, students in intervention schools outperformed comparison peers academically by 11-percentile points and were more likely to exhibit positive social behaviors like interpersonal problem solving and perspective taking (Durlak et al., 2011). A second meta-analysis of later follow up (6 months to 18 years later) found that SEL interventions had long-term impacts on a variety of outcomes, including positive peer relationships and less involvement with the justice system (Taylor et al., 2017).

Teachers influence students' social and emotional experiences in the classroom through their instruction, relationships with students, and implicit learning that happens when students witness how teachers manage their own thoughts and feelings (Schonert-Reichl, 2017). Jones and colleagues (2017) described the importance of teachers helping students practice basic SEL skills in a variety of contexts. For example, teachers might introduce sentence stems for students to use when agreeing and disagreeing as a first step towards students agreeing or disagreeing with peer claims related to science content. This bridging might be critical in middle childhood, when students prepare for the transition to middle school and adolescence, both of which are associated with increasingly complex social and emotional experiences (Eccles & Roeser, 2011).

About the Intervention

The current level of knowledge points to a need to bring together science and SEL instruction. The increasing number of states adopting science standards informed by NGSS has increased the urgency of developing evidence-based programming for teachers tasked with implementation. In response, our research team developed a new project-based learning curriculum called Connect Science (CS) as part of an IES Innovation and Development grant. The CS professional development experience prepares teachers to embed explicit instruction on social and emotional skills with science. The science lessons align with NGSS disciplinary core ideas in earth and physical science (PS3 and ESS3) and engage students in using NGSS SEPs.

The curriculum development project culminated in a randomized controlled trial (RCT), providing a unique opportunity to examine whether CS supported students' use of argumentation skills. In the current study, both intervention and comparison classrooms implemented lessons and activities designed to engage students in scientific discourse. Students in CS classrooms experienced a sequence of foundational SEL and science lessons prior to the observed discussions, allowing for comparison of discussions across the two groups.

Research Questions

Our exploration of small group science discussions in fourth grade classrooms addressed two research questions: 1) across both groups, how did students use argumentation practices and social gestures during small group discussions in science? and 2) how did discussions in CS classrooms differ from those in comparison classrooms? By first analyzing how students engaged in scientific argumentation across a variety of contexts, the study provides insight into the nature of rich discussions that bring young students into the sociocultural community of scientists. After describing the range of conversations, we investigate differences between the groups to identify if and when students leveraged social and emotional skills towards authentic science learning.

Method

Using an exploratory sequential mixed methods design, we described and coded transcripts of 14 conversations between 2-5 students in seven fourth grade classrooms (two conversations in each classroom; 4 classrooms in the intervention group). We began with qualitative analysis of conversational elements and themes. Next, we generated a dataset of frequency codes quantifying students' use of argumentation practices, social gestures, and communicative approaches. Findings from analytic memos and qualitative analysis were compared against descriptive statistics. Finally, we explored systematic differences in student conversations between the intervention and comparison group.

Participants

The research team partnered with a large, urban school district in the southeast United States transitioning to NGSS to evaluate Connect Science (CS). Administrators and researchers recruited fourth grade science teachers to participate via email. 32 teachers were randomly assigned at the school level to either attend professional development and implement CS during the 2017-18 school year, (n=18) or to use their typical science instruction using district-provided kits to address the same set of standards (n=14). The research team conducted observations in each classroom, and teachers and students completed surveys about their experiences in science. Teachers received program training and materials at no cost in addition to a stipend for participating in the research. Teachers in the waitlist comparison group were invited to attend the professional development in summer 2018 and implement the curriculum during the 2018-19 school year with developer support.

Sample Selection

We used a purposeful sampling method that maximized the amount of possible data to explore while leveraging the group equivalence established by random assignment. Three criteria were applied to the full sample in the CS RCT: 1) availability of observation data; 2) clear audio of two small group discussions; and 3) fall 2017 implementation of the energy science unit. Table 1 summarizes the demographic characteristics of the seven classrooms (4 intervention, 3 comparison) that met these criteria.

A check for representativeness found that the teachers sampled did not differ from the full RCT sample by education, years of teaching experience, recent professional development in SEL or science, gender, or ethnicity, (all *p*s > .1). School and classroom data were used to compare the selected sample to the full sample on enrollment, class size, prior achievement, socioeconomic composition, and departmentalized instruction. Classrooms differed from the full RCT sample on three demographic variables. The seven sampled classrooms had fewer Black students, more Latinx students, and more students designated as English learners than those in the full sample (all *p*s<.05).

Description of Connect Science

The CS program consists of 30 lessons that teach science, SEL skills, and civic engagement. The 8-step service-learning experience was modeled after the KIDS as Planners Framework (2011). Lessons were sequenced so that students first learned and practiced communication skills then used those skills to discuss complex ideas in science. First, teachers guide students in discovering energy and resource problems in their community. Students progress to identifying potential solutions, then finally engage in a service-learning project to implement a solution to their chosen problem. Teachers received five days of professional development, curriculum materials, and coaching throughout implementation.

Data Collection

Observation data was collected during a four-week window during fall 2017. Classrooms were video and audio recorded during a science lesson and small-group discussion task. On-site research assistants gathered data, which was then transmitted to the research team electronically.

Classroom Context

Science teachers in CS classrooms led a structured sequence of SEL lessons prior to the observation. Students began by collaboratively generating norms for their classroom (CS 1.1). Teachers then led a discussion about what it looks and feels like to listen to others, and students practiced active listening by paraphrasing (CS 1.3). Next, students learned sentence stems for agreeing and disagreeing with others respectfully (CS 1.4). Later, students built on these basic skills by discussing the more complex process of showing respect for multiple perspectives. After reading a book about two friends who resolve a conflict that arises from their differing cultural experiences, students brainstormed questions they could ask to better understand those with differing beliefs and ideas (CS 2.6). Since the sequence intertwined SEL and science lessons, students could practice skills in low-stakes situations prior to using them during science learning.

The curriculum also provided resources like student worksheets and sample anchor charts to reinforce SEL skills and concepts during science lessons.

Teachers in the comparison group used their existing curricula, which included materials and lesson guides from science kits for teaching about electric circuits and natural resources. Although two of the three teachers in the comparison group reported receiving SEL-focused PD within the last three years, they were not given guidance on how to integrate those SEL strategies with their science instruction.

Observation Context

Intervention teachers were observed while enacting CS 2.8, "Energy for the Future," which engaged students in a discussion about tradeoffs between various energy sources. The lesson began by introducing a table of pros and cons associated with renewable (e.g. hydropower, solar power) and non-renewable (e.g., coal, petroleum) energy sources. Next, students were asked to sort the list "from those we should use the least to those we should use the most in the future," (Connect Science Team, 2017). Teachers asked students to explain their reasoning to a partner using information from the pros and cons table to support their claims. Students were encouraged to rearrange their cards if their perspectives changed as a result of their discussion, and then to report final decisions to the group. In these classrooms, eight conversations between 2-5 students yielded 50 minutes of audio ($M = 6.25$, $SD=3.78$).

Comparison teachers were asked to identify a lesson that included science instruction and an activity that required students to use comparison (e.g., pros and cons) to talk about lesson content in small groups for the classroom observation. The content covered in these classes included: structural and behavioral adaptations of plants and animals; movement and relationships between objects in space; and pros and cons of technology (see results for more description of these lessons and activities). In comparison classrooms, six conversations between 3-5 students yielded 60 minutes of audio ($M = 10.00$, SD=3.19).

Quantitative Data Collection

School, teacher, and classroom demographics were provided by the district. Teachers reported on their years of experience, whether they were departmentalized instructors (i.e., only taught science) and other descriptive information.

Procedure

Research assistants followed a standardized protocol for conducting classroom observations. The camera was placed so that the teacher and as many students as possible were visible for the duration of the lesson. The researcher moved the camera to focus on a single group of students and placed audio recorders with groups when they began their discussions. Recordings of sufficient quality to understand and differentiate between speakers were transcribed by Rev Audio & Transcription. Transcripts were compared to corresponding audio and (when available) video data to clarify speakers and improve precision. The final set of transcripts consisted of fourteen conversations (two per classroom), each between a group of 2-5 students. After preliminary analyses, codes were transformed into quantitative data reflecting the frequency with which students used different scientific argumentation practices, social gestures, and communicative approaches.

Description of Coding Approach

We used an approach to assessing the normative pragmatics of science conversations similar to that used by Nielsen (2012). The procedure for applying codes to the transcripts included four steps (described and defined in Table 2). The first step was to define the two units of analysis: turns and interactions. A *turn* was the smallest unit of analysis and defined as everything said by a single speaker until another participant spoke, at which point, a new turn began. An *interaction* was defined as a cluster of thematically related turns.

Step 2 of the coding approach involved identifying claims. Claims were defined using Kuhn & colleagues' model of idea units in argumentation: "any assertion made with justification," (Kuhn, Zillmer, Crowell, & Zavala, 2013). Claims were coded, then sub-coded as justified if the speaker included a rationale during the same turn. For Step 3, turns were coded for: content (i.e. science-related or logistical), argumentation practices (i.e., evidence and questioning) and social gestures (i.e., agreement and disagreement). In Step 4, each interaction was coded by communicative approach: 1) dialogic-interactive; 2) authoritative-interactive; 3) dialogic-noninteractive; or 4) authoritative-noninteractive (Mortimer & Scott, 2003).

Qualitative Analysis

Coders wrote analytic memos to describing the argumentation and social skills observed during student conversations. The first phase of analysis aligned with Ryu & Lombardi, (2015) deeming any verbal contribution to a discussion as evidence of engagement in social learning. As such, we considered all claims, evidence, and questioning in our analysis regardless of content relevance. Audio recordings were transcribed and coded. Then, the research team engaged in an iterative process of identifying, confirming, and refining emergent themes. Figure XX shows an excerpt demonstrating how codes were applied to transcripts and counted.

Quantitative Analysis

Qualitative codes reflecting the frequency of argumentation practices, social gestures, and communicative approaches skills in student conversations were transformed into quantitative data, generating two analytic datasets (see Table 2). The first dataset includes 806 turns describing the content, argumentation practices, and social gestures in each transcript (M = 163.93, *SD*=101.44). The second dataset includes 283 interactions, (M = 22.64, *SD*=12.68) each categorized by communicative approach. Descriptive analyses quantified the observed patterns in scientific discourse and argumentation across classrooms. Chi-squared analyses were used to compare use of argumentation practices and social gestures between the intervention and comparison groups.

Results

Teachers enacted a science lesson that included a student discussion component, which served as the backdrop for observation and analysis of two research questions: 1) across both groups, how did students use argumentation practices and social gestures during small group discussions in science? and 2) how did discussions in CS classrooms differ from those in comparison classrooms? In Table 3, the column labeled “RQ1” summarizes discussions that occurred in all seven classrooms. The columns labeled “RQ2” compare results between intervention and comparison group classrooms.

In the next section, we provide key details to situate the results in the instructional context that preceded student discussions. Next, we describe the NGSS argumentation practices, social gestures, and communicative approaches across all seven classrooms. Finally, we evaluate how findings differed between the intervention and comparison group.

Discussion Tasks and Related Content Across All Classrooms

We operationalized "productive science conversation" as those where students used argumentation practices and social gestures to have a conversation related to science content that engaged all group members. Student discussions could be characterized as either: 1) about science content (renewable and non-renewable resources, movement and relationships between objects in space, pros and cons of technology), 2) about logistics (assigning roles, identifying materials, clarifying teacher expectations), or 3) off-task. Certain discussion tasks led to more science-related conversations whereas other tasks led to conversations about logistics. Aspects of the tasks set the stage for different types of conversations, as seen in this excerpt from an intervention classroom:

> Ms. Jones: *Open up your bag, take out your cards. You're going to look at this pro and con sheet. And you're going to talk with your partners about which [energy source] you think would be the best…that you guys think we should use for the future.*

In Ms. Jones' classroom, 84% of turns were science-related, 11% concerned logistics, and 5% were off-task. The discussion task described in the CS manual gave students a well-articulated discussion question without a single "right" answer and access to reference materials. Ms. Spencer, one of the comparison teachers, gave her students a similar task:

> Ms. Spencer: *We want to make sure that people are agreeing on our sort. Remember our two categories…behavioral and structural…You can use your Venn diagram as a tool to help if you would like.*

In Ms. Spencer's classroom, 70% of turns were science-related, 25% were about logistics, and 5% were off-task. Ms. Jones and Ms. Spencer presented a clear question, provided reference materials, and emphasized the need for students to reconcile differences of opinion. In contrast, another comparison classroom gave the assignment below:

> Ms. Hurst: *Your question was what kind of patterns did you see in space? …think about what you did today and use words and pictures to show me what you learned. Now, some of the words you might want to use - rotate, orbit, revolve…you have ten minutes to do that.*

Nearly three-quarters (74%) of the turns in Ms. Hurst's classroom were about logistics, while 20% were science-related. Again, off-task turns were the least common (6%). Ms. Hurst's instructions provided students a broad discussion topic and required them to produce a physical product. She listed vocabulary that students might use, but did not provide reference materials or link to prior lessons. Assignments that asked a focused, well-articulated question and included reference materials set the stage for more content-focused discussions.

Use of Argumentation Practices & Social Gestures Observed Across All Classrooms

Several patterns emerged in how students approached the assignment. We describe the range of ways in which students used argumentation practices and social gestures during their conversations based on the synthesis of qualitative themes and descriptive analyses.

Argumentation Practices

Instances of students making and justifying claims, using evidence, and asking questions were analyzed to understand how students in all seven classrooms engaged in NGSS argumentation practices. Results are reported in the RQ1 column of the argumentation practices section of Table 3.

Making claims. A claim functions as the beginning of an argument, and students used them to introduce new ideas. Students made an average of 19 claims per conversation, (SD = 10.65) but only 26% of claims included justification. Claims were often followed by another student responding directly, as seen in Ms. Grace's class below:

Ann: *The one that's used the least is coal*

Erin: *Coal, like the little rocks?*

Student turns that initiated an interaction by presenting a new idea all met the behavioral definition of the NGSS practice of making a claim, but varied in the extent to which they contributed to building scientific knowledge. On average, 52% (SD = .50*)* of claims were science-related. Science-related claims provided evidence of students' content understanding. Claims prompted continued interaction when they were brief and specific (e.g., "*Coal, it helps us get energy, too;" "I think they're going to use hydropower most;" "Well, gasoline and oil are running out").* Long or abstract claims were often interrupted or ignored by other members of the group.

Teacher participation in discussions helped students produce more complex science-related claims, as seen in the interaction below about animal adaptations:

Ms. Spencer: *What's something that for sure is going to go in structural?*

Rodrigo: *The fox.*

Ms. Spencer: *What about the fox is a structural adaptation?*

Diana: *He has thick fur, so he can camouflage.*

Ms. Spencer's first question elicits a correct (though unjustified) response from Rodrigo, confirming his understanding of structural and behavioral adaptations. Her follow up question prompted the group to build on Rodrigo's claim. In response, Diana made a more sophisticated claim that included justification.

More than one third of claims (39%, *SD*=.49) were related to logistics. These claims helped groups determine how to complete the assignment. For example, in Ms. Green's class, Lily began an interaction by making, ("*I think Krista should go…")* and justifying, (*"…she hasn't talked any*,") a logistical claim. Despite the lack of science content, Lily made a valuable contribution that helped manage social participation in the discussion.

Fewer than ten percent of claims were unrelated to the content and the assignment, and were categorized as off-task, (9%, *SD*=.29). Some off-task claims referenced the recording equipment in the classroom, (e.g., *"Hey, they're recording us,")* while others covered a wide range of topics, (e.g., "*I feel like I have a golden necklace, look at my necklace,"; "I love science, it's the most funnest part of the year,").* Off-task turns were included in our analysis because they met the behavioral definition of the NGSS practice of making claims. However, compared to science-related and logistical turns, they were not nutritive to science learning. Such claims often led to brief off-task interactions, or were ignored by other group members.

Using evidence. Students used three types of evidence: empirical, generalized, and personal (see the mutually exclusive sub-codes under "Using evidence" in Table 3). More than half (66%, *SD*=.47) of students' evidence use was empirical. In intervention classrooms, evidence often came from the pros and cons table provided for reference. For example, after Jessica suggested that one of the energy sources was "*good*," John responded: "*No, it's not. Look, releases carbon dioxide when burned."* John expressed his disagreement with Jessica's claim and used empirical evidence from the reference material to support his point of view.

Students also supported claims with generalized evidence of what some undefined set of people think or do ("*Yeah, people don't use it that much").* Generalized evidence was the second most common (20%, *SD*=.40). The remaining 14% (SD = .35) of evidence was based on personal preferences ("*I just like staring at them,"* referring to wind turbines).

Asking questions. Questions represent a bid to add another voice to an interaction. When students asked peers questions, they created opportunities to build understanding. Questions often prompted students to supply evidence for a claim:

John: *This one is the greatest.*

Max: *No, I mean how?*

John: *Because look, it just uses the sun's electrical power.*

Students asked an average of 12 questions per discussion, (SD = 9.94). Overall, conversations contained more questions about the science content than logistics, $\chi^2(1,806) = 8.94$, $p = .003$. However, logistical questions were sometimes useful, like when Chase asked “*Can I talk*?”

Some teachers used questions to help students understand the assignment. For example, Ms. Corbett introduced the discussion task using a series of questions:

Ms. Corbett: *So, does every group have to agree with what each other is saying?*

Class: *No.*

Ms. Corbett: *No! Eric said that the wind turbine is going to be the one used the least. Do you have to write this down?*

Class: *No.*

Ms. Corbett: *No, you're going to put your cards in order with your group as to which one you think will be used the least in the future to the one you think will be used the most.*

Ms. Corbett prompted students to indicate their understanding of the assignment by emphasizing key components of the discussions task with questions. Students in her class were prepared to begin their conversations about energy sources; 86% of their questions sought to build knowledge about energy sources, while 10% were inquiries about the task. The remaining 4% of questions were off-task.

Social Gestures

Expressions of agreement and disagreement, assertive speech, and prosocial speech characterized how students used social gestures across the fourteen conversations. Results are reported in the RQ1 column of the social gestures section of Table 3. Assertive speech (turns that interrupted another speaker or included a directive) was, on average, the most common social gesture (M = 10.36, *SD*=8.82). Next in prevalence was agreement (M = 7.57, *SD*=4.70), followed by disagreement (M = 7.21, *SD*=9.20). Turns that included a peer's name, gave a compliment, or included please or thank you were categorized as prosocial; prosocial speech was the least frequent social gesture (M = 4.43, *SD*=6.93).

Agreement and disagreement. When students stated whether they agreed with a peer's claim, they offered their point of view to the group's collective understanding. On average, students expressed agreement and disagreement equally, ($t(13)=0.22$, $p=.83$). However, the two gestures prompted distinct responses. Agreement often led to brief interactions about a single idea, as seen in an example from Ms. Woodward's class:

Mason: *I also think that wind could be good.*

Christian: *Yeah, wind could be good.*

Although Christian makes a social contribution by agreeing, he echoes Mason's claim rather than building on the group's understanding. Conversely, expressions of disagreement introduced a conflicting viewpoint for the group to resolve. The excerpt below illustrates how a disagreement about the size of the sun prompted further discussion:

Shane: *Because the sun is not really that big.*

Robert: *The sun's pretty big. Like way bigger.*

Shane: *The sun is the same size, as probably like the earth.*

Robert: *No, it's bigger! …didn't you hear Ms. Hurst say it's bigger? It's way bigger.*

Rather than agreeing with the claim that "*the sun is not really that big,*" Robert used evidence to support his disagreement. Although his evidence may seem unsophisticated, he draws from his empirical observations of Ms. Hurst's instruction to justify his perspective.

Assertive speech. Students interrupted each other and used directives more than they used other social gestures. The use of assertive statements in conjunction with other social gestures helped students to manage participation in the discussion. After two students in Ms. Green's class continually interrupted and spoke over each other, Amanda suggested, "*If someone's talking, just let them talk.*" During the same conversation, Meredith responded to a complicated claim by saying, "*Okay, hold on…we need to write this down.*"

Prosocial speech. Students' use of names, compliments, and please or thank you were infrequent relative to other social gestures, with two notable exceptions: one conversation in Ms. Hurst's class included 24 examples of prosocial speech, and one conversation in Ms. Green's classroom included 15. In the conversation from Ms. Hurst's classroom, the majority (67%) of prosocial speech related to logistics:

Lisa: *You're doing so well, Ruby and Giana. I wish I could be like you.*

Giana: *Thank you so much. Thank you, Lisa.*

Other examples include when Jennifer made an error, and apologized: "*I'm sorry, Ruby. I messed up,"* and her compliment: "*The sun looks really good."* While prosocial speech made for a polite, friendly conversation between the students, it did little to support science learning.

The prosocial speech that took place during the conversation in Ms. Green's classroom was primarily science-related (60%). Students addressed peers by name, such as when Amanda said: "*I disagree with Paul because let's say there are owls or squirrels or birds that live in that tree…it's like if someone just came to your house when you weren't there, and knocked down your house."* Later, Krista made a similar contribution: "*I agree with Amanda, and I disagree with Paul."* Students used prosocial speech to specify who they were addressing, allowing group members to acknowledge the viewpoints of others while making their own contributions.

Conversational profiles of social gesture use. Across the seven classrooms, distinct patterns emerged reflecting the extent to which students used social gestures, resulting in three conversational profiles: balanced, disconnected, and highly assertive. Differences in the amount and type of social gestures used across conversations revealed the range in students' approaches to the social dimension of the task (see Figure 1).

Balanced conversations included frequent use of a variety of social gestures and were longer than disconnected and highly assertive conversations ($t(12)=-3.13$, $p=.009$). Nearly two thirds of social gestures were expressions of disagreement (31%, $SD=.09$), and assertive speech (31%, *SD*=.11). The remaining gestures were expressions of agreement, (28%, *SD*=.10) and prosocial speech (10%, *SD*=.06). Six of the fourteen student discussions fit this profile.

Both discussion groups in Ms. Corbett's class had balanced conversations. One example took place between four students: Jordan, Brynn, Charlie, and Xander. Jordan began by making a claim that he continued to build on throughout the discussion: "*Well this is how it should be, but I don't think people are going to do that. They're probably going to use solar, nuclear… people are going to get so advanced and think they're getting smarter when they're really being stupider…I'd rely on old techniques the most*." Brynn responded by voicing her disagreement, ("*I think they're going to use hydropower the most,"*). Jordan and Brynn engaged in a lively back-and-forth about their views, supported by Xander, who made frequent social contributions despite not making his scientific point of view clear. When the fourth member of the group, Charlie, got the group off-task, Xander said: "*Okay, we're supposed to be working together. You don't even have your cards out. Stop distracting us, we're trying to work*." The rest of group then continued their discussion about energy sources. Here, Xander's assertiveness allowed the group to stay on task despite one member failing to contribute.

In contrast, disconnected conversations were brief, and students used argumentation skills without social supports. Five of the fourteen conversations fit this description. For example, consider the set of six interactions below that comprise nearly one third (29%) of a brief discussion in Ms. Spencer's class:

Jared: *Hawks have sharp claws that kill their prey*.

Casey: *What is this?*

Molly: *Bear?*

Kiera: *A artic fox has…*

Molly: *Insects are shaped like a leaf so predators think they are real leaves.*

Jared: *A rosebush has thorns to…where's this go?*

Molly: *Frogs have long strong legs to hop really far*.

Even though all four group members participated, the discussion lacks evidence of students listening or responding to each other. Group members didn't explicitly agree or disagree, and each contribution began a new, independent line of inquiry.

The remaining three conversations were highly assertive. These discussions were similar to disconnected conversations in length, but were dominated by the presence of assertive speech, (55%, *SD*=.17). Although agreement was the next most common gesture, (25%, *SD*=.16) the limited use of explicit expressions of disagreement (3%, *SD*=.03) did not balance the frequent assertive statements. As seen in Ms. Woodward's class:

Kamren: *Guys, stop! We need to work.*

Travis: *He's not working, he didn't even highlight.*

Ari: *This all doesn't matter.*

Kamren: *We're supposed to be working, come on.*

Travis: *I already picked the most highest one.*

Ari: *Well, you chose the ugly thing bro.*

Unlike when Xander used assertive speech to stop Charlie from derailing the discussion in Ms. Corbett's class, Kamren's attempts to keep the group on task and Travis' bid to redirect the conversation towards the science content were less effective in the absence of other social gestures. When the most prevalent social gestures used in a conversation were interruptions and use of directives, students did not have a productive dialog.

Communicative Approaches Across All Classrooms

Interactions were categorized as one of four communicative approaches: 1) dialogic-interactive; 2) authoritative-interactive; 3) dialogic-noninteractive, and 4) authoritative-noninteractive (Mortimer & Scott, 2003). Scientific argumentation prioritizes the dialogic-interactive (D-I) communicative approach and therefore, we expected that conversations with evidence of students using both argumentation practices and social gestures use would be primarily D-I. Results are reported in the RQ1 column of the communicative approaches section of Table 3.

Interactions that engaged multiple voices accounted for 73% of the interactions in the data. A chi-square test of independence revealed that within the portion of the interactions coded as interactive, more were dialogic than authoritative, $\chi^2(1{,}338) = 72.81$, $p<.001$. More than half (54%, $SD=.50$) of interactions consisted of multiple voices expressing multiple points of view (D-I). Conversations in intervention classrooms, like the example below from Ms. Grace's class, consisted of D-I interactions with connections to the lesson content:

Leslie: *Every single day people use...*

Erin: *Gasoline to fill up their cars.*

Sierra: *Yeah, yeah, we might run out of it.*

Regardless of content-relevance, opportunities to practice D-I conversation can help students become comfortable with managing the type of dialog that required to build knowledge through scientific discourse and argumentation (Berland, 2011). For example, the interaction below occurred in Ms. Hurst's class, where students were asked to "*use words and pictures to show me what you learned":*

Giana: *You have to do the writing. I know how to do this.*

Lisa: *Here, I'll do the writing.*

Jennifer: *I'll do the outlining.*

Here, different voices respectfully communicated their ideas about task roles. The students used D-I interactions to reach agreement on how to best complete the assignment rather than talking about what they learned about the relationships between objects in space.

Typical discourse patterns in elementary science prioritize multiple voices endorsing a single point of view (Tippett, 2009). Authoritative-interaction (A-I) communication was the second most common, and accounted for nearly one quarter (22%, *SD*=.39) of interactions. One example of an A-I interaction took place in Ms. Green's class:

Luke: *It's bad because people are digging up iron and copper for these utensils, which you can't make again. It's just gone, used for wiring. Just shorten that up.*

Marco: *Use too many resources.*

Luke: *Uses too many resources.*

Although the interaction engaged two students, it did not represent an exchange of ideas, as Marco merely echoed Luke's claim.

Discussion Characteristics in Intervention Versus Comparison Classrooms

The second research question compared differences between students' use of argumentation practices and social gestures between intervention and comparison groups. Results showed no difference in the frequency of argumentation practice or social gesture use between groups even though conversations in comparison classrooms, (M = 10.00, SD=3.19) were longer than those in intervention classrooms, (M = 6.25, SD=3.38), $t(12) = 2.10$, p=.03. However, the students in CS classrooms tended to have science-related conversations, while those in comparison classrooms talked about logistics (see RQ2 column of argumentation practices and social gestures sections of Table 3). These comparisons are presented in Figure 2.

Students in intervention classrooms made more science-related claims (69%, SD=.46) than those in comparison classrooms (40%, SD=.49), $\chi^2(1,806) = 21.60$, $p<.001$. In comparison classrooms, more than half of claims (54%, SD=.50) were logistical, which was less than the 17% (SD = .37) in intervention classrooms, $\chi^2(1,806) = 37.21$, $p<.001$. Both groups used evidence to discuss science content, but students in comparison classrooms used evidence to support claims about logistics (15%, SD=.36) significantly more than the intervention group (4%, SD=.20), $\chi^2(1,806) = 4.88$, p=.02. Nearly two-thirds of questions in intervention classrooms (65%, SD=.48) were about the science content, more than the 48% (SD = .49) in comparison classrooms, $\chi^2(1,806)=10.92$, $p<.001$. Finally, students in comparison classrooms asked more logistical questions (53%, $SD=.50$) than those in intervention classrooms, (19%, $SD=.40$), $\chi^2(1,806) = 18.35$, $p<.001$.

A similar pattern emerged in how the two groups used social gestures. Students in CS classrooms used all four of the observed social gestures to discuss science content more than students in comparison classrooms (all *p*s<.05). Conversely, students in comparison classrooms expressed agreement, disagreement, and used assertive speech in discussion about logistics more than students in intervention classrooms (all *p*s<.01).

Discussion

Researchers can support elementary educators rising to the ambitious goals of NGSS by identifying ways to integrate SEL and science instruction. We described how fourteen groups of 2-5 students in seven fourth grade classrooms used argumentation practices (making claims, using evidence, and asking questions) and social gestures (agreement, disagreement, assertive speech, and prosocial speech) during small group discussions in science. We also compared how students' use of argumentation practices and social gestures differed between two groups: four intervention classrooms implementing Connect Science, an NGSS-aligned curriculum that integrates SEL; and three comparison classrooms implementing their typical science instruction. Our findings describe how students navigated the challenge of learning science through discourse. Furthermore, the group comparison assisted in identifying conditions that encouraged productive science conversations among students in ways consistent with NGSS.

Common Discourse Practices Across All Classrooms

The first research question sought to determine whether students were able to engage in productive science conversations, which we defined as discussions including an explicit focus on science content, use of argumentation practices and social gestures, and predominantly dialogic-interactive (D-I) conversations. The conversations that met these criteria demonstrate that fourth graders are capable of engaging in high-quality scientific discourse under certain conditions. For example, tasks must be designed with clear discussion outcomes in mind. Students need sufficient understanding of how content relates to their world, and they need to know how to use social gestures to sustain the dialog. When these conditions were not met, conversations drifted toward the logistics of the assignment. When students did not use social gestures effectively, they had disconnected (i.e., two-way but unrelated to each other) or highly assertive conversations (i.e., one-way with students talking over each other). Subsequently, student conversations focused less on discussing and developing their ideas in ways that lead to deep-level thinking.

Claims, Justification, and Use of Evidence

Scientific argumentation described by NGSS centers around the practice of making and justifying claims with evidence (NRC, 2012b). Though claims were common, only one in four of claims were justified with evidence. Inclusion of reference materials seemed to push students towards supporting their claims with empirical evidence. For example, providing a handout with pros and cons of different energy sources prompted students to evaluate sources independently and objectively.

Distributing a graphic organizer illustrating different animals and adaptations gave students specific examples to add to the discussion. Our findings extend the value of reference materials to an upper elementary context, as prior research with middle (Berland, 2011) and high schoolers (Nielsen, 2012) has demonstrated that reference materials encourage students to use evidence to support their claims.

Asking questions. Asking questions during discussions prompted group engagement. From a sociocultural perspective, questioning can function as a psychological tool that pushes students to think at a more sophisticated level (Hackling, Smith, & Murcia, 2010; Vygotsky, 1978). When students asked questions, they pressed peers to explain their thinking in a way that promoted deep-level learning. Those who used questioning effectively (i.e., to better understand a peer's claim) challenged others to unpack their claims and examine their reasoning. Questions were also inherently social, requiring that students listen to one another to build upon the group's knowledge and understanding. Questioning also points to activation of a developmental process where students progress from one-sided conversation to dialogic conversations, moving them closer to the analytic conversations exemplified by scientists (Duschl & Osborne, 2002).

Using Social Gestures in Science Discussions

Groups that frequently used of a variety of social gestures elevated their conversations, facilitated sustained discussion, and managed participation. Expressions of agreement and disagreement, assertive speech, and prosocial speech contributed to the rhythm of the discussion and allowed students to deepen their conversations about science. Several patterns emerged in how discussion groups used social gestures. Students interrupted each other and used directives across conversations, but these potentially delicate turns were productive when tempered by explicit expressions of agreement or disagreement and prosocial speech. When assertive speech dominated the conversation and students had limited time, the resulting conversations did little to build scientific knowledge. For example, a student in a highly assertive group might interrupt a peer to share a conflicting claim rather than making their claim after an explicit expression of disagreement.

The presence of prosocial language suggests a positive sense of community in the classroom that supports student learning (Jones et al., 2017). Although students' use of prosocial language indicated positive social development, prosocial speech was rarely on topic. The ideal balance between assertive and prosocial speech involved students using peers' names and giving compliments in a way that embedded the science content. For instance, one student responded to a peer's claim by saying, "*Oh, I like that idea, that makes sense!*" Productive science discussions involve a combination of prosocial and assertive speech such that the prosocial speech maintains a sense of community and connection while resolving disagreements that lead to collective science learning. When assertive speech was combined with other social gestures, students elevated the conversation to talk about scientific ideas without personalizing their position. We observed multiple cases of young students meeting this lofty goal. Nevertheless, groups that did not use social gestures effectively had less productive conversations. Students appeared to be talking *at* each other rather than having meaningful interactions.

Communicative Approaches

Students' communicative approaches resembled the expectations for scientific argumentation described by NGSS. Interactions almost always engaged multiple voices, and individual bids for a response seldom went unanswered. Dialogic-interactive interactions, (D-I, multiple points of view presented by multiple voices, Mortimer & Scott, 2003) the most useful for collectively building knowledge, were the most prevalent type of interaction. However, the next most common communicative approach was authoritative-interactive, (A-I) wherein multiple voices discussed a single idea. This finding highlights the tendency for students to revert to authoritative discussion of facts, and the need for continued support and scaffolding to promote primarily D-I conversations in science.

Findings from Group Comparison

Teachers implementing CS set the stage for scientific discussion differently than the comparison teachers. The groups did not differ in the quantity of argumentation practices and social gestures used. Instead, distinct patterns in *how* students used argumentation practices and social gestures emerged between the intervention and comparison group.

Students in intervention classrooms used argumentation practices and social gestures to talk about science. As directed in the CS implementation manual, teachers gave a well-articulated discussion question and provided reference materials. In intervention classrooms where students appropriately used social gestures, the task elicited scientific argumentation as defined by NGSS in as little as five minutes. Since existing work shows that teachers have less time for science than other subjects, (Penfield & Lee, 2010) it's important to highlight possibilities for maximizing collaborative learning with limited instructional time. We posit that students' adept use of social gestures reflected earlier explicit instruction and low-stakes practice of these skills as directed by the CS manual. Students also had visual supports (e.g., anchor charts with sentence stems) of past SEL lessons available during their science discussions.

Students in comparison classrooms spent most of their time talking about logistics. Comparison group teachers tended to ask vague reflection questions and expected students to produce a physical product (like a poster or worksheet) in addition to discussing science ideas. Even though students were given more time for discussion in comparison classrooms, their conversations tended to focus on the logistics of the assignment. These findings identify ways of supporting young students as they learn how to make and justify claims, use evidence, and ask questions to learn in science. They also demonstrate the necessity of social skills for groups to function. By analyzing conditions that led to rich discussions in science, we add to the growing body of literature integrating content instruction and social and emotional skills development.

Implications for Practitioners

Teacher-centered instruction dominates most typical elementary science classrooms, (Reiser, 2013) with teachers positioned as the authority on what knowledge is valued or "correct." These conditions create classrooms where students are seldom challenged to think like scientists. NGSS calls for teachers to shift to a student-centered approach while also establishing boundaries that keep conversations focused. Our findings call attention to the importance of well-designed tasks, integration with social and emotional instruction, and authentic content understanding as important for elevating predominantly D-I conversations to instances of productive argumentation from evidence.

Productive, content-focused conversations occurred in classrooms where the task met three criteria: 1) a clearly-articulated question with no "right" answer; 2) provision of reference materials; and 3) collaborative knowledge building as the outcome for the activity (rather than a physical product). These findings reflect prior research identifying the types of discussion activities in science that prompt students to engage in inquiry learning (Kuhn et al., 2017). When using less-defined discussion tasks, teachers should anticipate that students will have interactions about logistics. Allowing sufficient time for both negotiation and completion of an assignment allows for authentic practice using social skills while talking about science. However, with the limited instructional time available for science, that additional time might not be available. Some teachers used questions to redirect conversations, for example: *"What's your evidence for that?";* "*Hey Krista, what do you think?"; "Is there any one right answer?"* Questions like these modeled the use of science-related argumentation practices and guided student discussions with dialogic (rather than authoritative) communication.

Limitations and Future Directions

Two limitations warrant mention and consideration in future research. First, despite a relatively large sample for qualitative research, we have limited information about the individual students in each discussion. Future work incorporating student demographics would provide a more nuanced view of their educational experiences in science. Our data were also limited to conversations from a single time point. Research suggests that neither communicative competence nor engagement in science develop linearly, but in “fits and starts,” (Ryu & Lombardi, 2015). The same activities in one classroom could yield different conversations on another date, even with the same student sample (Berland, 2011; Kuhn et al., 2017). Researchers could generate a more in-depth understanding of the classroom culture surrounding scientific discussions by observing multiple time points.

Concluding Comments

Students enjoy opportunities to interact with their peers. This paper describes how conversations can build science knowledge and communication skills among young students. By analyzing science tasks and corresponding instruction, we identified strategies that were associated with more productive, content-relevant discussions. Sufficient foundational instruction and the right materials can bring intentionality to peer interactions so that students can use their social and emotional skills to make claims, use evidence, and ask questions to build collective science knowledge.

Table 1

Participant Characteristics

	Count/Mean (*SD*)		
	Total (*n* = 7)	Intervention – Ms. Corbett, Ms. Grace, Ms. Jones, Ms. Woodward (*n* = 4)	Comparison – Ms. Green, Ms. Hurst. Ms. Spencer (*n* = 3)
Schools			
Enrollment	240(68)	246(70)	234(81)
Prior 4th grade science achievement	.26(.22)	.19(.08)	.36(.33)
Percent economically disadvantaged	.51(.21)	.57(.11)	.43(.31)
Teachers			
Years of experience	7.43(5.38)	8.50(5.07)	6.00(6.56)
Master’s degree	5	3	2
Recent SEL professional development	5	3	2
Recent science professional development	4	3	1
Female	7	7	7
White	7	7	7
Classrooms			
Class size	23(1)	23(1)	22(2)
Departmentalized science instruction	5	3	2
Percent female	.47(.10)	.48(.11)	.45(.10)
Percent English learners*	.40(.28)	.42(.25)	.39(.37)
Percent White	.26(.29)	.13(.07)	.43(.42)
Percent Black*	.18(.14)	.24(.17)	.10(.03)
Percent Latinx*	.41(.27)	.41(.27)	.33(.33)
Percent Asian	.07(.08)	.05(.08)	.09(.08)

Note. * differs significantly from RCT sample mean ($p < .05$)

Table 2

Coding Steps, Definitions, Examples, and Frequencies

Step	Code	Mutually exclusive sub-code(s)	Definition	Example Quote(s)	Average Frequency per Conversation (SD)
1. Define units of analysis.					
	Turn		Everything said by one individual until another speaks.	"You're wasting tons of paper." "We should do this one. Look, it says, 'reduces carbon dioxide when burned, releases another gas.' That's not good."	163.93 (101.44)
	Interaction		Cluster of turns related to a single claim.	"Rodrigo: Rabbits have large ears so they can hear above danger. Luca: The body? Rodrigo: Oh yeah, they use the body."	22.64 (12.68)
2. Identify claims.					
	Making claims		Any assertion made by an individual.	"So that means that hydropower is bad because it could kill fish".	18.86 (10.65)
		Justified	Inclusion of rationale in the same turn as a claim.	"So that means that hydropower is bad because it could kill fish."	4.86 (2.93)
3. Code turns.					
	Category				
	Content	Science-related	Related to presented science topic.	"And coal and wind turbines. People use all of these."	56.36 (37.07)
		Logistics*	Related to materials or expectations for assignment.	"Oh, you were supposed to bring your highlighter?"	20.57 (30.03)
		Off-task	Unrelated to science content or the assigned task.	"Do you want to have a sleep over this weekend?"	5.36 (6.16)

Step	Code	Mutually exclusive sub-code(s)	Definition	Example Quote(s)	Average Frequency per Conversation (SD)
			Argumentation practices		
	Using evidence		Grounds for belief or disbelief of a claim.	“This is one of those things that causes pollution.”	11.00 (11.01)
		Empirical	Evidence based on observation of phenomena.	“Okay, so the petroleum…it says here that it's often used.”	7.14 (9.40)
		Generalized*	Evidence based on an undefined person or group	“They're probably even going to forget even how to use bottom few.”	2.14 (2.14)
		Personal	Evidence based on personal preference	“I think that one’s bad.”	1.50 (1.65)
	Asking questions		A turn that explicitly solicits a response.	“Chase, what do you think?”	12.43 (9.94)
			Social gestures		
	Expressing agreement		Expression of a similar opinion.	“I agree with Luke.”	7.57 (4.70)
	Expressing disagreement		Expression of a difference of opinion.	“Well, I don't know about that.”	7.21 (9.20)
	Using prosocial speech*		Name use; compliments; please/thank you	“Thank you so much. Thank you, Lisa.” “That looks really good!”	4.43 (6.93)
	Using assertive speech*		Interruption of another speaker; use of directives	“Dylan: Natural gas is like fire or something like that and- Troy: It's easy to catch on fire.” “Listen to the bad things about it.”	10.36 (8.82)

Step	Code	Mutually exclusive sub-code(s)	Definition	Example Quote(s)	Average Frequency per Conversation (SD)
6. Code interactions.					
	Communicative approach	Dialogic-Interactive (D-I)	Multiple points of view presented by multiple voices.	"Luke: Danger. It just produces, kind of, danger. Marco: Oh yeah, like pollution. Luke: Well also you could get shocked."	12.21 (5.66)
		Authoritative-Interactive (A-I)	One point of view presented by multiple voices	"Mason: I also think that wind could be good. Christian: Yeah, wind could be good."	4.29 (3.52)
		Dialogic-Non-interactive (D-N)	Multiple points of view presented by one voice.	"We're just supposed to highlight the ones that we're interested in? We can do that."	1.57 (1.74)
		Authoritative-Non-interactive (A-N)	One point of view represented by one voice.	"Solar panel's the second least."	4.36 (3.50)

Note. * Emergent code developed from early analysis; all other codes were determined a priori.

Table 3

Descriptive Statistics of Coding Results Across All Classrooms and Between Groups

Code / Mutually exclusive sub-code	Percent (*SD*) Total (*RQ1*)	Intervention (*RQ2*)	Comparison	χ^2
Turn content				
Science-related	.55(.50)	.70(.46)	.45(.50)	49.03***
Logistics	.36(.48)	.15(.36)	.49(.50)	92.27***
Off-task	.09(.29)	.15(.35)	.06(.23)	17.58***
	Argumentation practices			
Making claims				
Justified	.26(.44)	.25(.44)	.26(.44)	.05
Science-related	.52(.50)	.69(.46)	.40(.49)	21.60***
Logistics	.39(.49)	.17(.37)	.54(.50)	37.21***
Off-task	.09(.29)	.14(.35)	.06(.23)	5.09*
Using evidence				
Empirical	.66(.47)	.58(.50)	.74(.44)	4.31*
Generalized	.20(.40)	.30(.46)	.11(.32)	7.93**
Personal	.14(.35)	.12(.34)	.15(.35)	.27
Science-related	.82(.38)	.86(.35)	.79(.41)	1.32
Logistics	.10(.30)	.04(.20)	.15(.36)	4.88*
Off-task	.08(.27)	.10(.30)	.06(.24)	.67
Asking questions				
Science-related	.48(.50)	.65(.48)	.48(.49)	10.92**
Logistics	.41(.49)	.19(.40)	.53(.50)	18.35***
Off-task	.11(.32)	.16(.37)	.09(.29)	2.03
	Social gestures			
Expressing agreement				
Science-related	.75(.43)	.86(.35)	.68(.47)	4.82*
Logistics	.17(.38)	.02(.14)	.30(.46)	15.07***
Off-task	.07(.25)	.12(.33)	.02(.13)	4.47*
Expressing disagreement				
Science-related	.79(.41)	.94(.24)	.72(.45)	6.92**
Logistics	.21(.41)	.06(.24)	.28(.45)	6.92**
Off-task	0	0	0	0
Using prosocial speech				
Science-related	.36(.48)	.71(.49)	.31(.47)	4.45*
Logistics	.56(.50)	.29(.49)	.60(.49)	2.50
Off-task	.08(.27)	0	.09(.29)	.69
Using assertive speech				
Science-related	.40(.49)	.53(.50)	.31(.47)	6.69*
Logistics	.50(.50)	.29(.46)	.63(.48)	16.01***
Off-task	.10(.31)	.18(.39)	.06(.23)	5.87*
	Communicative approaches			
Interaction				
Dialogic-Interactive (D-I)	.54(.50)	.60(.49)	.50(.50)	2.55
Authoritative-Interactive (A-I)	.19(.39)	.16(.37)	.20(.40)	.66
Dialogic-Non-interactive (D-N)	.08(.27)	.07(.26)	.08(.27)	.06
Authoritative-Non-interactive (A-N)	.19(.40)	.16(.37)	.21(.41)	1.08

Note. * p<.05; ** p<.01; *** p<.001

Figure 1

Excerpts from a Student Conversation Demonstrating How Codes Were Applied to Turns and Interactions

Turn	Transcript		Turn Codes	Interaction Code
1	Camila:	You know at the gas station that they used? That's natural gas. It's for your car.	Science-related claim with justification	Dialogic-Interactive
2	Troy:	No, it's gasoline.	Science-related disagreement	
3	Camila:	It's natural gas.	Science-related disagreement	
4	Troy:	I disagree with you because it's not natural gas, it's gasoline.	Science-related disagreement supported by empirical evidence	
5	Camila:	Yeah, it's natural gas and they dig it from the ground.	Science-related disagreement with generalized evidence	
6	Dylan:	Well, natural gas is different from-	Science-related disagreement	
7	Troy:	From gasoline.	Science-related assertive speech	
8	Dylan:	Yeah.	Science-related agreement	
14	Troy:	I think we use petroleum too much.	Science-related claim (not justified)	Authoritative-Noninteractive
15	Dylan:	Natural gas is like fire or something.	Science-related claim (not justified)	Dialogic-Interactive
16	Troy:	It's easy to catch on fire.	Science-related agreement	
17	Dylan:	Yeah. Because you rub two sticks together to make natural gas.	Science-related agreement with generalized evidence	
18	Troy:	Well, I don't know about that.	Science-related disagreement	
19	Dylan:	Yeah, to make fire you rub two sticks together.	Science related disagreement with generalized evidence	
39	Dominic:	I haven't even said anything yet because of y'all always talking.	Logistical claim with justification	Authoritative-Interactive
40	Camila:	Say some things.	Logistical assertive speech	

Figure 2

Composition of Gestures Use Across All Conversations Representing Three Conversational Profiles

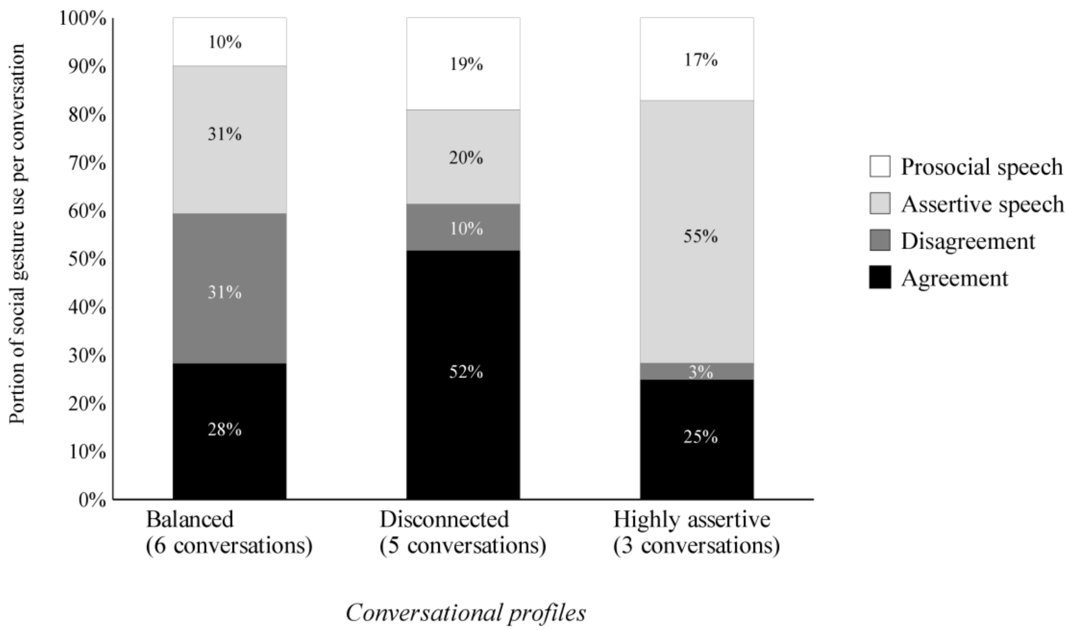
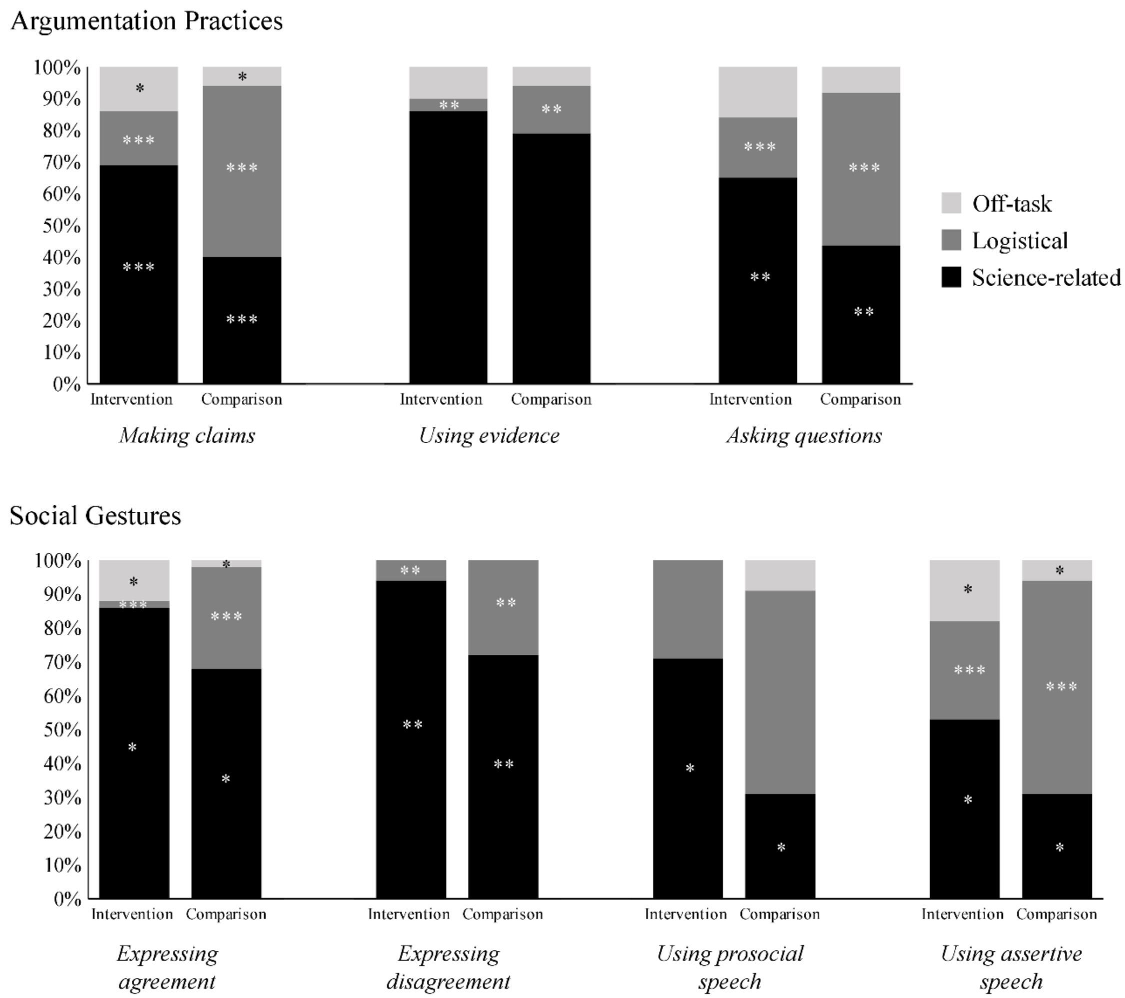


Figure 3

Differences in Science-Related, Logistical, and Off-Task Use of Argumentation Practices and Social Gestures Between Connect Science and Comparison Classrooms



References

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

Berland, L. K. (2011). Explaining variation in how classroom communities adapt the practice of scientific argumentation. *Journal of the Learning Sciences*, *20*(4), 625–664.

CASEL. (2012). *Effective social and emotional learning programs: Preschool and elementary edition*. Chicago, IL.

Cazden, C. (2001). *Classroom discourse: The language of teaching and learning*. Portsmouth, NH: Heinemann.

Colley, C., & Windschitl, M. (2016). Rigor in elementary science students' discourse: The role of responsiveness and supportive conditions for talk. *Science Education*, *100*(6), 1009–1038. https://doi.org/10.1002/sce.21243

Driver, R., Newton, P., & Osborne, J. (2000). Establishing the norms of scientific argumentation in classrooms. *Science Education*, *84*(3), 287.

Durlak, J. A., Weissberg, R. P., Dymnicki, A. B., Taylor, R. D., & Schellinger, K. B. (2011). The impact of enhancing students' social and emotional learning: A meta-analysis of school-based universal interventions. *Child Development*, *82*(1), 405–432.

Duschl, R. A., & Osborne, J. (2002). Supporting and promoting argumentation discourse in science education. *Studies in Science Education*, *38*(1), 39–72.

Eccles, J. S., & Roeser, R. W. (2011). Schools as developmental contexts during adolescence. *Journal of Research on Adolescence*, *21*(1), 225–241.

Erduran, S. & Jiménez-Aleixandre, M. P. (2008). Argumentation in science education: *Perspectives from classroom-based research*. Dordrecht, the Netherlands: Springer.

Hackling, M., Smith, P., & Murcia, K. (2010). Talking science: Developing a discourse of inquiry. *Teaching Science*, *56*(1), 17–22.

Hanuscin, D. L., & Zangori, L. (2016). Developing practical knowledge of the Next Generation Science Standards in elementary science teacher education. *Journal of Science Teacher Education*, *27*(8), 799–818. https://doi.org/10.1007/s10972-016-9489-9

Harris, C. J., Penuel, W. R., D'Angelo, C. M., DeBarger, A. H., Gallagher, L. P., Kennedy, C. A., & Krajcik, J. S. (2015). Impact of project-based curriculum materials on student learning in science: Results of a randomized controlled trial. *Journal of Research in Science Teaching*, *52*(10), 1362–1385. https://doi.org/10.1002/tea.21263

Hayes, K. N., & Trexler, C. J. (2016). Testing predictors of instructional practice in elementary science education: The significant role of accountability. *Science Education*, *100*(2), 266–289. https://doi.org/10.1002/sce.21206

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. a. (2007). Scaffolding and achievement in problem-based and inquiry learning: A response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107. https://doi.org/10.1080/00461520701263368

Jones, S. M., Barnes, S. P., Bailey, R., & Doolittle, E. J. (2017). Promoting social and emotional competencies in elementary school. *Future of Children*, *27*(1), 49.

Kam, C., Greenberg, M. T., & Walls, C. T. (2003). Examining the role of implementation quality in school-based prevention using the PATHS curriculum. *Prevention Science*, *4*(1), 55–63.

Kids as planners: A guide to strengthening students, schools and communities through service-learning. (2011). Lewiston, ME: Harkins Consulting, LLC.

Kochenderfer-Ladd, B., & Ladd, G. W. (2016). Integrating academic and social-emotional learning in classroom interactions. In K. R. Wentzel & G. B. Ramani (Eds.), *Handbook of Social Influences in School Contexts: Social-Emotional, Motivation, and Cognitive Outcomes* (pp. 349–366). Routledge. https://doi.org/10.4324/9781315769929.ch22

Kuhn, D., Zillmer, N., Crowell, A., & Zavala, J. (2013). Developing norms of argumentation: Metacognitive, epistemological, and social dimensions of developing argumentive competence. *Cognition and Instruction*, *31*(4), 456–496.

Lee, O., Miller, E. C., & Januszyk, R. (2014). Next Generation Science Standards: All standards, all students. *Journal of Science Teacher Education, 25*(2), 223–233.

Lemke, J.L. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J.R. Martin and R. Veel (Eds.), *Reading Science: Critical and Functional Perspectives on Discourses of Science* (pp. 87–113). New York: Routledge.

Lewis, K. M., Vuchinich, S., Ji, P., Dubois, D. L., Acock, A. C., Bavarian, N., … Flay, B. R. (2016). Effects of the Positive Action program on indicators of positive youth development among urban youth. *Applied Developmental Science*, *20*(1), 16–28.

Manz, E. (2015). Representing student argumentation as functionally emergent from scientific activity. *Review of Educational Research*, *85*(4), 553–590.

Mortimer, E. F., & Scott, P. H. (2003). *Meaning making in secondary science classrooms*. Philadelphia, PA: Open University Press.

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

National Research Council. (2012b). *Education for life and work: Developing transferable knowledge and skills in the 21st century*. Washington, DC: The National Academies Press.

National Research Council. (2015). *Guide to implementing the Next Generation Science Standards*. Washington, DC: National Academies Press.

Nielsen, J. A. (2012). Science in discussions: An analysis of the use of science content in socioscientific discussions. *Science Education*, *96*(3), 428–456.

Pasley, J. D., Trygstad, P. J., & Banilower, E. R. (2016). *What does "implementing the NGSS" mean? Operationalizing the science practices for K–12 classrooms*. Horizon Research, Inc.: Chapel Hilll, NC.

Penfield, R. D., & Lee, O. (2010). Test-based accountability: Potential benefits and pitfalls of science assessment with student diversity. *Journal of Research in Science Teaching*, *47*(1), 6–24. https://doi.org/10.1002/tea.20307

Reiser, B. J. (2013). *What professional development strategies are needed for successful* *implementation of the Next Generation Science Standards?* Presented at the K-12 Center at ETS Invitational Research Symposium on Science Assessment.

Rimm-Kaufman, S. E., & Hulleman, C. S. (2015). SEL in elementary school settings: Identifying mechanisms that matter. *Handbook of Social and Emotional Learning: Research and Practice.*, 151–166. https://doi.org/10.1007/s11240-008-9485-7

Rogoff, B., & Toma, C. (1997). Shared thinking: Community and institutional variations. *Discourse Processes*, *23*, 471–497. https://doi.org/10.1080/01638539709545000

Ryu, S., & Lombardi, D. (2015). Coding classroom interactions for collective and individual engagement. *Educational Psychologist*, *50*(1), 70–83.

Schonert-Reichl, K. A. (2017). Social and emotional learning and teachers. *Future of Children*, *27*(1), 137–155.

Taylor, R. D., Oberle, E., Durlak, J. A., & Weissberg, R. P. (2017). Promoting positive youth development through school-based social and emotional learning interventions: A meta-analysis of follow-up effects. *Child Development*, *88*(4), 1156–1171.

Tippett, C. (2009). Argumentation: The language of science. *Journal of Elementary Science Education*, *21*(1), 17–25.

Trygstad, P. J., Smith, P. S., Banilower, E. R., & Nelson, M. M. (2013). *The status of elementary science education: Are we ready for the Next Generation Science Standards?* Horizon Research, Inc.: Chapel Hilll, NC.

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

Investigating the Role of Supportive Classrooms for the Next Generation of Students in Elementary Science

Ashley Hunt, University of Virginia

Sara E. Rimm-Kaufman, University of Virginia

Jose Olaís, Northern Arizona University

Abstract

This study examined foundational elements of a neuroscientific learning model (Hohnen & Murphy, 2016) by measuring how student perception of classroom climate related to their concurrent engagement in science. We also explored how the proportion of English Learners (ELs) in the classroom moderated the relation between student-reported climate and engagement. Based on survey data from 832 students in 39 classrooms, multilevel regression analyses found that student perceptions of positive social support in the classroom explained a significant portion of the variance in both behavioral and social engagement in science. Negative social experiences, however, did not relate to engagement. Linguistic diversity moderated the association between positive social support and academic engagement. In classrooms with ELs, especially those that enrolled a majority of ELs, a more supportive climate predicted significantly higher behavioral (b=.44, p<.001) and social engagement in science (b=.88, p<.001). Results highlight the importance of a positive social climate in for effective implementation of the Next Generation Science Standards.

Keywords: engagement; academic engagement; classroom climate; English learners; Next Generation Science Standards

Manuscript 2: Investigating the Role of Supportive Classrooms for the Next Generation of Students in Elementary Science

The Next Generation Science Standards (NGSS) call for a shift towards a model of instruction where students' personal histories and experiences with the natural world act as the "raw materials" of science teaching (Larkin, 2020; Penuel & Reiser, 2018). The disciplinary core ideas, science and engineering practices, and crosscutting concepts of NGSS push educators to challenge *all* students to develop conceptual knowledge and habits of mind of scientists. Classrooms are no longer factories for knowledge acquisition; they are sociocultural environments where students learn about academic content, others, and themselves, (Gutierrez, 2008). Student-centered science instruction that creates more equitable learning experiences can be challenging to enact well. Children who do not feel comfortable sharing ideas, asking questions, or expressing disagreement might feel at a disadvantage in this type of environment (Olsen, 2008; Osterman, 2000; Ryan & Pintrich, 2007). Students might struggle to share their ideas and questions out of fear for social rejection (Pruitt, 2014). One way to better understand and accommodate this struggle is by incorporating knowledge about physiological experiences of stress into psychological models of learning.

Most educators are familiar with the concept of fight or flight, but many might not realize how physiological stress responses can inhibit students' ability to learn. Four commonly observed behavioral responses to the physiological experience of stress include fight, flight, freeze, and appease (Hammond, 2015). In the context of an elementary science classroom, fight and flight create extreme and behaviorally obvious impediments to instruction. Freeze and appease, however, might be more prevalent and harder to detect among students. For example, imagine an English learner (EL) student confronted with a question that includes high-level vocabulary that responds with only a blank stare. What a teacher might see as refusal to participate could actually reflect experiences of distress beyond the expected struggle of new learning. In contrast, appeasement takes the form of behavioral compliance. To an observer or teacher, students may appear engaged, but in fact, they are just being compliant (Rimm-Kaufman, Baroody, Larsen, Curby & Abry, 2014). Using active, student-centered science instruction makes it more difficult for students to go unnoticed. Instructional opportunities that challenge students to make meaning of content in active, visible ways give teachers better information about student understanding.

When an individual experiences stress, automatic physiological processes pull mental resources away from the cognitive demands of learning to prepare their bodies to fight or flee (Pawlak et al., 2003). It is important to keep in mind that children who live and learn surrounded by an unfamiliar language or those who have experienced discrimination tend to perceive the environment as more threatening, (Hammond, 2015). Subsequently, students living in foreign or dangerous environments face physical limitations to managing classroom stress while trying to learn. Since active engagement during science teaching serves as both a desired behavior and a potential stressor, educators must reach a delicate balance between challenging students and meeting their individual academic, social, and emotional needs. Research suggests that students who feel emotionally supported are more academically engaged, which in turn produces more authentic and deep learning due to sustained participation (Rimm-Kaufman, Baroody, Larsen, Curby, & Abry, 2014). Ensuring that the classroom climate is positive, supportive, and helpful might be critical for helping struggling students manage stress and motivating them to participate in challenging science tasks. It is plausible that positive, supportive classrooms may be even more important for students more likely to experience distress in the classroom due to a language barrier.

Drawing on an interdisciplinary model of learning informed by advances in neuroscience, (Hohnen & Murphy, 2016), this study investigated how student perception of classroom climate related to their feelings of behavioral and social engagement during science learning. First, we examined the relation among student report of positive social support and negative social experiences in the classroom and their engagement with the expectation that more positive environments would relate to more engaged students. Next, we explored how the interaction between climate and the proportion of EL students in the classroom related to engagement. Based on research indicating that EL students might be more sensitive to physiological stress responses than their English-fluent peers, we expected that a positive climate would be particularly important for engaging students in classrooms that enrolled more EL students.

A Neuroscientific Model of Academic Engagement and Optimum Learning

In educational psychology, academic engagement is viewed as a multi-dimensional construct of effortful participation that reflects a student's process of developing intrinsic motivation for continued learning (Skinner et al., 2008). The extent to which students reflect a motivation towards content understanding underlies the deep learning in science described by NGSS (National Research Council, 2012). Accordingly, high levels of engagement can serve as an indicator that the goals of NGSS have been met in a science classroom. Engagement is also associated with positive personal and educational outcomes including higher self-esteem, resilience, attendance, and graduation rates (Fredricks & Eccles, 2008; Ladd & Dinella, 2009). Recent decades have seen increasing efforts to leverage innovations in the study of brain development and function to inform educational practice. Despite many remaining questions about the brain, educators can learn from well-established understanding of certain key features and functions of the human nervous system.

In 2016, Hohnen and Murphy proposed a neurodevelopmental model that positions engagement as a critical midpoint in a hierarchy of learning contexts and outcomes they call the "optimum context for learning," (2016, p. 85). The model adapts Maslow's hierarchy of needs by associating behaviors that underlie engaged learning with physical structures and processes in the nervous system across development. Repetition and practice increase the speed of connections among brain areas needed to complete a certain task. Over the course of development, children increasingly draw on their past experiences to make decisions about how to behave as the brain steadily increases the speed of connections that have been established as safe and rewarding; in short, those that fire together wire together (Shatz, 1992). Therefore, if students have opportunities to experience meaningful engagement in science, they will become increasingly likely to continue participating.

Hohnen and Murphy's model (2016) integrates advances in neuroscience with corresponding investigations of learning and psychology to describe a developmental trajectory towards emotional well-being, empowerment, and equity. The foundation of the model identifies elements of the classroom context that activate base survival drives in the brain: to survive, and to seek social connection (Hammond, 2015). These drives are informed by information about the environment interpreted in the brain by a structure called the amygdala. Students' sense of safety, the quality of their relationships, and the sensitivity of their automatic stress responses set the neurological baseline for the classroom environment. When students interpret their environment as physically and socially safe, they enter a state of relaxed alertness ideal for learning (Gutierrez, 2008). The subsequent levels of the pyramid concern processes in the midbrain (or the emotional brain) and the forebrain (or the thinking brain). When children experience distress in the classroom, the survival drives alerted by perceived threats turn off power to the more sophisticated thinking brain (Pawlak et al., 2003).

This study examined three foundational levels of the model during a single time point (denoted in Figure 1 as elements below the dotted line). These analyses provide a snapshot of how elementary student perception of the learning context relates to two elements of their concurrent engagement: behavioral, or the extent to which students report actively participating in science activities; and social, the extent to which students report helping each other learn and solve content-related problems. Much of the research on students' academic engagement relies on teacher report measures (e.g., Hughes & Kwok, 2007; Valiente, Lemery-Chalfant, Swanson & Reiser, 2008). However, asking students about their perspective on academic engagement can contribute unique information about classroom dynamics (Rimm-Kaufman et al., 2014). The late elementary school years are also an important time to spark students' interest in science to support their persistence in the field, even as content becomes increasingly complex and specialized.

Classroom Climate and Supportiveness

An emotionally supportive classroom climate helps students feel more comfortable participating in the collective learning experience. This is especially true for students that struggle engage in learning due to underlying academic, social, or emotional factors, (Martin & Rimm-Kaufman, 2015). The brain works best under collaborative conditions that allow for interaction with others (Hanson, 2013; Zull, 2002). Cultivating a supportive climate focuses on meeting student needs, encouraging positive teacher and peer relationships, and creating varied opportunities for active participation in learning (Rimm-Kaufman & Hulleman, 2015; Sandilos, Rimm-Kaufman, & Cohen, 2017).

In the classroom context, the social climate plays an important role in determining whether students have the necessary physiological resources to engage. A positive, supportive classroom climate means students experience belonging and community relatedness frequently and overtly negative social interactions infrequently (Osterman, 2000). This might include connecting content to students' lives and experiences, deploying a variety of grouping strategies, giving ample time for questions and corrections, and holding consistently high expectations (Szpara & Ahmad, 2007; Woolfolk, 2012). Even children who show high levels of behavioral engagement at the beginning of the school year have demonstrated declines without emotional resources to motivate their continued participation (Skinner et al., 2008).

Supporting English Learners in Science

Students who receive EL services through their schools represent a large segment of the population. Nearly half of districts and schools in the U.S. enroll students requiring EL services, increasingly in classrooms alongside their English-fluent peers (Zehler et al., 2003; Larkin, 2020). In fall 2015, about 4.9 million public school students were identified as English language learners (de Bray et al., 2019). Though the population of emergent bilingual students share the experience of learning English, they represent a multitude of racial and cultural identities. An estimated three quarters of students who received EL services (or 3.8 million students) come from a Hispanic background (de Bray et al., 2019). Despite the prevalence of emergent bilingual students, language proficiency is often sidelined in race-centered considerations of educational equity (Sandilos, Baroody, Rimm-Kaufman, & Merritt, 2020).

Providing emergent bilingual students with equitable instruction is a critically important goal in science instruction. Trends have shifted away from different expectations for students in different academic tracks and toward more equitable access in instruction. This creates a unique challenge for teachers. It can be difficult for teachers to structure a student-centered learning environment that capitalizes on the advantages provided by language diversity. These conditions expose the power dynamics that underlie student experiences of mathematics instruction: emergent bilingual students are responsible for managing both their linguistic and conceptual understanding. Independent of any other individual characteristics, native English speakers experience a level of access to mathematics content that those still learning English do not. One promising path toward helping all students achieve their potential focuses on understanding the psychological and learning processes that contribute to positive outcomes for EL students (Sandilos et al., 2020).

By implementing programs that serve children from diverse backgrounds, (e.g., culturally responsive pedagogy, bilingual education, testing accommodations, translation services) schools can work towards increasing the equity of learning experiences. Ethnic and cultural minority students have reported feeling disconnected from their teachers and peers due to challenges with effectively communicating their ideas and feelings (Olsen, 2008). Furthermore, students who feel excluded from the school community may participate less in the classroom (Osterman, 2000). As ethnic, linguistic, and cultural diversity in the U.S. has increased, more schools and teachers are educating students who are new to English Furthermore, English Learners (ELs) experience substantial stress related to participation in school, (Woolfolk, 2012)

Approaching Equity Through Linguistic Diversity

The field needs more research on EL students and their learning experiences in science. Within the next decade, the majority of students in the U.S. will be current English learners or will have progressed through an English language learning curriculum (Ortman & Shin, 2011). There are more ELs in American classrooms now than have ever before with 9.4%, or an estimated 4.6 million students, in the 2014–2015 school year (National Center for Education Statistics, 2017). Increasingly, all educators need to be prepared to teach English learners. This is troubling given a substantial body of work indicating that pre-service teachers do not receive sufficient preparation for working with linguistically diverse populations (Lucas & Grinberg 2008). Effective EL instruction requires that teachers understand EL student’s linguistic, academic, and developmental background, understand the specific language demands of the tasks EL students will be required to complete in the classroom, and then help students meet these specific demands with the appropriate instructional scaffolds that are most likely to promote success (Banse, Palacios, Merritt & Rimm-Kaufman, 2017). One component of effective instruction for EL students is to establish a classroom culture that is prosocial in orientation and values helping others develop their ideas. This approach not only engage all students in learning but can be especially useful for students experiencing science instruction in a language that is new to them (Kibler, Elreda, Hemmler, Arbeit, Beeson & Johnson, 2019; LeClair, Doll, Osborn, & Jones, 2009).

When children learn science in an unfamiliar language, they must learn to recognize academic language and phrases to access science content (August et al., 2014; Lee et al., 2014). The argumentation example further illustrates this challenge: in common speech, “argument” refers to “an exchange of diverging views, *typically an angry one*,” (emphasis added). Prior to engaging in scientific argumentation, students need to understand that the conflict produced by the practice is intended to build knowledge, not escalate into heated disagreement. All students benefit from this distinction being made explicit, but the point might be crucial for non-native English speakers to view scientific argumentation as collaborative instead of antagonistic.

Teachers vary in their approach to teaching and managing their classrooms. Some aspects of teaching involve specific practices related to English Learners or using differentiated instruction to meet students where they are at in terms of readiness. Managing the social environment does not directly rely on an educators’ expertise with science or providing differentiated instruction for students with language needs. Instead, teachers can support all students by maintaining a positive and helpful social climate.

The Present Study

This research uses multi-level regression analyses of student-reported data measuring their perception of the classroom social climate and their own behavioral and social engagement in science while controlling for other important contributors to engagement (like academic achievement and demographic characteristics). We hypothesized that students who described working well with peers and experiencing limited negative social interactions would also report greater behavioral and social engagement during science learning. The second research question concerned our hypotheses about the value of social support for engaging young EL students in science. We expected that a positive, emotionally supportive classroom climate may relate to higher behavioral and social engagement among students in linguistically diverse classrooms.

Method

Data were gathered as part of a randomized controlled trial (RCT) of Connect Science, an NGSS-aligned curriculum that integrates science instruction, social-emotional learning (SEL), and service-learning. See the Connect Science efficacy study (Rimm-Kaufman, Merritt, Lapan, DeCoster, Hunt & Bowers, 2021) for full details of the intervention.

Procedure & Data Sources

The research team recruited 25 schools in a large, urban district in the southeastern US to participate in the Connect Science RCT during the 2017-18 school year. Teachers received professional learning and materials at no cost and a stipend for participating in the research. Teachers were randomly assigned to the intervention (n = 18) or waitlist comparison group (n = 14). The current study uses data from 39 classrooms where students completed academic engagement and science achievement measures during the RCT (n = 832). See Table 1 for further description of the sample.

Most of the participating teachers (n = 31) were white (91%, SD = .30) and female (91%, SD = .30). Teachers had an average of 10 years of teaching experience, (M = 9.53, SD = 6.89) and 24 held Master's degrees (75%, SD = .44). Nearly all the teachers (91%, SD = .30) reported attending SEL professional development independent of the concurrent RCT within the last three years.

On average, study schools reported that roughly half (52%, *SD* = .81) of their students were considered economically disadvantaged, and that 32% tested proficient on state science assessments in the prior school year (*SD* = .19). Classrooms enrolled an average of 22 students (*M* = 22, *SD* = 3, 49% female) and were ethnically diverse (39% African American, 29% white, 20% Latinx, 4% Multiracial, 4% Other). Classrooms were also linguistically diverse with about 5 of the 22 students in the average classroom designated as ELs (*M* = 22%, *SD* = .28, range 0-100%).

Researchers collected school-level district data, teacher-reported classroom data, and student measures of academic engagement and science achievement. District personnel provided information about enrollment, school information, and prior science achievement. Teachers reported on the demographic composition of their classroom at this time. Student survey data were collected during the fall of 2017. Teachers were provided a protocol for administering the surveys that included verbal instructions to give prior to distributing the paper and pencil surveys. Teachers were asked provide accommodations (e.g., reading items aloud, support from a translator) to students who typically received them.

This study investigates classrooms conditions from the student perspective and relies primarily on student-report data. Student surveys were developed based on thorough review of previously validated measures, pilot testing, and focus groups with fourth graders to check for understanding. Reliability analyses and item-rest correlations were used to identify items for removal, which resulted in the final measures of engagement and classroom supportiveness described below.

Outcome Measures

Academic engagement in science was assessed using a 10-item measure survey of students' feelings about learning (α = .82). Five items (α = .71) were adapted from an existing survey of behavioral engagement (Skinner et al., 2009). Students rated whether statements such as, "I try to do well in science" and "I pay attention in science class" were not at all true (1) to very true (4). The remaining five items (α = .78) were adapted from an existing measure of social engagement, (Patrick, Ryan, & Kaplan, 2007) a dimension of academic engagement that indicates the extent to which students have content-related social interactions. Students rated the truthfulness of items like "I share my ideas and materials with other kids in science," and "I answer questions about science in class," from not at all true (1) to very true (5).

Confirmatory factor analysis indicated that two-factor model was an adequate fit the data, RMSEA=.067, (90% CI [.055, .085]), CFI=.95, SRMR=.04. Standardized factor loadings for outcome survey items are presented in Table 2.

Other Student-Report Measures

Student self-report data about classroom climate was the independent variable of interest. We also measured student science knowledge, as academic performance is often highly correlated with academic engagement.

Classroom Climate. Student perception of classroom climate was measured using an existing 14-item measure (Developmental Studies Center, 2005). Students rated how much they agreed with statements on a scale from 1 (disagree a lot) through 5 (agree a lot). Items represented the extent of positive social support and the presence of negative social experiences in the classroom. Students rated how much they agreed with statements like, “When I’m having trouble with my schoolwork, at least one of my classmates will help,” and “Students in my class are mean to each other,” (α = .84). Negatively-worded items of were reverse-coded such that a higher number indicates fewer negative social experiences. Confirmatory factor analysis indicated that a two-factor model was a good fit to the data, RMSEA=.039, (90% CI [.026, .050]), CFI=.98, SRMR=.04. Standardized factor loadings for all items are presented in Table 2.

Science Achievement. Students completed an assessment of their understanding of the NGSS disciplinary ideas related to energy and resource use. The measure consisted of 13 multiple choice items (α = .68). Since this assessment was developed by the evaluation team, it has not been validated with other samples.

Classroom Measures

The additional measures included in the study were at the classroom level. Including known classroom characteristics in the model allowed for analysis of the relation between classroom climate and academic engagement in science.

EL population density. Teachers reported the number students in their classroom designated as receiving English-language learning services. We used this number and the reported classroom enrollment to calculate the proportion of EL students in each classroom.

Additional covariates. We included a selection of control variables in the analyses to reduce residual variance and account for expected correlations between classroom characteristics and study outcomes. This included whether the teacher reported attending professional development for SEL within the last three years, class size, and group assignment (intervention or comparison) in the Connect Science RCT.

Analyses

We began by conducting descriptive analysis and reviewing bivariate correlations among variables. Sample information is displayed in Table 1, and descriptive statistics for variables included in the final models are listed in Table 3. Following from Hox (2010), we conducted a multi-level random intercept model including key predictors to distinguish between individual and classroom-level contributions to students' academic engagement in science. Student survey measures of classroom climate and engagement are nested within classroom.

The first set of models (one with behavioral engagement as the outcome, the other with social engagement) analyzed classroom climate as a predictor of behavioral and social engagement in science using the following equations, where Level 1 refers to students and Level 2 refers to classrooms:

(1) Level 1: $Engagement_{ij} = \beta_{0j} + \beta_1 Positive_{1ij} + \beta_2 Negative_{2ij} + \beta_3 ScienceAch_{3ij} + \beta_4 \%EL_{4ij}$ $+ B_5 SELpd_{5ij} + \beta_5 ClassSize_{5ij} + \beta_6 Condition_{6ij} + e_{ij}$

Level 2: $B_{0j} = g_{00} + g_{01} Z_j + u_{0j}$

The second set of models tested for an interaction between the positive and negative classroom climate factors and EL population density (both at Level 2). Again, the same models were estimated separately for the behavioral and social engagement outcomes. This allowed for a moderation analysis that tested the extent to which the relation between the two elements of social climate and academic engagement varied by classroom EL population density, as represented in the equations below:

(2) Level 1: $Engagement_{ij} = \beta_{0j} + \beta_1 Positive_{1ij} + \beta_2 Negative_{2ij} + \beta_3 ScienceAch_{3ij} + \beta_4 \%EL\ X\ MeanPositive_{4ij} + B_5 SELpd_{5ij} + \beta_5 ClassSize_{5ij} + \beta_6 Condition_{6ij} + e_{ij}$

Level 2: $B_{0j} = g_{00} + g_{01} Z_j + u_{0j}$

(3) Level 1: $Engagement_{ij} = \beta_{0j} + \beta_1 Positive_{1ij} + \beta_2 Negative_{2ij} + \beta_3 ScienceAch_{3ij} + \beta_4 \%EL\ X\ MeanNegative_{4ij} + B_5 SELpd_{5ij} + \beta_5 ClassSize_{5ij} + \beta_6 Condition_{6ij} + e_{ij}$

Level 2: $B_{0j} = g_{00} + g_{01} Z_j + u_{0j}$

To account for a small amount of missing classroom climate and science achievement data (<5%), multiple imputation was used to estimate models.

Results

Descriptive Analyses

Descriptive statistics for all variables included in the fitted models are provided in Table 3. Overall, students in the sample reported neutral or relatively positive social climate. On a scale of 1 to 5, the classroom climate indicators averaged 3.54 (*SD* = .92) for positive social support and 3.47 (*SD* = 1.08) for negative social experiences (where higher numbers indicate fewer negative experiences). Students reported relatively high behavioral (*M* = 3.50, *SD* = .48) and social engagement (*M* = 3.72, *SD* = .92) in science, indicating that the average student in the sample perceived themselves as positively engaging in science learning "often." Positive social support, negative social experiences, and science achievement (*M* = .75, *SD* = .19) were moderately related to one another (correlation coefficients ranging from .10-.37) and the engagement measures (correlation coefficients ranging from .18-.40), where higher scores for classroom climate and achievement were associated with higher engagement.

The proportion of EL students in each classroom ranged from zero to 100%. On average, 5 students per classroom of 22 students (22%) were designated as ELs. Classroom EL population density was moderately related to both positive social support (.26, p<.001) and negative social experiences (.23, p<.001), indicating that classrooms with a higher proportion of EL students reported a more positive climate overall. Classrooms with more ELs were moderately less likely to be led by teachers with recent SEL professional development (-.29, p<.001), to be in the intervention group in the Connect Science RCT (.26, p<.001), and, to a lesser extent, to have higher enrollment (.08, p<.05). A small but significant negative association (-.12, p<.001) was found between EL population density and science achievement. No association was found between classroom EL population density and behavioral or social engagement in science.

Unconditional Model

The first model tested for between-class variation in the relation between classroom climate and academic engagement. Variance component model results revealed that variance in academic engagement was largely explained at the individual model. The intra-class correlation (ICC) for behavioral engagement was .05, which means that 5% of the variance in behavioral engagement was at the classroom level. The ICC for social engagement (.09) was higher: analyses found that 9% of the variance explained was at the classroom level. The ICCs for engagement align with previous studies considering student-report of social-emotional characteristics.

Regression Models Estimating the Relation Between Classroom Social Climate and Academic Engagement in Science

The next set of models measuring the associations between student-reported positive support and negative social experiences and engagement while controlling for EL population density, science achievement, class size, condition, and whether the teacher had recently attended SEL professional development are reported in Table 3. Student report of positive social support was positively predictive of behavioral (*b*=.19, *SE*=.02) and social engagement in science (*b*=.42, *SE*=.04). Presence of negative social experiences and EL population density were not significantly related to engagement.

Student experiences of positive social support explained a significant portion of the variation in behavioral and social engagement at the student level, but not the classroom level. The ICC for social engagement was .03, indicating that 3% of the variance in social engagement captured was at the classroom level. None of the variance in behavioral engagement was explained at the classroom level.

Regression Models Testing for an Interaction Between Positive Social Support and Classroom EL Population Density

The final set of models tested for an interaction between positive social support and classroom EL population density as reported in Table 3. These analyses found a significant interaction between positive social support and EL population density for both behavioral (b=.36, SE=.16) and social engagement, (b=.83, SE=.33). This finding indicates that, on average, a single point increase in positive social support was more strongly associated with engagement in classrooms that included EL students than in entirely EF classrooms. To elaborate on the interaction model results, we plotted three regression lines that grouped classrooms by EL population density:

1) *High:* Six classrooms enrolling 14% of students in the study (n = 112) enrolled a majority of EL students (>50%). In these classrooms, most or all students were non-fluent English speakers. On average, 16 of 21 students (M = .75, SD = .13) per classroom were ELs.

2) *Low:* Sixteen classrooms enrolling 42% of students in the study (n = 344) enrolled some EL students but were majority EF. On average, 6 of 22 students (M = .26, SD = .17) per classroom were ELs.

3) *None:* Eighteen classrooms enrolling 44% of students in the study (n = 376) did not enroll any EL students. 100% of the students in these classrooms were EF.

Plotting the marginal means by these categories highlighted the practical differences between the three types of classrooms (see Figure 1). Positive social support was significantly associated with higher behavioral and social engagement regardless of classroom language diversity. The magnitude of the interaction coefficients varied in two systematic ways. Coefficients were the highest in majority EL classrooms for both behavioral (b=.35, p<.001) and social engagement (b=.90, p<.001). Second, coefficients were larger for social engagement (b=.44-.90, all ps<.001) than for behavioral engagement (b=.17-.35, all ps<.05) across all three groups. None of the interactions between negative social experiences and EL population density were significant.

Discussion

The human brain is driven by two priorities: to evade threat, and to build social connections (Hammond, 2015; Hohnen & Murphy, 2016). The neurological processes that allow for survival and connection have implications for learning that have not been fully explored in childhood. As educators move towards student-centered instruction like that recommended by the NGSS, research on learning will advance if we consider how engagement reflects individual differences in students' underlying physiological experience. If students feel stress, they are less likely to engage. If students experience social connection and belonging, they will be more likely to engage.

This study investigated elements of a neuropsychological model of optimum learning by surveying students about their social and emotional experiences in fourth grade science. Two findings emerged. First, multilevel regression analyses found that positive social support related to behavioral and social engagement, while the presence of negative social experiences did not relate to engagement. Second, student perceptions of social support were most strongly associated with engagement in science classrooms with higher proportions of EL students. The NGSS Framework for K–12 Science Education states that "science is fundamentally a social enterprise, and scientific knowledge advances through collaboration and in the context of a social system with well-developed norms" (NRC, 2012, p. 27). The current study contributes to our understanding of aspects of that experience in middle childhood.

Behavioral and Social Engagement Largely Driven by Individual Perception

This study adds value to the existing body of work on academic engagement by focusing on student report of their own feelings and experiences in the science classroom. The multilevel analytic approach revealed that the explained variance in behavioral and social engagement was almost entirely at the individual rather than the classroom level. In other words, even in classrooms with an above average climate, individuals who perceived the environment as unsupportive still reported lower engagement. Teachers play an important role by maintaining a classroom climate that promotes (or inhibits) engagement (Braun, Zadzora, Miller, & Gest, 2019;

Gest, Madill, Zadzora, Miller, & Rodkin, 2014). The observed association between positive classroom climate and engagement highlights how teacher efforts to shape the classroom culture can support diverse classrooms. Effective teachers can harness their personal relationships with students to prompt consistent and active participation in learning (Zee, Koomen, & Van der Veen, 2013). For example, a teacher's awareness of students' prior knowledge can inform flexible grouping decisions that allow students to connect with peers in different educational contexts. These findings highlight how teacher-student relationships can provide critical insight into how to support individual students needs while remaining aware of the general social climate of the classroom.

Positive Classroom Climate Associated with More Engaged Students

Results aligned with prior findings regarding the association between positive peer and teacher-student relationships and academic engagement (Roorda et al., 2017). Student perception of positive social support contributed significantly to their behavior and social engagement in science. Over time, the presence of positive social support from teachers is associated with even greater behavioral engagement, suggesting an amplifying effect of positive social experiences and student engagement (Skinner et al., 2008). Students are subjected to less disciplinary action, set more ambitious educational goals, and express greater motivation towards participation in school when they perceive teachers and other adults at school as warm and caring (Murdock & Miller, 2003).

Contrary to expectation, higher incidence of negative social experiences did not relate to less engagement. There are a few possibilities for this lack of association. First, it is worth noting the correlation between students' perception of positive social support and negative social experiences (r = .37). Given that negative experiences are reverse scored, the relatively low magnitude of the correlation means that just because a classroom is high in positive support does not mean that it is low in negative support (and vice versa). This suggests the importance of measuring positive and negative experiences as two separate dimensions rather than the opposite poles of a single dimension. Second, this finding also suggests that students can tolerate a fair amount of negative feelings about their peers without it relating to their engagement in learning. Although surprising, this is useful information for future work that can examine other outcomes beside engagement to understand the consequences of students' negative social experiences and can continue to hone in on teaching practices that contribute to students' own feeling of social support. Understanding the classroom conditions that promote learning is an important way of supporting high quality implementation of NGSS standards requiring high levels of student engagement.

Social Support Matters More in Linguistically Diverse Classrooms

While sharing and evaluating scientific ideas can be challenging for all students, the work involves an additional cognitive demand for EL students. Experts in English language learning have largely focused on instructional approaches and accommodations and paid less attention to the role of social and emotional competencies that help students learn collaboratively (NRC, 2012). Furthermore, educational environments that foster positive social relationships may help students that are in the process of learning English to sustain their academic performance over time (Kim & Suárez-Orozco, 2015). We found that in classrooms where more than half of the students were ELs, social support related more strongly to behavioral and social engagement in science. In simpler terms, students in these classrooms were collectively more engaged than equally supportive classrooms enrolling only English-fluent students. On one hand, this finding is promising in that it suggests a positive climate might help students overall, and ELs in particular, engage more fully in science learning. On the other, it also indicates that EL students in less supportive environments might do little to engage in science learning despite individual interest or motivation. This begins to shed light on the role of social climate in classrooms where teachers face practical challenges with NGSS implementation due to increased linguistic diversity. Each EL is an individual with a unique linguistic and cultural background, making it challenging for teachers in science classrooms to accommodate language support and learning needs of many students at once. Our findings suggest that cultivating a climate characterized by positive social support might be a key goal for academically engaging larger, more diverse groups of ELs in student-centered science learning.

Limitations & Future Research

This study had two primary limitations. We did not have access to student-level data on potentially important variables, and the sample size was relatively small for the chosen analyses. Additionally, assessment of key variables was based on student surveys only. Although that has some benefit because we know we tapped into each student's actual social experience in school, the findings may have been more complete if we also had systematic observations of the social climate in the classroom. Collecting information from additional sources (e.g., observers, teachers and parents) would allow for greater precision in the accuracy of key constructs.

Based on these finding, we suggest that future work consider how SEL practices can be integrated with science to create a supportive, equitable learning environment for all students. For instance, many social and emotional learning programs build community building into classroom life (e.g., Meet Up in Harmony, Charters in RULER, Morning Meeting in Responsive Classroom). We recommend that teachers pay special attention to the ways that they support the development of positive relationship skills and encourage students to apply those skills during science instruction.

Yet another direction for future work comes from a strength of the present study – the focus on individual student's experience in the classroom from their own perspective. The U.S. is on track for increasingly diverse classrooms with expectations that by 2029, roughly 44% of students in classrooms will be White. The field needs future work that fully considers the heterogeneity within classrooms so that we are positioned better to meet the emotional and learning students of all students.

Closing Comments

Students experience, and thus need to manage and learn from, their emotions and interactions when engaging in the challenging, collaborative work of student-centered science instruction. Furthermore, NGSS places considerable language demand on ELs due to the increased focus on the language-intensive SEPs. The link between student-reported classroom climate and academic engagement highlights the importance of creating a positive context for science learning and communication. Our findings show that a socially supportive climate is associated with higher engagement in science for all students but might be especially important for those from diverse linguistic and cultural backgrounds. A classroom climate where all students feel supported might provide valuable social resources that help them meaningfully participate in authentic science exploration, communication, and problem solving.

Table 1

Participant Characteristics

	M(*SD*)						
		Group Assignment			*Classroom EL Population Density*		
	Total	Intervention	Comparison	*t*-value	High	Low	None
Students	*n* = 832	*n* = 412	*n* = 420		*n* = 112	*n* = 344	*n* = 376
Female	.49(.50)	.50(.50)	.48(.50)	-.70	.60(.49)	.46(.50)	.49(.50)
Classrooms	39	19	20		6	16	17
Enrollment	22(3)	22(4)	21(3)	-7.79***	21(6)	22(3)	22(2)
Percent female	.46(.10)	.49(.10)	.44(10)	-6.40***	.55(10)	.42(.11)	.47(.08)
Classroom EL population density (% ELs)	.22(.28)	.31(.31)	.13(.22)	-7.77***	.75(.13)	.26(.17)	0
Racial composition							
Black	.39(.29)	.28(.16)	.48(.35)	10.26***	.30(.23)	.33(.28)	.47(.30)
White	.29(.26)	.30(.24)	.27(.28)	-2.48**	.11(.09)	.28(.24)	.25(.30)
Latinx	.20(.21)	.26(.20)	.16(.21)	-6.08***	.42(.23)	.28(.19)	.06(.08)
Asian	.05(.06)	.05(.06)	.05(.06)	3.94***	.06(.06)	.05(.07)	.05(.05)
Multiracial	.04(.05)	.04(.04)	.03(.05)	-2.86**	.02(.02)	.04(.05)	.05(.05)
Other	.04(.08)	.07(.10)	.02(.04)	-9.03***	.09(.11)	.04(.08)	.03(.07)
Teachers	32	14	18	.37			
Gender (% female)	.91(.30)	.93(.27)	.89(.32)	-.37			
Percent white	.91(.30)	.93(.27)	.89(.32)	.37			
Percent black	.09(.30)	.07(.27)	.11(.32)	-.39			
Years of teaching experience	9.53(6.88)	10.07(7.30)	9.11(6.73)	-.40			
Obtained Master's degree	.75(.44)	.79(.43)	.72(.46)	-.37			
Recent SEL PD	.91(.30)	.93(.27)	.89(.32)	-.37			
Schools	25	12	13				
Enrollment	205(16)	231(21)	181(23)	-1.60			
Percent ELs	.19(.04)	.24(.05)	.15(.05)	-1.36			
Prior 4th grade science achievement	.32(.04)	.33(.05)	.31(.06)	-.25			
Prior 4th grade ELA achievement	.30(.03)	.30(.05)	.30(.05)	-.08			
Prior 4th grade math achievement	.30(.04)	.35(.04)	.26(.06)	-1.22			
Percent with IEPs	.12(.01)	.13(.01)	.12(.02)	-.08			
Percent economically disadvantaged	.50(.04)	.47(.05)	.54(.06)	.87			

Table 2

Confirmatory Factor Analyses for Student-Reported Classroom Climate and Academic Engagement in Science

Factor	Items	Standardized Loading	α
Classroom Climate*			.84
Positive social support			.82
	Students in my class are willing to go out of their way to help someone.	.66	
	My classmates care about my work just as much as their own.	.66	
	My class is like a family.	.67	
	Students in my class help each other learn.	.48	
	Students in my class help each other, even if they are not friends.	.50	
	Students in my class work together to solve problems.	.59	
	When someone in my class does well, everyone in the class feels good.	.69	
Negative social experiences			.80
	The students in my class don't really care about each other.	.60	
	A lot of students in my class like to put others down.	.57	
	Students in my class don't get along together very well.	.49	
	Students in my class just look out for themselves.	.64	
	Students in my class are mean to each other.	.41	
Academic Engagement in Science**			.82
Behavioral engagement			.71
	I try hard to do well in science.	.61	
	In science, I work as hard as I can.	.67	
	When I'm in science class, I participate in class discussions.	.71	
	I pay attention in science class.	.61	
	When I'm in science class, I listen very carefully.	.67	
Social engagement			.78
	During science class I explain how I work out science problems to other kids.	.50	
	I help other kids with science when they don't know what to do.	.45	
	I share my ideas and materials with other kids in science.	.57	
	In science class I help other kids learn.	.45	
	I answer questions about science in class.	.81	

*Data was a good fit to a two-factor model: RMSEA=.039, (90% CI [.026, .050]), CFI=.98, SRMR=.04.

**Data was an adequate fit to a two-factor model: RMSEA=.067, (90% CI [.055, .085]), CFI=.95, SRMR=.04.

Table 3

Descriptive Statistics and Correlations of Predictors and Engagement Outcomes

	1	2	3	4	5	6	7	8	9	10	11	12
Level 1 (Student)												
1. Positive social support	--											
2. Negative social experiences	.37***	--										
3. Science achievement	-.15***	.10*	--									
Level 2 (Classroom)												
4. Positive social support	.41***	.34***	.02	--								
5. Negative social experiences	.24***	.59***	.15***	.59***	--							
6. EL population density	.09*	.12**	-.09*	.26***	.23***	--						
7. Science achievement	.00	.16***	.52***	.02	.28***	-.12***	--					
8. Recent SEL PD	-.03	-.10**	.00	-.11**	-.20***	-.29***	.02	--				
9. Class size	-.04	.11**	.13***	-.08*	.24***	.08*	.23***	-.44***	--			
10. Condition	-.08	-.01	.19***	-.16***	.02	.26***	.40***	.08*	.07*	--		
Outcomes												
11. Behavioral engagement	.37***	.23***	.10*	.20***	.15***	.01	.03	.06	-.01	-.06	--	
12. Social engagement	.40***	.18***	.15***	.21***	.16***	.05	.09	.08*	.00	-.01	.56***	--
Mean (SD)	3.54 (.92)	3.47 (1.08)	.75 (.19)	3.54 (.39)	3.43 (.66)	0.21 (.26)	.76 (.09)	0.62 (.49)	20.93 (3.13)	.49 (.50)	3.50 (.48)	3.72 (.92)
Range	1-5	1-5	0-1	1-5	1-5	0-1	0-1	0-1	6-26	0-1	1-4	1-5

Note: * = *p*<.05; ** = *p*<.01; *** = *p*<.001

Table 4

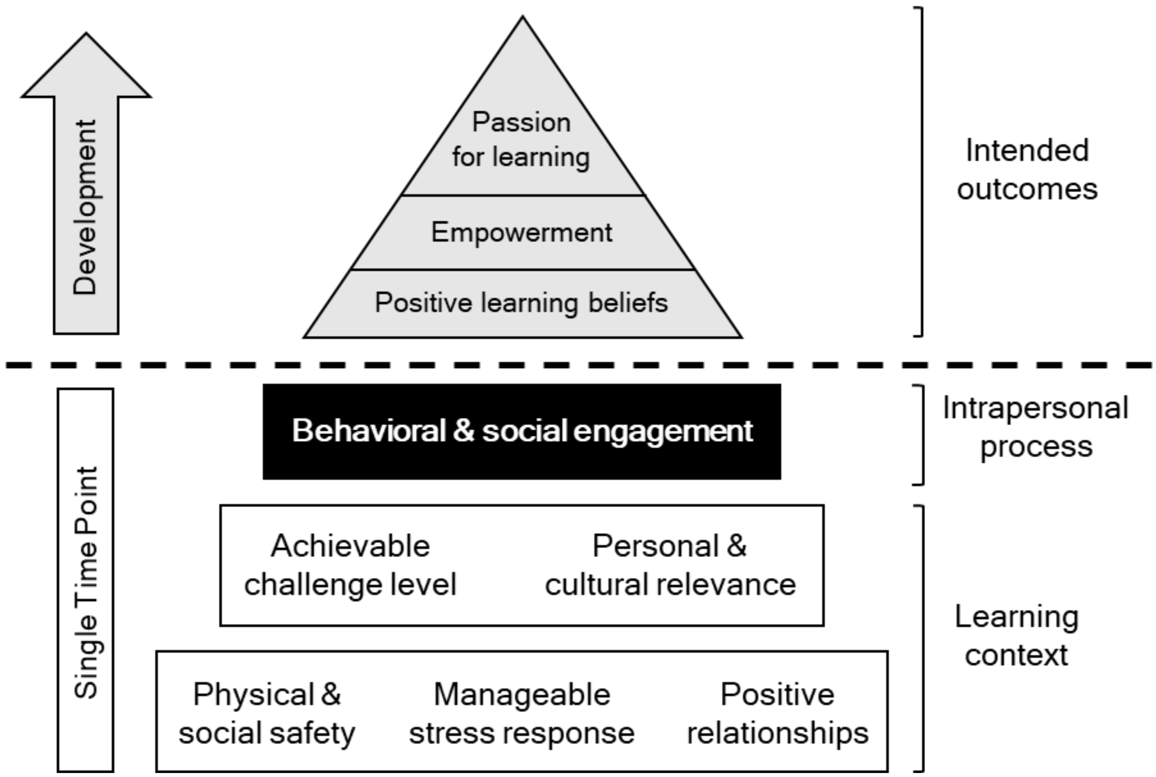
Model Results Explaining Behavioral and Social Engagement in Science

RQ1	**Behavioral Engagement**		**Social Engagement**	
	b	SE	*b*	SE
Level 1 (Student)				
Positive social support	.19***	.02	.42***	.04
Negative social experiences	.03	.02	-.02	.04
Science achievement	.48***	.12	1.12***	.20
Level 2 (Classroom)				
Positive social support	.04	.06	.11	.14
Negative social experiences	.02	.04	.04	.09
Science achievement	-.34	.24	-.21	.52
EL population density	.00	.08	.20	.17
Recent SEL PD	.15**	.06	.37**	.13
Class size	.01	.01	.02	.01
Condition	-.03	.04	-.03	.09
Constant	2.16***	.27	.37	.61
Level 1 Residual	.43	.01	.80	.02
Level 2 Error	.03	.0	.15	.04
ICC	.00		.03	

RQ2	**Behavioral Engagement**		**Social Engagement**		**Behavioral Engagement**		**Social Engagement**	
	b	SE	*b*	SE	*b*	SE	*b*	SE
Level 1 (Student)								
Positive social support	.18***	.02	.42***	.04	.18***	.02	.43***	.04
Negative social experiences	.03	.02	-.02	.04	.03	.02	-.02	.04
Science achievement	.41***	.10	1.08***	.20	.41***	.11	1.13***	.20
Level 2 (Classroom)								
Positive social support	-.06	.08	-.16	.17	.06	.06	.13	.14
Negative social experiences	.02	.04	.07	.08	-.03	.04	.00	.10
EL population density	-1.28**	.58	-2.84*	1.25	-.67	.39	-.41	.88
Recent SEL PD	.16**	.05	.40**	.12	.14*	.05	.36**	.13
Class size	.01	.01	.03*	.02	.01	.01	.01	.01
Condition	-.05	.04	-.04	.08	-.04	.04	-.02	.08
EL population density X positive support	**.36***	.16	**.83***	.33				
EL population density X negative experiences					.19	.10	.17	.23
Constant	2.22***	.26	.72	.56	2.12***	.25	.34	.57
Level 1 Residual	.43	.01	.80	.02	.43	.01	.80	.02
Level 2 Error	.00	.00	.12	.04	.00	1.04	.14	.04
ICC	.00		.02		.00		.03	

Figure 1

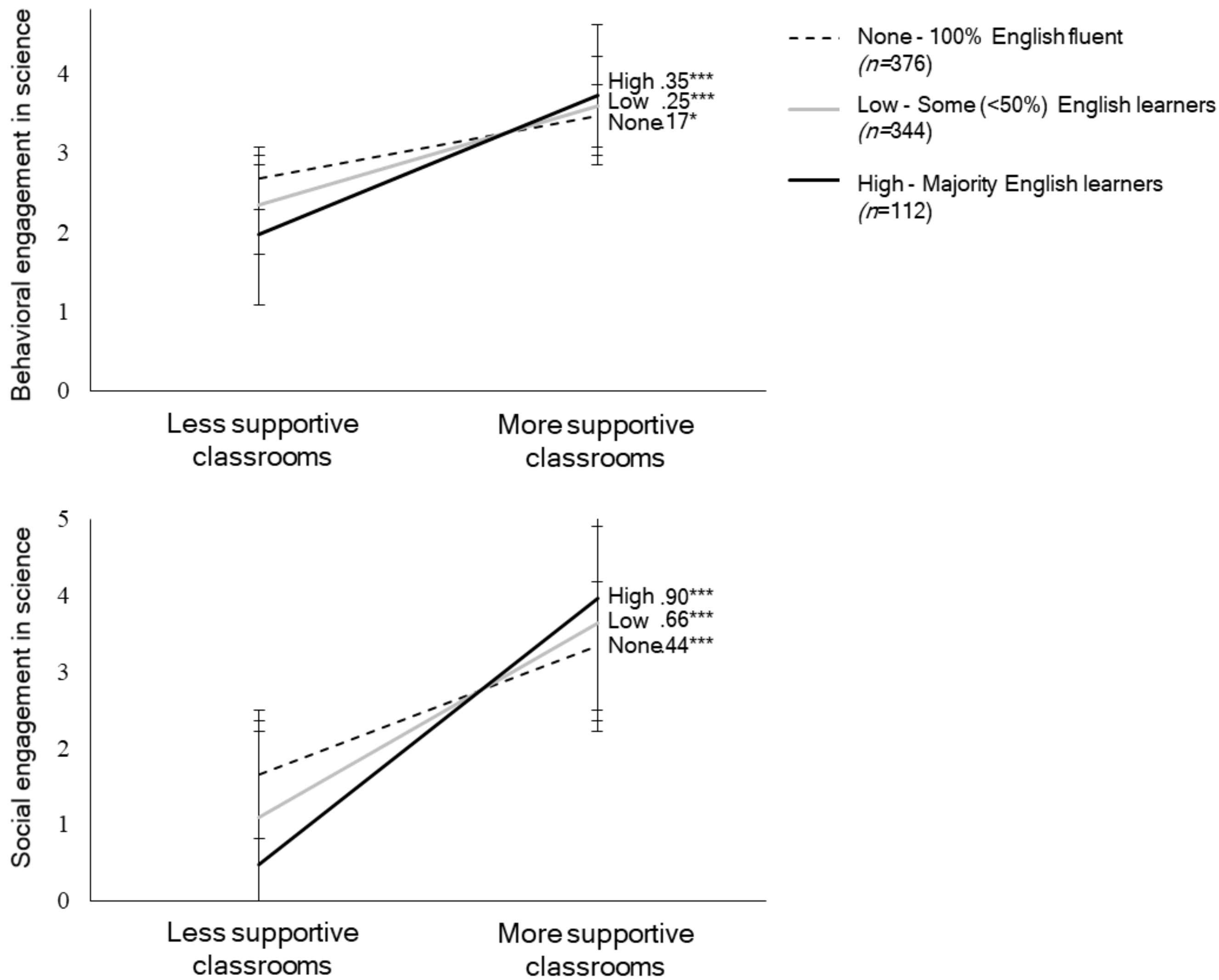
Theoretical Framework of Academic Engagement Adapted from the Optimum Context for Learning Model



Note. Original model described in Hohnen & Murphy, 2016. The current study examines elements below the dotted line.

Figure 2

Interaction Between Positive Social Support and Classroom EL Population Density Explaining Student Behavioral and Social Engagement



References

Banse, H. W., Palacios, N. A., Merritt, E. G., & Rimm-Kaufman, S. E. (2017). Scaffolding English language learners' mathematical talk in the context of Calendar Math. *Journal of Educational Research*, *110*(2), 199–208. https://doi.org/10.1080/00220671.2015.1075187

Braun, S. S., Zadzora, K. M., Miller, A. M., & Gest, S. D. (2019). Predicting elementary teachers' efforts to manage social dynamics from classroom composition, teacher characteristics, and the early year peer ecology. *Social Psychology of Education*, *22*(4), 795–817. https://doi.org/10.1007/s11218-019-09503-8

Gest, S. D., Madill, R. A., Zadzora, K. M., Miller, A. M., & Rodkin, P. C. (2014). Teacher Management of Elementary Classroom Social Dynamics: Associations With Changes in Student Adjustment. *Journal of Emotional and Behavioral Disorders*, *22*(2), 107–118. https://doi.org/10.1177/1063426613512677

Harkins, T., Merritt, E., Rimm-Kaufman, S.E., Hunt, A. & Bowers, N. (2019). *Connect Science*. Curriculum Manual, University of Virginia, Arizona State University, & Harkins Consulting, LLC, Charlottesville, Virginia.

Hohnen, B., & Murphy, T. (2016). *The optimum context for learning; drawing on neuroscience to inform best practice in the classroom*. *33*(1).

Kibler, A. K., Molloy Elreda, L., Hemmler, V. L., Arbeit, M. R., Beeson, R., & Johnson, H. E. (2019). Building linguistically integrated classroom communities: The role of teacher practices. *American Educational Research Journal*, *56*(3), 676–715. https://doi.org/10.3102/0002831218803872

Kim, H. Y., & Suárez-Orozco, C. (2015). The Language of Learning: The Academic Engagement of Newcomer Immigrant Youth. *Journal of Research on Adolescence, 25*(2), 229–245. https://doi.org/10.1111/jora.12130

Larkin, D.B. (2020). *Teaching Science in Diverse Classrooms: Real Science for Real Students* (1st ed.). Routledge. https://doi.org/10.4324/9780429199721

LeClair, C., Doll, B., Osborn, A., & Jones, K. (2009). English language learners' and non–English language learners' perceptions of the classroom environment. *Psychology in the Schools, 46*(6), 568–578. https://doi.org/10.1002/pits

Lucas, T., & Grinberg, J. (2008). Responding to the linguistic reality of mainstream class- rooms: Preparing all teachers to teach English language learners. In M. Cochran-Smith, S. Feiman-Nemser, J. McIntyre, & K. E. Demers (Eds.), *Handbook of research on teacher education: Enduring questions in changing contexts* (pp. 606–636). New York, NY: Routledge.

Martin, D. P., & Rimm-Kaufman, S. E. (2015). Do student self-efficacy and teacher-student interaction quality contribute to emotional and social engagement in fifth grade math? *Journal of School Psychology, 53*(5), 359–373. https://doi.org/10.1016/j.jsp.2015.07.001

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

Ortman, J. M., & Shin, H. B. (2011). Language projections: 2010 to 2020. *Annual Meetings of the American Sociological Association*, 1–23.

http://www.census.gov/hhes/socdemo/language/data/acs/Ortman_Shin_ASA2011_paper.pdf

Osterman, K. F. (2000). Students ' Need for Belonging in the School Community Author (s): Karen F . Osterman Source : Review of Educational Research , Vol . 70 , No . 3 (Autumn , 2000), pp . 323-367 Published by : American Educational Research Association Stable URL : http:/. *Review of Educational Research*, *70*(3), 323–367.

Penuel, W. R., & Reiser, B. J. (2018). Designing NGSS-aligned curriculum materials. *Committee to Revise America's Lab Report*.

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

Rimm-Kaufman, S. E., Baroody, A. E., Larsen, R. A., Curby, T. W., & Abry, T. (2014). To what extent do teacher-student interaction quality and student gender contribute to fifth graders' engagement in mathematics learning? *Journal of Educational Psychology*, *107*(1), 170–185. https://doi.org/10.1037/a0037252

Rimm-Kaufman, S. E., & Hulleman, C. S. (2015). SEL in elementary school settings: Identifying mechanisms that matter. *Handbook of Social and Emotional Learning: Research and Practice*, 151–166. http://search.ebscohost.com/login.aspx?direct=true&db=psyh&AN=2015-24776-010&site=ehost-live

Rimm-Kaufman, S.E., Merritt, E. G., Lapan, C., DeCoster, J., Hunt, A. & Bowers, N. (2021). Can service-learning boost science achievement, civic engagement, and social skills? A randomized controlled trial of Connect Science. *Journal of Applied Developmental Psychology*, 74. https://doi.org/10.1016/j.appdev.2020.101236.

Roorda, D. L., Koomen, H. M. Y., Spilt, J. L., Thijs, J. T., & Oort, F. J. (2013). Interpersonal behaviors and complementarity in interactions between teachers and kindergartners with a variety of externalizing and internalizing behaviors. *Journal of School Psychology*, *51*(1), 143–158. https://doi.org/10.1016/j.jsp.2012.12.001

Sandilos, L. E., Baroody, A. E., Rimm-Kaufman, S. E., & Merritt, E. G. (2020). English learners' achievement in mathematics and science: Examining the role of self-efficacy. *Journal of School Psychology*, *79*(June 2018), 1–15. https://doi.org/10.1016/j.jsp.2020.02.002

Sandilos, L. E., Rimm-Kaufman, S. E., & Cohen, J. J. (2017). Warmth and Demand: The Relation Between Students' Perceptions of the Classroom Environment and Achievement Growth. *Child Development*, *88*(4), 1321–1337. https://doi.org/10.1111/cdev.12685

Skinner, E., Furrer, C., Marchand, G., & Kindermann, T. (2008). Engagement and Disaffection in the Classroom: Part of a Larger Motivational Dynamic? *Journal of Educational Psychology*, *100*(4), 765–781. https://doi.org/10.1037/a0012840

Zee, M., Koomen, H. M. Y., & Van der Veen, I. (2013). Student-teacher relationship quality and academic adjustment in upper elementary school: The role of student personality. *Journal of School Psychology*, *51*(4), 517–533. https://doi.org/10.1016/j.jsp.2013.05.003

Do Reform Practices and Emotionally Supportive Interactions Promote Equity of Upper Elementary Mathematics Learning Experiences?

Ashley Hunt, University of Virginia

Sara E. Rimm-Kaufman, University of Virginia

Abstract

Does the social environment of elementary mathematics classrooms relate to equity of learner experiences and outcomes in the context of reform instruction? This study analyzed observations of fifth grade mathematics teachers' instruction for use of reform mathematics practices (using the Mathematics-Scan; M-Scan) and emotionally supportive interactions (using the Classroom Assessment Scoring System; CLASS). We used factor analyses to model a construct of equity-promoting mathematics instruction from the observation data, then analyzed the relation between teacher practice, student outcomes, and accessibility. Participants included 387 students (mean age = 10.46 years, *SD* = 0.38) and their teachers (*N* = 59) in a large suburban school district. Each classroom was observed three times during the 2010-11 school year. Multilevel regression analyses did not identify an association between use of equity-promoting mathematics practices and student outcomes. However, the characteristics of classrooms above and below the mean for observed use of reform mathematics practices and emotionally supportive interactions revealed that the schools where teachers were less likely to use equity-promoting mathematics instruction served a greater proportion of the low-income families in the district.

Manuscript 3: Do Reform Practices and Emotionally Supportive Interactions Promote Equity of Upper Elementary Mathematics Learning Experiences?

Standards and recommendations for practice from the National Council for Teachers of Mathematics (NCTM, 2000; 2014) and the Common Core State Standards (CCSSI, 2014) advance the frameworks for standards-based reform practices in mathematics. The contrast between traditional and reform instruction is stark. For example, many can call to mind a memory of a teacher correcting a calculation error by giving the correct answer or testing memorization of multiplication tables (NCTM, 2020; Stein, Remillard, & Smith, 2007). In a reform-oriented classroom the instructional choices that follow a student error are part of the learning process itself. Asking students to explain their thinking engages them in a process of inquiry towards a solution. Over time, high-quality, consistent standards-based reform instruction helps students associate mathematical knowledge with their lived experiences, which can springboard to deeper learning and, subsequently, future interest in mathematics education and careers. *Critical Conversations,* a set of research-based recommendations for implementing reform instruction, situates equity of student experiences and outcomes as a central guiding principle of reform instruction (NCTM, 2014). Whether implementation of standards-based reform practices produce the desired equity of mathematics learning experiences and outcomes remains to be seen.

Transitioning from teacher-centered instruction to a student-centered approach increases the frequency and intensity of social interactions in the classroom between students and with the teacher. Educators shape the social climate of their classroom in both explicit and implicit ways just as much as they determine instruction and assessment, (Jones, Barnes, Bailey & Doolittle, 2017). Teachers that cultivate a welcoming, supportive climate are likely to provide a better context for equitable learning experiences than those characterized by apathy or conflict (Hunt, Rimm-Kaufman, & Olais, under review). Recent decades have seen research and practice converge on the importance of this form of instruction in the field of social and emotional learning (SEL). However, recommendations for improving mathematics instruction rarely capitalize on the potential for also supporting the social and emotional development of students. For example, teachers often ask students to work together in pairs or small groups to solve complex problems. When activities are intentionally designed and well-managed, the communication that takes place during group work makes student ideas visible to themselves and each other. Teachers can guide continued discourse with students, scaffolding towards correct solutions (Pinter, Merritt, Berry, & Rimm-Kaufman, 2017). In a classroom where students feel their ideas are respected, disagreement can lead to collective investigation towards a solution. Without appropriate social supports, the same disagreement could cause distractions or inhibit students from speaking up.

Teachers vary in the degree to which they use reform mathematics practices and have emotionally supportive interactions with students. We have little evidence on whether the two sets of practices have a complementary effect on student outcomes. The present study investigates whether, when implemented together as implied by policy documents, reform mathematics instruction and emotionally supportive interactions relate to more equitable outcomes and experiences for elementary students.

Research Questions

This study used systematic classroom observation data collected as part of a longitudinal RCT of an SEL program called Responsive Classroom (RC, Rimm-Kaufman, et al., 2014). Data from a measure of reform mathematics instruction (M-Scan, Walkowiak, Berry, Meyer, Rimm-Kaufman, & Ottmar, 2014) and emotionally supportive interactions, (CLASS, Pianta, La Paro & Hamre, 2006) were used model a measure of equity-promoting mathematics instruction. Next, descriptive and multi-level regression analyses were used to assess the associations between equity-promoting instruction, accessibility, and student outcomes. The study addresses three research questions:

1) To what extent were teachers observed integrating reform mathematics and emotionally supportive practices?
2) Did greater use of equity-promoting mathematics instruction relate to higher student social skills and mathematics achievement?
3) Did students from diverse racial, ethnic, and socioeconomic backgrounds have access to equity-promoting mathematics instruction?

Literature Review

The current study takes a novel approach to using data generated by observing teacher use of reform mathematics and emotionally supportive practices during elementary mathematics instruction. Using a framework that positions inquiry learning as a force for promoting equity (Tang, el Turkey, Cilli-Turner, Savic, Karakok, & Plaxco, 2017), the review of literature highlights the theoretical overlap and potential for a synergistic relationship between reform mathematics practices and SEL. Finally, theory and knowledge of best practices for and promoting equity of mathematics learning experiences and outcomes probes how equity-promoting mathematics practices might effectively meet the individualized needs of students in diverse classrooms.

Conceptual Framework

This study adapts a framework described by Tang and colleagues (2017) to an elementary context to analyze the intersection of inquiry, equity, and mathematics instruction. The framework identifies four pathways through which reform mathematics practices promote equity: access, identity development, achievement, and power. Reflecting on these four pathways and how they relate to equity-promoting instruction sheds light on possible mechanisms by which reform mathematics practices contributes to equitable mathematics teaching and learning. In theory, equity-promoting instruction is valuable to students in that it helps them engage in learning and achieve effectively.

For equity-promoting instruction to translate to outcomes, students need access. *Access* refers to whether or not students are enrolled in classrooms with teachers that use reform instruction. Research has indicated the tracking and educator decisions determine whether students have access to, for example, advanced and honors-level coursework. *Identity development* refers to the extent to which students have opportunities to associate new concepts with their existing knowledge and identity. This extends to what type or level of content and level of mathematics education students reach based on preconceptions about who is or is not “good at math” (Aguirre, Mayfield-Ingram, &

Martin, 2013). *Achievement* describes the type of knowledge and learning experiences that are valued and praised in the classroom community. For example, if a learning environment prioritizes memorizing procedures and taking tests, achievement in that context means excelling with those specific skills. More equitable approaches draw on students' prior knowledge and experiences as the gateway to learning new concepts (Jacob, Hill & Corey, 2017). Replacing performance with other skills like asking good questions and sharing ideas as highly valued behaviors redefines achievement in the community. As students build confidence in their abilities, mathematics learning can become a dynamic and collaborative experience.

Finally, *power* emerges in mathematics classrooms by determining what gets taught to which children and in which schools. One expression of power in elementary mathematics is the dominance of English-only instruction and discussion even in contexts where emergent bilingual students work alongside bilingual and English-fluent peers (Larkin, 2020). At the upper elementary level, emergent bilingual students are exposed to increasingly complex academic content in their non-native language, (Lee, Quinn, & Valdes, 2013). Discouraging students from connecting conceptual knowledge to their linguistic identities devalues their ways of speaking and communicating, sending a message that mathematics knowledge is only valued when discussed in English.

Taken together, the Tang and colleagues framework (2017) proposes that when educators address issues of access, identity development, achievement, and power in their practice, students are empowered to take ownership of their learning. The framework fits with reform-based instruction in that outcomes are focused on deep learning rather than standardized assessment performance. The framework also implicitly requires a positive social climate where instruction can unfold as teachers use their relational knowledge to personalize instruction. Whether that adeptness translates to academic or developmental gains for students remains unknown.

Reform Mathematics Instruction

Knowledge and experience in mathematics is a highly valued trait in American society and is considered crucial for the academic and professional success 21st century of students (National Mathematics Advisory Panel, 2008; National Research Council, 2011). The NCTM guiding principles point to the importance of high expectations, adequate time, and consistent opportunities to learn as instructional components that enable students to be mathematically successful. Instead of one-size-fits-all practices or differing expectations and academic tracks, equitable access means accommodating differences within a single classroom to meet a common goal of high levels of learning by all students (NCTM, 2014). Creating multiple entry points to a mathematics concept allows more students to find connections between their lived experience and mathematical concepts (Schoenfeld, 2020). Meaningful learning experiences that incorporate challenging questions and opportunities to solve real-world problems have been found to predict greater engagement in mathematics, especially from middle childhood through adolescence (Marks, 2000). This type of instruction can be broadly conceptualized as the use of reform mathematics practices. The next section describes four aspects of reform mathematics teaching and learning: cognitive depth, discourse community, explanation and justification, and problem solving. These practices align with core practices of effective teaching as defined by NCTM (2014) and hinge on students' interpersonal skills and attitudes.

Cognitive Depth

The learning activities chosen and implemented by the teacher determine the level of cognitive demand placed on students during inquiry learning (Stein, Engle, Smith, & Hughes, 2008; Walkowiak, Berry, Meyer, Rimm-Kaufman, & Ottmar, 2014). Cognitive depth refers to the extent to which tasks are open-ended and the complexity of reasoning the teacher expects from students. Maintaining optimal cognitive depth of learning activities requires that teachers differentiate and match the needs of individual students (Pinter et al., 2018). Tasks that lead to deep cognitive engagement for one student may seem too easy for another, but it is important that all students have opportunities to experience productive struggle while learning (Schoenfeld, 2020).

One way to sustain cognitive depth throughout a lesson is to use student questions to build towards new knowledge (Stein et al., 2008). Teachers support student self-efficacy by modeling how to ask and answer questions in useful ways and probing students to provide support for their answers. A sense of self-efficacy can promote engagement and persistence with even the most challenging mathematics problems and tasks (Martin & Rimm-Kaufman, 2015). Recent NCTM instructional recommendations responds to appropriate cognitive depth in the context of procedural and conceptual understanding. Memorizing and successfully replicating procedures might show students have learned something, but it might not further expand their mathematical knowledge. In contrast, more effect approaches view conceptual understanding as foundational for building fluency with procedures, reminding students consider how new skills relate to what they have learned already (NCTM, 2014). One way this appears in many contexts is using scaffolding; teachers increase cognitive depth as they guide students towards new understanding.

Mathematics Discourse Community

Discourse refers to the extent of communication and discussion about mathematics ideas and procedures in the classroom. Teacher efforts to shape discussions and incorporate student ideas into conversation establish the classroom mathematics discourse community. Reform instruction places demands that students develop and effectively use mathematics language with their teacher and peers (Pinter et al., 2018). The community extends beyond technical aspects like getting a mathematics problem correct. The frequency with which students offer up their ideas to the group open them to affirmation or critique also reflects the discourse community (CCSSI, 2010; NCTM, 2000). When students are expected to listen and respond to others' ideas and manage disagreements, their success with those skills determine if group can reach consensus on mathematics ideas or remain disagreeable and chaotic.

Facilitating mathematical discourse can be challenging, especially in a classroom culture where students feel uninterested or disengaged with mathematics instruction. One way teachers stimulate the discourse communities in their classrooms is by situating mathematical discussions around issues of interest to students, including culturally relevant or social justice topics (Dominguez, 2011; Imm & Stylianou, 2012). Framing questions around student thinking is another approach; prompting students to talk through their understanding makes learning visible to both the student and the community. Mathematics discourse expands the learning process by making problem solving visible so that students can analyze, compare, and respond to different approaches.

Explanation and Justification

An extension of the classroom's discourse community is its normative approach to explaining and justifying ideas to others. This involves pressing students to provide reasons for their proposed solutions to problems by asking "how" and "why" questions (NCTM, 2000). The goal of this type of learning is not necessarily for students share a "correct" answer. Instead, the act of reflecting on ideas and connecting them to other concepts is a key mathematical habit of mind (Schoenfeld, 2020). Reform mathematics instruction uses questions to assess and advance students' reasoning and help them make sense about important mathematical ideas and relationships.

Problem Solving

Inquiry learning activities often do not begin with a clear solution method in mind. Instead, they are designed to challenge students to grapple with mathematics problems and apply their existing knowledge to developing new solutions (NCTM, 2000). Problem solving in an inquiry classroom promotes creative conceptual problem-solving rather than presenting mathematics as a set of procedures to memorize (Schoenfeld, 2020). Problem solving relies on effective use of tasks that promote reasoning. Effective teaching of mathematics engages students in solving problems in ways that have more than one correct approach. Learning through problem solving extends beyond the mathematics classroom. Connecting mathematical knowledge and thinking to real world problems can help students see mathematics as applicable to situations in school and their own lives (Smith et al., 2007).

Taken together, reform mathematics practices strive for classrooms where students can grapple with cognitively demanding concepts, talk with their peers about mathematics, clearly express and justify their ideas, and feel comfortable taking risks as they solve problems. However, simply providing more opportunities for students to work together might not be enough to move academic achievement without appropriate support for the underlying social skills (Ottmar, 2019).

Emotional Support in Upper Elementary Classrooms

Two decades of research support the idea that sensitive and responsive interactions between teachers and children and among students themselves are critical to student learning (Roorda, Jak, Zee, Oort, & Koomen, 2017). Emotional support refers to a teachers' responsiveness and warmth towards students as well as an awareness of students' interests and needs in addition to high levels of positive peer interactions and few incidents of bullying, teasing among peers (Pianta & Hamre, 2009). Teacher-student relationships are another key component of the emotional climate of the classroom. Even students described as "aggressive" are more likely to be accepted by peers when they have positive relationships with their teacher (Hughes, Cavell, & Willson, 2001). Beyond the teacher-student relationship, a positive climate within the classroom community is also associated with a number of positive learning outcomes (Gregory & Ripski, 2008). Research has also linked teacher-student closeness to gains in reading achievement (McCormick, O'Connor, Cappella & McClowry, 2013).

Creating a positive classroom climate can be a challenging prospect. One way to do this is with manualized SEL programs. Teachers in this study were randomly assigned to learn and implement the Responsive Classroom (RC) approach, a program shown to elevate the quality of teacher-student interactions in classroom environments (Abry, Rimm-Kaufman, Larsen & Brewer, 2013). The RC approach has a set of principle and practices designed to integrate social and academic learning. For instance, one principle is that skills such as cooperation, assertion, responsibility, empathy, and self-control are essential skills to learn. Another principle is that the process of learning is as important as the product of learning. RC practices emanate from these principles. For instance, a daily Morning Meeting is designed to create a sense of community among students as they start the day. Academic choice is designed to give teachers a structured way for students to exercise autonomy in the way that they learn. Like many programs, RC has changed over time and the principle and practices used in this study are described in a manual developed by Northeast Foundation for Children (2007; see Rimm-Kaufman et al., 2014 for more details). In this study, four constructs are used to conceptualize emotional support: positive and negative climate, teacher sensitivity, and regard for student perspectives.

Positive & Negative Climate

Respectful interactions and positive affect, communication, and relationships are characteristics indicative of a positive classroom climate. This also implies a limited experience of negatively charged interactions. There are varied positive benefits associated with lifting all student-teacher relationships versus focusing exclusively on negative relationships or problem behaviors. One way of doing this is to cultivate a broadly positive and emotionally supportive learning environment.

Teacher Sensitivity

Highly sensitive teachers are attuned to students' academic and social needs. Teacher sensitivity refers to an educator's awareness of and responsiveness to students' needs and concerns. Students' willingness to seek the teacher for comfort and assistance is also a behavioral indicator of teacher sensitivity. As students reach upper elementary school it becomes increasingly important for teachers to be proactive in their anticipation of individual challenges during learning (McCombs, Daniels, & Perry, 2008). Reform mathematics instruction places substantial social demands on both teachers and students. For students, sharing an idea means taking a social risk. Allowing students to direct how the class approaches problems can create surprises and frustration, especially for those used to a structured, procedural way of learning mathematics concepts.

Regard for Student Perspectives

When teachers frequently and effectively incorporate students' interests and ideas into their instruction, they demonstrate their regard for student perspectives. One way that teachers can demonstrate their regard for student perspectives is to allow for meaningful choice during learning. Students' ideas are the raw material of reform mathematics instruction; identifying those ideas is an important first step (Larkin, 2020)

Equity-Promoting Classroom Contexts

This study defines equity-promoting instruction as the integrated use of reform mathematics and emotionally supportive practices. Relatively few studies have examined the role of emotionally supportive practices in discipline specific contexts. In one such study, a warm and collaborative classroom climate at the beginning of the year was associated with higher mathematics achievement at the end of the year (Banse, Palacios, Curby, & Rimm-Kaufman, 2018). In another, student-report of teacher caring was found to relate to mathematics achievement and self-efficacy (Martin & Rimm-Kaufman, 2015). Such findings suggest the value of understanding the culture and climate of the classroom to forecast student outcomes.

Taken together, existing work on reform mathematics and emotional support practices suggests that we ask a series of questions about whether and how the two combine to produce positive student outcomes. The presence of an RCT on an SEL program designed to boost emotional support creates the ideal context to understand the combination of reform mathematics and emotionally supportive practices.

Method

This study uses data from a longitudinal RCT of the RC approach that took place in a large, suburban, Mid-Atlantic school district from 2008-11. The average family income in the district at that time was $103,010 and most elementary schools were relatively high performing (County, 2019). The University of Virginia Institutional Review Board for Social and Behavioral Sciences reviewed and approved all study procedures. Twenty-four schools were randomized into the intervention or waitlist comparison group, and teachers in the intervention schools received training in an SEL intervention called Responsive Classroom (RC; see Rimm-Kaufman et al., 2014 for a full description.). During the final year of the RCT, (2010-11) a subsample of fifth grade teachers, students, and families were invited to participate in additional classroom observations and data collection. Teachers and families completed surveys, and classrooms were observed during mathematics instruction for three one-hour windows over the course of the school year. The end of year standardized mathematics assessment was used as the outcome measure of achievement, and social skills were measured using an additional teacher-report survey.

Participants

The research team invited all schools participating in the RCT to join the additional data collection effort through letters and in-person meetings with administrators. Of the 24 schools in the RCT, 20 (83%) schools agreed to participate. The research team visited schools to recruit fifth grade mathematics teachers to participate; 59 teachers (79%) consented to participate. During the fall semester, students brought a consent form and family survey home. From the families that returned consent forms, the research team randomly selected an average of five students per classroom to be included in the study. Two constraints were applied to random selection: (a) to maintain an equal number of female and male participants and (b) to mirror the demographic profile of the whole school in terms of racial/ethnic composition and the percentage of students designated as ELs or FRPL recipients. This sampling resulted in 387 students from 61 classrooms (two teachers had two classroom sections) joining the study. Teachers were offered a $100 stipend and families received a $20 gift certificate for their participation. Standard district practices were used for family communication, and materials were translated into seven of the most spoken languages in the community.

Students missing mathematics pre- and post-test and the social skills measure were excluded, resulting in a final analytic sample of 363 fifth grade students in 61 classrooms with an average of 20 students per classroom (*SD* = 6). Descriptive statistics of school, classroom, student, and family characteristics are reported in Table 1. The sample included students enrolled in 20 district public schools, seven of which received Title I funds (*M* = 0.34, *SD* = 0.47) and all but two of which had made adequate yearly progress (AYP) in the prior three years (*M* = 0.93, *SD* = 0.26). Teachers had an average of 12 years of experience (*M* = 12.39, *SD* = 8.80), and more than half had earned a Master's degree (*M* = 0.66, *SD* = 0.47).

The student sample was balanced by gender (53% female) and racially diverse (38% White, 21% Asian American, 16% Hispanic, 13% Black, 9% multiracial, and 1% Native American or Hawaiian). 20% of students received special education services (*SD* = 0.40), and 32% qualified for free or reduced-price lunch (FRPL, *SD* = 0.47). This sample of students were linguistically diverse. One quarter (25%) were designated as currently receiving English learner (EL) services through the district, which was used as a proxy for identifying emergent bilingual students. English fluency for a portion of these students was identified using an assessment designed by the WIDA Consortium (assessment data was unavailable for 40% of students receiving EL services). Most students were in the 3-4 range on the WIDA assessment, indicating "Developing – Expanding" English fluency (41%, *SD* = 0.33). The remaining portion of students receiving EL services (19%, *SD* = 0.50) were in the 1-2 WIDA range, indicating "Entering – Beginning" English fluency.

Almost one fifth (18%, *SD* = 0.38) of the total student sample were identified as former recipients of EL services, which indicated their demonstrated fluency in English as a non-native speaker. Whether any of these students were fully bilingual is unknown. Based on parent report, the most common language in non-English speaking homes was Spanish (41%). More than one third of families (36%) indicated "other," which encompassed a global mix of languages (European, Middle Eastern, Southeast Asian, and African). The remaining families reported speaking Vietnamese (11%), Korean (6%), and Chinese (6%).

Procedures

Intervention schools in the RCT were in their third year of RC implementation during the study year (2010-11). Teachers in intervention schools had received two levels of training in the RC approach for two consecutive summers plus three days of school-based coaching and workshops during each school year. Students in the study had been in intervention classrooms for the two years prior to the study. Comparison schools implemented their "business-as-usual" approaches to SEL. Students in the control group typically received no exposure to RC, and teachers were directed to implement their "business-as-usual" approach to SEL. All teachers at RC schools received manuals and resources to support program implementation, which were also made available to teachers at comparison schools once the study ended. See Rimm-Kaufman et al., (2014) for more information.

Study data comes from four sources: school records, a family survey, classroom observations, and teacher surveys. The school district transmitted school record data and prior mathematics achievement to the research team in fall 2010. The family survey, classroom observations, and teacher surveys were collected over the course of the 2010-11 school year, as described below.

School Record Data

The district extracted, de-identified, and securely transmitted student record data for study participants to the research team. Variables provided by the district included prior and outcome year mathematics achievement data from standardized tests, English language learner program status and World- Class Instructional Design and Assessment (WIDA) level, free-reduced price lunch (FRPL) eligibility, gender, age, and race/ethnicity data for students.

Family Survey

Families (typically a parent) completed a survey providing information about the student's home and family life during the fall of 2010. The survey asked parents about the primary language spoken at home, parental education, and other indicators of socioeconomic status, as described below.

Classroom Observations

On-site researchers taped one hour of instruction in each classroom three times during the 2010-11 school year corresponding to three observational windows (Window 1: September to mid-November; Window 2: late November to mid-February; and Window 3: late February to April).

Observation videos were later coded by a team of research assistants trained and deemed reliable on two pre-existing observational measures of teacher practices: the

Mathematics-Scan (M-Scan; Berry, Rimm-Kaufman, Ottmar, Walkowiak, & Merritt, 2010), and the *Classroom Assessment Scoring System* (CLASS; Pianta, La Paro & Hamre, 2006). Coding for all three measures involved a rigorous, multi-step process involving familiarity, training, and checks for reliability and drift. See details below as well as Abry et al., 2013 and Ottmar et al., 2015 for complete information.

Measures

The M-Scan and CLASS were the observational measures included in these analyses. Data provided by the district also served as indicators in analyses for different characteristics like EL program status and FRPL eligibility. Finally, the student academic outcome was measured with a standardized assessment, and social skills were measured using a teacher-report survey.

Reform Mathematics Practices

The M-Scan is a measure of eight dimensions of reform teaching practices: Structure of the Lesson, Multiple Representations, Mathematical Tools, Cognitive Demand, Mathematical Discourse Community, Explanation & Justification, Problem Solving, and Connections & Applications (Berry, Rimm-Kaufman, Ottmar, Walkowiak, & Merritt, 2010). Dimensions are measured on a scale of 1-7, with higher scores indicating more frequent observation of NCTM-aligned mathematics teaching and learning.

An expert panel established the content validity of the M-SCAN dimensions. Subsequent analyses of variance components of the measure found that the largest portion of variance in scores was at the classroom level, indicating that the M-SCAN could differentiate between different levels of reform mathematics practice across classrooms when reliably implemented by trained coders. Finally, the study found correlations ranging from .45 to .79 between each of the eight M-SCAN dimensions and the *Reformed Teaching Observational Protocol*, (RTOP; Piburn et al., 2000) a previously validated observational measure of inquiry teaching practices (Walkowiak et al., 2014).

Researchers on the M-SCAN coding team attended a four-day training workshop. Training included videos, readings, and discussions of reform mathematics instruction each dimension of the measure in-depth (NCTM, 2000; NCTM, 2007). Research assistants coded six hour-long observation videos reached 80% agreement within one scale point before they began coding study data. Biweekly drift tests helped coders to maintain reliability throughout the study, with ICCs measuring inter-rater reliability >.94. Twenty percent of tapes were double-coded, and overall reliability was .83

Observation videos were broken into two 30-minute segments. Coders watched the first segment while taking notes on teacher practices and assigned "soft codes" at the midpoint of the lesson. They then repeated the process for the second half of the video. At the end of the hour-long lesson, coders consulted their soft codes in relation to the second segment and assigned a final rating to eight dimensions of mathematics instruction based on the M-SCAN coding manual (Berry et al., 2010).

Emotionally Supportive Interactions

The measure of emotionally supportive interactions was derived from the CLASS. Quality of teacher-student interactions were measured with ratings of ten dimensions on a scale of 1-7, with higher scores indicating more positive interactions (Pianta, La Paro &

Hamre, 2006). Dimension scores are aggregated into three domains: Emotional Support, Classroom Organization, and Instructional Support (Hamre, Pianta, Mashburn, & Downer, 2007). The three-factor structure of the CLASS measure is widely used in practice and research and correlates with other measures of classroom quality and positive academic outcomes (Pianta & LaParo et al., 2008).

Researchers on the CLASS coding team attended a two-day training workshop. Inter-rater reliability was assessed as recommended by the measure developers using the percent of coder agreement with a master codebook plus or minus one scale point, (Pianta & LaParo et al., 2008). Coders established inter-rater reliability exceeding 80% on sample videos following training. Once official coding began, the team held regular calibration meetings to compare results on randomly selected videos. Inter-rater reliability was assessed at six points during coding, with intraclass correlations ranging from .73-.85. Additionally, 10% of observations were randomly selected for double coding.

Two segments were coded from each hour-long mathematics lesson: minutes 0 to 15 and minutes 30 to 45. Coders took notes on classroom interactions, focusing primarily on teacher behaviors throughout each 15-minute segment. Coders assigned numerical ratings to each dimension based on their notes of behavioral markers and knowledge of manual definitions. The emotional support domain score was calculated by averaging four of these scores across time points: positive climate, negative climate, teacher sensitivity, and regard for student perspectives.

Sociodemographic Factors and Other Covariates

The district provided prior year mathematics achievement, which was measured using scores from the Standards of Learning assessment, the same state standardized test as the outcome achievement measure (Virginia Department of Education, 2010). The district also indicated which students were current and former recipients of EL services, school Title 1 status, whether the school had made AYP in the previous three years, and school-level data about student eligibility for FRPL and EL services.

Family socioeconomic status was measured using school record and family survey data. The family survey included a 7-item questionnaire about educational and financial resources in the home. Questions included: “Does your child have access to a computer?” and “How many bedrooms does your home have?” A vector including school Title I status and family survey responses from was included in model estimation as an indicator of family SES ($\alpha = 0.78$). The family SES indicator included school Title I status, student FRPL eligibility, parental education, whether the student had access to computer, and the number of bedrooms in their home.

Student Outcomes

Social Skills. Teachers reported on social skills for the students in their class using the *Social Skills Improvement System* (SSIS) measure at the end of the school year (Gresham & Elliott, 2008). Teachers were asked to reflect on the student’s behavior during the prior two months and answer questions about how frequently they observed a list of twenty behaviors like “this student speaks in an appropriate tone of voice” and “this student invites others to join in activities.” Teachers rated each item on a scale from “never” (1) to “very often” (4; $\alpha = .90$).

Mathematics Achievement. The standardized state mathematics assessment was used as the academic outcome measure. The 2011 fifth-grade Standards of Learning (SOL) assessment, (Virginia Department of Education, 2011) measured student mathematics achievement at the end of the year. Scale scores were continuous ranging from 200 to 600, with a score of 400 or higher indicating proficiency. The mathematics SOL consisted of 50 multiple choice items (Grade 5, $\alpha = 0.88$) assessing students' procedural knowledge and conceptual understanding of computation and estimation (e.g., order of operations), measurement (e.g., volume), geometry (e.g., classifying angles and figures), and probability and statistics (e.g., mean, median, mode). Students receiving EL services have the option to complete a “Plain English” version of the assessment.

Analyses

Analyses consisted of three primary steps. The first used factor analyses to construct a measure of observed use of reform mathematics and emotionally supportive practices in each classroom based on M-Scan and CLASS data. The second step used multilevel regression analyses to test whether observed use of reform practices and emotional support related to student outcomes. Finally, comparative analyses examined the accessibility of classrooms where teachers were most frequently observed using equity-promoting instruction with a focus on low-income families and children from historically marginalized backgrounds. All analyses were conducted using Stata (v 15.0).

Descriptive Statistics and Bivariate Correlations

Analyses began with standard review of the descriptive statistics all study variables. This included testing for outliers and patterns of missing data. An advantage of the dataset was its completeness; no adjustments for missing data were necessary for model estimation.

Factor Analyses of M-Scan Reform Practices & CLASS Emotional Support

The CLASS dimension of emotional support has a history of use as a standalone predictor (Martin & Rimm-Kaufman, 2015), and previous M-Scan analyses have employed a one-factor solution to model teacher use of NCTM-aligned mathematics practices (Ottmar et al., 2015). Exploratory factor analyses of observation data from the full RCT were used to construct the measures of reform mathematics practices and emotionally supportive interactions tested in the later models.

The CLASS emotional support domain score was used independently due to the conceptual overlap between Classroom Organization, Instructional Support, and the M-Scan. The emotional score is a composite of four dimensions: (a) Positive Climate; (b) Teacher Sensitivity, (c) Regard for Student Perspectives, and (d) Negative Climate. High scores (i.e., 6-7) indicate observations of mathematics instruction including multiple instances of interactions reflecting the first three dimensions and limited examples of the fourth.Factor analyses using the *sem* command generated factor scores based on 450 M-Scan and CLASS observations of 107 classrooms during the final year of the RCT.

Multilevel Regression Model Development and Estimation

Students in the study were nested in classrooms, which were nested in schools that were randomly assigned to a condition in the RC evaluation. To account for the nesting of the data, we conducted a multi-level random slope-intercept model to estimate contributors to students' mathematics achievement and social skills at each level (Hox, 2010). Observed reform mathematics practices and emotionally supportive interactions, prior mathematics achievement, and RC evaluation condition were included as key predictors of interest in each model. Following descriptive analyses and review of correlations among variables, a series of models tested the role of potential covariates. Student gender, teacher years of experience, and class size were included in final models.

Descriptive Analysis of Accessibility of Equity-Promoting Instruction

Following model estimation, as series of basic descriptive analyses were conducted to explore the characteristics of schools and classrooms where teachers were observed using equity-promoting mathematics practices more frequently. T-tests were conducted comparing school, student, and family characteristics between classrooms in above and below the mean for observed use of reform mathematics practices and emotionally supportive interactions.

Results

School, classroom, family, and student data collected during a longitudinal evaluation of the RC approach were used to investigate three research questions:

1) To what extent were teachers observed integrating reform mathematics and emotionally supportive practices?
2) Did greater use of equity-promoting mathematics practices relate to higher student social skills and mathematics achievement?

3) Did students from diverse racial, ethnic, and socioeconomic backgrounds have access to equity-promoting mathematics instruction?

This section begins with an overview of descriptive and correlational analyses of school, classroom, and family characteristics (summarized in Tables 1 and 2). The results continue with a description of factor analyses of the M-Scan and CLASS observational data (see Figure 1). Multilevel regression model development and estimation are outlined next, including a summary of results as they relate to equity-promoting mathematics instruction and RC program implementation. Finally, the section concludes with results that compare classrooms where teachers were above and below the mean for observed use of reform practices and emotionally supportive interactions (see Figure 3).

Descriptive and Correlational Analyses

Descriptive statistics and correlational analyses are presented in Table 2. On average, 35% of students in study schools were eligible for FRPL (*SD* = 0.24) and 31% received EL services (*SD* = 0.18). Most schools (92%, *SD* = 0.28) had met AYP for the last three years. Analyses of family survey data found that nearly all students (94%, *SD* = 0.23) had access to a computer to use at home and that most parents (85%, *SD* = 0.35) expected their child to obtain a two or four-year post-secondary degree. Roughly half of mothers (51%, *SD* = 0.50) and fathers (48%, *SD* = 0.50) held at least an Associate’s or Bachelor’s degree. Very few students had a mother (6%, *SD* = 0.23) and/or father (5%, *SD* = 0.22) that had not graduated from high school. School and family data provided a variety of income indicators to include in the model but were highly related and certain child-level data missing for a small number of students (<2%). To avoid multicollinearity and ensure parsimony of the model, a vector including school Title I status, student FRPL eligibility, parental education, whether the home had three or more bedrooms, and whether the student had access to a computer at home and a selection of survey as a single indicator of family SES ($\alpha = 0.78$).

Results show that the RC approach had a moderate positive association with emotionally supportive interactions, ($r = 0.25$, $p < .05$) but not the use of reform mathematics practices ($r = 0.07$, $p > .05$). Schools in the comparison condition were more likely to have met AYP ($r = 0.23$, $p < .05$) and to enroll higher percentages of students eligible for FRPL ($r = 0.13$, $p < .05$) and EL services ($r = 0.15$, $p < .05$). Assignment to the intervention condition was also associated with slightly smaller class size ($r = -0.17$, $p < .05$). Treatment condition was not significantly correlated with mathematics achievement in the prior year ($r = -0.07$) or the outcome year, ($r = 0.04$) nor was it related to social skills ($r = -0.06$, all ps $> .05$).

Factor Analyses of M-Scan and CLASS Observation Data

A series of structural equation models (SEM) were conducted on a randomly selected half of the observation data, then the final, best-fitting model (see Figure 1) was confirmed on the full dataset. First, independent models treating reform practices and emotionally supportive interactions as separate constructs were tested, establishing that the M-Scan was not a good fit to the data as a single measure, (RMSEA = 0.148, CFI = .872). A two-factor model allowing the M-Scan and the emotional support domain of the CLASS to covary yielded slightly better fit statistics, (RMSEA = .094, CFI = .904). This model also revealed low factor loadings for several M-Scan subconstructs (structure of the lesson, multiple representations, connections and applications, and student use of tools) and one CLASS emotional support dimension (negative climate). A third model included only M-Scan dimensions with factor loadings above 0.65, which led to another slight improvement in model fit (RMSEA = .087, CFI = .959). A final, more restricted model was tested with the four M-Scan dimensions (cognitive depth, mathematical discourse, explanation and justification, and problem solving) and three of the four CLASS emotional support dimensions (positive climate, teacher sensitivity, and regard for student perspectives), thus excluding the negative climate dimension. Doing so did not improve model fit, (RMSEA = .108, CFI = 0.954).

Based on results obtained from the randomly selected half of data, a final model was fit to the full dataset (shown in Figure 1). The final SEM model included the four M-Scan dimensions with factor loadings above 0.65 (cognitive demand, mathematical discourse community, explanation and justification, and problem solving) as a reform practices factor, and the four CLASS dimensions as an emotional support factor (positive climate, teacher sensitivity, regard for student perspectives, and negative climate). The model indicated that the two-factor solution allowing the two factors to share variance was a good fit to the data (RMSEA = .052, CFI = .980).

Multilevel Regression Analyses of Student Outcomes

Multilevel regression analyses were used to estimate whether exposure to equity-promoting mathematics practices over the course of a school year related to improve student outcomes while controlling for prior mathematics achievement, gender, teacher years of experience, class size, and RC intervention condition. Neither reform mathematics practices nor emotionally supportive interactions were found to be significantly related to student mathematics achievement based on the annual state standardized assessment or social skills based on the SSIS.

Accessibility of Equity-Promoting Mathematics Instruction

To answer the third research question, a series of t-tests were conducted comparing teachers whose observed use of equity-promoting practices were above and below the mean for the sample (see Figure 2). Some characteristics only differed by use of one set of practices. Teachers that used an above average amount of reform practices were less likely to work in Title I schools (40% vs. 30%, $p <.05$) with significantly fewer FRPL-eligible students (32% vs. 38%, $p < .05$). Students in above-average reform practice use classrooms were significantly more likely to live in a home with three or more bedrooms (58% vs. 42%, $p < .05$). Finally, students in classrooms where teachers used more than the average amount of observed emotional support were significantly more likely to live with two parents or other adult caregivers (91% vs. 82%, $p < .05$). Parental education and student computer access differed by both observed reform practice and emotional support following a similar pattern: students in classrooms where teacher use of reform mathematics practices and emotional support were above average were significantly more likely to have college-educated parents (53-56% vs. 40-46% , ps $< .05$) and to have access to a computer at home (97-98% vs. 90-92%, ps $< .05$).

Discussion

Recommendations for high-quality, equitable mathematics instruction encourage teaching and learning that incorporates students' cultural, racial, and linguistic identities into the cultivation of a community of learners (Chao, Murray & Gutiérrez, 2014; NCTM, 2014). This study showed that data from the M-Scan and CLASS observational tools were psychometrically appropriate to use as a measure of equity-promoting mathematics instruction. However, multi-level regression models did not indicate a significant association between equity-promoting mathematics instruction and student outcomes. Comparisons between schools and classrooms where teachers were observed using more equity-promoting instruction highlighted a gap in the accessibility of emotionally supportive reform mathematics teaching for already marginalized communities in the district. These results make several contributions to the growing literature on the integration of SEL with academic instruction. The analytic approach suggests the value of combining classroom observations of mathematics and emotional support as an index of equity-promoting practice in upper elementary classrooms. Further, the results reveal inequities in which students experience the EP mathematics practices in their classroom in that those students from demographic groups that have the greatest need for such practices appear to have less access to classrooms using those approaches.

Validation of a Measure of Equity-Promoting Mathematics Practice

Frameworks for high-quality instruction share characteristics across content areas and developmental stages. The combined use of M-Scan and CLASS observational data in this study provides unique insight into the overlap between reform mathematics and SEL program practices: when modeled as a standardized distribution, classrooms tended to have higher scores for emotionally supportive interactions than use of reform mathematics practices. Additionally, observers indicated more limited range in the use of reform practices, with most teachers using them infrequently excepting several higher performers. Importantly, prior research has indicated that it is the regular use of innovative practices, not simply knowledge of them, that associate with student outcomes (Linley, 2008; Peterson & Seligman, 2004). Identifying areas of overlap across instructional frameworks can streamline how educators are trained and supported in accommodating the needs of their increasingly diverse classrooms.

The psychometric properties and model fit of the equity-promoting practice measure support the viability of this approach to reconfiguring existing data from systematic coding of classroom observations to highlight how a set of practices appear in combination. The measure constructed for this study represents an early step in the process of identifying the active ingredients of instruction the meets the equity goals described by NCTM and how they emerge in the real world. Educators increasingly have access to pre-service and professional learning that prioritizes equity and positive social development, but the extent to which these values are visible in their practice remains difficult to ascertain. Findings from this study indicate that future research or evaluation of teacher instruction could incorporate similar methods to better understand the complexity of student-centered, emotionally supportive reform mathematics instruction in practice.

Regression Results

Prior research has indicated that student characteristics like engagement and positive affect relate to improved student outcomes (Quinlan et al., 2019), but less is known about how teacher practices targeting these emotionally salient outcomes relate to academic achievement. Though correlational analyses found a small positive relation between equity-promoting mathematics instruction and social skills, the association did not extend to more complex models. Despite employing a sophisticated measurement approach that captured teacher instruction over the course of a school year, regression analyses did not find the expected association between equity-promoting mathematics practice and student achievement or social skills.

The associations between equity-promoting practice, outcomes, and RC program conditions highlight the complexity of creating optimal mathematics learning conditions for all students. Elementary school students tend to spend most of the day with one teacher, and the quality of the learning environment is heavily influenced by that teacher's manner of directing the classroom social climate (Opdenakker & Van Damme, 2006).

Accessibility of Equity-Promoting Practices in Schools and Classrooms

Analyses of school, student, and family characteristics shed light on our understanding of which students had access to teachers that were frequently observed using both reform mathematics instruction and emotionally supportive interactions in their practice. Specifically, schools serving lower income communities in the district were less likely employ teachers that were below average in observed use of reform practices and emotionally supportive interactions. In this sample of 20 schools, seven were eligible for Title 1 funds. Typically, those funds are used for math or reading specialists to increase the quality of instruction. Despite the access to additional resources, teachers at Title 1 schools were less likely to use reform mathematics instruction. These findings align with prior research indicating that the tendency towards teacher use of repetitive, minimally demanding mathematics learning activities and assignments is greater in schools that serve a high number of students living in poverty (Education Trust, 2015).

Limitations and Future Directions

Three limitations to this study warrant mention. First, the student race data provided by the district was relatively restricted and imprecise. The backgrounds of students identified as “multiracial” were not specified, and “Hispanic” students were racially identified as explicitly not “White” or “Black.” These data storage decisions obscure the true diversity of racial identities represented in the sample. In the context of promoting educational equity, this is problematic. For example, one recent study of the racialized experiences of Latinx children indicated that greater incidence of mental health challenges was associated with discrimination based on darker skin color (Calzada, Kim, & O’Gara, 2019). It is the responsibility of districts and researchers to design data collection systems that acknowledge the diversity of cultural and ethnic identities present within their communities to better serve all students and families. Given the limitations in these data, we were unable to pursue some of the more nuanced questions about race and ethnicity that could be a part of this work.

The study was also limited by the nature of the student outcome data. Mathematics achievement was measured using standardized assessments, which have been shown to be biased towards more affluent and white students (Caraballo, 2014).

Furthermore, social skills competency was informed by teacher report exclusively. Collecting additional indicators from multiple reporters would provide a more comprehensive picture of student mathematics understanding and social skills competency. Future research should continue to elucidate methods for measuring teacher practice and progress with meeting ambitious standards for mathematics instruction.

Conclusion

As the field of SEL increasingly pushes to integrate social development with content learning, development and validation of ways to measure complex instructional practices are increasingly necessary. The idea that emotionally supportive learning environments might relate to improved outcomes is not surprising or unexpected. However, mathematics instruction, especially in diverse contexts, does not often incorporate explicit supports for students' social and emotional development. The findings in this study provide observational evidence of the variation in teacher use of equity-promoting mathematics practices. The challenge for scholars and policymakers, then, is to provide educators with the tools, support, and resources necessary for mathematics learning experiences that inspire young mathematicians to continue finding and solving problems.

Table 1

Participant Characteristics

	Full Sample		Intervention		Comparison		*t*
	N	Mean *(SD)*	N	Mean *(SD)*	N	Mean *(SD)*	
Schools	20		12		8		
Condition		0.52 (0.5)		1 (0)		0 (0)	
School receives Title 1 funds	7	0.34 (0.47)	6	0.47 (0.5)	1	0.20 (0.4)	5.63***
School met AYP in 3 prior years	18	0.93 (0.26)	10	0.87 (0.34)	8	0.99 (0.11)	-4.49***
Percentage of students that receive FRPL		0.34 (0.24)		0.37 (0.69)		0.31 (0.24)	2.45*
Percentage of students receiving EL services		0.31 (0.18)		0.34 (0.18)		0.28 (0.18)	2.97**
Classrooms	61		31		28		
Teacher years of experience		12.39 (8.8)		11.86 (7.87)		12.94 (9.67)	-1.16
Teacher education (Master's degree = 1)		0.66 (0.47)		0.7 (0.46)		0.62 (0.49)	1.69*
Class size		20.34 (5.49)		19.43 (4.66)		21.27 (6.12)	-2.98**
Students	363		189		174		
Gender		0.53 (0.5)		0.53 (0.5)		0.53 (0.5)	-0.05
Prior math achievement		503.45 (67.43)		499.09 (71.3)		508.19 (62.83)	-1.29
Receiving EL services	91	0.25 (0.36)	50	0.22 (0.42)	41	0.15 (0.36)	0.63
WIDA Level 1 – 2 (Entering – Beginning)	17	0.19 (0.50)	11	0.36 (0.48)	6	0.07 (0.26)	0.25
WIDA Level 3 – 4 (Developing – Expanding)	37	0.41 (0.33)	18	0.42 (0.37)	19	0.07 (0.26)	1.26
WIDA Level not specified	37	0.40 (0.30)	21	0.11 (0.32)	16	0.09 (0.29)	0.60
Formerly received EL services	65	0.18 (0.38)	34	0.18 (0.39)	31	0.18 (0.38)	-0.04
Race/Ethnicity							
White	139	0.38 (0.49)	65	0.34 (0.48)	74	0.43 (0.5)	-1.59
Asian American	75	0.21 (0.41)	38	0.2 (0.4)	37	0.21 (0.41)	-0.27
Hispanic American	59	0.16 (0.37)	34	0.18 (0.39)	25	0.14 (0.35)	0.93
Black	49	0.13 (0.34)	25	0.13 (0.34)	24	0.14 (0.35)	-0.16
Multiracial	32	0.09 (0.28)	22	0.12 (0.32)	10	0.06 (0.23)	1.98*
Native American or Hawaiian	3	0.01 (0.09)	2	0.01 (0.1)	1	0.01 (0.08)	-0.51
Protected groups							
Special education	71	0.2 (0.4)	41	0.22 (0.41)	30	0.17 (0.38)	1.07
Low income (FRPL)	101	0.32 (0.47)	52	0.33 (0.47)	49	0.32 (0.47)	0.17
Family & Home Information							
Mother did not graduate high school	21	0.06 (0.23)	10	0.05 (0.22)	11	0.06 (0.24)	-0.42
Father did not graduate high school	18	0.05 (0.22)	10	0.05 (0.22)	8	0.05 (0.21)	-0.30
Mother has an Associate or Bachelor's degree	184	0.51 (0.5)	93	0.49 (0.5)	91	0.52 (0.5)	-0.59
Father has an Associate or Bachelor's degree	173	0.48 (0.5)	84	0.44 (0.5)	89	0.51 (0.5)	-1.28
Expect student will complete a post-secondary degree	310	0.85 (0.35)	154	0.81 (0.39)	156	0.9 (0.31)	-2.21*
Access to a computer at home	299	0.94 (0.23)	149	0.94 (0.24)	150	0.95 (0.22)	-0.47
Home has 3+ bedrooms	184	0.51 (0.5)	91	0.48 (0.5)	93	0.53 (0.5)	-1.01

Table 2

Correlations and Descriptive Statistics of Model Predictors and Outcomes

	1	2	3	4	5	6	7	8	9	10	11	12	13
1. Condition (RC = 1)	--												
2. Emotional support score	0.07	--											
3. Reform practices score	0.25*	0.47*	--										
4. Family SES[a]	-0.12*	0.13*	0.22*	--									
5. School met AYP	-0.23*	0.11*	0.22*	0.23*	--								
6. School-level FRPL%	0.13*	-0.08	-0.23*	-0.72*	0.25*	--							
7. School-level EL%	0.15*	-0.07	-0.10	-0.69*	0.07	0.92*	--						
8. Teacher years of experience	-0.06	-0.10	-0.03	0.06	0.13*	-0.11*	-0.04	--					
9. Teacher education (Master's degree = 1)	0.09	-0.08	0.12*	-0.07	0.02	0.09	0.15*	-0.03	--				
10. Prior year math achievement	-0.07	0.19*	0.08	0.35*	0.12*	-0.30*	-0.30*	-0.02	-0.08	--			
11. Current student EL status	0.03	-0.09	-0.05	-0.35*	0.05	0.22*	0.26*	-0.06	0.05	-0.30*	--		
12. Outcome year math achievement	0.04	0.15*	0.04	0.14*	0.13*	-0.15*	-0.13*	-0.13*	-0.02	0.65*	-0.15*	--	
13. Teacher-reported social skills	-0.06	0.13*	0.07	0.13*	0.16*	-0.06	-0.04	-0.02	-0.13*	0.32*	-0.10	0.31*	--
Mean *(SD)*	0.49 *(0.50)*	3.98 *(0.45)*	4.00 *(0.65)*	0.22 *(0.26)*	0.08 *(0.28)*	0.35 *(0.24)*	0.31 *(0.18)*	11.76 *(8.41)*	0.66 *(0.48)*	501.49 *(64.31)*	0.25 *(0.43)*	542.20 *(56.50)*	3.52 *(0.45)*
Min-Max	0 - 1	1 -6	1 -6	0 - 1	0 - 1	0 -1	0 -1	1 - 38	0 - 1	200 – 600	0 - 1	200 - 600	1 - 4

Note. * = $p<.05$; ** = $p<.01$

[a] Family SES reflects the combined means of: school Title I status; student FRPL eligibility; parental education; whether parents expect student to obtain a post-secondary degree; whether the home has three or more bedrooms; and whether the student has access to a computer at home.

Table 3

Regression Results Estimating Effect of Equity-Promoting Mathematics Instruction on Social Skills and Mathematics Achievement

	Social Skills (SSIS)		Math Achievement (SOL)	
	b	SE	*b*	SE
Level 1 (Student)				
Prior mathematics achievement	0.00***	0.00	0.55***	0.04
Gender	0.24***	0.05	-9.12	4.23
Level 2 (Classroom)				
Reform mathematics practices	0.07	0.04	6.73	5.48
Emotionally supportive interactions	-0.01	0.07	-3.18	7.82
Teacher years of experience	-0.00	0.01	-1.11**	0.37
Class size	0.01	0.01	-0.30	0.48
Level 3 (School)				
Condition	-0.05	0.08	7.98	7.96
Constant	2.45***	0.26	0.37	0.61
Level 1 Residual	0.13	0.03	1512.02	223.69
Level 2 Error	0.04	0.01	194.49	153.26
Level 3 Error	0.00	0.01	82.83	124.80
School ICC	0.02	0.03	0.05	0.07
Teacher | School ICC	0.25	0.07	0.15	0.07

Note. * = p<.05; ** = p<.01

Figure 1

Fitted SEM Model for Measure of Equity-Promoting Mathematics Instruction

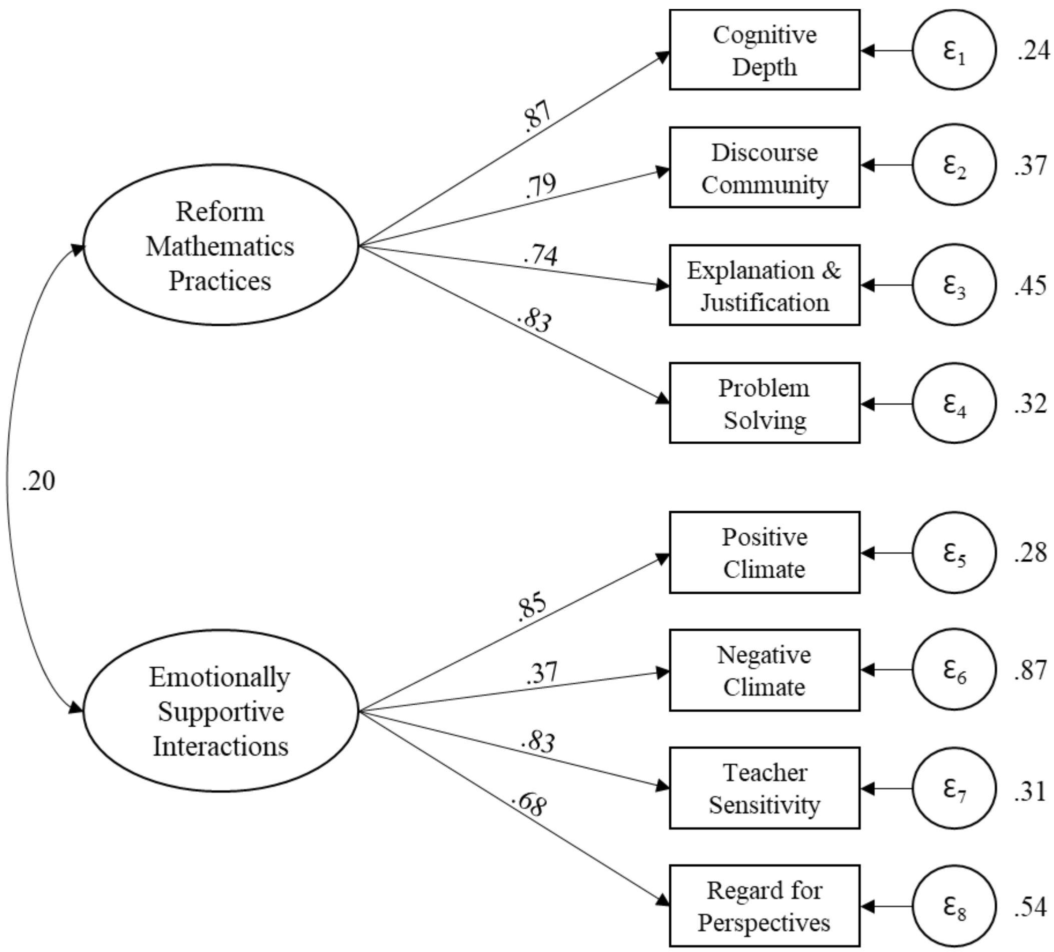
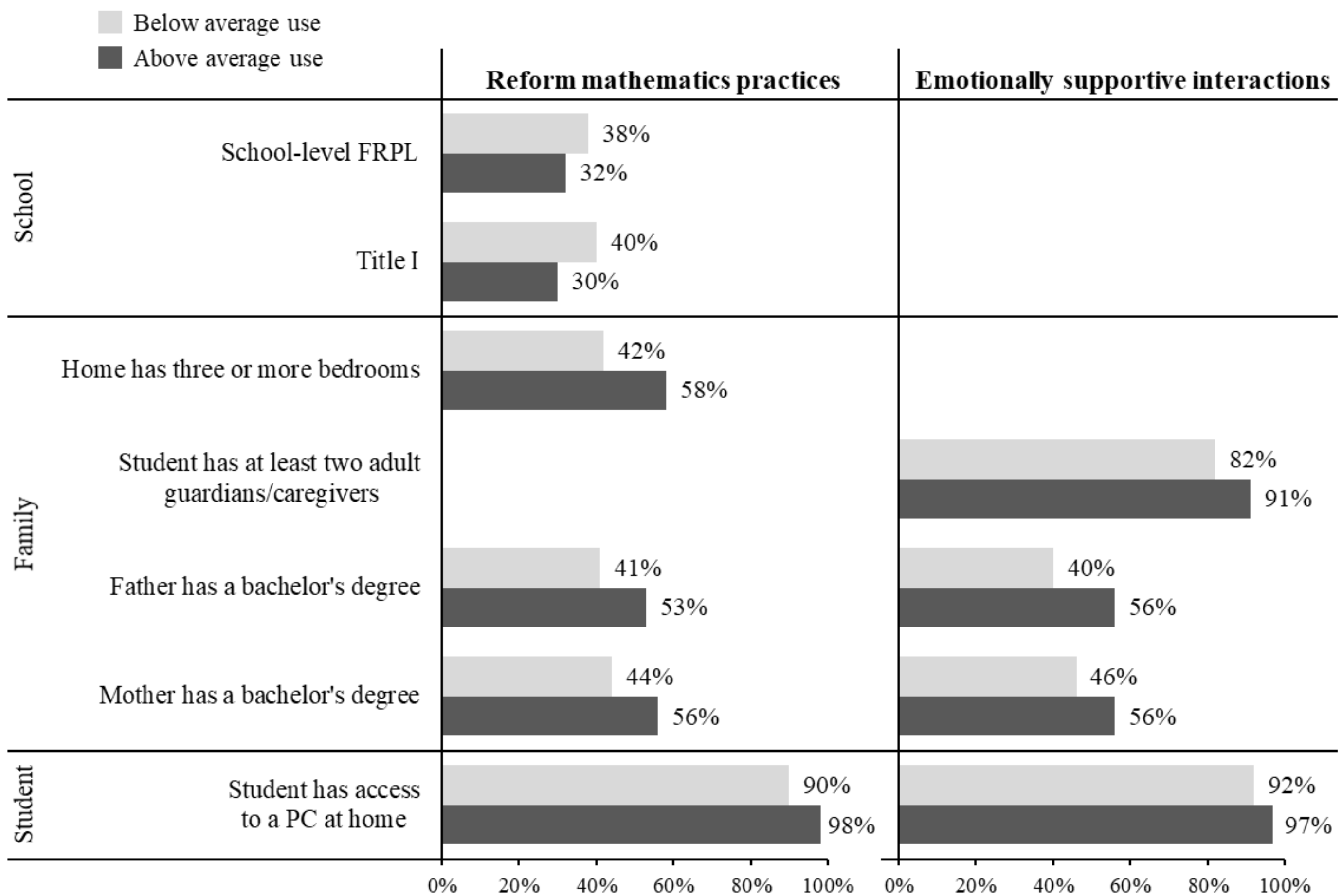


Figure 2

Comparing School, Family, and Student Characteristics in Classrooms Above and Below the Average Observed Use of Practices



References

Abry, T., Rimm-Kaufman, S. E., Larsen, R. A., & Brewer, A. J. (2013). The influence of fidelity of implementation on teacher-student interaction quality in the context of a randomized controlled trial of the Responsive Classroom approach. *Journal of School Psychology*, *51*(4), 437–453. https://doi.org/10.1016/j.jsp.2013.03.001

Banse, H. W., Palacios, N. A., Curby, T. W., & Rimm-Kaufman, S. E. (2018). How should fifth-grade mathematics teachers start the school year? Relations between teacher-student interactions and mathematics instruction over one year. *Teachers College Record*, *120*(6).

Calzada, E. J., Kim, Y., & O'Gara, J. L. (2019). Skin color as a predictor of mental health in young Latinx children. *Social Science and Medicine*, *238*(August 2018), 112467. https://doi.org/10.1016/j.socscimed.2019.112467

Caraballo, L. (2014). The student as an assemblage of success: Constructing multiple identities amidst classes, gendered, and raced discourses of achievement. *Curriculum and Teaching Dialogue*, *16*(1), 103–123.

Common Core State Standards Initiative. (2014). *Common Core State Standards Initiative: Preparing America's students for college and career.*

Dominguez, H. (2011). Using what matters to students in bilingual mathematics problems. *Educational Studies in Mathematics*, 76, 305–328.

Hamre, B. K., Pianta, R. C., Mashburn, A., & Downer, J. (2007). *Building and validating a theoretical model of classroom effects in over 4,000 early childhood and elementary classrooms*. Retrieved from the Foundation for Child Development website: http://www.fcd-us.org/resources/building-science-classrooms- application-class-framework-over-4000-us-early-childhood-and-e

Henderson Pinter, H., Merritt, E. G., Berry, R. Q., & Rimm-Kaufman, S. E. (2018). The importance of structure, clarity, representation, and language in elementary mathematics instruction. *Investigations in Mathematics Learning*, *10*(2), 106–127. https://doi.org/10.1080/19477503.2017.1375354

Hughes, J. N., Cavell, T. A., & Willson, V. (2001). Further support for the developmental significance of the quality of the teacher–student relationship. *Journal of School Psychology*, 39(*4*), 289-301. doi: 10.1016/S0022-4405(01)00074-7

Imm, K., & Stylianou, D. A. (2012). Talking mathematically: An analysis of discourse communities. *The Journal of Mathematical Behavior*, 31(*1*), 130–148. doi:10.1016/j.jmathb.2011.10.001

Jacob, R., Hill, H., & Corey, D. (2017). The Impact of a Professional Development Program on Teachers' Mathematical Knowledge for Teaching, Instruction, and Student Achievement. *Journal of Research on Educational Effectiveness*, *10*(2), 379–407. https://doi.org/10.1080/19345747.2016.1273411

Jones, S. M., Barnes, S. P., Bailey, R., & Doolittle, E. J. (2017). Promoting social and emotional competencies in elementary school. *Future of Children*, *27*(1), 49–72. https://doi.org/10.1353/foc.2017.0003

Larkin, D.B. (2020). *Teaching Science in Diverse Classrooms: Real Science for Real Students* (1st ed.). Routledge. https://doi.org/10.4324/9780429199721

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

Marks, H.M. (2000) Student engagement in instructional activity: Patterns in the elementary, middle, and high school years. *American Educational Research Journal* 37(*1*), 153-184.

Martin, D. P., & Rimm-Kaufman, S. E. (2015). Do student self-efficacy and teacher-student interaction quality contribute to emotional and social engagement in fifth grade math? *Journal of School Psychology*, *53*(5), 359–373. https://doi.org/10.1016/j.jsp.2015.07.001

McCombs, B. L., Daniels, D. H., & Perry, K. E. (2008). Children's and teachers' perceptions of learner-centered practices, and student motivation: Implications for early schooling. *The Elementary School Journal*, 109(*1*), 16–35. doi:10.1086/592365

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

National Council of Teachers of Mathematics. *Principles to actions: ensuring mathematical success for all*. Reston, VA.

McCormick, M. P., O'Connor, E. E., Cappella, E., & McClowry, S. G. (2013). Teacher–child relationships and academic achievement: A multilevel propensity score model approach. *Journal of School Psychology*, 51(*5*), 611–624.

doi:10.1016/j.jsp.2013.05.001

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

Ottmar, E. (2019). The effects of deeper learning opportunities on student achievement: Examining differential pathways. *Psychology in the Schools*, *56*(5), 840–855. https://doi.org/10.1002/pits.22237

Ottmar, E. R., Rimm-Kaufman, S. E., Berry, R. Q., & Larsen, R. A. (2013). Does the responsive classroom approach affect the use of standards-based mathematics teaching practices?: Results from a randomized controlled trial. *Elementary School Journal*, *113*(3), 434–457. https://doi.org/10.1086/668768

Quinlan, D., Vella-Brodrick, D. A., Gray, A., & Swain, N. (2019). Teachers matter: Student outcomes following a strengths intervention are mediated by teacher strengths spotting. *Journal of Happiness Studies*, *20*(8), 2507–2523. https://doi.org/10.1007/s10902-018-0051-7

Roorda, D. L., Jak, S., Zee, M., Oort, F. J., & Koomen, H. M. Y. (2017). Affective teacher-student relationships and students’ engagement and achievement: A meta-analytic update and test of the mediating role of engagement. *School Psychology Review*, *46*(3), 239–261. https://doi.org/10.17105/SPR-2017-0035.V46-3

Schoenfeld, A. H. (2020). Mathematical practices, in theory and practice. *ZDM Mathematics Education*, *52*(6), 1163–1175. https://doi.org/10.1007/s11858-020-01162-w

Stein, M. K., Remillard, J., and Smith, M. S. (2007). How curriculum influences student learning. In F. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (pp. 319–369). Greenwich, CT: Information Age Publishing.

Walkowiak, T. A., Berry, R. Q., Meyer, J. P., Rimm-Kaufman, S. E., & Ottmar, E. R. (2014). Introducing an observational measure of standards-based mathematics teaching practices: Evidence of validity and score reliability. *Educational Studies in Mathematics*, *85*(1), 109–128. https://doi.org/10.1007/s10649-013-9499-x

APPENDIX: MANUSCRIPT 3 MEASURES

Social Social Skills Improvement System (SSIS)

Please read the following items and think about this student's behavior during the past month or two. Decide how often this student does the behavior described.

Statement	**Never**	**Rarely**	**Sometimes**	**Very often**
1. Follows your directions.	1	2	3	4
2. Says "please."	1	2	3	4
3. Completes tasks without bothering others.	1	2	3	4
4. Makes friends easily.	1	2	3	4
5. Responds well when others start a conversation or activity.	1	2	3	4
6. Participates appropriately in class.	1	2	3	4
7. Speaks in appropriate tone of voice.	1	2	3	4
8. Pays attention to your instructions.	1	2	3	4
9. Interacts well with other children.	1	2	3	4
10. Takes turns in conversations.	1	2	3	4
11. Joins activities that have already started.	1	2	3	4

Statement	Never	Rarely	Sometimes	Very often
12. Says “thank you.”	1	2	3	4
13. Ignores classmates when they are distracting.	1	2	3	4
14. Invites others to join in activities.	1	2	3	4
15. Makes eye contact when talking.	1	2	3	4
16. Participates in games or group activities.	1	2	3	4
17. Follows classroom rules.	1	2	3	4
18. Starts conversations with peers.	1	2	3	4
19. Uses gestures or body appropriately with others.	1	2	3	4
20. Introduces himself/herself to others.	1	2	3	4

21. I have had the above named student in my math class for (please circle one):

(a) almost all of the school year up until this point
(b) half of the school year
(c) a few months
(d) a few weeks
(e) a few days

Mathematics-Scan (M-Scan)

Structure of the Lesson

1. Structure of the Lesson: The extent to which the design of the lesson is organized to be conceptually coherent such that activities are related mathematically and build on one another in a logical manner. NOTE 1: Ratings of observations should take into account interruptions for procedural activities that are not part of the instructional unit, when these interruptions consume a non-trivial amount of time. NOTE 2: If a warm-up (i.e., brief, possibly unrelated segment at beginning of class) does not interfere with overall flow of lesson, it should not count against the overall score for this measure.

	Low (1, 2)	**Medium (3, 4, 5)**	**High (6, 7)**
Logical Sequence	Overall, the components of the math lesson do not appear to be logically organized.	Some components of the math lesson are logically organized, but others do not seem to fit.	All components of the math lesson are logically organized.
Mathematical Coherence	The components of the lesson are not mathematically connected.	Some components of the lesson are mathematically connected.	All components of the lesson are mathematically connected and coherent.
Promotion of Deeper Understanding	The lesson does not lead students to a deeper understanding of the mathematical concepts.	The lesson leads students toward partial depth of understanding mathematical concepts or the lesson leads students toward depth some of the time.	The lesson leads students to a deeper understanding of the concept.

Use of Representations

2. Use of Representations: The extent to which the lesson promotes the use of and translation among multiple representations (pictures, graphs, symbols, words) to illustrate ideas and concepts. The use of and translation among representations should allow students to make sense of mathematical ideas or extend what they already understand. NOTE: Dimension includes both exposure (by teacher or curriculum) and use by students. As outlined in NCTM's *Principles and Standards for School Mathematics (2000),* "students in grades 3-5 should continue to develop the habit of representing problems and ideas to support and extend their reasoning. Such representations help to portray, clarify, or extend a mathematical idea" (p. 206).

	Low (1, 2)	**Medium (3, 4, 5)**	**High (6, 7)**
Presence of Representations	Teacher and/or students rarely use more than one representation of a mathematical concept.	Teacher and/or students sometimes use more than one representation of a mathematical concept.	Teacher and/or students often use more than one representation for a mathematical concept.
Teacher Translation among Representations	For the representation(s) used, the teacher does not make connections to concepts or between representations. (i.e., procedural approach to use of representations).	For the representation(s) used, the teacher makes some connections to concepts and between representations.	For the representation(s) used, the teacher often makes connections to concepts and between representations.
Student Translation among Representations	Students do not translate between representations.	Students sometimes translate back and forth between representations. They *do not* explain their representations.	Students translate back and forth between representations. They also explain their representations at times.

Students' Use of Mathematical Tools

3. Use of Mathematical Tools: The extent to which the lesson affords students the opportunity to use appropriate mathematical tools (e.g., calculators, pattern blocks, fraction strips, counters, virtual tools) and that these tools enable them to represent abstract mathematical ideas. NOTE: When students use equipment and/or objects to collect data that are later used in exploring mathematical ideas, the equipment/objects are not considered to be mathematical tools unless they are also explicitly used to develop the mathematical ideas.

	Low (1, 2)	Medium (3, 4, 5)	High (6, 7)
Opportunity to Use Tools	Students do not use tools and/or are only permitted to use tools for help with procedural skills.	Students sometimes use tools to investigate concepts and solve problems.	Students often use tools to investigate concepts and solve problems.
Depth of Use	Connections are never made between tools and mathematical concepts.	Connections are sometimes made between tools and mathematical concepts.	Connections are often made between tools and mathematical concepts.

Cognitive Demand

4. Cognitive Demand: Cognitive demand refers to command of the central concepts or "big ideas" of the discipline, generalization from specific instances to larger concepts, and connections and relationships among mathematics concepts. This dimension considers two aspects of cognitive demand: task selection and teacher enactment. That is, it considers the extent to which the selected task is cognitively demanding and the extent to which the teacher consistently and effectively promotes cognitive depth (Stein & Lane, 1996).

	Low (1, 2)	Medium (3, 4, 5)	High (6, 7)
Task Selection	The tasks of the lesson are focused on memorization or procedures without connections to underlying concepts.	Some of the tasks are focused on memorization or procedures without connections to underlying concepts, *and* some of the tasks are focused on procedures with connections to underlying concepts or non-algorithmic, complex thinking.	The majority of the tasks of the lesson are focused on procedures with connections to underlying concepts or non-algorithmic, complex thinking.
	None of the tasks are open-ended.	Some of the tasks are open-ended.	Most of the tasks are open-ended.
Teacher Enactment	The teacher rarely provides feedback, modeling, or examples that promote complex thinking by students.	The teacher sometimes provides feedback, modeling, or examples that promote complex thinking by students.	The teacher often provides feedback, modeling, or examples that promote complex thinking by students.
	The teacher rarely encourages students to make conceptual connections.	The teacher sometimes encourages students to make conceptual connections.	The teacher often encourages students to make conceptual connections.

Explanation and Justification

6. Explanation and Justification: The extent to which the teacher expects and students provide explanations/justifications, both orally and on written assignments. NOTE: Simply "showing your work" on written assignments – i.e., writing the steps involved in calculating an answer – does not constitute an explanation.

	Low (1, 2)	Medium (3, 4, 5)	High (6, 7)
Presence of Explanation and Justification	Students rarely provide explanations or justify their reasoning. Teachers rarely ask "what, how, why" questions or otherwise solicit student explanations/justifications.	Students sometimes provide explanations and/or justify their reasoning. Teachers sometimes ask "what, how, why" questions or otherwise solicit student explanations/justifications.	Students often provide explanations and/or justify their reasoning. Teachers often ask "what, how, why" questions or otherwise solicit student explanations/justifications.
Depth of Explanation and Justification (procedural and conceptual)	Student explanations often focus on procedural steps *and rarely* include conceptual understanding of the topic(s).	Student explanations sometimes focus on procedural steps *and sometimes* include conceptual understanding of the topic(s).	Student explanations rarely focus on procedural steps *and often* focus on conceptual understanding of the topic(s).

Problem Solving

7. Problem Solving: The extent to which instructional activities enable students to identify, apply and adapt a variety of strategies to solve problems. The extent to which the problems that students solve are complex and allow for multiple solutions. NOTE: This dimension focuses more on the nature of the activity/task, rather than the enactment. To receive a "High" rating, problems should not be routine or algorithmic; they should consistently require novel, challenging, and/or creative thinking.

[1] Student formulation of problems can improve the score for this domain, but scores should not decrease if this is not present.

[2] Student formulation of problems may involve students extending/following up on problems not originally formulated by students

	Low (1, 2)	Medium (3, 4, 5)	High (6, 7)
Students' Engagement with Problems	Students rarely engage in problems that allow them to grapple with mathematical concepts. Students often work on exercises for which they are practicing an already learned procedure.	Students sometimes engage in problems that allow them to grapple with mathematical concepts. Students sometimes work on exercises for which they are practicing an already learned procedure.	Students often engage in problems that allow them to grapple with mathematical concepts. Students rarely work on exercises for which they are practicing an already learned procedure.
Presence of Problem Solving with Multiple Strategies	Classroom activities encourage only one strategy to solve each problem.	Classroom activities sometimes encourage multiple strategies to solve each problem.	Classroom activities often encourage multiple strategies to solve each problem.
Student Formulation of Problems *(when applicable)*[1, 2]	If students formulate problems, they are generally procedural.	If students formulate problems, they are sometimes solved with multiple strategies.	If students formulate problems, they are generally solved with multiple strategies.

Connections/Applications

8. Connections/Applications: The extent to which the lesson helps students connect mathematics to other mathematical concepts, their own experience, to the world around them, and to other disciplines. The extent to which the lesson helps students apply mathematics to real world contexts and to problems in other disciplines. NOTE: The experiences may be teacher-generated or student-generated, but they should relate to the students' actual life situations.

	Low (1, 2)	Medium (3, 4, 5)	High (6, 7)
Connections	Meaningful connections between mathematics learned in the classroom and other math concepts, experiences, disciplines, and the world are rarely made.	Meaningful connections between mathematics learned in the classroom and other math concepts, experiences, disciplines, and the world are sometimes made.	Meaningful connections between mathematics learned in the classroom and other math concepts, experiences, disciplines, and the world are often made.
Applications	Students are never asked to apply the math they learn to the world around them.	Students are sometimes asked to apply the math they learn to the world around them.	Students are often asked to apply the math they learn to the world around them.
	The class work is not relevant to students' lives.	The class work is potentially relevant to the students' lives.	The class work is relevant to the students' lives.

Mathematical Accuracy

9. Mathematical Accuracy The extent to which the mathematical concepts are presented clearly and accurately throughout the lesson. The extent to which student misconceptions are present, and whether teachers handle student misconceptions in a way that clarifies conceptual understanding.

	Low (1, 2)	Medium (3, 4, 5)	High (6, 7)
Accuracy in Teacher Presentation	A few of the concepts and procedures presented to the students by the teacher are mathematically accurate. Most of the concepts and procedures are mathematically inaccurate.	Most of the concepts and procedures presented to students by the teacher are mathematically accurate, but on a few occasions, the concepts and procedures are mathematically inaccurate.	The concepts and procedures presented to the students by the teacher are mathematically accurate.
Clarity of Mathematical Concepts	The mathematical concepts are not articulated clearly by the teacher. There is ambiguity in presentation of key mathematical concepts.	The mathematical concepts may be articulated with some clarity by the teacher. There is some ambiguity in presentation of key mathematical concepts.	The tasks in the lesson allow students to transfer the mathematics to future mathematical experiences.
Responsiveness to Student Mathematical Thinking	Student misconceptions are obvious in the lesson. The misconceptions are not noticed or addressed appropriately by the teacher. By the end of the lesson, students appear to retain misconceptions.	One or more student misconceptions are observed during the lesson. The misconceptions may have been addressed by the teacher, and student understanding was improved due to teacher responses.	Student misconceptions may or may not have been observed during the lesson. The teacher clarified all misconceptions that students had during the lesson.

Classroom Assessment System (CLASS)

EMOTIONAL SUPPORT

Positive Climate

Positive Climate reflects the overall emotional tone of the classroom and the connection between teachers and students. The warmth of the teacher's interactions with students and the teacher's display of enjoyment and respect of students during instruction as well as social conversations are included in this rating. Interactions among peers should be considered in this rating.

Negative Climate

Negative Climate reflects the overall level of expressed negativity in the classroom. Teacher negativity (e.g., anger, sarcasm, irritability) as well as peer negativity (arguing, aggression, victimization, bullying) should be considered in this rating. The quality, severity, and intensity of expressed negativity are important.

Teacher Sensitivity

Teacher Sensitivity encompasses the teacher's responsivity to students' needs and awareness of students' level of academic and emotional functioning. The extent to which the teacher is available as a secure base (allowing students to actively explore and learn and being there to provide comfort, reassurance, and encouragement) should be included in this rating.

Regard for Student Perspectives

Regard for Student Perspectives captures the degree to which the teacher's interactions with students and classroom activities place an emphasis on students' interests, motivations, and points of view. The teacher's flexibility within activities and ability to demonstrate respect for students' autonomy to participate in and initiate activities should be considered under this rating.

Positive Climate

Reflects the overall emotional tone of the classroom and the connection between teachers and students. The warmth of the teacher's interactions with students and the teacher's display of enjoyment and respect of students during instruction as well as social conversations are included in this rating. Interactions among peers should be considered in this rating.

	Low (1,2)	Mid (3,4,5)	High (6,7)
Relationships	There are few, if any, indications that the teacher enjoys warm, supportive relationships with students.	There are some indications that the teacher enjoys warm, supportive relationships with students.	There are many indications that the teacher enjoys warm, supportive relationships with students.
Positive Affect	There are no or few displays of joint laughter or smiles, genuine praise, or physical/verbal affection among teachers and students.	There are times of joint laughter and smiling, genuine praise, or physical affection, but other times when these are absent.	There is frequent joint smiling and laughter, genuine praise, and/or physical affection among the teacher and students.
Respect	Teacher rarely, if ever, demonstrates respect for the students.	Some of the teacher's interactions with the students connote respect, while others do not.	Teacher consistently demonstrates respect for the students.
Positive peer interactions	Students rarely, if ever, engage in positive interactions with one another.	Although there is not clear evidence of a strong emotional connection among students, there is an underlying positive tone to the interactions.	Students are clearly positively connected to one another.

Negative Climate

Reflects the overall level of expressed negativity in the classroom. Teacher negativity (e.g., anger, sarcasm, irritability) as well as peer negativity (arguing, aggression, victimization, bullying) should be considered in this rating. The quality, severity, and intensity of expressed negativity are important.[3]

	Low (1,2)	Mid (3,4,5)	High (6,7)
Negative Affect	**Teacher and students do not display strong negative affect and only rarely, if ever, displays more mild negativity**	**Classroom is characterized by moderate displays of irritability, anger, or other negative affect.** by teacher and/or peers	Classroom is characterized by frequent and intense irritability, anger, or other negative affect by teacher and/or peers
Punitive Control	Teacher does not yell or make threats to establish control	Teacher occasionally uses expressed negativity such as threats or yelling to establish control	Teacher frequently yells at students or makes threats to establish control
Sarcasm/ Disrespect	Teacher is not sarcastic or disrespectful	Teacher is sometimes sarcastic or disrespectful	Teacher is frequently sarcastic or disrespectful
Negativity not Connected Events	Any mild negativity is clearly tied to a specific situation, is quickly alleviated, and does not rise above this very mild level	Mild negativity may appear disconnected from any specific event in the classroom	Negativity is often disconnected from any specific event
Negativity Escalates	If mild negativity is observed, it does not escalate to higher levels	Mild negativity sometimes escalates into more moderate intensity negativity	Mild negativity often escalates into more moderate intensity negativity
Severe Negativity	There are no instances of severe negativity between teachers and students	There are no instances of severe negativity between teachers and students	There are instances of severe negativity between the teacher and students or among students

Teacher Sensitivity

Encompasses the teacher's responsivity to students' needs and awareness of students' level of academic and emotional functioning. The extent to which the teacher is available as a secure base (allowing students to actively explore and learn and being there to provide comfort, reassurance, and encouragement) should be included in this rating.

	Low (1,2)	Mid (3,4,5)	High (6,7)
Responsive	Teacher is unresponsive or dismissive to students	Teacher is sometimes responsive to students, but at other times is more dismissive or unresponsive	Teacher is consistently responsive to students
Notices when Students Need Assistance	Teacher consistently fails to notice when students need extra support or assistance	Teacher sometimes notices when students need extra support or assistance	Teacher consistently notices when students need extra support or assistance
Appropriate Activities	Teacher consistently provides activities or speaks at levels inconsistent with student needs and abilities	Teacher sometimes provides activities or speaks at levels inconsistent with student needs and abilities	Teacher provides activities or speaks at levels consistent with student needs and abilities
Addresses Problems	Teacher is ineffective at helping students	Teacher generally tries to respond to students who approach with problems or questions, but is not consistently effective at addressing these problems	Teacher is consistently effective in addressing students' questions, concerns, and problems
Students Seek Support	Students do not seek out the teacher's support or guidance	Students sometimes seek out the teacher's support or guidance	Students appear comfortable approaching the teacher for support or guidance
Student Comfort	When the teacher asks questions, few students respond or raise their hands	When teachers ask questions, some students respond or raise their hands	Students appear comfortable sharing their ideas with the teacher and respond freely to questions

Regard for Student Perspectives

Captures the degree to which the teacher's interactions with students and classroom activities place an emphasis on students' interests, motivations, and points of view. The teacher's flexibility within activities and ability to demonstrate respect for students' autonomy to participate in and initiate activities should be considered under this rating.

	Low (1,2)	Mid (3,4,5)	High (6,7)
Flexibility and Student Focus	Teacher is rigid, inflexible, and controlling in her plans and/or rarely "goes with the flow" of students' ideas and the classroom activities are very teacher-driven	Teacher may follow students' lead during some periods and be more controlling during others	Teacher is flexible in her plans, "goes with the flow" of students' ideas, and organizes instruction around students' interests
Support of Autonomy	Teacher does not support student autonomy within the context of either structured or unstructured lessons and activities	Teacher sometimes provides support for student autonomy but at other times fails to do so	Teacher makes an effort to maximize students' abilities to be autonomous within the context of both structured and unstructured lessons and activities
Student Expression	There are few opportunities for student talk and expression	There are periods in which there is a lot of student talk and expression, but other times when teacher talk predominates	There are many opportunities for student talk and expression
Student Responsibility	Students do not have clear and real responsibilities/roles within the classroom	Students have jobs/roles within the classroom, but these provide only moderate or occasional opportunities for true responsibility	Students have clear and real responsibilities/roles within the classroom
Peer Interaction Encouraged	Teacher rarely if ever encourages students to interact with one another	Teacher allows, but does not actively encourage students to interact with one another	Teacher actively encourages students to interact with one another
Restriction of movement	Teacher is highly controlling of students' movement and placement during activities	The teacher is somewhat controlling of students' movement and placement during activities	Students have freedom of movement and placement during activities