Understanding and Preparing for Ambitious Elementary Mathematics Instruction

A Dissertation

Presented to

The Faculty of the Curry School of Education

University of Virginia

In Partial Fulfillment

of the Requirements for the Degree

Doctor of Philosophy

by Rebekah Berlin

August, 2019

Rebekah Berlin
Department of Curriculum, Instruction, and Special Education
Curry School of Education
University of Virginia
Charlottesville, Virginia

APPROVAL OF THE DISSERTATION

The dissertation ("Understanding and Preparing for Ambitious Elementary Mathematics Instruction") has been approved by the Graduate Faculty of the Curry School of Education in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

____
Name of Chair (Dr. Julie Cohen)

____
Committee Member (Dr. Peter Youngs)

____
Committee Member (Dr. Robert Berry)

____
Committee Member (Dr. Bridget Hamre)

____
Committee Member (Dr. Vivian Wong)

Date: ______________________________

Acknowledgements

I could not have done this work without the support of many people. I would especially like to thank the members of my committee: Robert, Peter, Bridget, and Vivian. These manuscripts are significantly stronger thanks to your generosity of time, resources, and feedback. Julie, your support and mentorship were foundational to these manuscripts, but more importantly, my broader development as a scholar.

TABLE OF CONTENTS

Elements

Linking Document: Understanding and Preparing for Ambitious Elementary Mathematics Instruction

Training and supporting teachers to teach mathematics ambitiously remains a challenge for the field. Ambitious mathematics instruction promotes rich conceptual understanding, procedural fluency, and flexible problem solving (Lampert, Beasley, Ghousseini, Kazemi, & Franke, 2010; Ball, Hill, & Bass, 2005). Yet, many students remain in mathematics classrooms that prioritize rote learning (Lithner, 2015; Hassan & Quershi, 2018), limiting students' ability to transfer mathematical understandings into novel contexts (Mayer, 2002). The challenge facing the field is twofold: (1) Preparing mathematics teachers is complex work that is not well-understood or measured in systematic ways (2) Researchers and practitioners alike need measures that get at the multifaceted nature of math instruction, otherwise it is impossible to track growth over time and provide teachers with ongoing support. Yet, assessing instructional quality is complex and there is not yet consensus on how to define instructional quality in mathematics (Charalambous & Praetorius, 2018).

There are several measurement challenges that limit current understanding of teacher preparation. The scope of teacher preparation makes it difficult to measure and disentangle the effects of particular components. Mathematics teacher preparation includes experiences in methods courses and in field placements. There are multiple faculty members who support and instruct teacher candidates, including methods instructors, university supervisors, and cooperating teachers. Because of the volume of sites and actors that influence teacher candidate learning, the level of support teacher candidates receive may vary widely from candidate to candidate, even within a single program. In addition, candidates' methods courses and fieldwork often provide candidates with conflicting viewpoints about how they should teach mathematics meaning that in some cases, the effects of one part of teacher preparation may be muted or washed out by another part of teacher preparation (Clift & Brady, 2005; Zeichner, 2010).

Mathematics learning in K-12 classrooms also occurs in a complex ecosystem. Characteristics of schools, teachers, and teaching all interact to support, to varying degrees, students' mathematics learning (e.g., Hill, Blazar, Lynch, 2015; Hill, Rowan, & Ball, 2005; Jung, Brown, & Karp, 2014) For example, a highly-resourced school may provide teachers with quality mathematics curricular supports or mathematics professional development, while another school may not provide teachers with any curriculum or professional development (Balfanz & Byrnes, 2006). In addition, teachers have varying mathematical beliefs, content knowledge, pedagogical content knowledge, and proficiency with instructional practices, which can each impact their mathematics instruction in substantive ways (Hill, Rowan, & Ball, 2005; Ma, 1999; Beilock, Gunderson, Ramirez, & Levine, 2010; Wilkins, 2008).

Current approaches to conceptualizing and measuring teaching and learning in mathematics teacher preparation and K-12 mathematics classrooms foreground some of these complexities and obscure others. For example, many researchers studying mathematics teacher preparation conceptualize teacher candidate learning as exposure to particular topics (e.g., Blömeke & Kaiser, 2012; Wang & Tang, 2013). These researchers use course syllabi and self-reports from teacher candidates to gauge candidates' exposure to a variety of mathematical and pedagogical topics. They leverage these data to examine associations between content exposure and outcomes of interest, including teacher candidate knowledge. This approach foregrounds the content to which teacher candidates are exposed throughout their preparation but obscures the interactive contextualized nature of candidate learning. Some of the aspects of preparation that are obscured include: 1) the pedagogies used by teacher educators (Grossman et al., 2009; Lampert et al., 2013, Ghousseini & Herbst, 2016), 2) the characteristics of the field placement or methods course in which the exposure occurred, and 3) candidate characteristics, such as prior knowledge, that might influence the way candidates participate in or make sense of content exposure.

There are both positive and negative consequences of the current approach. This type of broad topical conceptualization of mathematics teacher candidate learning allows for measures that can be used at a large scale without a major investment of resources and yields data that facilitate comparisons across teacher preparation programs (Blömeke & Kaiser, 2012). At the same time, the field is in need of more nuanced and detailed information about the types of experiences teacher preparation programs provide if we are to develop more teacher candidates ready to engage in high-quality mathematics instruction.

Conceptual and measurement issues are equally relevant in the K-12 space. One central conceptual challenge in K-12 mathematics education is the extent to which instructional quality in mathematics is subject-specific or content-generic (Charalambous & Kyriakides, 2017; Charalambous & Praetorius, 2018). Some researchers conceptualize mathematics teaching as subject-specific, or, fundamentally distinct from teaching other subjects (Hill and Grossman, 2013; Franke, Kazemi, Battey, 2007). Other researchers conceptualize high-quality instruction as comprised of practices that support content learning across different subjects, including mathematics (Pianta & Hamre, 2009). These different lenses have yielded different measures of mathematics teaching. These include mathematics-specific measures, designed for exclusive use in mathematics classrooms to capture subject-specific nuances of content teaching (Hill et al., 2008; Walkowiak, Berry, Meyer, Rimm-Kaufman, & Ottmar, 2013) and content-generic measures that can assess instructional quality in any classroom, regardless of the subject being taught (Hamre et al., 2013; Danielson Group, 2018).

Distinct measurement approaches highlight and obscure different aspects of teaching and learning in mathematics classrooms. For example, while a subject-specific tool like the Mathematical Quality of Instruction (Hill et al., 2008) can be used to glean important information about the mathematical depth or accuracy of instructional explanations in a lesson, it cannot provide information on the emotional tenor of teacher-student interactions. Likewise, because the Framework for Teaching (Danielson Group, 2018) was designed to evaluate the extent to which a teacher employs constructivist pedagogies in any content area, it does not yield information on the quality of mathematical discourse or the mathematical depth of a curricular task.

As a result of these different conceptualizations and measures, the field has gained two orthogonal streams of insights about mathematics instruction: insights about the ways in which content-generic practices support students' mathematics learning and insights about mathematics-specific practices support students' mathematics learning. For example, using subject-specific lenses, researchers have shown that teachers' content errors are negatively associated with their students' mathematics achievement gains (Blazar, 2015). Researchers taking a content-generic approach have shown that increases in the degree to which teaching practices support students' developmental needs, such as the need for autonomy, is associated with students' achievement (Allen, Pianta, Gregory, Mikami, & Lun, 2011). They found these interactions were not more important in one subject area than they were in another. That is, supportive interactions supported student achievement in mathematics as well as science, social studies, and language arts classrooms. While each of these approaches yields discrete pieces of information about the importance of particular subject-generic or mathematics-specific practices, there is little information about how subject-specific and content-generic aspects of mathematics teaching work together to support students' mathematical learning. Thus, one of the major consequences of these distinct conceptualizations and lines of inquiry into mathematics teaching is that there is not yet consensus on how to define or support teachers in improving their instructional quality in mathematics (Charalambous & Praetorius, 2018).

The goal of this dissertation is to surface issues of conceptualization and measurement in mathematics education and mathematics teacher education. Paper 1, "Toward a Systems Lens in Elementary Mathematics Teacher Preparation," uses Cultural Historical Activity Theory to inform a more holistic conception of elementary teacher education than is commonly used in large scale studies of mathematics teacher preparation. While prior research has provided the field with evidence about the importance of discrete experiences in mathematics teacher education, there are not yet studies that explore the ways in which these fit together. In this study, I use multivariate path analysis to consider how multiple sites, actors, and events across mathematics teacher preparation interact to form a system of learning for teacher candidates. By exploring interactive effects in the model, I highlight the importance of considering both the volume *and* the nature of learning opportunities in teacher preparation when studying the development of teacher candidate's mathematical knowledge and beliefs.

In particular, multivariate path analysis indicates there is significant variability in candidates' perceptions of their experiences and that these perceptions moderate the relationship between their learning experiences and their mathematical knowledge and beliefs at the end of their teacher preparation program. Specifically, increases in the number of mathematics content courses a candidate takes are only associated with increased knowledge when a candidate also perceives they had a positive experience learning mathematics. Similarly, more time and more reported opportunities to learn in field placements show greater associations with candidate beliefs when candidates were paired with a cooperating teacher that they perceived of as highly supportive. Discussion highlights the importance of particular roles and sites within teacher preparation systems and the affordances of a systems lens in quantitative analysis of teacher preparation.

Paper two, "Understanding Instructional Quality Through a Relational Lens," zooms in on a particular measure of instructional quality used in mathematics classrooms, the Classroom Assessment Scoring System (CLASS; Pianta, Hamre, & Mintz, 2012). The CLASS is a standardized observation protocol that suggests that high-quality lessons are distinguished by the tenor and frequency of classroom interactions. Because the CLASS focuses on interactions, rather than the specifics of a particular content area, it can be used across subjects, from language arts to mathematics to science. While many previous studies have used CLASS as a measure of instructional quality, to date, no work has examined the affordances and constraints of using the content-agnostic CLASS to examine instructional quality in mathematics lessons. This close qualitative analysis of three mathematics lessons highlights the importance of including practices that cut across content areas in measurement of instructional quality in mathematics classrooms. In addition, this paper is the first to highlight aspects of instruction in mathematics classrooms that are obscured by the CLASS. Discussion highlights how a relational lens foregrounds particular instructional aspects and marginalizes others. This paper has been published in *ZDM: Mathematics Education.*[1]

The third paper takes a more expansive view of instructional quality in mathematics than is used in current literature, one that includes *both* mathematics-specific and content-generic teaching practices. In particular, I consider the ways in which these different facets of teaching may interact with one another. This paper leverages data from mathematics-specific and content-generic observation rubrics to explore the convergence of classroom environments that support College and Career Ready (CCR) mathematics learning and classroom environments that support social and emotional learning. Two findings emerge from my multilevel latent profile analysis: 1) In this sample, there was never evidence of consistent CCR-aligned mathematical engagement absent an engaging, emotionally supportive learning environment and 2) in the lessons observed in this study, students in different classrooms had substantively

[1] This is an international journal. Therefore, the in-text citations and reference list follow a format required by the journal and do not conform to APA guidelines.

different opportunities to develop social, emotional, and mathematical competencies. Together, these findings suggest there may be substantial overlap between policy initiatives designed to improve students' college and career readiness in mathematics and their social and emotional learning.

In summary, the three papers presented here take a broader view of preparing for and understanding mathematics instruction than is typical in extant literature. Each of these papers encourages readers to hold multiple conceptions of teaching and learning in tandem, in order to consider the ways they can be leveraged to form a more holistic understanding of mathematics teacher preparation and K-12 mathematics instruction. This more expansive view both builds upon prior conceptualizations and surfaces previously obscured complexities in service of improving preparation, development, and support for ambitious mathematics teaching.

References

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

Balfanz, R., & Byrnes, V. (2006). Closing the mathematics achievement gap in high-poverty middle schools: Enablers and constraints. *Journal of Education for Students Placed at Risk*, *11*(2), 143-159.

Ball, D. L., Hill, H. C., & Bass, H. (2005). Knowing mathematics for teaching: Who knows mathematics well enough to teach third grade, and how can we decide?. *American Educator, 29*, 14-22.

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, *107*(5), 1860-1863.

Blazar, D. (2015). Effective teaching in elementary mathematics: Identifying classroom practices that support student achievement. *Economics of Education Review*, *48*, 16-29.

Blömeke, S., & Kaiser, G. (2012). Homogeneity or heterogeneity? Profiles of opportunities to learn in primary teacher education and their relationship to cultural context and outcomes. *ZDM: Mathematics Education*, *44*(3), 249-264.

Charalambous, C. Y., & Praetorius, A. K. (2018). Studying mathematics instruction through different lenses: Setting the ground for understanding instructional quality more comprehensively. *ZDM: Mathematics Education*, *50*, 1-12. https://doi.org/10.1007/s11858-018-0914-8

Charalambous, C. Y., & Kyriakides, E. (2017). Working at the nexus of generic and content-specific teaching practices: An exploratory study based on TIMSS secondary analyses. *The Elementary School Journal*, *117*(3), 423-454.

Clift, R. T., & Brady, P. (2005). Research on methods courses and field experiences. In M. Cochran-Smith & K Zeichner (Eds.) *Studying teacher education: The report of the AERA panel on research and teacher education*, (pp. 309 – 424). New York: Routledge.

Danielson Group (2018). *The framework*. Retrieved from https://www.danielsongroup.org/framework/

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

Ghousseini, H., & Herbst, P. (2016). Pedagogies of practice and opportunities to learn about classroom mathematics discussions. *Journal of Mathematics Teacher Education, 19*(1), 79-103.

Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, *111*(9), 2055-2100.

Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S. M., ... & Brackett, M. A. (2013). Teaching through interactions: Testing a developmental framework of teacher effectiveness in over 4,000 classrooms. *The Elementary School Journal*, *113*(4), 461-487.

Hassan, M. M., & Qureshi, A. N. (2018). Disrupting the rote learning loop: CS majors iterating over learning modules with an adaptive educational hypermedia. In R. Nkambou, R. Azevedo, J. Vassileva (Eds.), *International conference on intelligent tutoring systems* (pp. 66-77). Switzerland: Springer.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, *39*(4), 372-400.

Hill, H. C., Blazar, D., & Lynch, K. (2015). Resources for teaching: Examining personal and institutional predictors of high-quality instruction. *AERA Open*, *1*(4). https://doi.org/10.1177/2332858415617703

Hill, H., & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review*, *83*(2), 371-384.

Hill, H. C., Rowan, B., & Ball, D. L. (2005). Effects of teachers' mathematical knowledge for teaching on student achievement. *American Educational Research Journal*, *42*(2), 371-406.

Jung, E., Brown, E. T., & Karp, K. S. (2014). Role of teacher characteristics and school resources in early mathematics learning. *Learning Environments Research*, *17*(2), 209-228.

Lampert, M., Beasley, H., Ghousseini, H., Kazemi, E., & Franke, M. (2010). Using designed instructional activities to enable novices to manage ambitious mathematics teaching. In M. K. Stein & L. Kucan (Eds.), *Instructional explanations in the disciplines* (pp. 129-141). Boston, MA: Springer.

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., ... &

Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education, 64*, 226-243.

Lithner, J. (2015). Learning mathematics by creative or imitative reasoning. In S. J. Cho (Ed.), *Selected regular lectures from the 12th international congress on mathematical education* (pp. 487-506). Switzerland: Springer.

Ma, L. (1999). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States.* Mahwah, NJ: Erlbaum.

Mayer, R. E. (2002). Rote versus meaningful learning. *Theory Into Practice*, *41*(4), 226-232.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*(2), 109-119.

Pianta, R. C., Hamre, B. K., & Mintz S. (2012) *Classroom Assessment Scoring System Upper Elementary Manual*. Charlottesville, VA: Teachstone.

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.

Wang, T. Y., & Tang, S. J. (2013). Profiles of opportunities to learn for TEDS-M future secondary mathematics teachers. *International Journal of Science and Mathematics Education*, *11*(4), 847-877.

Wilkins, J. L. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education*, *11*(2), 139-164.

Zeichner, K. (2010). Rethinking the connections between campus courses and field experiences in college-and university-based teacher education. *Journal of Teacher Education*, *61*(1-2), 89-99.

Toward a Systems Approach to Understanding Teacher Preparation: Investigating the Development of Elementary Teacher Candidate's Mathematical Knowledge and Beliefs

Rebekah Berlin

1. Introduction

Many elementary teacher preparation programs (TPPs) seek to prepare candidates to engage in ambitious mathematics instruction. Ambitious mathematics instruction supports students in engaging in rigorous, conceptually rich mathematics tasks and has been associated with student achievement in mathematics (Hiebert and Grouws, 2007; National Mathematics Advisory Panel, 2008). Extant literature suggests that preparedness to engage in ambitious elementary mathematics instruction is multifaceted and includes teacher candidates' knowledge, beliefs, and skills (Ball & Forzani, 2009; Grossman, Hammerness, MacDonald, 2009; Drageset, 2010). However, there is markedly less scholarship that highlights how teacher candidates develop the requisite knowledge, beliefs, and skills in the context of TPPs.

A review of existing literature on preparing elementary teachers to engage in ambitious mathematics instruction reveals a common approach. In the majority of studies, researchers focus on a single component of elementary mathematics teacher preparation and its effect on teacher candidate knowledge, beliefs, or practice. Some examine the associations between course-taking and teacher candidate knowledge (Schmidt, Houang, Cogan, 2011). Others focus on the relationship between particular pedagogical experiences, such as video analysis, and candidates' developing mathematical beliefs (Philipp et al., 2007). There are few studies, however, that look *across* components of teacher preparation to examine the way candidates' experiences, perceptions, knowledge and beliefs interact as the candidate moves through the teacher preparation system.

This paper addresses this gap in the literature by applying a systems lens to elementary mathematics teacher preparation. While previous scholarship has illuminated promising practices in mathematics teacher education, the ways in which these fit together to create an effective learning system remains unknown. I argue that a turn toward sociocultural frameworks in quantitative research would allow for a broader more expansive lens that would enable researchers to consider how multiple sites, actors, and events across teacher candidate preparation interact to form a system of learning for teacher candidates.

2. Literature Review

2.1 Ambitious Mathematics Instruction

Ambitious mathematics instruction supports students in reaching ambitious mathematics learning goals that include procedural fluency, flexible reasoning, deep conceptual understanding, and the development of productive mathematical dispositions (Kazemi, Franke, Lampert, 2009; Lampert, 2001; Lampert et al., 2009; Jackson & Cobb, 2010). This type of instruction develops students' independent reasoning so that they can strategically choose which processes to use while solving problems (Lampert et al., 2009; Munter, 2014) and their ability to articulate complex mathematical arguments and critiques (Jackson & Cobb, 2010 Lampert et al., 2009). Scholars have asserted that ambitious mathematics instruction is comprised of learnable practices (Lampert & Graziani, 2009), such as eliciting students' mathematical thinking (Kazemi et al., 2009) and adjusting in real time to what students do and say while engaged in mathematical problem solving (Kazemi, Franke, Lampert, 2009; Fennema, Franke, Carpenter, & Carey, 1993).

2.2 Preparedness to Engage in Ambitious Mathematics Instruction

2.2.1 Mathematical knowledge for teaching. Extant literature has focused on the importance of preparing teacher candidates for the work of mathematics teaching (Ma, 1999; Ball & Forzani, 2009). These studies stress the importance of teaching candidates' ability to actually *enact* teaching practices in authentic classroom contexts (Grossman, Hammerness, MacDonald, 2009; Lampert, 2010; Lampert et al., 2013; MacDonald, Kazemi, Kavanagh, 2013). Researchers have begun to codify the knowledge and skills necessary for planning high-quality mathematics lessons and the moment-to-moment classroom decisions embedded in teacher practice (Ball & Bass, 2002; Ball, Lubienski, & Mewborn, 2001; Ma, 1999). These researchers suggest acquiring this body of knowledge and skills, most commonly termed "Mathematical Knowledge for Teaching" (MKT), necessitates thoughtful training (Ball & Bass, 2002; Ball & Forzani, 2009; Ma, 1999).

Research on the associations between MKT and student outcomes indicates that MKT is an important indicator of preparedness to engage in ambitious mathematics instruction. Scholars have identified significant associations between MKT and high quality instructional practices (Hill, Blazar, Lynch, 2015; Hill, Umland, Litke, & Kapitula, 2012; Kunter, Klusmann, Baumert, Richter, Voss, & Hachfield, 2013). Additionally, Hill, Rowan, and Ball (2005) found that teachers' MKT scores significantly predicted the size of their students' gain scores and that the effect size of higher MKT scores was equivalent to up to three extra weeks of mathematics instruction. Promisingly, researchers in Germany have demonstrated that this type of knowledge is malleable within the context of TPPs (Kleickmann et al., 2013). Taken together, these studies indicate that a teacher's MKT is a strong predictor of their ability to enact ambitious mathematics instruction and can be developed during teacher preparation.

2.2.2 Mathematical beliefs. In addition to teacher knowledge, many scholars emphasize that teachers' mathematics beliefs are an important part of teacher candidates' preparedness to engage in ambitious mathematics instruction (Ball, Lubienski, & Mewborn, 2001; MacDonald, Kazemi, Kavanagh, 2013; Pajares, 1992). Teachers' mathematics beliefs are multi-dimensional and include their conception of themselves as a learner, doer, and teacher of mathematics, as well as their beliefs about how mathematics ought to be taught (Wilson & Cooney, 2002; Ren & Smith, 2017).

Teacher's conceptions of themselves as learners, doers, and teachers of mathematics have been associated with many important outcomes. For example, researchers have found that a teacher's negative attitude toward mathematics is inversely associated with their assessment of their ability to do mathematics (Geist, 2015), their mathematics teaching efficacy (Swars, Daane, & Giesen, 2010), as well as their students' attitudes toward mathematics and mathematics achievement (Beilock, Gunderson, Ramirez, & Levine, 2010). Research also suggests that teachers who espouse beliefs in their ability to teach mathematics are more persistent with students (Nurlu, 2015) and intend to spend more instructional time on mathematics (Geist, 2015).

Ample evidence also exists supporting claims that teachers' mathematical pedagogical beliefs are relevant to their mathematics instruction. Pedagogical beliefs are the ideas a teacher holds about how children construct mathematical knowledge and the teacher's role in this process (Swars, Smith, Smith, Hart, 2009). In particular, pedagogical beliefs probe whether a teacher believes children benefit from more from the direct transmission of discrete skills or a cognition-based approach that prioritizes deep conceptual understanding and connections to other mathematical ideas over rote memorization (Staub & Stern, 2002; Peterson, Fennema, Carpenter, & Loaf, 1989). These divergent orientations are associated with differences in instructional decision-making, instructional practices, use of curriculum materials (Leder & Forgasz, 2002; Maasz & Schloglmann, 2009; Philipp, 2007; Wilkins, 2008; Wilson & Cooney 2002), and uptake from professional development (Phillip, 2007; Swars et al., 2009). Productive pedagogical beliefs, such as an orientation toward cognition-based approaches to mathematics learning, are associated with increases in students' mathematics achievement (Staub & Stern, 2002; Peterson et al., 1989).

Several studies highlight the mechanisms through which pedagogical beliefs shape teacher practice. Qualitative studies document that two teachers with identical knowledge may ultimately teach in very different ways because of their beliefs about how mathematical knowledge ought to be constructed (Charalambous, 2015; Clark et al., 2014; Ernest, 1989; Pajares, 1992; Ross, McDougall, Hogaboam-Gray, & LeSage, 2003; Voss et al. 2013). This qualitative knowledge base is complemented by quantitative studies suggesting beliefs mediate the relationship between practitioner knowledge and practice in elementary and middle school settings (Campbell & Malkus, 2014; Drageset, 2010; Wilkins, 2008). The associations between practicing teachers' beliefs and multiple facets of teacher practice and student learning suggest that understanding teacher candidate's beliefs is an essential part of understanding their preparedness to engage in ambitious mathematics instruction.

2.2.3 Interrelatedness of knowledge and beliefs. Knowledge and beliefs are closely related. Indeed, some scholars believe they cannot be truly separated (Liljedhal, et al., 2009). Researchers have found strong associations between pre-service teachers' MKT and mathematical beliefs in samples of pre-service teachers (Blömeke, Buchholtz, Suhl, Kaiser, 2014) and practicing teachers (Drageset, 2010; Ren & Smith, 2017; Wilson & Cooney, 2002). Other work has underscored that while it is important to distinguish between a teacher's beliefs about themselves in relation to mathematics and their beliefs about mathematics pedagogy (Wilson & Cooney, 2002), the two are closely associated (Wilkins, 2008). Though the exact nature of the relationship between prospective teachers' knowledge and beliefs remains unknown, what is clear across these studies is that they are linked. Despite this, the majority of studies that focus on the development of knowledge and beliefs in pre-service contexts tend to focus on their development in isolation, that is, focusing on either knowledge *or* beliefs (e.g., Hill et al., 2015; Schmidt et al., 2011; Geist, 2015). This may contribute to the notion that these are separate rather than interrelated constructs.

2.3 Contexts for Development in Elementary Mathematics Teacher Preparation

The majority of the literature describing the development of prospective teachers' mathematical pedagogical content knowledge and beliefs coalesces around three major learning sites: mathematics courses, mathematics methods courses, and field placements. Each of these contexts is reviewed individually below. However, these sites should also be considered in light of a robust body of evidence supporting the importance of alignment across two of these sites—methods and fieldwork (Boyd, Grossman, Lankford, Loeb, & Wyckoff, 2009; Clift & Brady; 2005; Hamerness, 2006; Mewborn, 2000; Tatto, 1996; Swars et al., 2009; Zeichner, 2010; Feiman-Nemser & Buchmann, 1985; Hamerness & Darling-Hammond, 2005).

2.3.1 Mathematics content courses. Many scholars have examined whether taking mathematics courses is related to the work of mathematics teaching. Mathematics courses are of particular interest because they present a possible avenue to address a common challenge facing programs that want to prepare elementary teachers for ambitious mathematics instruction; there can be wide variability in elementary candidates' mathematical knowledge (Charalambous, Panaoura, & Philippou, 2009). Mathematics course work can be viewed as an important way to develop content knowledge. For elementary mathematics teachers, content knowledge includes the ability to calculate correct answers to elementary mathematics problems. Topics range from counting and cardinality to operations with whole and rationale numbers. Content courses may also be a site where teacher candidates intuit ideas about how mathematics should be taught (Lortie, 1975).

There is conflicting evidence on the value of mathematics content courses. Some scholars have found strong associations between the volume of mathematics courses teacher candidates have taken and their mathematical content knowledge (Drageset, 2010; Schmidt, Houang, Cogan, 2011; Evans, 2011), MKT (Hill, Charalambous, & Chin, 2018; Schmidt et al., 2011), student outcomes (Boyd et al., 2009; Schmidt et al., 2011), and beliefs about mathematics (Matthews & Seaman 2007; Ren & Smith, 2017; Smith, Swars, Smith, Hart, & Haardörfer, 2012; Swars, Smith, Smith, & Hart, 2009; Wilkins, 2008). Other researchers have found disconfirming evidence of the positive influence of mathematics courses (Ball, Lubienski, and Mewborn, 2001; Hill, 2010; Qian & Youngs, 2015).[2]

Charalambous and colleagues (2009) documented heterogeneity in candidates' experiences of mathematics courses. Specifically, for some teacher candidates, these courses spurred positive developments. Some candidates reported more productive epistemological beliefs, increases in efficacy beliefs, and more positive attitudes toward mathematics after completing content courses. For others, content courses were the site of negative experiences that led to detrimental beliefs. These candidates stated that their experiences in mathematics courses led them to believe mathematics was a monotonous discipline made up of discrete rules and procedures, that their struggles in the course made them questions their fitness to teach elementary mathematics, and that they felt increased anxiety about mathematics. This finding may offer insight into how to interpret the conflicting evidence above.

The majority of studies on the relationship between mathematics content courses and teacher knowledge and beliefs listed above consider only the volume or mathematical focus of content courses a candidate takes (Schmidt et al., 2011; Hill, 2010; Boyd et al., 2009). Very few consider, as Charalambous and colleagues (2009) did, the *nature* of the candidate's experience in courses, examining whether the course was experienced in a fashion that facilitated candidate growth. Were this considered, clearer associations

---

[2] While some of these studies do not clarify whether the term "mathematics content courses" describes courses that are specifically designed for teachers or whether they refer to traditional college mathematics courses (Schmidt et al., 2011; Hill, Charalambous, & Chin, 2018; Boyd et al., 2009; Hill, 2010), others are more explicit. Some focus on mathematics content courses for teachers (Drageset, 2010; Matthews & Seaman 2007; Smith, et al., 2012; Swars, et al., 2009) and others on college mathematics courses (Drageset, 2010; Evans, 2011; Ren & Smith, 2017; Wilkins, 2008; Qian & Youngs, 2015).

between mathematics courses and teacher knowledge and beliefs might emerge. Without this information, it is difficult to determine whether or how many mathematics courses should be required in elementary TPPs.

2.3.2 Mathematics methods courses. With few exceptions (e.g., Qian & Youngs, 2015) researchers have found positive associations between mathematics methods courses and important outcomes for teacher candidates in the United States. The number of mathematics methods courses taken is associated with teachers' MKT (Hill et al., 2018; Swars Smith, Smith, & Hart, 2009), decreased mathematics anxiety (Harper & Daane, 1998; Swars et al., 2009), and positive attitudes toward mathematics (Gresham, 2007; Jong & Hodges, 2015; Philipp et al., 2007; Ren & Smith, 2017). In a review of the literature, Clift and Brady (2005) documented 10 qualitative studies that identified math methods courses as the site of important shifts in teachers' beliefs about mathematics, orientation toward reform instruction, decrease in math anxiety, and confidence in their ability to teach mathematics.

Several researchers have examined the ways in which MKT might be developed in methods courses and uncovered promising pedagogies. For example, a growing body of work provides theoretical arguments for pedagogies of practice (Grossman et al., 2009). Pedagogies of practice include representations, decompositions, and approximations of practice. Representations of practice are examples of the work of teaching and can include video footage, written cases, and lesson plans. Decompositions of practice involve the breaking apart of representations so that individual components of a practice are available for analysis. For example, a methods instructor might decompose the process of launching a complex mathematics task into a series of four steps (Jackson, Shahan, Gibbons, & Cobb, 2012) so that novices can analyze and rehearse them individually. Approximations of practice allow candidates to rehearse in a highly supported environment instructional practices, such as leading a mathematics routine, they will use as classroom teachers (Kazemi & Waege, 2015; Lampert & Graziani, 2009; Lampert et al., 2013; McDonald, Kazemi, Kavanagh, 2013). While these pedagogies have great conceptual promise, to date, no one has empirically examined their impact on any teacher candidate knowledge, beliefs, or skills.

2.3.3 Field placements. Much of the literature on field experiences in teacher education emphasizes that it is not so much the presence, but the characteristics, of the field placement, that impacts future teacher practice (Ronfeldt, 2012; Boyd et al., 2009). Beyond the characteristics of the school where students are placed, much research has focused on the role and quality of the cooperating teaching and university supervisor. Unfortunately, there is wide variability in the quality of feedback teacher candidates receive from their cooperating teachers and university supervisors (Borko & Mayfield, 1995; Frykholm, 1998; Solomon, Erikson, Smestad, Rodal, & Bjerke, 2017). There is also substantial variability in the types of practice candidates observe (Jong & Hodges, 2015). Some hypothesize this is due to resource constraints. For example, there are often not enough cooperating teachers that model the type of instruction valued in the TPP for every candidate to receive a high-quality placement (Borko & Mayfield, 1995; Philipp et al., 2007). Similarly, university supervisors are frequently assigned to support elementary teacher candidates in subjects in which the supervisor has no training (Borko & Mayfield, 1995; Zeichner, 2010). Research has also emphasized the importance of teacher candidate perception of and trust in their cooperating teacher and university supervisor (Goodman, 1984; Mewborn, 2000) for stabilizing teacher candidate beliefs.

2.4 Toward a Systems Approach

Across studies in each of these three sites — mathematics content courses, mathematics methods courses, and teacher candidate fieldwork— the predominant goal was to highlight the relationship between two constructs of interest. Most of the analyses focused on the role of a specific component of teacher preparation, such as the volume or type of courses taken, in developing teacher knowledge or beliefs. There are valid reasons for this type of research. By zooming in to focus on a particular aspect of teacher preparation, several of the qualitative studies above were able to provide nuanced descriptions of the conditions in which particular experiences were and were not beneficial for teacher candidates.

The quantitative studies tended to approach teacher preparation through a framework of "inputs and outputs." In these studies, researchers used descriptive techniques to highlight the significance and magnitude of associations between specific variables of interest after controlling for other teacher candidate or program characteristics (e.g., Schmidt et al., 2011, Drageset, 2010, Ren & Smith 2017). This approach has given the field several individual puzzle pieces that provide discrete, and at times conflicting, information about teacher preparation.

Though there is value in zooming in to understand the relationship between experiences in and outcomes of teacher preparation, it may be that by only focusing on single associations of interest, other important relationships are obscured. Hill, Charalambous, and Chin (2018) assert that the field is in need of studies that use an expanded view and consider multiple constructs in tandem, particularly those that explore interactive effects. They also suggest that while in isolation certain variables may appear significant, they may not remain significant when other variables are included. Though Hill and colleagues (2018) are speaking about research on practicing teachers, their message is salient to research on teacher preparation. Until researchers begin to consider the ways in which these puzzle pieces (i.e., methods courses, field experiences, perceptions of TPP faculty) fit together to form the broader landscape of teacher preparation, the conclusions they draw are inherently limited.

It may be that an alternate framework is better suited to studying the complex ecology of teacher preparation than one of inputs and outputs. Scholars of activity systems analysis posit that human learning activity cannot be reduced and separated into mutually exclusive variables while still reflecting the richness of real-world experience (Yamagata-Lynch, 2010). Rather than focus on the impact of a particular quantifiable component of teacher education, scholars working within the framework of Cultural Historical Activity Theory (CHAT) shift the unit of analysis to the teacher preparation system *as experienced by* the teacher candidate (Engeström, 1993). In this framework mental meaning making and observable activity are inextricably intertwined (El'Konin, 1993). Therefore, analyzing the teacher preparation system on the personal plane (Rogoff, 2008), or at the level of elementary teacher candidates, requires considering the ways in which teacher candidates' perceptions and experiences interact with one another.

Through this lens, each teacher candidate is an actor in a system of learning driving toward a multifaceted "object" or goal, in this case multidimensional preparedness to engage in ambitious mathematics instruction (Foot, 2014; Zeichner, Payne, & Brayko, 2015). Candidates must be considered in light of multiple characteristics including their personal and mathematical history, their knowledge, and their beliefs. CHAT analyses also include a community of significant others (Engeström, 1999). Here, this includes members of the teacher preparation community with whom teacher candidates interact. This community divides the labor of teacher preparation (Foot, 2014; Engeström, 1999). Therefore, expertise is distributed across the system (Zeichner, 2015). For example, methods instructors, university supervisors, and cooperating teachers each contribute to preparedness in different settings and ways. Under the CHAT framework, each member of the system employs various material and conceptual tools to pursue the desired object of the system (Douglas, 2010; Engeström, 1999), which are most effective when they are aligned to the object of the system (Douglas, 2010). For example, a methods instructor might utilize pedagogical tools, such as decompositions of practice in their methods course, in order to develop in candidates the pedagogical content knowledge required to facilitate discussions on children's solution strategies (Ghousseini & Herbst, 2016).

The CHAT framework also emphasizes that teacher candidates' learning does not occur in a vacuum. Rather candidates regularly have to navigate between multiple overlapping activity systems, such as those in their coursework and field placements (Jahreie & Otteson, 2010). This work of "boundary crossing" can be difficult for novices because each activity system has its own goals and candidate development may not be the sole priority (Zeichner, 2010). For example, in a field placement a cooperating teacher may desire to support a candidate's development. However, this cooperating teacher is also likely focused on the important work of attending to their K-12 students' development and may also be held accountable to other district and school initiatives that cause them to diverge from the framework for mathematics teaching presented in candidates methods courses (Zeichner et al., 2015).

CHAT also attends to notions of power and privilege and how these are built over time (Foot, 2014). For example, Zeichner and colleagues (2015) use a CHAT framework to suggest that candidates' sense making about the differences between coursework and fieldwork is made more difficult by power imbalances across overlapping activity systems. They argue that TPPs often imply that there is a hierarchy of expertise where the perspective of methods instructors is more valuable than that of university supervisors or cooperating teachers. They suggest this hierarchy is detrimental to both the growth of candidates and the health of the TPP. In a CHAT framework, such tensions are viewed as beneficial because they are catalysts for growth (Engeström 2001), and indeed, multiple scholars have leveraged the conceptual tools of CHAT to subvert traditional hierarchies in teacher education (Anagnostopoulos, Smith, & Basmadjian, 2007; Norton-Meier & Drake, 2010; Zeichner, 2010).

Finally, CHAT attends to the processes candidates engage in to make meaning of each of the experiences they have during their TPP. These include those that take place in mathematics content courses, mathematics methods courses, and field placements. What candidates take from each experience is mediated by their prior experiences and beliefs, and simultaneously informs alterations to their beliefs and knowledge (Vygotsky, 1978; Vygotsky, 1987). Because the CHAT framework better represents the lived experiences of teacher candidates, it decreases the likelihood of omission of critical constructs or the interactions among constructs.

Quantitative methods are limited in their ability to capture the richness or real-time evolution of activity systems. For this reason, researchers using a CHAT framework almost exclusively use qualitative methods (Yamagata-Lynch, 2010). However, Yamagata-Lynch (2010) notes the importance of complimentary quantitative studies, if researchers are interested in examining whether conceptualizations of activity systems hold across broader populations. Increasingly, researchers are utilizing structural equation modeling to create analytic models that are more reflective of participants' experiences, and associated internal negotiation of these experiences, than traditional quantitative research (Stage & Wells, 2014). In the context of teacher preparation, this type of quantitative analysis could be particularly helpful in examining the ways in which aspects of TPPs (i.e. experiences and personnel across different sites) and teacher candidates themselves (i.e. perceptions, knowledge, and beliefs) work together to form a learning system.

Presently, the field is saddled with the task of cobbling together an approximation of quality elementary teacher preparation in mathematics from the myriad isolated experiences researchers have found relate to teacher candidates' knowledge and beliefs. Thus, studies that focus not only on *what* facets of teacher preparation are important, but also on *how* these facets interact with one another are required. To capture a global portrait of the interacting agents and experiences that make up the teacher preparation learning system, this paper addresses the following research question:

- To what extent and in what ways are TCs' perceptions of their background experiences, experiences in teacher preparation, and faculty in their preparation program associated with their preparedness to teach elementary mathematics?

This study is among the first to take a holistic approach to understanding the preparation and preparedness of hundreds of elementary teacher candidates across multiple institutions. Rather than focus on the development of individual facets of preparedness to engage in ambitious mathematics instruction, such as an exclusive examination of teacher candidate knowledge or teacher candidate beliefs, I explore their development in tandem. Likewise, instead of looking for associations between specific features of teacher preparation, such as methods courses, and outcomes, I look *across* coursework and field experiences to capture the broader teacher preparation system. Finally, I consider the teacher candidates themselves, using their perceptions as a way to probe the nature of their experiences during teacher preparation and the meaning they make from those experiences. By shifting the unit of analysis away from program characteristics to the unit where preparation actually occurs—preparation as experienced by the teacher candidate—important information about the nature of the experiences that support candidate preparedness is revealed.

3. Method

3.1 Sample

The data presented here are drawn from a larger ongoing longitudinal study of elementary teacher preparation. Participants in this study were pre-service elementary teaching candidates in four traditional TPPs at public universities in the United States (48% response rate; *n = 220; Program 1 = 26, Program 2 = 77, Program 3 = 50, Program 4 = 67)*. Information about the structure of each preparation program is listed in Table 1. Typical of most TPPs in the United States (Sleeter, 2001), most candidates identified as white (87%) and female (95%).

Table 1. Descriptive Information for Elementary Teacher Education Programs in Study

	Program 1	*Program 2*	*Program 3*	*Program 4*
Length of Program	5-year BS/MA program	5-year BA program plus MA credits	5-year BS/MA program	4-year BA program
Approximate Annual Number of Elementary Graduates	40	200	60	200
Required Course Sequence	Yes	Yes	No	Yes
Required Math Methods Courses	1 math methods	2 math methods	1 math methods	1 math methods
Pre-Student Teaching Field Experience	6 hours/week for 3 semesters	4 hours/week for 2 semesters	1 day/week for 1 semester	6 hours/week for 1 semester
Length of Student Teaching	12 weeks	30 weeks	15 weeks	15 weeks
Length of Lead Responsibility for Teaching	5 weeks	10 weeks	8 weeks	8 weeks
Timing of Student Teaching	Spring of 4th year	Fall and spring of 5th year	Fall of 5th year	Spring of 4th year

3.2 Measures

Scales were drawn from two measures. The first, the Elementary Teacher Candidate Survey, included multiple scales that probed teacher candidate's backgrounds and experiences during teacher preparation. The second measure was the Mathematical Knowledge for Teaching Number Concepts and Operations measure (Hill, Ball, & Schilling, 2008). Both were administered online to teacher candidates near the end of or immediately following graduation from their TPP. Teacher candidates were paid $25 for each measure they completed.

3.2.1 Mathematical knowledge for teaching. The Mathematical Knowledge for Teaching (MKT) measure focuses on various parts of subject specific teacher knowledge (e.g., the ability to identify the mathematical misconceptions causing common student errors or whether a non-traditional student solution strategy generalizes). The research team administered the Elementary Numbers Concepts and Operations – Content Knowledge form. This domain was chosen because it is the largest curriculum focus in the United States (National Mathematics Advisory Panel, 2008). Responses were scored using the 2-PL Item Response Theory (IRT) method. Teaching candidate's IRT scores were used as the outcome in the subsequent analyses. IRT scores ranges from -2.75 to 2.41 (M = 0.30, SD = 0.83).

3.2.2 Mathematical beliefs. A teacher candidate's mathematical beliefs are multifaceted and include their beliefs about the ways mathematics should be taught as well as their conceptions of themselves as a learner, doer, and teacher of mathematics. Questions regarding candidate's mathematical beliefs in the Elementary Teacher Candidate Survey were modified from existing surveys (Dweck, Chiu, & Hong, 1995; Schmidt, Cogan, & Houang, 2011; Vacc & Bright, 1999) and interview protocols (Drake, 2006; Dweck, Chiu, & Hong, 1995). Respondents were asked to respond on a four-point Likert scale indicating the extent to which they agreed with several statements.

Statements probing teacher candidate's conceptions of themselves as a learner, doer, and teacher of mathematics included: "Even if I work hard, I will not teach math as well as I will most subjects," "I'm not the type to do well in mathematics," and "I understand math concepts well enough to be effective in teaching math." After negative items were reverse coded, the five items were averaged to create a scale score ($M = 3.30$, $SD = 0.45$, $\alpha = 0.75$).

Four items were used to probe teacher candidates' mathematical pedagogical beliefs. These items are intended to measure a teacher candidate's orientation toward ambitious mathematics teaching practices – in particular those that prioritize deep conceptual understanding over rote memorization. Some of these items probed general ambitious instructional practices that are important to developing conceptual understanding in mathematics (e.g., "It is helpful for pupils to discuss different ways to solve particular problems" and "In addition to getting a right answer in mathematics, it is important to understand why the answer is correct"). Other practices were mathematics-specific, in that they pertain to mathematics teaching and not to the teaching of other subjects (e.g., "Non-standard procedures in mathematics should be discouraged because they can interfere with the correct learning procedure" (reverse coded)). Items were again averaged to create a scale score ($M = 3.41$, $SD = 0.31$, $\alpha = 0.45$).

3.2.3 Opportunities to learn (OTL). Survey items about OTL in teacher preparation were modified from the New York City Pathways study (Boyd et al., 2009) and the Teacher Education and Development Study in Mathematics (Schmidt, Cogan, & Houang, 2011). Participants were asked to indicate whether or not they had had the opportunity to learn about or attempt teaching practices including: "Design high cognitive demand mathematics tasks for students," "Facilitate students' use of manipulatives in doing mathematics," and "Identify and respond to common patterns of student thinking in mathematics (e.g., strategies, misconceptions)."

Previous literature suggests that content cannot be divorced from pedagogy in teacher education (Ghousseini & Herbst, 2016). Therefore, for each topic, candidates were also required to indicate whether or not they had engaged with this topic through particular pedagogies of practice in their methods courses (Grossman et al., 2009; Lampert et a., 2013). They indicated whether they had engaged with representations and decompositions of specific mathematics practices through written and video cases. They also reported the number of these practices they had the opportunity to approximate, or rehearse with a peer in their methods course. Each opportunity to engage with a topic through a practice-based pedagogy candidates reported was assigned one point. Items were summed to create a total score for OTL in methods courses (*M=10.73, SD= 4.86)*.

Candidates also indicated whether they had exposure to the same practices in their field placements. They reported whether they had "observed a teacher use this practice with students" as well as whether they had "received feedback on their attempts to use this practice with students." Each of these reported OTL was assigned one point and summed to create a total score for OTL in field placements (*M=12.40, SD= 5.40)*.

3.2.4 Teacher candidate's perceptions of quality. Teacher candidates were asked to rate on a four-point Likert scale the extent to which they agreed with several statements regarding their university supervisor and cooperating teacher. Scores of 1 or 2 indicated disagreement; scores of 3 or 4 indicated agreement. Negative items were reverse coded. Items were first averaged to create scale scores and then dichotomized to indicate whether the teaching candidate perceived this faculty member to be a high-quality mentor or not. Average scores over 2.5 were assigned a value of "1" to indicate that, on average, the candidate perceived this faculty member to be a high-quality mentor.

Average scores below 2.5 were assigned a value of "0" to indicate that, on average, the teaching candidate did not perceive this faculty member to be a high-quality mentor.

The five-item Perceptions of Support from University Supervisors scale (*M=3.57, SD = 0.51,* α = 0.85*)* probed candidates' perceptions of support from their supervisor as well as the extent to which this support aligned with other components of the preparation system. The scale included the following statements: "My supervisor was available to talk with me when I had questions or concerns about teaching," "My supervisor observed me on a regular basis," "My supervisor gave me useful feedback on my teaching," "My supervisor provided feedback that was aligned with the theories and methods advocated in my methods courses," and "My supervisor and cooperating teacher held similar ideas about teaching and learning." 210 candidates indicated they perceived their supervisor to be a high-quality mentor. Only 10 candidates indicated that they perceived their supervisor was not.

The seven-item Perceptions of Support from Cooperating Teacher scale (*M=3.38, SD = 0.43,* α = 0.72*)* also explored perceptions of support and alignment. This scale included statements such as "My cooperating teacher is an excellent teacher and a worthy role model," "My cooperating teacher taught in ways that were quite different from the methods I was learning in my university courses (reverse coded)," and "My cooperating teacher gave me useful feedback." 206 candidates indicated their cooperating teacher was a high-quality mentor. 14 candidates indicated that their cooperating was not a high-quality mentor.

3.2.5 Teacher candidates' mathematics histories. Respondents indicated the number of courses they took in each of the following topics in high school, college, and graduate school: Algebra I & II, Geometry, Statistics (Probability), Pre-Calculus/Calculus, and Trigonometry. These were summed to create a total count of mathematics coursework for each participant. Responses ranged from one to eleven mathematics courses ($M = 6.38$, $SD = 2.10$). In addition, teacher candidates indicated whether they agreed with the statement, "I have had mostly positive experiences learning mathematics."

Participants also self-reported their high school grade point average (GPA) as well as the scale their district used to compute GPAs. I divided GPA by the GPA scale to create a scaled GPA for each respondent ($M = 0.93$, $SD = 0.08$).

3.3 Analysis

A path diagram (see Figure 1 for conceptual model) allows for simultaneous exploration of the relationship between teacher candidates' background experiences, perceptions of experiences in teacher preparation, perceptions of faculty in their preparation program, and multiple aspects of preparedness to teach elementary mathematics. A multivariate approach, or an analytic design that includes multiple theoretically associated outcome variables, was chosen given the depth of literature on the interrelatedness of teacher candidate knowledge and beliefs and the goal of examining a holistic conception of teacher candidate preparedness. The model has three outcomes: (a) teacher candidate MKT; (b) teacher candidate beliefs about themselves as a doer, learner, and teacher of mathematics; and (c) teacher candidate beliefs about mathematics pedagogy. I also correlated the residuals for each of these three endogenous variables assuming that some of the variance not accounted for by my predictors was likely related between the outcomes. In order to model the system of teacher preparation, I included teacher candidate background experiences, teacher candidate reports of OTL in methods courses, and teacher candidate reports of OTL in their field placement.

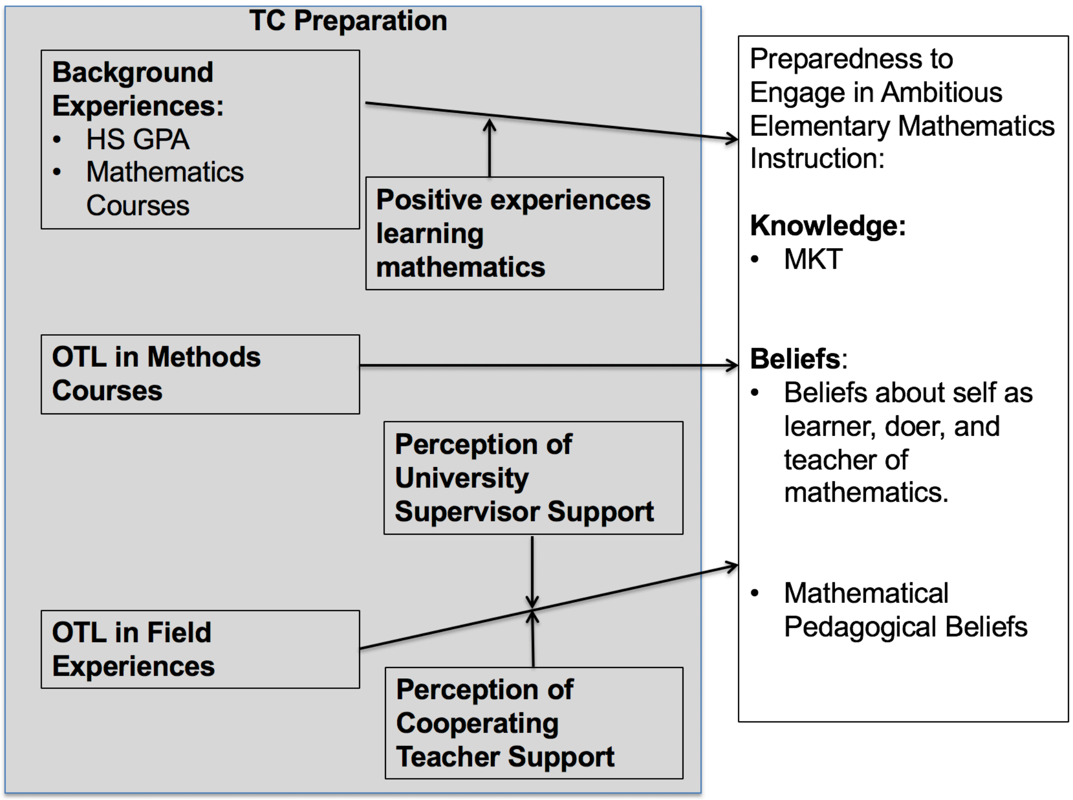


Figure 1. Hypothesized framework for path analysis.

In order to capture the broader "learning system" a candidate participates in during teacher preparation, I attended not only to candidates' experiences, but also to candidates' perceptions of those experiences. This allowed me to capture not just the number but also what candidates perceived as the quality of their experiences in their TPP. To do this, I generated several interaction terms. The first was designed to see if the association between the volume of mathematics courses a teacher candidate reported taking and their knowledge and beliefs was moderated by whether the they reported having "mostly positive experiences learning mathematics." I used this to investigate whether the differential impact of positive and negative experiences in mathematics courses documented qualitatively by Charalambous and colleagues (2009) held at a larger scale.

A substantial body of literature suggests significant heterogeneity in teacher candidates' experiences in their field placements (Borko & Mayfield, 1995; Solomon et al., 2017; Zeichner et al., 2015). I wanted to see whether certain types of experiences during fieldwork were differentially associated with candidates' developing knowledge and beliefs. For example, I hypothesized that additional OTL during fieldwork might not be associated with greater preparedness if the teacher candidate reports a lack of support in the field or little alignment between fieldwork and methods. To test this, I generated two interaction terms to see if the association between the number of teacher candidate reported OTL in their field placement and teacher candidate knowledge and beliefs was moderated by their perceptions of whether their university supervisor and cooperating teacher were high-quality mentors.

Despite the fact that teacher candidates were nested into four different TPPs, I did not account for this in my analysis for two reasons. First, four clusters is below the recommended level-two sample size for multi-level modeling (Maas & Hox, 2005). Second, after calculating intraclass correlation coefficients (ICC) from unconditional two-level models for each outcome variable, I found that very little of the variance was accounted for at the program level (ICCs all < 0.05; Raudenbush and Bryk, 2002).[3]

[3] As a sensitivity analysis, after initial estimation I reran my model with program fixed effects. Parameter estimates were robust to inclusion of program fixed effects. This is consistent with prior literature suggesting there are more similarities than differences between teacher preparation programs (e.g., Boyd, Lankford, Loeb, Rockoff, & Wyckoff, 2008).

To evaluate model fit, I examined the comparative fit index (CFI), the root-mean-square error of approximation (RMSEA), standardized root mean squared residual (SRMR) and the Tucker-Lewis index (TLI). CFI and TLI values > 0.90 and SRMR and RMSEA < 0.08 indicate acceptable fit (Hu & Bentler, 1999; Kline, 2005). All models were estimated using the maximum likelihood option in Stata 14.

4. Results

4.1 Model Fit

Several factors point to good model fit with the data. First, the chi-square value was not significant ($\chi^2(2) = 1.44$, $p = 0.49$). Second, fit indices suggest strong model fit (CFI = 1.00, TLI = 1.05, RMSEA = 0.00, SRMR = 0.01). In addition, R-squared values indicate the model helps to explain variation in the endogenous variables. The model explains 18% of the variance in teacher candidate MKT scores; 35% of the variation in teacher candidate beliefs about themselves as a doer, learner, and teacher of mathematics (DLT_M); 15% of the variance in their beliefs about mathematics pedagogy (M_Ped); and 52% of the variance in the model overall.

4.2 Estimates

4.2.1 Overview Standardized parameter estimates for are presented in Table 1.[4] In addition to the estimates in Table 2, the correlation between the residuals of the two belief measures was significant ($p < 0.001$). The correlations between the error terms of the other endogenous variables were not significant. In addition, neither the interaction terms nor the main effect for the perceived quality of university supervisor were

[4] I do not highlight the main effects for the dichotomous variables in the results or discussion section. This is because the point estimates for the individual dichotomous variables that are also included in interaction terms do not yield meaningful interpretations.

significant. For this reason, I dropped them in favor of a more parsimonious model. Likelihood ratio tests indicated that the reduced model did not differ significantly from the original model ($p > 0.10$).

Table 2

Standardized parameter estimates and standard errors

	MKT	DLT_M	M_Ped
HS GPA	0.15**	-0.01	-0.04
	(0.06)	*(0.06)*	*(0.07)*
Number of Mathematics Courses	-0.09	-0.32**	-0.03
	(0.13)	*(0.11)*	*(0.13)*
Mathematics Courses x Positive	0.39**	0.76***	0.19
Experiences Learning Mathematics	*(0.16)*	*(0.14)*	*(0.17)*
OTL Method	0.17**	0.12*	0.20**
	(0.07)	*(0.06)*	*(0.07)*
OTL Field	0.11	-0.74**	-0.65**
	(0.28)	*(0.25)*	*(0.29)*
OTL Field x Perception of CT Quality	-0.20	0.94**	0.91**
	(0.33)	*(0.28)*	*(0.33)*

Note. * $p < 0.10$, ** $p < 0.05$, *** $p < 0.001$, ns = non-significant

4.2.2 Mathematics histories. These data indicate the impact of taking more mathematics content courses is highly divergent, depending on the quality of the experience. For example, an increase in the number of mathematics courses a teacher candidate reported taking was *negatively* associated with a teacher candidate's belief in their ability to do, learn, and teach mathematics when the candidate reported they had *not* had “mostly positive experiences learning mathematics.” However, for every one standard deviation increase in the number of math courses a candidate reported taking when the candidate also reported having had *positive* experiences learning mathematics, there was an associated 0.44 standard deviation increase in their belief in their ability to do, learn, and teach mathematics. There was also an associated 0.30 standard deviation increase in their MKT score. Put plainly, there were different associations between the number of mathematics courses a candidate reported taking and their knowledge and beliefs, depending on whether they reported their experiences learning mathematics were positive or negative. There was no association between mathematics course taking and candidates’ pedagogical beliefs, regardless of whether they had positive or negative experiences learning mathematics. In addition, candidates who reported higher high school GPAs also had significantly higher MKT scores (0.16, $p = 0.01$). GPAs were not associated with either type of belief.

4.2.3 OTL in methods courses. These data show that increased opportunities to engage with representations, decompositions, and approximations of mathematics teaching practices in methods courses are associated with higher MKT scores (0.17, $p = 0.01$) as well as more productive pedagogical beliefs (0.20, $p < 0.001$). Increased exposure to pedagogies of practice is also positively associated with teacher candidates’ beliefs about their ability to do learn and teach mathematics, though this relationship is only marginally significant (0.12, $p = 0.06$).

4.2.4 OTL in field placements. In this sample, when a candidate reported that their cooperating teacher was a high-quality mentor, increased opportunities to observe, attempt, and get feedback on mathematics teaching practices during their field experience were associated with more productive beliefs about mathematics pedagogy and their ability to do, learn, and teach mathematics. Specifically, for every one standard deviation increase in candidates' reported OTL in their field placements when they also reported their cooperating teacher as a high-quality mentor, there was a 0.26 standard deviation increase in their pedagogical beliefs and a 0.20 standard deviation increase in their beliefs in their ability to do, learn, and teach mathematics.

For candidates who did not report having a high-quality mentor, increased opportunities to observe, attempt, and get feedback on mathematics teaching practices were *negatively* associated with their mathematical beliefs. These candidates were less likely to report they were capable of doing, learning, or teaching mathematics (-0.74, p = 0.03), just as they were less likely to endorse pedagogical strategies that promote conceptual understanding over rote memorization (-0.65, p = 0.02). OTL in the field were not associated with MKT, regardless of whether candidates reported their cooperating teacher was a high-quality mentor.

5. Discussion

5.1 Overview

The goal of this study was to leverage the conceptual underpinnings of CHAT to understand the ways in which teacher candidates' perceptions and experiences within elementary teacher preparation interact to form a system of learning that supports preparedness to engage in ambitious mathematics instruction. Results from the path analysis demonstrate how reported experiences in methods courses and field placements work in complementary ways to support candidate preparedness. Results also highlight the extent to which relationships between reported experiences (i.e., content courses taken, OTL in fieldwork) and candidate's knowledge and beliefs vary by other reported characteristics of those experiences.

Most prior work on opportunities to learn considers only the volume of OTL (e.g., Drageset, 2010; Schmidt et al., 2011; Evans, 2011). There are affordances of this approach. For example, it is possible to enact policies that regulate the volume of specific OTL. TPPs can require specific courses or topics to be covered in coursework. However, the significant interaction terms in this analysis suggest that when conceptualizing a *system* of teacher preparation, it is important to attend to both the volume and the quality of teacher candidates' experiences.

These data suggest that understanding the quality of individual teacher candidates' experiences is critical to understanding the relationship between preparation and preparedness. Only when teacher candidates provided evidence that they had high quality and well-supported OTL were their experiences positively associated with preparedness to engage in ambitious elementary mathematics instruction. A systems approach revealed that these findings were not unique to a single aspect of teacher preparation. Instead, they held true across the broader learning system. Taking mathematics content courses was positively associated with preparedness *only* when candidates also reported positive experiences learning mathematics. Opportunities to observe and attempt teaching practices in candidates' field placements were positively associated with preparedness *only* when candidates reported their cooperating teacher was a high-quality mentor. In methods courses increased OTL through well-supported practice-based pedagogies were associated with increased preparedness. These results highlight that teacher candidates' experiences in teacher preparation cannot be assumed to be uniform and that understanding the quality of these experiences is essential to understanding the relationship between preparation and preparedness.

This discussion moves through components of the system of teacher candidate learning—candidates' mathematics histories, OTL in mathematics methods courses, and OTL during field placements—using CHAT as a means to interpret results and provide suggestions for improving elementary mathematics teacher preparation. Next, it provides a synthesis of findings at the system level and outlines the affordances of a systems approach to quantitative analysis of teacher preparation. It concludes with an outline of study limitations and overarching conclusions.

5.2 Mathematics Histories

The volume of mathematics content courses a teacher candidate took was positively associated with their preparedness to engage in ambitious mathematics instruction only when the candidate also reported having positive experiences learning mathematics. In the case of teacher candidates' beliefs about their ability to do, learn, and teach mathematics, additional mathematics content courses were actually *negatively* associated with these beliefs when the candidate reported they had not had positive experiences learning mathematics. This confirms across hundreds of candidates what Charalambous, Paunaoura, and Philippou (2009) documented qualitatively in a small sample—negative experiences learning mathematics are associated with negative beliefs about mathematics. In addition, the relationship between candidates' mathematics histories and their preparedness to teach elementary mathematics ambitiously is evidence that the preparation system functions at the individual level—not just at the programmatic level—and that preparation begins long before candidates enter an official TPP. Because CHAT assumes candidates' histories are not uniform and that this variability is vital for understanding their later preparedness, this finding supports using a holistic lens such as CHAT to analyze novices' learning.

5.3 Pedagogies of Practice in Mathematics Methods Courses

A major finding from this study is that the use of practice-based pedagogies in mathematics methods courses is positively associated with multiple facets of candidates' preparedness to engage in ambitious elementary mathematics instruction. Prior literature has theorized that practice-based pedagogies may be critical to solving the problem of enactment in teacher preparation, or, the gap between the educational theory that is often the focus of methods courses and the actual knowledge, beliefs, and skills that constitute preparedness to support student growth in real classrooms (Ball & Forzani, 2009; Grossman et al., 2009; Lampert et al., 2013; Kennedy, 1999). This is the first study to provide empirical evidence that supports this claim and to show that this theory holds at a large scale in the context of mathematics teacher education.

The strong association between exposure to practice-based pedagogies in mathematics methods courses and candidates' preparedness to teach elementary mathematics suggests that the pedagogies of practice framework could be a powerful lens through which teacher educators could reflect and improve the quality of the OTL they afford teacher candidates. Ghousseini and Herbst (2016) provide a model of how individual mathematics teacher educators can engage in this type of analysis of their teaching to determine whether and how the pedagogies they employ contribute to various facets of candidate development. Viewed through a CHAT lens, this type of self-inquiry affords the opportunity for teacher educators to transform the tools they use to support teacher candidates' development so that these tools become better aligned to their object—candidates' preparedness to teach mathematics ambitiously (Douglas, 2010).

5.4 High-Quality Mentors

These data also suggest that the impact of a field placement depends equally on the perceived quality of the cooperating teacher as it does on the types of OTL a candidate is afforded while student teaching. Consistent with prior literature (Borko & Mayfield, 1995, Hamerness & Darling Hammond, 2005; Zeichner, 2010), these results reveal wide variability in the experiences teacher candidates report during student teaching. They also concretize the impact of this variability in the unique context of mathematics teacher preparation: when a candidate is placed with a cooperating teacher that the candidate reports is not a high-quality mentor, increased OTL may actively harm a candidate's developing mathematical beliefs.

The CHAT framework offers a lens through which to interpret the divergent impact cooperating teachers have on the relationship between a candidate's OTL in the field and their mathematical beliefs. Using CHAT, scholars have suggested teacher preparation is comprised of overlapping activity systems, such as those at the university-based program and the school in which a candidate is placed (Jahreie & Otteson, 2010; Zeichner et al., 2015). Scholars also indicate that navigating between the divergent goals and rules of multiple activity systems is difficult for novices (Zeichner et al., 2015). Indeed, the candidates in this study who did not report having a high-quality mentor whose teaching and support were aligned to what candidates were learning in their methods courses may have fallen into the "two worlds pitfall" described by Feiman-

Nemser and Buchmann (1985). Researchers suggest that candidates left to independently make sense of the disconnect between their methods courses and field are less likely to apply what they learn in either setting, do not learn as much as better supported peers, and often report feeling confused, frustrated, and doubtful of their ability to become a teacher (Hamerness & Darling-Hammond, 2005). This may explain the negative association between increased OTL in the field and both types of mathematical beliefs when candidates report they were paired with a mentor who did not teach, provide feedback, or allow candidates to attempt practices in ways that were aligned with what candidates were learning in other parts of their program.

In contrast, other candidates in this sample report having high-quality mentors who were knowledgeable about the candidate's teacher education program, were a worthy role model, taught in ways that were aligned to what was taught in the candidate's methods courses, and allowed the candidate to try out practices they learn in their methods courses. Being placed with this type of cooperating teacher may have meant that the burden of boundary crossing between the activity systems of the university and school placement did not fall solely upon the teacher candidate. In addition, candidates who reported that their cooperating teacher was a high-quality mentor also indicated that their cooperating teacher was supportive – giving helpful feedback, in the room when the candidate taught, and regularly holding useful meetings to discuss teaching. Studies that are not specific to the preparation of mathematics teachers, but discuss teacher preparation in general, point out that increased support and coherence of this type can lead to increased teacher candidate take up of important teaching practices; candidates are not left to simply sink or swim in student teaching (Zeichner, 2010; Hamerness &

Darling-Hammond, 2005). The results of this study add to this body of literature by showing across hundreds of teacher candidates the importance of this type of coherence and support from a cooperating teacher in the specific context of mathematics teacher preparation – being paired with a high-quality mentor is of critical importance to an elementary teacher's developing identity as a mathematics educator because of its impact on candidates' mathematical beliefs.

In light of these findings, TPPs may find the cooperating teacher role a particularly important lever for candidate's development. In this sample, a cooperating teacher proved effective when they were both actively involved in the field placement and supported teacher candidates in ways that aligned with what candidates were learning methods courses. Because these skills are the result of knowledge sharing and dialogue between cooperating teachers and university faculty, the results of this study suggest the value of investing in communication with cooperating teachers more than they suggest a particular selection model. Several scholars have outlined the ways in which TPPs have successfully created "third spaces," where members of overlapping activity systems such as methods instructors, university supervisors, and cooperating teachers come together to engage in synergistic work that supports candidates' development (Anagnostopoulos, Smith, & Basmadjian, 2007; Norton-Meier & Drake, 2010; Zeichner, 2010). TPPs may also want to reconsider the role of the university supervisor given that in this study, as in others before it (Borko & Mayfield, 1995; Hamerness & Darling-Hammond, 2005, Zeichner & Miller, 1997), the reported actions of university supervisors did not appear related to candidates' development.

5.5 Affordances of a Systems Lens

Analyzing teacher candidate learning at the system level yielded findings that would not have emerged had this analysis focused instead on the relationship between isolated experiences in teacher preparation and a single dimension of preparedness. This approach revealed that at the system level, cooperating teachers played as important a role as methods instructors in candidates' holistic development. For example, the effect of having a high-quality mentor teacher on the relationship between OTL and both types of candidates' mathematical beliefs was positive and significant. This was not the case for experiences in methods courses. OTL in methods courses were only positively and significantly associated with MKT and productive pedagogical beliefs; they were not significantly associated with candidates' beliefs about their ability to do learn and teach mathematics. On the other hand, unlike in field placements, OTL during methods courses were associated with teacher candidates' MKT. One way to interpret these results is that both university faculty and mentor teachers have distinct but equally valuable knowledge and skills that contribute in different ways to teacher candidates' broader development. These findings show empirically that, as Zeichner and colleagues (2015) theorized, critical expertise is evenly distributed across the teacher candidate learning system.

There may be elements of the context in which methods instructors and cooperating teachers operate that explain why methods courses and field placements attend to different aspects of teacher development. Methods courses comprise an activity system where the sole object is teacher candidate learning. Therefore they may afford candidates the time and space to acquire the granular pedagogical content knowledge and mathematics content knowledge that is assessed on the MKT. Cooperating teachers, on the other hand, are placed in the difficult situation of balancing multiple competing priorities: their students' learning, the teacher candidate's development, and the expectations of their administration (Zeichner, 2010). K-12 classroom teachers rightly need to design lessons with their students in mind, so the opportunities they afford candidates may be more haphazard than those designed by methods instructors. In addition, with so much happening in an elementary classroom, novices may have difficulty knowing what facets of practice to attend to when they are observing or attempting mathematics teaching. Methods courses quiet this noise, allowing candidates to deliberately focus on one facet of teaching at a time (Grossman et al., 2009; Lampert et al., 2013).

Rather than lament these tensions inherent to fieldwork, scholars of CHAT suggest that these "contradictions" precipitate development (Engeström 2001; Foot, 2014). Through this interpretation, it may be precisely *because* of the fast pace and competing demands inherent in field placements that these experiences are associated with candidates' beliefs that they are capable of doing the complex work of mathematics teaching. The tensions of fieldwork may come with benefits for candidates that more distilled experiences in mathematics methods courses do not. This is not to suggest that experiences in methods courses do not contribute to candidates' broader development; the results of this study clearly show their association to MKT and pedagogical beliefs. Instead, these results highlight the affordances of using a CHAT lens because it shifts implications from an either/or binary (e.g., field experiences as a source of contradictions *or* candidate growth), to a both/also view (Zeichner, 2010).

Instead of treating field and methods courses as distinct learning environments, TPPs can leverage the expertise of teacher educators in both spaces to improve the broader teacher learning system (Zeichner et al., 2015). For example, these data suggest cooperating teachers may be able to provide methods instructors valuable insights into the unique contexts and demands teacher candidates face in their placement. Similarly, methods instructors might be able to point cooperating teachers to key mathematics pedagogical content to focus on in their feedback to novices, and in this way experiences in field placements may attend more substantively to candidates' developing MKT. Importantly, however, this idea that TPPs should intentionally bring together and seek ways to amplify the voices of both methods instructors and cooperating teachers would not have emerged were the work of either examined in isolation. It is only through analyzing candidates' learning at the system level that opportunities for this type of productive exchange can be identified.

5.6 Limitations

There are several limitations to this analysis. First only approximately half of the 455 teacher candidates contacted completed both the MKT and the Teacher Candidate Survey. There are likely important differences between candidates who chose to respond and those that chose not to respond. Therefore, these findings cannot be generalized beyond the 220 teacher candidates in this sample.

Second, the data in this study all rely on teacher candidate self-reports, not observations of actual experiences. It possible that a candidates' exposure to various OTL about mathematics teaching during teacher preparation was different than what they reported. It is also possible that they misremembered or misreported their prior course taking or high school GPA. Further, grading scales and course rigor vary across districts so high school GPAs are themselves a limited measure of prior achievement. Despite these limitations, this analysis provides a foundation for multiple lines of future research.

Future researchers could build from this analysis to delve further into the quality of teacher candidates' experiences. Just as teaching quality has proved essential to understanding student learning in K-12 classrooms, it is likely that instructional quality in methods and content courses plays an equally important role in the context of teacher preparation. Additionally, in this analysis, teacher candidates' perceptions proved important to understanding the associations between their experiences during fieldwork and their preparedness. It is probable that their perceptions of their experiences during methods courses are equally as important. Data in this study were limited to reports of the quantity of particular experiences in methods courses; there was no data on candidate perceptions of the quality of those experiences. It is possible that there was substantial heterogeneity in candidates' experiences in methods courses that mutes the average effect presented here.

6. Conclusion

Taken together the findings from this path analysis, which account for both the reported quantity and the perceived quality of candidate experiences, may shed light on the conflicting findings in prior literature. There is little agreement in extant literature about the relationships between facets of teacher preparation and candidate knowledge and beliefs. It may be that the reason there are positive associations between the two in some samples and no associations in others has more to do with the quality of teacher candidate experiences than with whether or not a candidate was exposed to a particular opportunity to learn.

These data also emphasize the utility of using a “systems” lens, such as CHAT, to conceptualize teacher preparation when collecting and analyzing data. The associations found here could not have been uncovered were data on teacher candidate backgrounds, experiences, perceptions of those experiences, knowledge, and beliefs not collected and analyzed simultaneously. A systems approach also prompted the use of structural equation modeling. This methodology is rarely used in teacher preparation literature, yet it has the potential, as shown in the present study, to highlight the highly contextualized and interdependent nature of candidates’ experiences in teacher preparation. This is the first study to use multivariate path analysis as a means to intentionally highlight the variability inherent in teacher preparation. As such, it marks a turn toward a more expansive use of quantitative methodology, better suited to reflect at a large scale the complexity of preparing teachers that has been documented by qualitative researchers for decades.

References

Anagnostopoulos, D., Smith, E. R., & Basmadjian, K. (2007). Bridging the university-school divide. *Journal of Teacher Education*, *58*(2), 138-152.

Ball, D. L., & Bass, H. (2002). *Toward a practice-based theory of mathematical knowledge for teaching*. Retrieved from http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.132.7284&rep=rep1&type=pdf#page=13

Ball, D. L., & Forzani, F. M. (2009). The work of teaching and the challenge for teacher education. *Journal of Teacher Education*, *60*(5), 497-511.

Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teachers' mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching*, (4th ed.), (pp. 433-456). New York: Macmillan

Beilock, S. L., Gunderson, E. A., Ramirez, G., & Levine, S. C. (2010). Female teachers' math anxiety affects girls' math achievement. *Proceedings of the National Academy of Sciences*, *107*(5), 1860-1863.

Blömeke, S., Buchholtz, N., Suhl, U., & Kaiser, G. (2014). Resolving the chicken-or-egg causality dilemma: The longitudinal interplay of teacher knowledge and teacher beliefs. *Teaching and Teacher education*, *37*, 130-139.

Borko, H., & Mayfield, V. (1995). The roles of the cooperating teacher and university supervisor in learning to teach. *Teaching and Teacher Education*, *11*(5), 501-518.

Boyd, D. J., Grossman, P. L., Lankford, H., Loeb, S., & Wyckoff, J. (2009). Teacher preparation and student achievement. *Educational Evaluation and Policy Analysis*, *31*(4), 416-440.

Campbell, P. F., & Malkus, N. N. (2014). The mathematical knowledge and beliefs of elementary mathematics specialist-coaches. *ZDM*, *46*(2), 213-225.

Charalambous, C. Y. (2015). Working at the intersection of teacher knowledge, teacher beliefs, and teaching practice: A multiple-case study. *Journal of Mathematics Teacher Education*, *18*(5), 427-445.

Charalambous, C. Y., Panaoura, A., & Philippou, G. (2009). Using the history of mathematics to induce changes in pre-service teachers' beliefs and attitudes: Insights from evaluating a teacher education program. *Educational Studies in Mathematics*, *71*(2), 161.

Clark, L. M., DePiper, J. N., Frank, T. J., Nishio, M., Campbell, P. F., Smith, T. M., ... & Choi, Y. (2014). Teacher characteristics associated with mathematics teachers' beliefs and awareness of their students' mathematical dispositions. *Journal for Research in Mathematics Education*, *45*(2), 246-284.

Clift, R. T., & Brady, P. (2005). Research on methods courses and field experiences. In M. Cochran-Smith & K Zeichner (Eds.) *Studying teacher education: The report of the AERA panel on research and teacher education*, (pp. 309 – 424). New York: Routledge.

Douglas, A. (2010). What and how do student teachers learn from working in different social situations of development within the same school?. In V. Ellis, A. Edwards, & P. Smagorinsky (Eds.), *Cultural-historical perspectives on teacher education and development* (pp. 196-211). London, England: Routledge.

Drageset, O. G. (2010). The interplay between the beliefs and the knowledge of mathematics teachers. *Mathematics Teacher Education and Development*, *12*(1), 30-49.

Drake, C. (2006). Turning points: Using teachers' mathematics life stories to understand the implementation of mathematics education reform. *Journal of Mathematics Teacher Education*, *9*(6), 579-608.

Dweck, C. S., Chiu, C. Y., & Hong, Y. Y. (1995). Implicit theories and their role in judgments and reactions: A word from two perspectives. *Psychological Inquiry*, *6*(4), 267-285.

El'konin, B. D. (1993). The nature of human action. *Journal of Russian & East European Psychology*, *31*(3), 22-46.

Engeström, Y. (1993). Developmental studies of work as a testbench of activity theory: The case of primary care medical practice. In S. Chaiklin, & J. Lave (Eds.), *Understanding practice: Perspectives on activity and context* (pp. 64–103). New York: Cambridge University Press.

Engeström, Y. (2001). Expansive learning at work: Toward an activity theoretical reconceptualization. *Journal of Education and Work*, *14*(1), 133-156.

Ernest, P. (1989). The impact of beliefs on the teaching of mathematics. In P. Ernest (Ed.), *Mathematics teaching: The state of the art* (249 – 254). London: Falmer Press.

Evans, B. R. (2011). Content knowledge, attitudes, and self-Eefficacy in the mathematics new york city teaching fellows (NYCTF) program. *School Science and Mathematics*, *111*(5), 225-235.

Feiman-Nemser, S., & Buchmann, M. (1985). Pitfalls of experience in teacher preparation. *Teachers College Record*, *87*(1), 53-65.

Fennema, E., Franke, M. L., Carpenter, T. P., & Carey, D. A. (1993). Using children's mathematical knowledge in instruction. *American Educational Research Journal*, *30*(3), 555-583.

Foot, K. A. (2014). Cultural-historical activity theory: Exploring a theory to inform practice and research. *Journal of Human Behavior in the Social Environment*, *24*(3), 329-347.

Frykholm, J. A. (1998). Beyond supervision: Learning to teach mathematics in community. *Teaching and Teacher Education*, *14*(3), 305-322.

Geist, E. (2015). Math anxiety and the "math gap": How attitudes toward mathematics disadvantages students as early as preschool. *Education*, *135*(3), 328-336.

Ghousseini, H., & Herbst, P. (2016). Pedagogies of practice and opportunities to learn about classroom mathematics discussions. *Journal of Mathematics Teacher Education*, *19*(1), 79-103.

Goodman, J. (1984). Reflection and teacher education: A case study and theoretical analysis. *Interchange*, *15*(3), 9-26.

Gresham, G. (2007). A study of mathematics anxiety in pre-service teachers. *Early Childhood Education Journal*, *35*(2), 181-188.

Grossman, P., Hammerness, K., & McDonald, M. (2009). Redefining teaching, reimagining teacher education. *Teachers and Teaching: Theory and Practice*, *15*(2), 273-289.

Hammerness, K. (2006). From coherence in theory to coherence in practice. *Teachers*

College Record, *108*(7), 1241.

Hammerness, K., Darling-Hammond, L., Bransford, J., Berliner, D., Cochran-Smith, M., McDonald, M., & Zeichner, K. (2005). How teachers learn and develop. In L. Darling-Hammond & J. Bransford (Eds.), *Preparing teachers for a changing world* (pp. 358-389). San Francisco, CA: Jossey-Bass.

Harper, N. W., & Daane, C. J. (1998). Causes and reduction of math anxiety in preservice elementary teachers. *Action in Teacher Education*, *19*(4), 29-38.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*, (pp. 371-404). Charlotte, NC: Information Age Publishing.

Hill, H. C., Ball, D. L., & Schilling, S. G. (2008). Unpacking pedagogical content knowledge: Conceptualizing and measuring teachers' topic-specific knowledge of students. *Journal for Research in Mathematics Education*, 372-400.

Hill, H. C. (2010). The nature and predictors of elementary teachers' mathematical knowledge for teaching. *Journal for Research in Mathematics Education*, 513-545.

Hill, H. C., Blazar, D., & Lynch, K. (2015). Resources for teaching: Examining personal and institutional predictors of high-quality instruction. *AERA Open*, *1*(4). https://doi.org/10.1177/2332858415617703

Hill, H. C., Charalambous, C. Y., & Chin, M. J. (2018). Teacher Characteristics and Student Learning in Mathematics: A Comprehensive Assessment. *Educational Policy*, 0895904818755468

Hill, H. C., Umland, K., Litke, E., & Kapitula, L. R. (2012). Teacher quality and quality teaching: Examining the relationship of a teacher assessment to practice. *American Journal of Education*, *118*(4), 489-519.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55.

Jackson, K., & Cobb, P. (2010). Refining a vision of ambitious mathematics instruction to address issues of equity. In *annual meeting of the American Educational Research Association, Denver, CO*.

Jahreie, C., & Ottesen, E. (2010). Learning to become a teacher: Participation across spheres for learning. In V. Ellis, A. Edwards, & P. Smagorinsky (Eds.), *Cultural-historical perspectives on teacher education and development* (pp. 196-211). London, England: Routledge.

Jong, C., & Hodges, T. E. (2015). Assessing attitudes toward mathematics across teacher education contexts. *Journal of Mathematics Teacher Education*, *18*(5), 407-425.

Kazemi, E., Franke, M., & Lampert, M. (2009). Developing pedagogies in teacher education to support novice teachers' ability to enact ambitious instruction. In *Crossing divides: Proceedings of the 32nd annual conference of the Mathematics Education Research Group of Australasia* (Vol. 1, pp. 12-30). Adelaide, SA: MERGA.

Kazemi, E., & Wæge, K. (2015). Learning to Teach within Practice-Based Methods Courses. *Mathematics Teacher Education and Development*, *17*(2), 125-145.

Kennedy, M. (1999). The role of preservice teacher education. In L. Darling-Hammond

& G. Sykes (Eds.), *Teaching as the learning profession: Handbook of policy and practice* (pp. 54-85). San Francisco: Jossey-Bass.

Kleickmann, T., Richter, D., Kunter, M., Elsner, J., Besser, M., Krauss, S., & Baumert, J. (2013). Teachers' content knowledge and pedagogical content knowledge: The role of structural differences in teacher education. *Journal of Teacher Education*, *64*(1), 90-106.

Kline, R., (2005). *Principles and practice of structural equation modeling, third edition.* NYC: The Guildford Press

Kunter, M., Klusmann, U., Baumert, J., Richter, D., Voss, T., & Hachfeld, A. (2013). Professional competence of teachers: Effects on instructional quality and student development. *Journal of Educational Psychology*, *105*(3), 805.

Lampert, M. (2010). Learning teaching in, from, and for practice: What do we mean?. *Journal of teacher education*, *61*(1-2), 21-34.

Lampert, M. (2001). *Teaching problems and the problems of teaching.* New Haven, CT: Yale University Press.

Lampert, M., Franke, M. L., Kazemi, E., Ghousseini, H., Turrou, A. C., Beasley, H., ... & Crowe, K. (2013). Keeping it complex: Using rehearsals to support novice teacher learning of ambitious teaching. *Journal of Teacher Education*, *64*(3), 226-243.

Lampert, M., & Graziani, F. (2009). Instructional activities as a tool for teachers' and teacher educators' learning. *The Elementary School Journal*, *109*(5), 491-509.

Leder, G. C., & Forgasz, H. J. (2002). Measuring mathematical beliefs and their impact on the learning of mathematics: A new approach. In G. C. Leder, E. Pehkonen, & G. Törner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 95-113). Dordrecht: Springer.

Liljedahl, P., Durand-Guerrier, V., Winsløw, C., Bloch, I., Huckstep, P., Rowland, T., ... & Davis, Z. (2009). Components of mathematics teacher training. In E. Ruharma & D. L. Ball (Eds.), *The professional education and development of teachers of mathematics* (pp. 25-33). Boston, MA: Springer.

Lortie, D. C. (1975). *Schoolteacher: A sociological study.* Chicago: University of Chicago Press.

Ma, L. (2010). *Knowing and teaching elementary mathematics: Teachers' understanding of fundamental mathematics in China and the United States*. Abingdon, UK: Routledge.

Maas, C. J., & Hox, J. J. (2005). Sufficient sample sizes for multilevel modeling. *Methodology: European Journal of Research Methods for the Behavioral and Social Sciences*, *1*(3), 86.

Maasz, J., & Schlöglmann, W. (Eds.). (2009). *Beliefs and attitudes in mathematics education*. Rotterdam: Sense Publishers.

Matthews, M. E., & Seaman, W. I. (2007). The Effects of Different Undergraduate Mathematics Courses on the Content Knowledge and Attitude towards Mathematics of Preservice Elementary Teachers. *Issues in the Undergraduate Mathematics Preparation of School Teachers*, *1*.

McDonald, M., Kazemi, E., & Kavanagh, S. S. (2013). Core practices and pedagogies of teacher education: A call for a common language and collective activity. *Journal of Teacher Education*, *64*(5), 378-386.

Mewborn, D. S. (2000). Learning to teach elementary mathematics: Ecological elements of a field experience. *Journal of Mathematics Teacher Education*, *3*(1), 27-46.

Munter, C. (2014). Developing visions of high-quality mathematics instruction. *Journal for Research in Mathematics Education*, *45*(5), 584-635.

National Mathematics Advisory Panel. (2008). *Foundations for Success: Reports of the Task Groups and Subcommittees of the National Mathematics Advisory Panel*. Retrieved from https://www2.ed.gov/about/bdscomm/list/mathpanel/report/final-report.pdf

Norton-Meier, L., & Drake, C. (2010). When third space is more than the library: The complexities of theorizing ad learning to use family and community resources to teach elementary literacy and mathematics. In V. Ellis, A. Edwards, & P. Smagorinsky (Eds.), *Cultural-historical perspectives on teacher education and development* (pp. 196-211). London, England: Routledge.

Nurlu, Ö. (2015). Investigation of teachers' mathematics teaching self-efficacy. *International Electronic Journal of Elementary Education*, *8*(1), 21.

Pajares, M. F. (1992). Teachers' beliefs and educational research: Cleaning up a messy construct. *Review of Educational Research*, *62*(3), 307-332.

Philipp, R. A. (2007). Mathematics teachers' beliefs and affect. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning*, (pp. 257-315). Charlotte, NC: Information Age Publishing.

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

Qian, H., & Youngs, P. (2015). The effect of teacher education programs on future elementary mathematics teachers' knowledge: A five-country analysis using TEDS-M data. *Journal of Mathematics Teacher Education*, *19*(4), 371-396.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). CA: Sage.

Ren, L., & Smith, W. M. Teacher characteristics and contextual factors: Links to early primary teachers' mathematical beliefs and attitudes. *Journal of Mathematics Teacher Education*. https://doi.org/10.1007/s10857-017-9365-3

Ronfeldt, M. (2012). Where should student teachers learn to teach? Effects of field placement school characteristics on teacher retention and effectiveness. *Educational Evaluation and Policy Analysis*, *34*(1), 3-26.

Rogoff, B. (2008). Observing sociocultural activity on three planes: Participatory appropriation, guided participation, and apprenticeship. In P. J. Murphy & J. Soler (Eds.), *Pedagogy and practice: Culture and identities* (pp. 58-74). CA: SAGE Publications

Ross, J. A., McDougall, D., Hogaboam-Gray, A., & LeSage, A. (2003). A survey measuring elementary teachers' implementation of standards-based mathematics teaching. *Journal for Research in Mathematics Education*, 344-363.

Schmidt, W. H., Houang, R., & Cogan, L. S. (2011). Preparing future math teachers. *Science*, *332*(6035), 1266-1267.

Sleeter, C. E. (2001). Preparing teachers for culturally diverse schools: Research and the overwhelming presence of whiteness. *Journal of teacher education*, *52*(2), 94-106.

Smith, M. E., Swars, S. L., Smith, S. Z., Hart, L. C., & Haardörfer, R. (2012). Effects of an additional mathematics content course on elementary teachers' mathematical beliefs and knowledge for teaching. *Action in Teacher Education*, *34*(4), 336-348.

Solomon, Y., Eriksen, E., Smestad, B., Rodal, C., & Bjerke, A. H. (2017). Prospective teachers navigating intersecting communities of practice: early school placement. *Journal of Mathematics Teacher Education*, *20*(2), 141-158.

Stage, F. K., & Wells, R. S. (2014). Critical quantitative inquiry in context. *New Directions for Institutional Research*, *2013*(158), 1-7.

Staub, F. C., & Stern, E. (2002). The nature of teachers' pedagogical content beliefs matters for students' achievement gains: Quasi-experimental evidence from elementary mathematics. *Journal of educational psychology*, *94*(2), 344.

Swars, S. L., Daane, C. J., & Giesen, J. (2010). Mathematics anxiety and mathematics teacher efficacy: What is the relationship in elementary preservice teachers? *School Science and Mathematics, 106*(7), 306-315.

Swars, S. L., Smith, S. Z., Smith, M. E., & Hart, L. C. (2009). A longitudinal study of effects of a developmental teacher preparation program on elementary prospective teachers' mathematics beliefs. *Journal of Mathematics Teacher Education*, *12*(1), 47-66.

Tatto, M. T. (1996). Examining values and beliefs about teaching diverse students: Understanding the challenges for teacher education. *Educational evaluation and policy analysis*, *18*(2), 155-180.

Vacc, N. N., & Bright, G. W. (1999). Elementary preservice teachers' changing beliefs and instructional use of children's mathematical thinking. *Journal for Research in Mathematics Education*, *30*(1), 89-110.

Voss, T., Kleickmann, T., Kunter, M., & Hachfeld, A. (2013). Mathematics teachers' beliefs. In M. Kunter, J. Baumert, W. Blum, U. Klusmann, S. Krauss, & M. Neubrand (Eds.), *Cognitive activation in the mathematics classroom and professional competence of teachers* (pp. 249-271). Dordrecht: Springer.

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

Vygotsky, L. S. (1987). The problem and the method of investigation. In R. W. Rieber, & A. S. Carton (Eds.), *The collected works of L. S. Vygotsky: Volume 1 problems of general psychology* (pp. 43–51). New York: Plenum Press.

Wilkins, J. L. (2008). The relationship among elementary teachers' content knowledge, attitudes, beliefs, and practices. *Journal of Mathematics Teacher Education*, *11*(2), 139-164.

Wilson M., & Cooney T. (2002). Mathematics teacher change and developments. In G. C. Leder, E. Pehkonen, & G. Törner (Eds.), *Beliefs: A hidden variable in mathematics education?* (pp. 127-147). Dordrecht: Springer.

Yamagata-Lynch, L. C. (2010). *Activity systems analysis methods: Understanding complex learning environments*. New York: Springer Science & Business Media.

Zeichner, K. (2010). Rethinking the connections between campus courses and field experiences in college-and university-based teacher education. *Journal of Teacher Education*, *61*(1-2), 89-99.

Zeichner, K., Payne, K. A., & Brayko, K. (2015). Democratizing teacher education. *Journal of Teacher Education*, *66*(2), 122-135.

Understanding Instructional Quality Through a Relational Lens

Rebekah Berlin and Julie Cohen

Published in *ZDM: Mathematics Education*

DOI: 10.1007/s11858-018-0940-6

Author's Note

The authors wish to thank Dr. Bridget Hamre for her feedback on this manuscript.

1. Introduction

There is not consensus on the best way to conceptualize and measure high-quality mathematics teaching. This is evident in the recent proliferation of mathematics-specific classroom observation tools (English and Kirshner 2015). The expanding landscape of observation rubrics can make it difficult for researchers and practitioners to determine which tool best suits their purposes. One factor complicating this decision is that different protocols emphasize different dimensions of instructional quality in mathematics lessons (Kane and Staiger 2012).

Most observation protocols used in mathematics classrooms are made up of scales that measure context, content, and/or subject-specific content. Scales designed to measure behavior management, use of instructional time, and other features of the classroom environment which influence the extent to which students can access content-related learning opportunities provide information about the extent to which there is a *context* that supports learning (Bell et al. 2012, Danielson Group 2017; Pianta and Hamre 2009). Scales designed to measure the teaching of academic *content* capture practices that pertain to the teaching of mathematics content but are not mathematics-specific. Content-focused scales include teacher feedback practices, questioning, and/or connections to prior academic material. Contextual and content-focused practices are undoubtedly important in mathematics classrooms, but they are also important in teaching language arts, science, and social studies.

In contrast to scales that assess practices we might expect to observe across content areas, *subject-specific* scales are designed to measure things we expect *only* to happen during mathematics instruction. These practices include the mathematical substance of teacher explanations and multiple representations of mathematical content (Charalambous and Litke this issue; Walkowiak et al. this issue). They pertain to teaching mathematics and not other subjects.

Observation protocols designed to measure instruction using only measures of content and context are content-generic, meaning they can be used in any classroom, regardless of the subject being taught (see Figure 1). Other protocols have been designed exclusively for use during mathematics lessons. While mathematics-specific protocols can, in theory, include measures of context, content, and subject-specific content, most focus on practices related to the teaching of mathematics content. Some of these content-focused practices may be useful across subject areas, but many are exclusive to the teaching of mathematics.

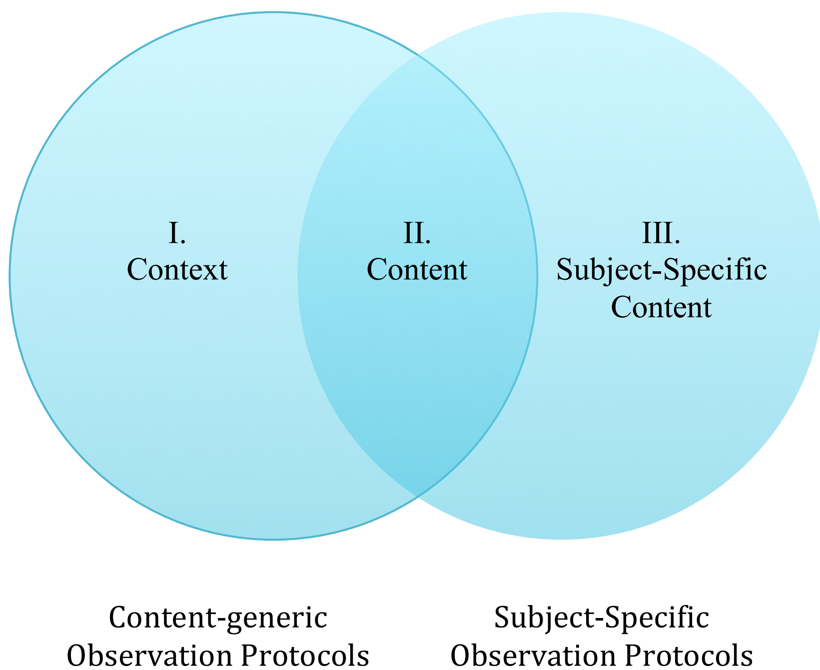


Figure 1. Different lenses for measuring mathematics instruction.

Recent work highlights the difference between mathematics-specific and content-generic protocols. Multiple measurement studies offer evidence that content-generic and mathematics-specific observational protocols capture distinct facets of instruction in mathematics classrooms (Blazar et al. 2017; McClellan et al. 2013; Walkowiak et al. 2014) and may require different types of rater expertise (Hill et al. 2012). Theoretical work has explored the impact of using different lenses. Hill and Grossman (2013) warn that general observation rubrics miss key subject-specific aspects of instruction and argue that when districts use content generic tools, teachers are deprived of feedback on important subject-specific practices. For example, mathematics teachers might receive information on how much time students spent engaged in academic work rather than on the mathematical depth of the task in which students were engaged.

In this paper, we argue that if districts or researchers focus solely on subject-specific aspects of mathematics instruction, they too will miss vital indicators of quality that may contribute to student learning of mathematics content. While it is critical to capture the nuances of mathematics teaching and learning using subject specific tools, there are also important aspects of classrooms obscured by such tools. In particular, relational aspects of quality instruction in mathematics classrooms have been shown to support student engagement in and learning of mathematics (Hamre and Pianta 2005; Kane and Staiger 2012; Mashburn et al. 2008; Walkowiak et al. 2014). Students may be better able to learn mathematics content if teachers foster warm classroom environments and effectively redirect off-task behavior. These kinds of practices are rarely featured in mathematics-specific observational measures.

To illustrate the importance of including scales designed to capture context and content in measures of instructional quality in mathematics classrooms we engaged in a close analysis of three upper elementary mathematics lessons. We analyzed each with a widely used subject-generic instrument, the Classroom Assessment Scoring System Upper Elementary (CLASS UE; Pianta et al. 2012). These data suggest there are compelling reasons to consider subject-generic practices in conceptualizations of high-quality mathematics instruction in mathematics classrooms.

Our approach in this paper is distinct from those in extant literature on the CLASS. The CLASS has been used to examine instructional quality in mathematics classrooms (Allen et al. 2011; Mashburn et al. 2008; Kane and Staiger 2012; Hamre and Pianta 2005; Walkowiak et al. 2014), but these studies highlight the explanatory power of the tool's domains and dimensions. Other authors have contrasted the CLASS with other frameworks to highlight areas of commonality and uniqueness (Blazar et al. 2017; McClellan et al. 2013, Walkowiak et al. 2014). To date, no studies have engaged in an in-depth analysis of what the CLASS alone reveals and obscures about instructional quality in mathematics classrooms. Ours is the first to treat the tool as the sole unit of inquiry.

2. Theoretical Underpinnings and Empirical Findings Related to CLASS UE

2.1 Tool Domains and their Theoretical and Empirical Foundations

The CLASS UE is based on developmental theory, which suggests that the interactions children have with adults and peers drive learning and social development (Bronfenbrenner and Morris 1998). A relational lens suggests that a child's behavior in the classroom cannot be understood outside of the relationship between child-level and classroom-level processes (Slavin et al. 2003). From this perspective, proximal classroom processes, or the relationship between micro (within child) and macro (environmental) level processes, not isolated events, are the primary driver of academic and emotional development (Ford and Lerner 1992). This relational lens underpins all parts of the CLASS UE in that the measure focuses exclusively on the frequency, depth, and duration of teacher-child and child-child interactions.

	Emotional Support	Classroom Organization	Instructional Support
Dimensions	Positive Climate Teacher Sensitivity Regard for Adolescent Perspectives	Behavior Management Productivity Negative Climate	Instructional Learning Formats Content Understanding Analysis and Inquiry Quality of Feedback Instructional Dialogue
	Student Engagement		

Figure 2: The four domains and 12 dimensions of UE and Secondary.[5]

The CLASS UE is divided into four domains: Emotional Support, Classroom Organization, Instructional Support, and Student Engagement (see Figure 2). It is important to note that while early analyses provided empirical support for this conceptualization (Bell et al. 2012; Hafen et al. 2015; Hamre et al. 2013), findings from more complex analyses suggest alternate structures of the CLASS dimensions (Hamre et al. 2014; Kane and Staiger 2012; McCaffrey et al. 2015). Although determining the best-fitting latent structure of the tool remains an open empirical question, to be consistent with user-facing scoring, training, and rating documents, we have organized the theoretical underpinnings according to the four-domain structure. The theoretical underpinnings for each domain were drawn from an extensive literature review. While briefly outlined below, they are discussed in greater detail in the CLASS UE Manual (Pianta et al. 2012). Existing validity arguments for the CLASS also include support for these conceptual domains (see Bell et al. 2012 for an outline of the target domain, empirical evidence of the appropriateness of the scoring rules and the tool as an adequate representation of teaching quality). In each domain, there are a subset of more specific classroom-level dimensions. Each dimension is scored on a 1 (low) to 7 (high) scale. Scores of 1 and 2 are

---

5 Readers interested in the indicators and behavioral markers nested under each dimension can contact Teachstone Training, LLC.

considered “low,” scores of 3, 4, and 5 are at the “middle” level, and scores of 6 and 7 are considered “high.” Raters score all 12 dimensions separately.

2.1.1 Emotional Support. The Emotional Support domain was drawn from research demonstrating that student success is fostered by feelings of relatedness to adults and classmates, opportunities for autonomy and choice in classroom activities, and interactions that promote a sense of competence (e.g., Allen et al. 1994; Allen et al. 2002; Ryan and Deci 2000). Literature documents the importance of teacher-student relationships for multiple student outcomes, including increased academic achievement, enhanced school motivation, and improved classroom behavior (Skinner et al.1998). In particular, relationships that are characterized by a balance of challenge and support seem to promote positive student outcomes (Eccles 2004; Sandilos et al. 2017).

The broader Emotional Support domain is comprised of three specific classroom dimensions that are scored as individual practices: Positive Climate, Teacher Sensitivity, and Regard for Student Perspectives. Positive Climate measures “the enjoyment and emotional connection that teachers have with students, as well as the nature of peer interactions” (Pianta et al. 2012; p. 2). Teacher Sensitivity assesses “the level of teachers’ responsiveness to the academic and social/emotional needs” of individual students (Ibid, p. 2). Regard for Student Perspectives foregrounds student choice in classroom decision-making. Within each dimension, raters are asked to score specific behavioral indicators that attend to fine-grained aspects of interactions. These include: Relationships, defined by specific behaviors such as physical proximity, peer interactions, shared positive affect, and social conversation; Positive Affect, defined by behavioral markers such as smiling, laughter, and enthusiasm; and Student Comfort, defined by behavioral indicators such as students take risks, participate freely, and seek support and guidance.

2.1.2 Classroom Organization. The Classroom Organization domain includes three dimensions: Behavior Management, Productivity, and Negative Climate. Compliant student behavior, efficient behavioral redirections, and minimal downtime and transitions characterize classrooms with strong Classroom Organization. These markers were drawn from theoretical work by developmental and ecological psychologists suggesting children develop divergent self-regulatory behaviors in different environments based on how adults manage time and behavior (Raver 2004; Kounin 1970). The authors also drew from constructivist theories on student engagement (Bowman and Stott 1994; Bruner 1996; Vygotsky 1978) as well as from empirical evidence that behavior and time management are associated with academic growth (Brophy and Evertson 1976; Good and Grouws 1977; Hoy and Weinstein 2006).

The Behavior Management dimension focuses on student behavior, the presence of specific proactive behavior management strategies, and the effectiveness and efficiency of behavioral redirections. The Productivity rubric assesses the degree to which learning time is utilized. Specifically this dimension focuses raters on classroom routines, teacher preparedness, and clarity of instructions. Negative Climate evaluates the levels of anger, hostility, and/or disrespect in a classroom as evidenced by teacher or student behaviors such as yelling, punitive consequences, or sarcasm. While Negative Climate was originally hypothesized to load onto the Emotional Support domain (e.g., Hamre et al. 2007; Hamre et al. 2014), more recent evidence drawn from samples with older students suggests it loads more strongly onto Classroom Organization (Hafen et al. 2015). The authors posit this may be because an increased Negative

Climate could cause or be the result of classroom disruptions captured under the Behavior Management dimension.

2.1.3 Instructional Support. Based upon research that suggests the ways in which teachers represent content to children may affect student learning, the Instructional Support domain focuses on the instructional strategies teachers use to support children's cognitive and linguistic development (Taylor et al. 2003). The dimensions under this domain draw from literature on the positive association between varied instructional modalities and student engagement (Yair 2000), the positive relationship between immediate, specific, contingent feedback and student outcomes (e.g., Butler 1987; Brophy 1981; Marzano et al. 2001), and the importance of higher-order thinking skills and metacognition (e.g., Bransford et al. 2000; Davidson and Sternberg 2003; Marzano et al. 2001). In addition, research suggests that specific pedagogical strategies are instrumental in supporting student learning. These include: breaking new material into small steps (Bransford et al. 2000), connecting new knowledge to prior knowledge and real world examples (Lee 2007; Tharp and Gallimore 1988; Levin and Pressley 1981), numerous examples and opportunities to practice (Rosenshine 1995), providing students with a strong base of factual knowledge and skills that build toward "big ideas" in the larger academic discipline (Bransford et al. 2000), and highlighting similarities and differences between examples (Marzano et al. 2001).

The Instructional Support Domain includes five dimensions. Instructional Learning Formats measures how teachers facilitate learning activities to maximize student engagement. Content Understanding assesses how teachers engage students in the key ideas in an academic discipline. Analysis and Inquiry focuses on the degree to which teachers promote higher-order thinking skills such as hypothesis testing and the application of knowledge and skills in a wide array of contexts. Quality of Feedback assesses whether teacher feedback pushes students to extend their understanding of concepts and skills. Finally, Instructional Dialogue foregrounds the ways teachers engage students in rich, academic questioning and discussion. Indicators nested within the above domains include "Learning Targets/Organization," which focuses raters on behaviors such as "clear learning targets", "previews", "reorientation/summary statements" (p. 63); "Opportunity for Practice of Procedures and Skills" which directs rater attention to "supervised practice" and "independent practice" (p. 71); and "Scaffolding" where the behavioral markers are "Assistance," "Hints," and "Prompting completion and thought processes" (p. 89).

2.1.4 Student Engagement. The final domain in the CLASS UE is Student Engagement. It assesses how actively students participate in classroom activities by analyzing whether children ask questions, volunteer ideas, look at the teacher, and focus on the academic task at hand. This domain was added to the tool because of a National Research Council report (2003) that highlighted the positive association between student engagement and student outcomes.

2.2 Prior Empirical Use

2.2.1 Prior Use Across Subjects. Substantial work documents the substantial associations between the types of interactions highlighted by the CLASS and key child outcomes in preK-12 settings. Across grade levels, teachers' instructional interactions have consistently predicted student academic and language outcomes, and emotional interactions have predicted the development of students' social skills (e.g., Allen et al. 2013; Mashburn et al. 2008; Parkarinen et al. 2010). Specifically, prior work has found that struggling or "high-risk" students perform similarly to their "low-risk" peers when they are placed in classrooms with high emotional and instructional support, but significantly worse than their peers when they are placed in less supportive classrooms (Hamre and Pianta 2005). Classrooms with improved teacher-student interactions are associated with increases in student achievement across subjects (Allen et al. 2011).

There are consistent classroom trends in studies using the CLASS across a range of contexts and diverse populations of students (Downer et al. 2012). Synthesizing evidence from multiple studies, Pianta and Hamre (2009) note many preschool and elementary school classrooms have high levels of emotional support, but low levels of instructional support. They also find that many students spend a large amount of time without the opportunity to engage in any learning activity: 42% of the time in preschool to 30% of the time in fifth grade (Ibid 2009). While these trends characterize classrooms in the United States, there is ongoing research looking at the use of the CLASS in international contexts (e.g., Hu et al. 2016; Levya et al. 2015).

Given the developmental lens of the CLASS, there are different versions of the tool for different age groups. While the Pre-K, K-3, Upper Elementary (UE), and Secondary tools share similar domains, the Infant and Toddler tools have different foci. The Infant tool is made up of a single domain, Responsive Caregiving, and the Toddler tool is made of two domains, Emotional and Behavioral Support and Engaged Support for Learning.

The version used in this paper, the CLASS UE, was used in the Measures of Effective Teaching Study (MET; Kane and Staiger 2012). According to the CLASS UE manual (Pianta et al. 2012), psychometric evidence from the MET show acceptable model fit for the three-factor model (RMSEA .11, CFI = .91; Acock 2013; Hair et al. 2010) and shows that each dimension loads strongly onto its associated domain (loadings range from .76 to .96). Domain-level Cronbach's alphas ranged from 0.87-0.92 indicating high internal consistency. Analysis of double coded videos demonstrates that raters were able to score an exact or adjacent score in 68% to 95% of the double coded videos, depending on the domain. Data from the MET study demonstrated a positive correlation (*r* = 0.25) between teachers' CLASS scores and value-added estimates of their effects on student achievement.[6]

2.2.2 CLASS as a Measure of Mathematics Instruction Extant work on the CLASS focuses on the practices as outcomes for interventions (e.g., Allen et al. 2011), as a measure of instructional quality used across subjects (e.g., Mashburn et al. 2008; Kane and Staiger 2012), or as a complement to subject specific tools (Hamre and Pianta 2005; Walkowiak et al. 2014). While CLASS has been used as the sole instrument to measure instructional quality in mathematics classrooms (e.g., Bell et al. 2012), these studies focus more on measurement issues and the general quality of interactions in the *context* of mathematics classrooms rather than squarely on the *mathematical* quality of instruction.

Ours is the first study to engage in a detailed qualitative analysis of a small number of lessons to illustrate what is highlighted and what is obscured when a subject generic lens like the CLASS is applied to mathematics classrooms. To concretize and extend theoretical work detailing the limitations of using content-generic tools, we engage in a close examination of three upper elementary mathematics lessons. We ask:

1. *What do ratings from the CLASS UE make visible about instructional quality in mathematics lessons?*

2. *What do ratings from the CLASS UE obscure about instructional quality in mathematics lessons?*

---

[6] Outside of the MET study, increases in CLASS scores have been shown to predict student achievement scores including a nine percentile term increase in student test scores (Allen et al. 2013) and a 0.16 standard deviation increase in student achievement Allen et al. 2012).

3. Methods

In the present analysis, we viewed three fourth-grade lessons from the National Center of Teaching Effectiveness video library. For more information on these lessons see Charalambous and Praetorius (this issue). We watched one lesson each from Mr. Smith's, Ms. Young's, and Ms. Jones' classrooms using the CLASS UE rubrics[7]. CLASS UE requires raters to collect evidence on a range of behavioral indicators and weigh the overall composition of evidence when scoring a particular domain of a classroom. According to CLASS UE protocol, we collected evidence under the three to five behavioral indicators nested in each dimension and aggregated these into a dimension level score at the end of the lesson. See Figure 3 for an example of a dimension face page, which provides an overview, but not the actual scoring guidance, for a dimension.

[7] All teacher's names are pseudonyms.

Content Understanding

Content Understanding refers to both the depth of lesson content and the approaches used to help students comprehend the framework, key ideas, and procedures[11] in an academic discipline. At a high level, this refers to interactions among the teacher and students that lead to an integrated understanding of facts, skills, concepts, and principles.

	Low (1,2)	Mid (3,4,5)	High (6,7)
Depth of understanding • Emphasis on meaningful relationships among facts, skills, and concepts • Real world connections • Multiple and varied perspectives	The focus of the class is primarily on presenting discrete pieces of topically related information; broad, organizing ideas are not presented.	The focus of the class is sometimes on meaningful discussion and explanation of broad, organizing ideas, while at other times, it is focused on discrete pieces of topically related information.	The focus of the class is on encouraging deep understanding of content through the provision of meaningful, interactive discussion and explanation of broad, organizing ideas.
Communication of concepts and procedures • Essential components identified • Conditions for how and when to use the concept and/or procedure • Multiple and varied examples • Contrasting non-examples	Class discussion and materials fail to effectively communicate the essential attributes of concepts/procedures to students.	Class discussion and materials communicate a few of the essential attributes of concepts/procedures but examples are limited in scope or not consistently provided.	Class discussion and materials consistently and effectively communicate the essential attributes of concepts/procedures to students.
Background knowledge and misconceptions • Attention to prior knowledge • Explicit integration of new information • Attention to misconceptions • Students share knowledge and make connections	There is little effort made to elicit or acknowledge students' background knowledge or misconceptions or to integrate previously learned material when presenting new information.	There are some attempts to elicit and/or acknowledge students' background knowledge or misconceptions or to integrate information with previously learned material, but these moments are limited in depth or provided inconsistently.	New concepts/procedures/ broad ideas are consistently linked to students' prior knowledge in ways that advance understanding and clarify misconceptions.
Transmission of content knowledge and procedures[12] • Clear and accurate definitions • Effective clarifications • Effective rephrasing	Content/procedural knowledge is inaccurate or not presented clearly.	Content/procedural knowledge is sometimes effectively and accurately communicated to students; at other times, information is confusing and/or inaccurate.	Content/procedural knowledge is effectively and accurately communicated to students.
Opportunity for practice of procedures and skills[13] • Supervised practice • Independent practice	Students simply receive information about procedures and skills and do not have opportunities to practice procedures or skills relevant to the content area of the lesson.	The teacher occasionally incorporates opportunities for supervised or independent practice of procedures and skills relevant to the content area of the lesson.	The teacher regularly incorporates opportunities for supervised or independent practice of procedures and skills relevant to the content area of the lesson.

Figure 3: Content Understanding Dimension Face Page.

Each observer was trained and certified as a reliable CLASS rater. The CLASS UE manual specifies that video observations should be rated in 15-20 minute cycles. Therefore, we divided each of the three videos into segments of equal length. For video one (total time 38 minutes), we rated two segments; for video two (total time 68 minutes), we rated four segments; for video three (total time 56 minutes), we rated three segments. While watching each video, each rater took notes into the CLASS UE Score Sheet, categorizing observations into the 12 dimensions under their associated behavioral indicators in real time. Following the end of each segment, raters paused the video and immediately rated the cycles. Segments were rated within a 10-minute window on each of the 12 dimensions. Finally, after the last rating cycle for each video, we composited each score by averaging scores across cycles to arrive at a single score for each dimension for the observation period. Dimension scores were averaged to provide domain level scores after reverse coding Negative Climate. Finally, after each video, we created analytic memos detailing what was highlighted and obscured in using the CLASS to rate upper elementary mathematics instruction.

It is important to note that because of the number of cycles we observed, neither we, nor our readers, can make generalizations about individual teacher effectiveness. Due to the instability of ratings of single lessons, the manual explicitly states if the CLASS is being used to measure teacher quality, it must be through "multiple lessons, and ideally […] across multiple class sections" (Pianta et al. 2012; p. 8). Therefore, the results and discussion below are merely meant to ground our discussion of the tool in concrete examples and to provide readers a snapshot of the types of classroom evidence captured with the CLASS as compared to other observational measures.

4. Results

The three lessons varied in terms of the quality of instruction, as measured by the CLASS (see Table 1 for aggregated dimension and domain level scores).

Table 1: Average Dimension and Domain Scores

		Mr. Smith	Ms. Young	Ms. Jones
Emotional Support	Positive climate	3	3	6
	Teacher sensitivity	3	4	6
	Regard for student perspectives	2	4	4
	Domain Average	**3**	**3**	**5**
Classroom Organization	Behavior management	7	4	7
	Productivity	6	6	7
	Negative climate	1	2	1
	Domain Average	**6**	**5**	**7**
Instructional Support	Instructional learning formats	5	6	6
	Content understanding	3	6	5
	Analysis and inquiry	1	4	2
	Quality of feedback	2	5	4
	Instructional dialogue	1	5	3
	Domain Average	**2**	**5**	**4**
	Student engagement	3	4	6

Note: Negative Climate was reverse coded before being averaged. All scores have been averaged across segments and are rounded to the nearest whole number for ease of interpretation. For some, such as Ms. Young on Emotional Support, the domain average of the averaged rounded dimension scores (3, 4, 4) is not equivalent to the domain average of the unrounded dimension level scores averaged across segments (2.75, 3.5, 3.75). The rounded numbers are presented only for ease of interpretation.

All three scored highest on the Classroom Organization domain, and two of the three lessons scored the lowest on the Instructional support domain. Ms. Young's instruction scored at the mid-level across the four domains. Ms. Jones' instruction was consistently at the mid and high level. Mr. Smith had the most varied portrait of instruction, with domain levels scores ranging from low (instructional support) to high (classroom organization).

4.1 Mr. Smith.

Averaged across dimensions, across segments, and rounded to the nearest whole number, Mr. Smith's classroom received a score of 3 for Emotional Support, 6 for Classroom Organization, 2 for Instructional Support, and 3 for Student Engagement. The classroom's Emotional Support score of 3 places it in the lower end of the mid range. This score reflects that there was occasional, but inconsistent evidence of emotional support throughout the video. For example, despite a few instances of shared positive affect, such as a joke about acute angles, both Mr. Smith's and his students' affects were flat for the majority of the video. Mr. Smith occasionally connected material to common terms in students' life such as when he related acute angles to being "cute and tiny," and obtuse angles to being "obese." Though Mr. Smith sporadically appeared to scan the classroom, he spent the majority of the lesson pacing the front of the classroom and never noticed a student's raised hand or students whispering, "What are we supposed to do?" to one another. The lesson was tightly teacher controlled, and he did not provide students with authentic choices, opportunities for meaningful peer interactions, or opportunities for leadership and responsibility.

The classroom's aggregated score was 6 for Classroom Organization because little instructional time was lost due to student behavior. There were occasional instances where productivity of the classroom slowed because Mr. Smith was writing out a problem by hand or distributing materials inefficiently. There was only one instance of Negative Climate, when students laughed at another student at the board.

The classroom scored a 2 for instructional support. There was evidence of clear learning targets and multiple modalities for instruction, for example the lesson included both auditory, through the form of Mr. Smith's lecture, and kinetic, such as when students had the opportunity to circle the correct type of angle at the Smart Board, ways to engage with the lesson material.

However, there was little evidence of depth, higher-order thinking, quality feedback, instructional dialogue, or opportunities for students to independently engage with the lesson material. Most tasks were rote in nature. For example, students were asked to come to the Smart Board and use the protractor tool to open an angle to the number of degrees Mr. Smith provided or to come to the Smart Board and choose whether an angle was acute, right, or obtuse.

Finally, Student Engagement was rated as 3. There was a group of students off task for the majority of the video, whispering and laughing amongst themselves. Several students appeared compliant and on task, however, they did not seem actively engaged. Students yawned throughout the lesson and did not demonstrate active listening behaviors.

4.2 Ms. Young

Ms. Young's classroom scored a 3 for Emotional Support, 5 for Classroom Organization, 5 for Instructional Support, and 4 for Student Engagement. Though there was little evidence of teacher warmth or shared positive affect throughout the video, students demonstrated comfort with Ms. Young, approaching her to ask questions, show their work, and suggest alternate solution strategies. Ms. Young demonstrated mixed awareness of and responsiveness to students' academic needs. She circulated throughout the room and checked in with almost every student individually about their academic progress during small group work. She provided supportive feedback to some students but chastised others for not working and did not offer them instructional support. At times, she demonstrated Regard for Student Perspectives such as when she anchored abstract mathematics problems in scenarios students could relate to (equal groups became "apples in boxes"), and allowed students to work in groups and choose their own materials to solve mathematics problems. At other times, she restricted student autonomy by telling students they were not allowed to get their own materials and not to argue with her about certain solution strategies.

Ms. Young's classroom scored a 5 for Classroom Organization because while there were clear and consistently enforced expectations for student behavior when students were on the carpet, instructional time was lost to student behavior during small group work and to a long transition from desks to carpet. There were also repeated instances of Negative Climate throughout the video. Ms. Young made comments CLASS classifies as sarcastic and derogatory such as, "Thank you for disrupting the lesson throughout the day" and "You don't have the worksheet. People are asked to do it in their journal, and they're doing it in their journal. And you're sitting down there sucking your finger." There were also several instances of mild irritability and a few of punitive control such as when Ms. Young threatened to send various students away from the group or out of the room. She eventually sent them into the hallway.

Of the three classrooms reviewed, Ms. Young's classroom scored highest for Instructional Support. Ms. Young outlined clear learning targets and the lesson was aligned to these goals. She actively facilitated student learning through a variety of modalities, strategies, and materials. Students were allowed to choose between proving the relationship between the factors in two multiplication problems through a variety of materials including graph paper, cubes, and diagrams. Lesson activities consistently focused students on independently discovering meaningful relationships between concepts and procedures, such as those between representations of multiplication and between factors. Ms. Young provided open-ended tasks and consistently pushed students to explain their cognitive processes and approaches by stating that knowing the answer to a problem was not enough, and that each student should be "justifying that your answer is true." Students received extensive practice time.

Scores on the Instructional Support domain indicated that, despite these strengths, there was substantial evidence of student confusion throughout small group time. Rather than providing encouragement, affirmation, or support for struggling students, Ms. Young often chastised students for their incorrect responses and pace. Though the tasks she presented were open-ended, often her dialogue with students limited engagement with the task so that students may have experienced tasks as close-ended. For example, there were multiple occasions where she explicitly told students the steps to complete in order to create the visual she wanted them to share on the carpet. Student Engagement was mixed throughout the video resulting in an aggregate mid-range score. Most students appeared actively engaged during the opening and closing that took place on the carpet, but many students appeared distracted and disengaged during the group work at their desks.

4.3 Ms. Jones

Ms. Jones's classroom scored a 5 for Emotional Support, a 7 for Classroom Organization, a 4 for Instructional Support, and a 6 for Classroom Engagement. Under the Emotional Support domain there was consistent evidence of relationships, positive communication, respectful language, and student comfort throughout the lesson. Ms. Jones displayed sensitivity circulating around the room, anticipating and circumventing problems with sharing materials, group work, and lesson content. For example, when there were not enough scissors and rulers for students to use, Ms. Jones explained the system tables would use to share them to ensure every member of the group got equal access. There was little evidence, however, of authentic student autonomy or leadership, and no evidence of meaningful peer interactions, until the end of the lesson when students worked in groups cutting apart circles to represent multiplication as equal groups of fractional parts.

This classroom scored highest for Classroom Organization of the three lessons, because there was no evidence of negativity and the classroom was highly productive. Ms. Jones used behavior management strategies such as positive behavior narration, hand signals, and quick redirections. No time was lost to student misbehavior.

Ms. Jones's classroom scored a 4 for Instructional Support. The lesson had several strengths in this domain. In every segment, the lesson was aligned to the learning targets, and lesson material was presented through a variety of engaging materials. For example, students represented three ways to multiply fractions by a whole number on a three panel foldable. One of the methods involved using construction paper circles, cutting them into equal groups, and using repeated addition to find the total. Ms. Jones clearly presented lesson content, breaking down strategies for multiplying fractions into crisply delineated steps. She built on student background knowledge by connecting multiplying fractions to students' knowledge of repeated addition. She first had students represent 2x2 as 2+2, 2x3 as 2+2+2, before they represented 5 x ¾ as ¾ + ¾ + ¾ + ¾ + ¾. She also explicitly reviewed a strategy students had already learned to multiply a fraction by a whole number, before exposing them to new strategies. Additionally Ms. Jones anticipated student misunderstandings by asking questions like "can I just put R3?" so that students had to explain to her why she needed to write a remainder as a fractional part.

Despite these strengths, there was limited evidence of higher-order thinking or quality teacher-student and student-student dialogue for the first two-thirds of the lesson. Talk was heavily teacher-directed. Sometimes she engaged in substantive feedback loops with students or provided scaffolds to those who struggled, such as when she prompted a student 1 x 4 is what, now 7 x 4 is what? At other times, however, her feedback was perfunctory; she often simply exclaimed, "Good!" and at other times she ignored incorrect responses. Most students appeared actively engaged throughout the lesson. Students were manipulating materials, asking and answering questions, and sharing ideas with the teacher. This placed the classroom in the upper range of Classroom Engagement.

5. Discussion

As is clear in the interactions described above, there are certain aspects of mathematics instruction that are foregrounded or marginalized when lessons are scored with the CLASS. Below, we argue that certain foci of the CLASS, which are missing from many mathematics-specific tools, offer essential information to those trying to understand instructional quality in mathematics. We also detail aspects of instruction in mathematics that are not captured by the CLASS.

5.1 Aspects of Instruction Highlighted by the CLASS

5.1.1 Facets of mathematics instruction. CLASS highlights aspects of high-quality teaching of academic content under the Instructional Support domain. For example, Content Understanding and Analysis and Inquiry focus raters in mathematics classrooms on the ways content is represented and students are able to engage with academic content. Importantly, while these aspects of instruction are relevant in mathematics classrooms, these practices are not *unique* to the teaching of mathematics.

Evidence from scales that measure the nature of instructional activities is illustrative of the way the CLASS is able to highlight meaningful differences in mathematics instruction, while only capturing practices that can be used across content areas. Within the CLASS framework, higher scoring instruction contains open-ended tasks allowing students to explore relationships between ideas. One of the reasons Mr. Smith received a low score on the Instructional Support domain is because his lesson relied on discrete questions with a single correct answer (e.g., “What type of angle is this?”). Ms. Jones, on the other hand, scored in the midrange because she posed a mix of open and close-ended tasks. Like Mr. Smith, she asked students several close-ended questions. However, in the third segment, she gave students several minutes to complete a task that allowed for student choice. She first asked students to generate equations where a fraction with a denominator of four was multiplied by a whole number. Because not all students chose the same equation to model, there were multiple opportunities to discuss how to represent different products as both “improper fractions” and mixed numbers. Ms. Jones also capitalized on different student equations to explore how fractional pieces can be grouped to show whole numbers (eight fourths as two wholes).

A teacher can provide student choice and open-ended tasks in mathematics, language arts, science, or social studies classrooms; these practices are not limited to mathematics instruction. Nevertheless, a focus on general content practices reveals important features of mathematics instruction. While a mathematics-specific tool may have provided different insights about the mathematical quality of the instructional explanations Ms. Jones provided, the CLASS still captured important variation (e.g., a two point difference in Instructional Support) in the types of mathematical reasoning and representations students were exposed to across the two classrooms.

5.1.2 Interaction between content and context. CLASS highlights the interaction between the content students are exposed to and the context in which that exposure occurs. In the CLASS framework, content is captured primarily through the Instructional Support domain. Different facets of context are measured through the Emotional Support, Classroom Organization, and Student Engagement domains. Many mathematics-focused tools do not have indicators to assess contextual factors such as student engagement or the emotional tenor of classroom interactions (Walkowiak et al. 2014) that influence the extent to which students can access these learning opportunities.

Ms. Young's classroom is particularly illustrative of the importance of capturing the relationship between content and context when assessing mathematics instruction. Of the three lessons analyzed, Ms. Young presented students with the greatest opportunity to engage with deep, rigorous mathematical tasks. There was evidence of high quality discourse about mathematical relationships, including those between 30 x 4 and 15 x 8, and broad organizing ideas such as why, when multiplying, doubling a factor doubles the product. These are reflected in a high score on the Content Understanding dimension. While, as documented in the results section, there was room for improvement in the consistency of the academic supports she provided students, analysis focused on content reveals a promising portrait of mathematics instruction.

Content without context, however, does not paint a full portrait of the interactions in her classroom. Students did not consistently take the opportunities Ms. Young provided. Several students used group work time to socialize, throw manipulatives at one another, or build patterned towers of cubes, ignoring Ms. Young's redirections. This was reflected in lowered Behavior Management and Student Engagement scores because for the average student in the classroom, a large segment of the lesson was not spent on mathematics. Similarly, a chaotic transition from students' desks to the adjacent carpet resulted in lost instructional time and lowered the classroom's Productivity score during that segment. Put simply, the quality of the mathematical tasks Ms. Young presented may have mattered little because many students did not fully engage with them.

Along the same lines, there were multiple instances captured under the Negative Climate dimension where Ms. Young limited children's ability to engage with content. While she engaged in extended mathematical discourse with some students and asked them to share their work with the class, when other students provided incomplete or incorrect reasoning, she responded to them by saying, "No," "Don't argue," "You cannot be a part of the discussion," and "Go sit down." She sent some students out of the classroom or to the back of the classroom where she largely ignored them. In one of her only interactions with this group of students, she reminded one student that the reason he was struggling in this class was because he "refused to complete" his work on Monday. She did not offer to assist him and told him that he only had five minutes to complete it. The unequal distribution of materials, teacher time, instructional support, and warm interactions in this classroom, may have lead students to believe that mathematics is a discipline for a chosen few, not for all students in the classroom.

In classrooms like Ms. Young's, there are marked implications of excluding contextual practices that are shared across content areas from measurement of mathematics teaching. The absence of data on contextual factors may skew the conclusions researchers and practitioners draw from content-focused data. For example, were a school administrator to review only Ms. Young's scores under the Instructional Support domain, they might assume her development should focus on improving the way she responds to students' mathematical misunderstandings and errors. Using ratings from the full spectrum of CLASS dimensions however, this administrator might instead choose to focus on how to increase Ms. Young's ability to reduce the off task behavior in her classroom or how to build positive relationships with struggling students. Similarly, in research settings, classrooms like Ms. Young's may cloud the relationship between mathematics-specific teaching practices and student learning if researchers do not consider contextual factors in their measurement of mathematics teaching. While arguably a mathematics-specific tool would have picked up additional information on the content Ms. Young presented, this does not alter the fact that contextual factors in her classroom are likely impacting students' mathematical learning. Only a protocol that includes subject-generic practices such as those in the CLASS can provide this information.

5.2 Aspects of instruction obscured by the CLASS

5.2.1 Mathematics-specific content and teaching practices. As Hill and Grossman (2013) conjectured, the general lens of CLASS obscures nuances of mathematical instruction. More broadly, ratings from the CLASS do not indicate that mathematics was taught at all. Because of this, lesson segments can receive high scores in the Instructional Support domain, regardless of the presence or quality of the mathematics in the segment, if other general pedagogical practices are observed. For example, in Ms. Jones's video, the first nine minutes did not contain any mathematics; students were constructing a foldable they were going to use throughout the lesson. She provided detailed explicit instruction about how to fold the construction paper, created a visual on the board to illustrate where she wanted students to write their name and what they should title it, and modeled the procedure with student materials. While all these constitute high quality general practices captured in the Instructional Learning Format dimension, they do not relate to mathematics. This example suggests that some scores on dimensions within the Instructional Support domain could be "inflated" by explicit instruction on myriad non-mathematical topics. This could potentially mislead users of the CLASS about the quality of mathematics instruction in a classroom.

Similarly, the Quality of Feedback and Instructional Dialogue dimensions capture general practices of classroom discourse, regardless of their mathematical substance. Thus, CLASS may classify comments of differential mathematical significance similarly. For example, one criterion of mid-range evidence of the "facilitation strategies" indicator in the Instructional Dialogue dimension is that "the teacher and/or fellow students sometimes acknowledge students' comments and repeat or extend these in ways that affirm their observations and/or recast the information in a more complex form" (Pianta et al. 2012; p. 99). Therefore, Mr. Smith's pattern of repeating student responses and adding an affirmative comment such as "Less than a right [angle]. Okay!" was counted as evidence of an equal weight to a more mathematically substantive comment from Ms. Young. When a student struggled to articulate the way he had transformed his array, Ms. Young stated, "[after cutting the original array in half] so you know you have two rectangles, and you move one of the rectangle down here to create a longer rectangle with one longer dimension and a short dimension. So now you have – this one has doubled and this side has been reduced." Ms. Young's comment used precise mathematical language to affirm a student and rephrase their resonse in academic language. Mr. Smith's "Okay!" while also affirming, did not add depth or mathematical richness to his student's understanding of angels. Ms. Young ultimately had a greater frequency of dialogue, which resulted in her having an overall higher score, however, at the evidence level, these particular interactions were viewed identically through the lens of CLASS.

Relatedly, CLASS does not focus on precise mathematical language. Thus, statements like Ms. Jones's "four over four" instead of "four-fourths," or "I want you to have an equal sign and your final result" instead of "I want you to show your two fractions are equivalent" were not considered as evidence. It is likely that were this same lesson observed with a mathematics-*specific* lens such as the Mathematical Quality of Instruction (MQI) tool, these differences in mathematical discourse across the three lessons would be captured under the "Mathematical

Language" and "Imprecision in Language and Notation" codes (see Charalambous and Litke, this issue). In summary, precisely as Hill and Grossman (2013) predicted, there are some aspects of high quality mathematics instruction the CLASS will not provide users information about.

5.2.2 Teaching mathematical concepts and procedures. Importantly, CLASS does not take a pedagogical stance on mathematics instruction. That is, neither procedural nor conceptual mathematics instruction is privileged. As such, the CLASS UE obscures distinctions between teaching focused on mathematical procedures and teaching focused on mathematical concepts.

Ms. Jones's classroom was characterized by exchanges focused on executing mathematical procedures, such as the one below:

> *Ms. Jones*: Very good. So I take 15 and I put inside. It becomes my dividend. And 4 becomes – what is that word that we use for the number that's outside the box? Raise your hand. What is that word that we use, Student R?
>
> *Student:* The divisor.
>
> *Ms. Jones:* Divisor. So 15 becomes my dividend, and 4 becomes my divisor, and I divide it out. Does 4 go into 1?
>
> *Multiple students:* No.
>
> *Ms. Jones:* No. So I put a zero. How many times does 4 go into 15?

Ms. Jones focuses only on the name and order of components of the process for long division. She does not explain why she is taking any of the above steps.

In contrast to the procedural exchanges highlighted in Ms. Jones's lesson, there were frequent interactions focused on mathematical concepts in Ms. Young's classroom. For example, she and a student explored why 16x6 is equivalent to 16x3+16x3:

Ms. Young: So Student C is saying that 48 plus 48 will give us 96, and that will be the same thing as 16 times 6 is 96. Yes, do you have another way of explaining it, Student C? I saw your hand up.

Student: You can instead drawing *[inaudible]*, you can just draw six boxes.

Ms. Young: We can draw 6 boxes showing the 3 and the 3. So if you combine all of the boxes together, 1, 2, 3, 4, 5 – so that's 16, 16, 16, 16, 16, 16.

Student: And then you could just cut the middle off the one.

Ms. Young: And they say like I cut the middle of this one [separates three of the boxes from the remaining 3], and that would give me my three group of 16 and three group of 16.

Though both of these exchanges focus on operations, they differ considerably in mathematical substance. Ms. Jones's focuses on the steps for dividing a two-digit number, and Ms. Young's focuses on connecting a semi-concrete representation of multiplication to an abstract numerical one. CLASS is ambivalent to this difference. These interactions both count as mid-range evidence for the "communication of concepts and procedures" indicator under the Content Understanding dimension because in both exchanges the "teacher demonstrates sufficient knowledge of the material to support student learning at a level that meets the goals of the lesson" (Pianta et al. 2012; p. 74).

Both of these interactions would also count as high-level evidence of "building on student responses" indicator under the Quality of Feedback dimension. Indeed both teachers expand "students' initial responses or action in ways that provide additional information or clarification" (Ibid p. 92). Based on similar patterns across the lessons, Ms. Young's conceptual and Ms.

Jones's procedurally oriented lessons scored within one point of each other on the Instructional Support domain, though they diverged substantially in their approach to teaching mathematics.

A mathematics specific tool such as the Mathematics Scan (M-Scan), explicitly attends to these differences in language under "Depth" in its "Explanations and Justifications" dimension. As Hill and Grossman (2013) suggest, the differences between general and mathematics-specific tools have implications for providing teachers feedback. Coaches and administrators seeking to understand the volume of instructional time focused on mathematical procedures versus mathematical concepts could not gain this information from the CLASS UE.

6. Conclusion

These data suggest observation protocols that can be used across subjects, such as the CLASS, capture some, but not all, facets of instructional quality in mathematics classrooms. For example, our analysis of Ms. Jones' classroom corroborated Hill and Grossman's (2013) conjecture that high ratings on subject-generic dimensions such as Positive Climate or Productivity do not necessarily also indicate quality mathematics instruction. Rather, they provide a context in which quality mathematical engagement is possible.

What is also clear from our analysis is that subject-generic and mathematics-specific teaching practices interact in meaningful ways. Ms. Young's lesson demonstrated that even when high-quality mathematical opportunities are available, they may be of limited impact if students do not engage with them. While multiple indicators of quality mathematics-specific instruction including mathematical discourse, meaningful mathematical choices, and student-generated mathematical justifications, were present in her classroom, student behavior reduced the extent to which these occurrences were likely to impact student learning. Because the CLASS attends to *both* the content and contextual practices, users obtain a holistic understanding of classroom practices that likely impact student experiences.

These data suggest a strong rationale for including subject-generic practices in conceptualizations of high-quality instruction in mathematics classrooms. Confining the measurement of mathematics instruction only to practices that are unique to mathematics may push out important features of classrooms in which mathematics instruction occurs. When contextual factors such as whether a classroom is a safe, productive, and engaging place are not considered, users of observation tools risk misinterpreting the relationship between mathematics-specific practices and student learning. Of course, working from a completely content-generic perspective means that while observers will assess content instruction in mathematics classrooms, it will be with broader brush strokes than a mathematics-specific tool. Therefore, there are limitations of the exclusive use of *both* subject-generic and mathematics-specific tools. These data suggest that conceptions of high-quality instruction in mathematics classrooms likely need to include both subject-specific and content-generic practices.

References

Acock, A., (2013) *Discovering structural equation modeling using stata*. College Station, Texas: Stata Press.

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher-student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system-secondary. *School Psychology Review*, 42(1), 76.

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

Allen, J. P., Hauser, S. T., Bell, K. L., & O'Connor, T. G. (1994). Longitudinal assessment of autonomy and relatedness in adolescent-family interactions as predictors of adolescent ego development and self-esteem. *Child development*, 179-194.

Allen, J. , Marsh, P. , McFarland, C., McElhaney, K., Land, D., & Jodi, K. (2002). Attachment and autonomy as predictors of the development of social skills and delinquency during midadolescence. *Journal of Consulting and Clinical Psychology*, *70*(1), 56-66.

Bell, C. A., Gitomer, D. H., McCaffrey, D. F., Hamre, B. K., Pianta, R. C., & Qi, Y. (2012). An argument approach to observation protocol validity. *Educational Assessment*, *17*(2-3), 62-87.

Blazar, D., Braslow, D., Charalambous, C. Y., & Hill, H. C. (2017). Attending to general and mathematics-specific dimensions of teaching: Exploring factors across two observation instruments. *Educational Assessment*, *22*(2), 71-94.

Bowman, B., & Stott, F. (1994). Understanding development in a cultural context: The challenge for teachers. In B. L. Mallory & New, R. S. (Eds.), *Diversity and developmentally appropriate practices: Challenges for early childhood education* (pp. 19-34) New York, NY: Teachers College Press.

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

Bronfenbrenner, U., & Morris, P. A. (1998). The ecology of developmental processes. *Handbook of child psychology: Theoretical models of human development* (pp. 993-1028). Hoboken, NJ, US: John Wiley & Sons Inc.

Brophy, J. E., & Evertson, C. M. (1976). *Learning from teaching: A developmental perspective*. Allyn and Bacon.

Brophy, J. (1981). Teacher praise: A functional analysis. *Review of Educational Research*, *51*(1), 5-32.

Bruner, J. S. (1996). *The culture of education*. Cambridge, MA: Harvard University Press.

Butler, R. (1987). Task-involving and ego-involving properties of evaluation: Effects of different feedback conditions on motivational perceptions, interest, and performance. *Journal of Educational Psychology*, *79*(4), 474.

Charalambous, C., & Litke, E. (this issue). Studying instructional quality by using a content-specific lens: The case of the mathematical quality of instruction framework. *ZDM Mathematics Education.*

Danielson Group. (2017). *The framework.* http://www.danielsongroup.org/framework/

Davidson, J. E., & Sternberg, R. J. (2003). *The psychology of problem solving.*

Cambridge, UK: Cambridge University Press.

Downer, J. T., López, M. L., Grimm, K. J., Hamagami, A., Pianta, R. C., & Howes, C. (2012). Observations of teacher–child interactions in classrooms serving Latinos and dual language learners: Applicability of the Classroom Assessment Scoring System in diverse settings. *Early Childhood Research Quarterly*, *27*(1), 21-32.

Eccles, J. S. (2004). Schools, academic motivation, and stage-environment fit. In R. M. Lerner & L. Steinberg (Eds.), *Handbook of adolescent psychology*, (pp. 125-153). Hoboken, NJ: John Wiley

English, L. D., & Kirshner, D. (Eds.). (2015). *Handbook of international research in mathematics education*. New York, New York: Routledge.

Ford, D. H., & Lerner, R. M. (1992). *Developmental systems theory: An integrative approach.* Sage Publications, Inc.

Good, T. L., & Grouws, D. A. (1977). Teaching effects: A process-product study in fourth-grade mathematics classrooms. *Journal of Teacher Education*, *28*(3), 49-54.

Hafen, C. A., Hamre, B. K., Allen, J. P., Bell, C. A., Gitomer, D. H., & Pianta, R. C. (2015). Teaching through interactions in secondary school classrooms: Revisiting the factor structure and practical application of the Classroom Assessment Scoring System–Secondary. *The Journal of Early Adolescence*, *35*(5-6), 651-680.

Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., & Tatham, R. L. (1998). *Multivariate data analysis*. Upper Saddle River, NJ: Prentice hall.

Hamre, B., Hatfield, B., Pianta, R., & Jamil, F. (2014). Evidence for general and domain-specific elements of teacher–child interactions: Associations with preschool children's development. *Child Development*, *85*(3), 1257-1274.

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure?. *Child Development*, 76(5), 949-967.

Hamre, B. K., Pianta, R. C., Mashburn, A. J., & Downer, J. T. (2007). *Building a science of classrooms: Application of the CLASS framework in over 4,000 US early childhood and elementary classrooms.* Foundation for Childhood Development. https://www.fcd-us.org/building-a-science-of-classrooms-application-of-the-class-framework-in-over-4000-u-s-early-childhood-and-elementary-classrooms/

Hoy, A. W., & Weinstein, C. S. (2006). Student and teacher perspectives on classroom management. In C. Everston & C. Weinstein (Eds.), *Handbook of classroom management: Research, practice and contemporary issues* (pp. 181-222). Mahwah, NJ: Lawrence Erlbaum Associates.

Hill, H. C., Charalambous, C. Y., & Kraft, M. A. (2012). When rater reliability is not enough: Teacher observation systems and a case for the generalizability study. *Educational Researcher, 41*(2), 56-64.

Hill, H., & Grossman, P. (2013). Learning from teacher observations: Challenges and opportunities posed by new teacher evaluation systems. *Harvard Educational Review*, *83*(2), 371-384.

Hu, B. Y., Fan, X., Gu, C., & Yang, N. (2016). Applicability of the classroom assessment scoring system in Chinese preschools based on psychometric evidence. *Early Education and Development*, *27*(5), 714-734.

Kane, T. J., & Staiger, D. O. (2012). Gathering feedback for teaching: Combining high-quality observations with student surveys and achievement gains. Research Paper. MET Project. *Bill & Melinda Gates Foundation*.

Kounin, J. S. (1970). *Discipline and group management in classrooms*. Oxford, England: Holt, Rinehart & Winston.

Lee, C.D. (2007). *Culture, literacy, and learning.* New York, New York: Teachers College Press.

Levin, J.R., & Pressley, M. (1981). Improving children's prose comprehension: Selected strategies that seem to succeed. In C.M. Santa & B.L. Hayes (Eds.), *Children's prose comprehension: Research and practice (pp. 44-71).* Newark, Delaware: International Reading Association.

Leyva, D., Weiland, C., Barata, M., Yoshikawa, H., Snow, C., Treviño, E., & Rolla, A. (2015). Teacher–child interactions in Chile and their associations with prekindergarten outcomes. *Child Development*, *86*(3), 781-799.

Marzano, R. J., Pickering, D., & Pollock, J. E. (2001). *Classroom instruction that works: Research-based strategies for increasing student achievement*. Alexandria, VA: ASCD.

Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D., Burchinal, M., Early, D., & Howes, C. (2008). Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, 79(3), 732-749.

McCaffrey, D. F., Yuan, K., Savitsky, T. D., Lockwood, J. R., & Edelen, M. O. (2015). Uncovering multivariate structure in classroom observations in the presence of rater errors. *Educational Measurement: Issues and Practice*, *34*(2), 34-46.

McClellan, C., Donoghue, J., & Park, Y. S. (2013).

Commonality and uniqueness in teaching practice observation. Retrieved from http://www.clowderconsulting.com/wp-content/uploads/2016/01/Commonality-and-Uniqueness-in-Teaching-Practice-Observation_paper.pdf

National Research Council. (2003). *Engaging schools: Fostering high school students' motivation to learn*. National Academies Press.

Pakarinen, E., Lerkkanen, M. K., Poikkeus, A. M., Kiuru, N., Siekkinen, M., Rasku-Puttonen, H., & Nurmi, J. E. (2010). A validation of the classroom assessment scoring system in finnish kindergartens. *Early Education and Development*, *21*(1), 95-124.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational researcher*, 38(2), 109-119.

Pianta, R. C., Hamre, B. K., & Mintz S. (2012) *Classroom Assessment Scoring System Upper Elementary Manual*. Charlottesville, VA: Teachstone.

Raver, C. C. (2004). Placing emotional self-regulation in sociocultural and socioeconomic contexts. *Child development*, *75*(2), 346-353.

Rosenshine, B. (1995). Advances in research on instruction. *The Journal of Educational Research*, *88*(5), 262-268.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54-67.

Sandilos, L., Rimm-Kaufman, S., & Cohen, J. (2017). Warmth and demand: The relation between students' perceptions of the classroom environment and achievement growth. *Child Development, 88*(4).

Skinner, E. A., Zimmer-Gembeck, M. J., Connell, J. P., Eccles, J. S., & Wellborn, J. G.

(1998). Individual differences and the development of perceived control. *Monographs of the society for Research in Child Development*, i-231.

Slavin, R. E., Hurley, E. A., & Chamberlain, A. (2003). Cooperative learning and achievement: Theory and research. In W. M. Reynolds & G. E. Miller (Eds.) *Handbook of psychology* (pp. 177 – 198). Hoboken, New Jersey: John Wiley & Sons.

Taylor, B. M., Pearson, P. D., Peterson, D. S., & Rodriguez, M. C. (2003). Reading growth in high-poverty classrooms: The influence of teacher practices that encourage cognitive engagement in literacy learning. *The Elementary School Journal*, *104*(1), 3-28.

Tharp, R.G., & Gallimore, R. (1988). *Rousing minds to life: Teaching, learning, and schooling in social context.* Cambridge: Cambridge University Press.

Vygotsky, L. (1978). Interaction between learning and development. *Readings on the development of children*, *23*(3), 34-41.

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.

Yair, G. (2000). Educational battlefields in America: The tug-of-war over students' engagement with instruction. *Sociology of Education, 73*(4), 247-269.

The Convergence of Emotionally Supportive Learning Environments and College and Career Ready Mathematical Engagement in Upper Elementary Classrooms

Rebekah Berlin

Introduction

A third-grade teacher launches a mathematics task asking students to compare fractional quantities. After a class discussion of the task's key contextual features and mathematical ideas, the teacher provides time for individual reasoning. Some students use concrete materials to model the different quantities, others draw diagrams, while others represent their thinking through a series of inequalities, equations, and sentences. After several minutes, the teacher suggests students share, compare, and revise their findings with their table group. Finally, the teacher brings the class together to collectively identify mathematical relationships between students' different models. Before transitioning to the next subject, the teacher asks students to summarize their learning for the day and records their thinking on chart paper at the front of the classroom.

Literature suggests lessons like the one described here provide students opportunities to develop mathematical competencies necessary for college and career readiness (CCSSI, 2010; Cobb & Jackson, 2011; Kamin, 2016; Mishkind, 2014; Koestler, Felton, Bieda, & Otten, 2013; Schoenfeld, 2015)—often codified as College and Career Ready (CCR) mathematical practice standards. In the vignette above, students are provided an opportunity to develop Mathematical Practice Standard 1 (MP.1)—make sense of a task and persevere in solving it (Tennessee Department of Education, 2016)—and model with mathematics (MP.4; CCSSI, 2010). Likewise, students have an opportunity to reason abstractly and quantitatively (MP.2, Georgia Department of Education, 2019) and construct viable arguments and critique the reasoning of others (MP.3, Indiana Department of Education, 2014). The adoption of mathematical practice standards marked a shift from previous state standards, which focused only on mathematics content. By adopting mathematical practice standards, states for the first time planted a stake in the ground about the *processes* through which students should engage with required content to be considered mathematically proficient.

Recent work theorizes there are implicit expectations for social and emotional competencies of the classroom environment embedded in these CCR mathematical practice standards (Dana Center & CASEL, 2016; Dymnicki, Sambolt, & Kidron, 2013; Rimm-Kauffman & Youngs, in preparation). For example, making sense of a problem and persevering in solving it (MP.1) requires cognitive and emotional regulation (Denham & Brown, 2010; Hannula, 2006). Providing and/or receiving a constructive mathematical critique (MP.3) requires nuanced social awareness, relational skills, and self-management (Gest, Domitrovitch, & Welsh, 2005; Ginsburg-Block, Rohrbeck, Fantuzzo, 2006). Thus, a critical facet of supporting students in meeting the ambitious goals outlined in CCR mathematics standards may be supporting students' social and emotional learning.

Psychologists have documented how the qualities of a classroom environment can nurture or thwart students' ability to develop and exhibit social and emotional competencies (Frenzel, Pekrun, & Goetz, 2007; Reeve, 2006; Ryan & Deci, 2009; Patrick & Ryan, 2005; Sakiz, Pape, Woolfork-Hoy, 2012). Classroom environments that foster social and emotional learning are characterized by authentic opportunities for student autonomy (Reeve, 2006, Rimm-Kaufman & Hulleman, 2015; Urdan & Shoenfelder, 2006), a caring, supportive emotional climate (Hamre & Pianta, 2005; Reyes, Brackett, Rivers, White, & Salovey, 2012; Zins, Bloodworth, Weissberg, & Walberg, 2007), and productive student-student and teacher-student interactions (Patrick, Anderman, & Ryan, 2002; Pianta, Hamre, & Allen, 2012; Ryan & Patrick, 2001). Given the implicit social and emotional demands of mathematics practice standards, it is possible the degree to which a classroom environment supports the development of social and emotional competencies is more relevant to CCR mathematics teaching and learning than it may have been under previous standards.

Despite theories that the classroom environment may be associated with CCR mathematics teaching and learning, this relationship has not been examined empirically. Extant research on CCR-aligned mathematics teaching and learning focuses instead on standards (Cobb & Jackson, 2011; Dingman, Teuscher, Newton, & Kasmer, 2013; Porter, McMaken, Hwang, & Yang, 2011), curriculum and assessment (Polikoff, 2015; Shoenfeld, 2015), improving teacher content knowledge (Bausmith & Barry, 2011), and implementation and accountability efforts (Coburn, Hill & Spillane, 2016; Roth McDuffie, Drake, Choppin, Davis, Magaña, & Carson, 2017). While these each represent important elements of the shift to CCR standards, they do little to illuminate the characteristics of classroom environments where students do and do not engage with mathematics in the ways described in CCR standards. Focused squarely on this issue, this paper explores the following research questions:

1. What are the characteristics of classroom learning environments during lessons where raters do and do not observe CCR-aligned mathematical engagement?
2. In what ways do the classroom learning environments and students' engagement with CCR-aligned mathematics content vary within and across teachers?

In raising these questions, this article surfaces the potential convergence of two prevalent policy initiatives—college and career readiness and social and emotional learning—highlighting the ways in the which teaching practices that support each goal may be deeply intertwined.

Literature Review

College and Career Ready Mathematics

For many states, the adoption of CCR mathematics standards represented a substantive shift in expectations for what students had to know and do in order to be considered mathematically proficient (Dingman et al., 2013; Porter et al., 2011). The new standards outlined expectations for mathematics *content*—the specific concepts, procedures, and problem types students were expected to master at each grade level—and *practices*—the mathematical habits, processes, and dispositions students were expected to develop by engaging with mathematics content in particular ways.

Content. Compared to previous state standards, CCR content standards are designed to have greater *focus*—covering less content with greater depth in each grade level, *coherence*—progressions of mathematical understanding built within and across grades, and *balance*—attending equally to conceptual understanding, procedural fluency, and applied problem solving (Cobb & Jackson, 2011, Student Achievement Partners, 2016).

The impetus for creating more focused standards was an international comparison that showed curricula in high-performing countries covered only a few topics with great depth in each grade (Schmidt et al., 2001). Researchers found that in the United States, on the other hand, mathematics curricula tended to be a "mile wide and inch deep" (Porter et al., 2011, Schmidt et al., 2001). That is, students were expected to gain a cursory understanding of several topics, but rarely explored any topic with great depth. In light of these findings, CCR standards specify fewer topics covered each grade than previous state standards, but increase expectations for the level of conceptual understanding, procedural fluency, and problem solving required within each topic (Dingman et al., 2013; Porter et al., 2011; Student Achievement Partners, 2016).

CCR standards attend not just to what content should be covered, but also to how content is sequenced. To make standards more coherent, developers pulled from a large body of literature on learning progressions in mathematics that illustrate how students build mathematical understanding and fluency over time (Cobb & Jackson, 2011). This work was also informed by findings from the Third International Mathematics and Science Study, which suggested curricula organized through logical, hierarchical coverage of mathematics topics was a strong predictor of student achievement (Schmidt et al., 2001; Schmidt, Wang, & McKnight, 2005).

Finally, developers intended CCR standards to provide students equal opportunities to develop conceptual understanding, gain procedural skill and fluency, and apply mathematical knowledge to solve problems (Alberti, 2012). At the time the first CCR standards were written, many existing state standards and curricula were procedurally focused (Dingman et al., 2013; Hirsch & Reys, 2009). Analyses indicate that as a result, prior state standards were often lower in cognitive demand than current CCR standards (Porter et al., 2011). At the same time, there were several mathematics reform initiatives that emphasized problem solving and/or conceptual understanding at the cost of procedural fluency (Alberti, 2012; Schoenfeld, 2004). The standards were designed to address these imbalances.

Practices. CCR mathematics practice standards provide guidelines for the ways in which students interact with mathematics content. Teachers are expected to infuse these standards into the teaching of content at every grade level. The CCR practice standards were derived, in part, from the National Council of Teachers of Mathematics five Process Standards (NCTM, 2000) and focus on problem solving, reasoning, making mathematical connections, and representing and communicating mathematical ideas.

Across these standards, cognitive ownership is placed firmly with students. To support students in becoming proficient with these practices, teachers need to provide instructional scaffolds and explicitly foster the development of the habits, skills, and dispositions embedded in the standards. However, the practice standards make clear that the goal is for *students*, not teachers, to drive the mathematical work in the classroom.

For example, historically, a common discourse structure in mathematics classrooms has been: a) the teacher poses a question, b) a student responds, and c) the teacher evaluates the students' answer (Nathan, Eilam, & Kim, 2007; Schleppenbach, Perry, Miller, & Fang, 2007; Tainio & Laine, 2015). This structure, often called "Initiate, Respond, Evaluate," positions the teacher as expert and students as novices seeking to gain the teacher's knowledge and approval. In contrast, CCR mathematical practice standards explicitly outline different expectations for discourse in the mathematics classroom. MP.3 reads, "Students at all grades can listen or read the arguments of others, decide whether they make sense, and ask useful questions to clarify or improve the arguments" (Louisiana Department of Education, 2015, p. 6-7). Under CCR practice standards, *students*, not teachers, evaluate and build upon one another's responses in service of collective mathematical meaning making.

In the CCR practice standards, mathematical processes are outlined at both the individual and collective level (Koestler et al., 2013). For example, practice MP.6, "Attend to Precision," states that at the individual level, elementary students carefully specify units of measure, "calculate accurately and efficiently", and "express numerical answers with a degree of precision appropriate for the problem context" (CCSSI, 2010, p. 7). When students are engaged in collective mathematical work, this practice standard also states students should strive to "communicate precisely to others" using "clear definitions in discussion" stating "the meaning of the symbols they choose" and giving "carefully formulated explanations to each other" (Ibid, p. 7). Thus, mathematical practice standards articulate not only how students should monitor and drive their individual mathematical understanding, they also outline ways in which groups productively and collaboratively do the work of the discipline.

The Classroom Learning Environment

A physically and emotionally safe, predictable classroom learning environment creates a space where social, emotional, and academic competencies, such as those in the CCR standards, are enabled and developed (Hamre & Pianta, 2005; Zins & Elias, 2007). When students are in environments where they feel supported by their teacher and peers, they are more engaged, develop closer relationships, less fearful of making mistakes, are more likely to effectively communicate needs, and put forth greater effort (Hawkins, 1997; Lazarides, Gaspard, Dicke, 2019; Ryan & Patrick, 2001; Sakiz et al., 2012; Zins et al., 2007). Likewise, in classrooms where teachers effectively organize students' time and attention around challenging tasks and provide opportunities for autonomous decision-making, students are more likely to engage in extended periods of concentration, become intrinsically motivated, and develop self-management strategies (Shernoff, Csikszentmihalyi, Schneider, & Shernoff, 2003; Rathunde & Csikszentmihalyi, 2005; Turner & Meyer, 2004; Zins et al., 2007).

In order to foster such an environment, teachers must build positive relationships with and among students (Kiuru et al., 2015; McGrath & Van Bergen, 2015; Cornelius-White, 2007; Muller, 2001). They must maintain classroom norms and procedures that support students in becoming responsible for their own behavior (Charney, 1993; Marzano, Marzano, & Pickering, 2003; Egeberg, McConney, & Price, 2016; Pianta & Hamre, 2009). In addition, teachers who support high-quality classroom learning environments ensure constructive use of time and high student engagement (Hamre & Pianta, 2010; Pianta & Hamre, 2009).

Safe, productive, and emotionally classroom learning environments have been widely associated with student achievement in large-scale studies (Allen et al., 2013; Good & Grouws, 1977; Lockwood, Savitsky, & McCaffrey, 2015). Alternatively, in environments characterized by a lack of support, chaos, or negative emotional climate, students are less engaged, do not develop productive beliefs, and have lower academic achievement (Kunter, Baumart, & Köller, 2007; Patrick, Ryan, & Kaplan, 2007; Pianta, Belsky, Houts, & Morrison, 2007).

Researchers have found that interventions improving the extent to which the classroom learning environment supports the development of social and emotional competencies are associated with a variety of positive outcomes in mathematics classrooms (Ben-Avie et al., 2003). These include improved mathematics achievement for marginalized students (Cheema & Kitsantas, 2013), improved social and academic outcomes (Brock, Nishida, Chiong, Grimm, & Rimm-Kaufman, 2008; Flay, Allred, & Ordway, 2001; Rimm-Kaufman, Fan, Chiu, & You, 2007), and closer teacher-student relationships (Baroody, Rimm-Kaufman, Larsen, 2013). Together, these studies make a compelling case for the importance of safe, supportive classroom learning environments for mathematics achievement and other social-emotional outcomes that are relevant across disciplines.

These studies do not, however, illuminate whether safe, supportive classroom learning environments might be related to the ways in which children actually engage in mathematical work. For example, it is possible that in a turbulent classroom environment, students may not have the opportunity to demonstrate interwoven social, emotional, and mathematical competencies. When students begin a complex mathematics task as outlined in CCR practice standard MP.1, “by explaining to themselves the meaning of a problem and looking for entry points to its solution” and later choosing to “monitor and evaluate their progress and change course if necessary" and "check their answers to problems using a different method, and … continually ask themselves, 'Does this make sense?'" (CCSSI, 2010, p. 6), they draw on cognitive and emotional resources that may be depleted in an unpredictable classroom learning environment (CASEL, 2018; Durlak et al., 2011; Zins et al., 2007). During a lesson characterized by a more orderly, but unsupportive learning environment, a student may be able to engage in problem solving uninterrupted. Qualitative work, however, suggests that in such environments students may be unwilling to look for their own entry points into a task or share their developing thinking with others because they fear the social and emotional consequences of mathematical errors (Turner & Meyer, 2004).

In addition to limiting the extent to which students can demonstrate social and emotional competencies, unsupportive learning environment may limit their development and transfer (Curby, Brock, & Hamre, 2013; Kern & Clemens, 2007). Prior literature suggests that in some classrooms the classroom learning environment is consistently safe, productive, and supportive (Curby, Grimm, & Pianta, 2010). In such environments there are implicit supports for students' social, emotional, and academic development (Rimm-Kauffman & Youngs, in preparation). Due to consistent exposure, students begin to demonstrate social, emotional, and academic competencies they see modeled by their teachers and their peers (Becker & Domitrovich, 2011; Blazar & Kraft, 2017; Martin & Rimm-Kaufman, 2015). In other classrooms, however, the extent to which a classroom environment is well organized or emotionally supportive can vary considerably over the course of the year (Pianta & Hamre, 2009). This unpredictability is associated with reduced social, emotional, and academic outcomes over time (Curby et al., 2013; Curby et al., 2009). Given the potential association between the development of social and emotional competencies and students' ability to engage with mathematics in the ways outlined in

CCR standards, there is a need for studies that highlight variability in classroom learning environments and students' mathematical engagement between teachers and within classrooms over time. Recent research suggests the creation of empirical profiles of instruction provides an effective means for exploring this variability (Halpin & Kieffer, 2015; Keller, Becker, Frenzel, & Taxer, 2018).

In this article I examine the relationship between classroom learning environments and CCR-aligned mathematical engagement. By identifying characteristics of classroom learning environments where students do and do not engage in CCR-aligned mathematical work, I provide empirical evidence supporting one facet of the theorized overlap between social and emotional learning and CCR mathematics. I also investigate the extent to which classroom learning environments and students' CCR-mathematical engagement are stable across lessons taught by the same teacher as well as the extent to which there are meaningful differences on these two dimensions between teachers. These results have the potential to illuminate one of the ways that, over time, students in different classrooms may have divergent opportunities to develop social, emotional, and mathematical competencies. The findings, which identify commonalities and differences across teachers' instruction, also raises questions about the need for differentiated professional development with an emphasis on social and emotional learning for elementary mathematics teachers.

Method

Sample

These data are drawn from a larger study of content-focused professional development in a large urban school district (Cohen, Hutt, Berlin, & Wiseman, under review). The study included 49 third, fourth, and fifth grade mathematics teachers from 23 schools (see Table 1 for teacher information). Data collection took place during the 2016-17 and 2017-18 school years. Each teacher was filmed multiple times (*m = 8.6 lessons).* This resulted in a total sample of 419 videos of 30-minute mathematics lesson segments.

Table 1. Teacher demographics.

	2016-17	**2017-18**
Participants	27	35
% Female	70	60
% White	56	60
Experience (years)	4.5	6.3
% Teaching in High Poverty Status School	63	53

During Year 1 of data collection, the 2016-17 school year, 27 teachers participated in the study. Of these teachers, thirteen continued in the study for the second year. An additional 22 teachers joined the study in the second year. The average teaching experience for the teachers in this sample was 4.45 years during Year 1 and 6.27 years in Year 2. Like most teachers in the United States, teachers in this sample largely identified as female (70% Year 1, 60% Year 2) and White (56% Year 1, 60% Year 2; Cherng & Halpin, 2016; Hodgkinson, 2001; Lindsay & Hart, 2017). The majority of teachers in this sample taught in schools labeled with high poverty status (63% Year 1, 53% Year 2). There were no significant differences on these demographic variables between teachers that participated in both years of the study and teachers that participated in only one year of the study.

Measures

This analysis is drawn from scores assigned by certified raters on two observational rubrics. The first, the Classroom Assessment Scoring System (CLASS), is a widely used content-generic tool that captures different dimensions of classroom interactions (Hamre & Pianta, 2005). The second, the Instructional Practice Research Tool for Mathematics (IPRT-M), is a Common Core aligned mathematics-specific observation tool developed for the project (Cohen et al., under review).

CLASS. The CLASS Upper Elementary includes 12 dimensions of classroom interactions, each coded on a scale ranging from 1 (low) to 7 (high). The dimensions are organized into four domains: Emotional Support, Classroom Organization, Instructional Support, and Student Engagement. In order to understand the extent to which there is a supportive classroom learning environment during a mathematics lesson, this analysis includes scores from the dimensions of three of these four domains: Emotional Support, Classroom Organization, and Student Engagement.

The Emotional Support (ES) domain is comprised of three dimensions. The first, Positive Climate, measures mutual enjoyment and emotional connection between teachers and students as well as students and their peers. Teacher Sensitivity assesses teachers' responsiveness to students' academic, social, and emotional needs. Regard for Student Perspectives captures the degree to which teachers support student leadership and autonomy.

The domain Classroom Organization contains three dimensions: Behavior Management, Productivity, and Negative Climate. High-quality Behavior Management is characterized by teachers who are consistent, proactive, attend to positive behavior, and demonstrate low-reactivity. Their students know what to do, consistently meet teacher expectations, and independently manage behavior (e.g., peer redirection and problem solving). The Productivity dimension assesses the degree to which classroom routines, teacher preparedness, and clarity of instructions support student learning. Negative Climate evaluates the levels of negative affect, punitive control, and disrespect (e.g., humiliation, sarcasm, and exclusionary behavior) in a classroom.

The domain Student Engagement only contains one dimension. This dimension assesses the degree to which all student are actively participating (e.g., volunteering information, asking and answering questions) in the activity the teacher is facilitating. Lesson means and standard deviations for these seven dimensions are presented in Table 2 along with intraclass correlations (ICCs) indicating interrater-reliability for the 19% of mathematics lesson segments that were double scored.

Table 2. Means (M), Standard Deviations (*SD)*, and Intraclass Correlations (ICC)[8] for the seven CLASS dimensions, which pertain to the classroom learning environment.

Domain	Dimension	M *(SD)*	ICC
Emotional Support	Positive Climate	4.51 *(0.98)*	0.59
	Teacher Sensitivity	5.24 *(0.97)*	0.59
	Regard for Student Perspectives	2.97 *(1.07)*	0.73

[8] The CLASS certification process uses adjacent rather than exact scoring. Therefore ICCs based on exact scores from CLASS dimensions tend to be lower than those of other observation rubrics. The values presented here are consistent with those in other studies (Cohen, Ruzek, & Sandilos, 2018; Hamre et al., 2013)

Classroom	Behavior Management	6.05 *(1.01)*	0.68
Organization	Productivity	6.30 *(0.78)*	0.54
	Negative Climate	1.42 *(0.71)*	0.58
Student Engagement	Student Engagement	5.42 *(0.92)*	0.62

IPRT-M. This measure, developed specifically for this project, was adapted from Student Achievement Partners' Instructional Practice Guide.[9] Our district partner used the Instructional Practice Guide to develop professional development designed to promote CCR-aligned instruction across the district. They requested we design a measure aligned to their goals.[10]

The IPRT-M was explicitly drawn from the Common Core Standards for Mathematical Practice (CCSSI, 2010) and the Progressions for the Common Core Standards in Mathematics (Institute for Mathematics and Education, 2007). There are 8 rubrics: Coherence, Depth, Student Representations and Solution Strategies, Prompting Student Thinking, Responding to Misunderstanding, Opportunities to Engage with Mathematics, Opportunities to Justify and Critique, and Student Justifications and Critiques. Each rubric is scored on a four-point ordinal scale. A rating of 1 indicates no opportunity to engage with ambitious CCR-aligned mathematics content. A score of 2 indicates shallow or cursory opportunities.[11] A score of 3 indicates

[9] The Instructional Practice Guide can be viewed at the following website: https://achievethecore.org/category/1155/printable-versions

[10] Though this tool currently lacks published predictive validity evidence supporting its use (paper providing validity evidence linking scores on this tool to student outcomes is in preparation), this tool is used to coach thousands of teachers across the United States. Given the wide adoption of this conception of mathematics instruction by practioners in K-12 environments, it is important to understand characteristics of classroom environments that are and are not associated with this conception of mathematical engagement.

[11] The IPRT scoring guide defines cursory opportunities as those that do not elicit mathematical thinking (e.g., instead of an authentic opportunity to justify or critique a peer's reasoning, a

occasional opportunities to engage with ambitious CCR-aligned mathematics content and a score of 4 indicates consistent engagement with ambitious CCR-aligned mathematics content on that dimension. Therefore, a score of 4 is the only rating that indicates students are regularly engaging with mathematics in the ways outlined in CCR mathematics standards. Scores less than 4 signify students have limited opportunities for CCR-aligned engagement.

The Coherence and Depth scales are derived from the Progressions for the Common Core Standards in Mathematics and Standards for Mathematical Practice (IME, 2007; CCSSI, 2010), The IPRT-M Coherence rubric assesses the extent to which a teacher intentionally relates the current lesson to students' prior mathematical skills and knowledge. When scoring the Depth rubric, raters determine whether the teacher makes the depth of the mathematics in the lesson plain through the use of explicitly connected explanations, representations, tasks, and/or examples. Raters also account for whether the mathematics presented is clear and correct.

The remaining indicators were derived from the Common Core Standards for Mathematical Practice (CCSSI, 2018). The Student Representations and Solution Strategies rubric captures the degree to which understanding of mathematics content is supported through the strategic sharing of students' representations and solution methods. To score at the high end of this rubric, the teacher must support students in explicitly drawing mathematical connections between various representations and/or solution strategies. Prompting Student Thinking assesses the frequency with which the teacher poses questions and tasks that elicit mathematical reasoning and provide opportunities for productive struggle (Granberg, 2016; Kapur, 2014; Warshauer, 2014). Responding to Misunderstanding captures whether the teacher responds

---

teacher might provide a cursory opportunity saying, "Thumbs up if you agree, thumbs down if you don't!")

constructively to student misunderstandings—that is, with scaffolds[12] that offer specific, clear, mathematical support for the student to use re-engage with the problem and revise their thinking (Granber, 2016). When scoring Opportunities to Engage with Mathematics, raters determine the proportion of the lesson that the teacher provided opportunities for all students to work with and practice mathematics problems, tasks, and exercises. When a teacher prompts students to justify their thinking and/or critique the reasoning of others, this is scored under the Opportunities to Justify and Critique rubric. The mathematical depth, precision, and logic of student justifications and critiques are scored on the Student Justifications and Critiques rubric.

Raters were trained over a three-day period and certified when they scored an exact match to the master score for each scale on three out of four videos. Under the scoring procedure, raters watch and take notes on 30-minute segments of mathematics instruction before assigning scores. Table 3 provides inter-rater reliability for the 15% of lessons that were double scored as well as descriptive statistics for each rubric.

Table 3. Means (M), Standard Deviations (*SD)*, and Intraclass Correlations (ICC) for the 8 IPRT-M rubrics.

	M *(SD)*	ICC
Coherence	2.21 *(1.04)*	0.81
Depth	2.93 *(0.71)*	0.86
Student Representations and Solution Strategies	1.50 *(0.87)*	0.92
Prompting Student Thinking	2.19 *(0.50)*	0.92

[12] A teacher telling a student what to do is not considered a scaffold just as providing a student part or all of an answer is not considered a scaffold.

Responding to Student Misunderstanding	2.64 *(0.84)*	0.71
Opportunities to Engage with Mathematics	3.72 *(0.51)*	0.90
Opportunities for Justification and Critique	3.03 *(0.79)*	0.90
Student Justifications and Critiques	2.57 *(0.91)*	0.86

Analysis

Each video in the sample captures 30 minutes of mathematics instruction. While raters using the IRPT-M score 30-minute increments of instruction, CLASS raters score 15-minute increments. Therefore, CLASS segment scores were averaged to create lesson-level scores.

Research question 1: What are the characteristics of classroom learning environments during lessons where raters do and do not observe CCR-aligned mathematical engagement? To explore characteristics of classroom learning environments in lessons where raters do and do not observe CCR-aligned mathematical engagement, I used latent profile analysis (Magidson & Vermunt, 2004) to uncover groups of similar lessons, or lesson profiles, present in the sample of 422 videos. Though traditionally person rather than lesson centered, this analytic technique can also be used to examine the co-occurrence of particular lesson characteristics and determine whether there are patterns of co-occurrence pervasive enough that they can be used to sort lessons into different groups (Keller et al., 2018). In this case, I used indicators of (a) the classroom learning environment drawn from the CLASS and (b) teacher's and students' engagement with CCR mathematics content using ratings from the IPRT-M to create commonly occurring profiles of elementary mathematics instruction. The profiles allowed me to determine what types of classroom learning environments co-occur with CCR-aligned mathematical engagement.

The 419 lessons analyzed here are nested within 49 teachers. To account for this, I performed a multilevel latent profile analysis using “Type Is Twolevel Mixture” in Mplus 8 (Asparouhov & Muthén, 2008). In accordance with current recommendations for multilevel latent profile analysis, all data were standardized prior to analysis (Keller et al., 2018; Mäkikangas, Tolvanen, Aunola, Feldt, Mauno, & Kinnunen, 2018).

Research question 2: In what ways do the classroom learning environments and students’ engagement with CCR-aligned mathematics content present in mathematics lessons vary within and across teachers? To determine the consistency of particular mathematics lesson profiles within individual teachers, I ran models to identify whether it was possible to form level 2 teacher classes based on the relative frequency of level 1 lesson profile membership (Mäkikangas et al., 2018). Put plainly, these models examined whether it was possible to identify groups of teachers who tended to teach mathematics in consistent ways. These groups provide valuable information about common strengths and areas for growth in mathematics instruction for teachers in this sample. This teacher-level analysis also illuminates variability in the classroom learning environment and students’ CCR-aligned engagement within teachers. Together, these analyses provide important information about the extent to which students in different classrooms may have had divergent access to a supportive classroom learning environment and/or opportunities for CCR-aligned mathematical engagement over time.

Model specification. The process for identifying multilevel mixture models with profiles at level 1 and classes at level two consists of two phases (Henry & Muthén, 2010; Mäkikangas et al., 2018). During the first phase, the researcher identifies the correct number of level 1 (lesson) profiles starting from a one-profile solution working up until the correct number of profiles is determined. The correct number of profiles is identified using several fit indices—the sample adjusted Bayesian information criteria (SABIC, lower values indicate better model fit), the bootstrap likelihood ratio test (BLRT, p-value indicates the *k* profile solution is a better fit to the data than the *k-1* profile solution), entropy value (values close to 1 suggest greater distinction between profiles), the classification probabilities for most likely class membership, and the number and percent of lessons in each profile (Mäkikangas et al., 2018; McLachlan & Peel, 2000; Nylund, Asparouhov, & Muthén, 2007). The focus during extraction of the single level model should be on identifying a model that is both substantively meaningful and parsimonious so that after adding the additional parameters necessary for the multilevel model, the model can still converge (Henry & Muthén, 2010).

During the second phase, the researcher determines the correct number of level 2 (teacher) classes based on the frequency of level 1 profile membership. Several studies indicate that the lowest BIC should be used to determine the correct number of level 2 classes as other fit indices are not reliable for multilevel mixture models (Finch & French, 2014; Henry & Muthén, 2010; Yu & Park, 2014). Finally, graphical presentations of the final solution should be examined using standardized and raw data (Meyer & Morin, 2016).

Results

Based on observation ratings of their classroom learning environments and CCR-aligned mathematical engagement, the mathematics lessons in this sample can be separated into four distinct instructional profiles and teachers into three distinct instructional groups. Model identification, lesson-level profiles, and teacher groups are described in greater detail below.

Model Identification

Lesson-level profiles. Table 4 shows the fit indices for the one- through seven-profile solutions. These solutions describe the number of groups the 419 lessons can be categorized into based on patterns in their observation ratings. The eight-profile solution is not included in the table for two reasons. First, the best log likelihood failed to replicate for the eight-profile solution, even after increasing starting values several times. The eight-profile solution also included profiles with only one lesson. Together, this evidence suggests the eight-profile solution did not contain true profiles of teaching and was likely the result of chance.

As is common in latent profile analysis, the fit indices do not point to a single solution, or a "correct" number of categories into which the mathematics lessons should be sorted (Nylund et al., 2007). The BLRT and SABIC suggest the 7-profile solution best fits the data. That is, the lessons could be divided into 7 different groups based on the observation ratings. However, the entropy values and classification probabilities point to a smaller number of lesson profiles. For example, in the seven-class solution, there are some lessons with up to a 10% chance of misclassification whereas with four or less solutions there is only a 4% chance of misclassification.

Prioritizing parsimony in service of successfully running a multilevel model, I selected the four-profile solution. The SABIC declines sharply, and then begins to level out after the 4-profile solution. In addition, the entropy values and likelihood of classifying lessons into the correct profile drop with more than four profiles. This indicates that as the number of groups increases, the groups become less distinct from one another. Therefore, the 4-profile solution seemed the best fitting, most parsimonious solution. This means lessons in this sample can be sorted into four distinct groups based on the IPRT-M and CLASS observation ratings.

Table 4. Fit indices for the single-level profile solutions.

Number of	SABIC	Entropy	BLRT *p*	Classification Probabilities for	Lessons in each Profile

Extracted Profiles				Most Likely Class Membership	n *(%)*
1		-	-	-	419 (*100*)
2	16223.63	0.97	< 0.001	0.98-0.99	86 (*21*)
					333 (*79*)
3	15291.99	0.93	< 0.001	0.96-0.99	45 (*11*)
					227 (*54*)
					147 (*35*)
4	14699.22	0.94	< 0.001	0.96-0.99	189 (*45*)
					16 (*4*)
					80 (*19*)
					134 (*32*)
5	14483.80	0.91	<0.001	0.91-0.99	15 (*4*)
					96 (*23*)
					72 (*17*)
					153 (*37*)
					83 (*20*)
6	14216.76	0.92	<0.001	0.92-1.00	15 (*4*)
					110 (*26*)
					27 (*6*)
					51 (*12*)
					136 (*32*)

					80 *(19)*
7	14080.35	0.92	<0.001	0.90-1.00	15 (*4*)
					50 (*12*)
					21 (*5)*
					131 (*31*)
					109 (*26*)
					25 *(6*)
					68 (*16*)

Teacher groups. Next, in order to account for the nested data structure and to determine whether it was possible to identify groups of teachers who taught mathematics in similar ways to one another, I ran a series of multilevel models building on the four-profile solution. The BIC values for the four-profile and 1, 2, 3, and 4 class solutions are listed in Table 5. The BIC value points to the three-class, four-profile solution. Simply put, the mathematics teachers in this sample teach in consistent ways – specifically, there are four profiles of mathematics lessons, and three groups of teachers that emerge from these data.

Table 5. Fit indices for multilevel mixture models.

Extracted Classes	BIC
1 class, 4 profile	16329.5
2 class, 4 profile	16256.4
3 class, 4 profile	16244.6
3 class, 4 profile	16265.2

Lesson Profiles

Four lesson profiles emerged from these data (see Table 6): (1) Turbulent Classroom Learning Environment (CLE), Rare College and Career Ready Mathematical Engagement (CCR-M); (2) Inconsistent CLE, Infrequent CCR-M; (3) Orderly CLE, Infrequent CCR-M; (4) Supportive CLE, Consistent CCR-M. Figure 1 shows the standardized means on each dimension for each profile. Table 7 shows the raw means on each dimension for each profile. Each profile is described in greater depth below.

Table 6. Lesson Profile Numbers and Names

Lesson Profile	Classroom Learning Environment	College and Career Ready Mathematical Engagement
1	Turbulent	Rare
2	Inconsistent	Infrequent
3	Orderly	Infrequent
4	Supportive	Frequent

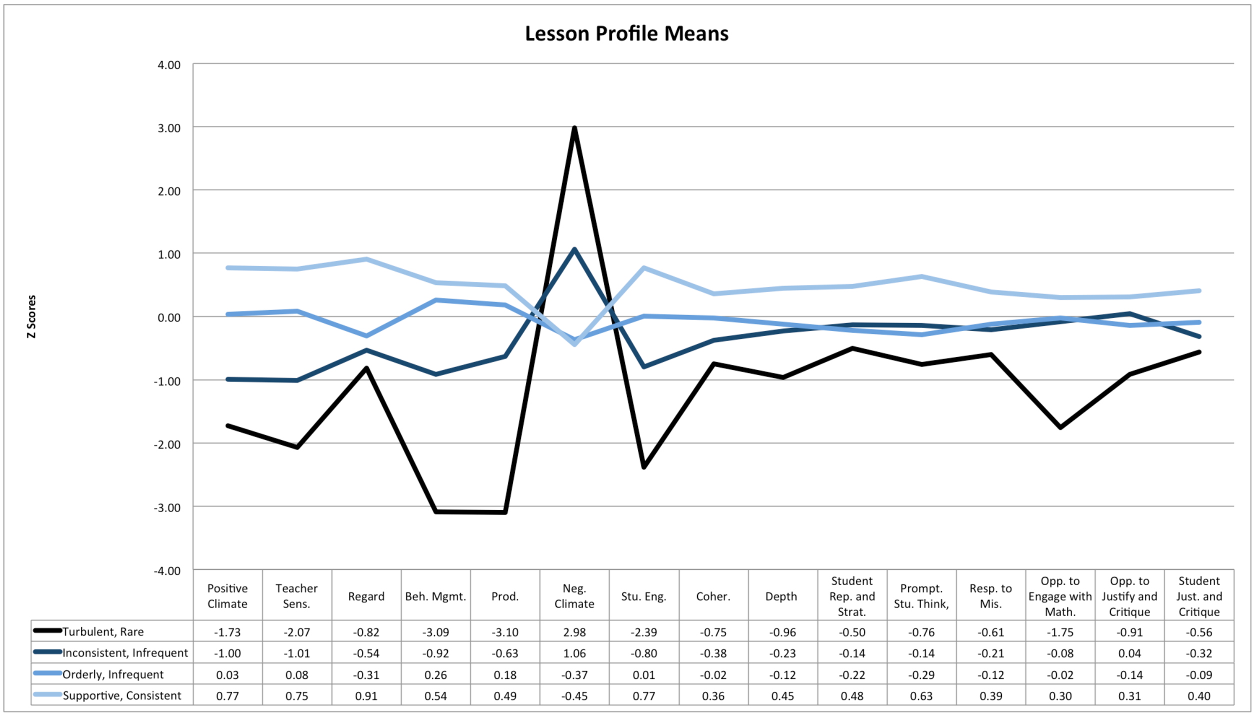


Figure 1. Mathematics lesson profiles and standardized profile means.

Table 7. Dimension Raw Means and Standard Deviations by Lesson Profile.

		Turbulent, Rare (n = 16)	Inconsistent, Infrequent (n = 76)	Orderly, Infrequent (n = 201)	Supportive, Consistent (n = 126)
CLASS (scale 1-7)	Positive Climate	2.81	3.52	4.53	5.31
		0.52	*0.69*	*0.67*	*0.77*
	Teacher Sensitivity	3.22	4.26	5.30	5.99
		0.69	*0.79*	*0.68*	*0.55*
	Regard for Student Perspectives	2.10	2.40	2.64	3.97
		0.70	*0.76*	*0.79*	*0.97*
	Behavior Management	2.94	5.10	6.32	6.62

		0.88	*0.74*	*0.57*	*0.49*
	Productivity	3.91	5.81	6.45	6.69
		0.87	*0.74*	*0.50*	*0.35*
	Negative Climate	3.54	2.20	1.16	1.10
		1.03	*0.54*	*0.28*	*0.32*
	Student Engagement	3.22	4.68	5.42	6.15
		0.73	*0.80*	*0.64*	*0.51*
IPRT-M (scale 1-4)	Coherence	1.44	1.83	2.19	2.59
		0.51	*0.82*	*1.04*	*1.08*
	Depth	2.25	2.75	2.83	3.27
		0.68	*0.70*	*0.68*	*0.61*
	Student Representations	1.06	1.37	1.30	1.94
	and Solution Strategies	*0.25*	*0.74*	*0.67*	*1.09*
	Prompting Student	1.81	2.12	2.05	2.51
	Thinking	*0.54*	*0.41*	*0.33*	*0.61*
	Responding to	2.13	2.45	2.54	2.97
	Misunderstanding	*0.89*	*0.82*	*0.79*	*0.83*
	Opportunities to Engage	2.81	3.69	3.70	3.88
	with Mathematics	*0.98*	*0.51*	*0.49*	*0.33*
	Opportunities to Justify and	2.31	3.07	2.91	3.30
	Critique	*0.87*	*077*	*0.84*	*0.59*
	Student Justification and	2.06	2.28	2.49	2.95
	Critique	*0.77*	*0.98*	*0.89*	*0.78*

Characteristics of classroom learning environments where raters do not observe CCR-aligned mathematical engagement. Raters did not observe consistent CCR-aligned mathematical engagement in lessons that were categorized into the first three profiles: (1) Turbulent CLE, Rare CCR-M; (2) Inconsistent CLE, Infrequent CCR-M; (3) Orderly CLE, Infrequent CCR-M.

The Turbulent CLE, Rare CCR-M profile contained lessons characterized by low and lower mid-range emotional support (Positive Climate, *m=2.81*; Teacher Sensitivity, *m =3.22*; and Regard for Student Perspectives, *m = 2.10*) and chaotic (Behavior Management, *m = 2.94*; Productivity, *m = 3.91*), negative classroom learning environments (Negative Climate, *m = 3.54*). During these lessons, students were either offered no opportunities for CCR-aligned mathematical engagement or infrequent and/or shallow opportunities (all IPRT-M ratings < 3). Four percent of the total lessons fell into this category.

The second profile, Inconsistent CLE, Infrequent CCR-M, contains lessons where both the learning environment and CCR-aligned mathematical engagement varied over the 30-minute segment (all average CLASS ratings, with the exception of productivity ranged from 2-6; six out of eight average IPRT-M scores < 3). On both the CLASS and the IPRT, mid-range scores (3-5, and 2-3, respectively) indicate that while there are sporadic examples of high-quality practice, they are not representative of the majority of the lesson.[13] Standardized scores on the Opportunities to Engage with Mathematics IPRT-M rubric show that students have substantially more opportunities to do mathematical work in lessons characterized as having an Inconsistent

---

[13] For example, a mid-range Behavior Management score on the CLASS indicates that a teacher employs a mixture of effective and ineffective behavior management strategies. In these lessons, there may be periods of chaos, though these do not last for the full lesson. Similarly, a mid-range Prompting Student Thinking Score is assigned to lessons where the teacher occasionally poses questions to elicit students' mathematical thinking. However, for the majority of the lesson they pose questions with simple right-or-wrong answers.

CLE than in categorized as having a Turbulent CLE (1.67 standard deviation difference in average scores between the two). Eighteen percent of the lessons fell in the Inconsistent CLE, Infrequent CCR-M profile.

The third profile, Orderly CLE, Infrequent CCR-M, showed a marked increase in the classroom learning environment over the Inconsistent CLE, Infrequent CCR-M profile (over 1 standard deviation increase in average scores on Positive Climate, Teacher Sensitivity, Behavior Management, and lack of Negative Climate). However, the average IPRT-M scores indicate that CCR-aligned mathematical engagement during lessons in the Orderly CLE, Infrequent CCR-M profile is indistinguishable from that during lessons in the Inconsistent CLE, Infrequent CCR-M profile. Indeed, with the exception of the Coherence indicator, between-group comparisons indicate there are no significant differences between the average IPRT-M scores for the Inconsistent CLE, Infrequent CCR-M and Orderly CLE, Infrequent CCR-M profiles ($p > 0.05$).

The learning environment present in the Orderly CLE, Infrequent CCR-M profile is notable because while CLASS scores suggest it is orderly, they do not indicate this classroom learning environment is particularly supportive or engaging. The average behavior management and productivity scores are in the high range for this profile. The average Emotional Support and Student Engagement scores, on the other hand, are in the low to mid range. This profile contained the most lessons (48%).

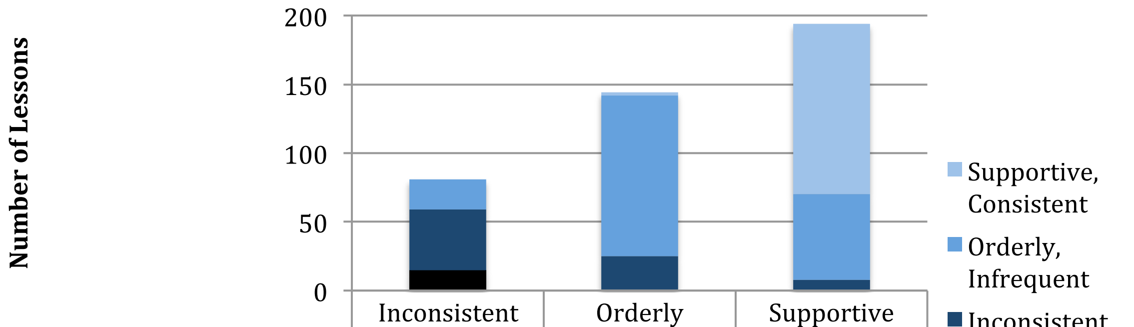
Across lessons with Inconsistent and Orderly learning environments, an interesting pattern emerges—in these lessons, teachers offered students opportunities for CCR- aligned mathematical engagement, but students did not take them. Both profiles had higher average scores on the Opportunities to Justify and Critique scale than on the Student Justifications and Critiques scale of the IPRT-M. This reveals that in Inconsistent and Orderly environments teachers offer authentic opportunities for students to justify their responses and/or critique the reasoning of others, but there was little student uptake of these opportunities. In addition, the low average scores on the IPRT-M Student Representations and Solution Strategies rubric (1.37 and 1.30, respectively) indicate there was virtually no evidence of students sharing solution strategies or mathematical representations with peers in lessons categorized into the Inconsistent CLE, Infrequent CCR-M or Orderly CLE, Infrequent CCR-M profiles. Unlike the ratings for Opportunities to Justify and Critique, the ratings for Student Representations and Solution Strategies do not illuminate whether students were offered opportunities to share representations and solution strategies. They only show that in these lessons students did not share their thinking in this way.

Characteristics of classroom learning environments where raters observe CCR-aligned mathematical engagement. The fourth profile, Supportive CLE, Consistent CCR-M, contained 30% of the lessons in the sample. The learning environments in these lessons are orderly like those in profile 3. Notably, they are also significantly more emotionally supportive and engaging than lessons in profile 3 ($p < 0.05$ for all indicators in Emotional Support and Student Engagement domains). According to the standardized scores, the greatest differences between the learning environments in profiles 3 (Orderly) and 4 (Supportive) were on Positive Climate, Regard for Student Perspectives, and Student Engagement (differences ranged from 0.74-1.22 standard deviations). These ratings indicate that compared to lessons in Orderly environments, during lessons in Supportive environments, there was greater evidence of warm interactions, opportunities for student autonomy, and active student participation. CCR-aligned mathematical engagement was also significantly higher during lessons in profile 4 than in all other profiles ($p < 0.05$ for all IPRT-M indicators). Notably, the largest differences in CCR-aligned mathematical engagement between the Supportive CLE, Consistent CCR-M and the other profiles are on indicators that most explicitly call on students' social and emotional competencies for engaging in collective mathematical work—Student Representations and Solution Strategies and Student Justification and Critique—and for engaging in individual productive struggle—Depth, Responding to Student Misunderstanding and Prompting Student Thinking.

In summary, there were marked differences between the characteristics of classroom learning environments where raters did and did not observe CCR-aligned engagement. Raters did not observe CCR-aligned engagement in classroom learning environment characterized by chaos and low emotional support. Importantly, raters also did not observe CCR-aligned instruction in classroom learning environments that are orderly and productive, but offer only inconsistent emotional support and student engagement. Raters only observed consistent opportunities for CCR-aligned mathematical engagement in environments that were safe and productive, *as well as* positive, supportive, and engaging.

Teacher Groups

Variation in classroom learning environments and CCR-aligned mathematical engagement across teachers. Level-two classes, based on profile membership, indicated that there were three groups of teachers in this sample (see Figure 2). The first and smallest group of teachers (n= 10, 20% of teachers in sample) did not contain any lessons where raters observed consistent CCR-aligned mathematical engagement. This group of teachers taught all the lessons from the Turbulent CLE, Rare CCR-M profile (profile 1) as well as the majority of lessons from the Inconsistent CLE, Infrequent CCR-M (profile 2). This group is termed "Inconsistent" to highlight that in these teacher's classrooms, students did not consistently have access to even an orderly learning environment.



	Inconsistent	Orderly	Supportive
Supportive, Consistent	0	2	124
Orderly, Infrequent	22	117	62
Inconsistent, Infrequent	44	24	8
Turbulent, Rare	15	1	0

Figure 2. Teacher groups created from lesson profile membership.

The second group, which comprised 35% of the teachers in the sample (n = 17), taught lessons characterized by infrequent engagement with ambitious mathematics content. The majority of these lessons (81%) were from the Orderly CLE, Infrequent CCR-M profile. Thus, this group was termed "Orderly" to suggest that while students in these classrooms were, for the most part, in safe and productive learning environments, they rarely (only 2 lessons) had access to more emotionally supportive classroom learning environments, just as they rarely had opportunities for consistent CCR-aligned mathematical engagement.

The third and largest group of teachers (n = 22, 45%) was termed "Supportive." The teachers in this group provided a safe and productive learning environment in almost all of their lessons (8 lessons taught by this group were from the Inconsistent CLE, Infrequent CCR-M profile). In 64% of these lessons taught by this group of teachers, students also had consistent opportunities for CCR-aligned mathematical engagement and were in a supportive, engaging classroom learning environment (lessons from the Supportive CLE, Consistent CCR-M profile).

Between-group comparisons of IPRT-M scores highlight similarities and differences between the three groups of teachers. Teachers in the Inconsistent and Orderly groups afforded students similar opportunities for CCR-aligned mathematical engagement. With one exception, there are no statistically distinguishable differences between the average IPRT-M scores for teachers in the Orderly group and teachers in the Inconsistent group ($p > 0.10$).[14] In contrast, the average score on each rubric of the IPRT-M for teachers in the Supportive group is significantly higher than the average scores of teachers in the Inconsistent or Orderly group ($p < 0.05$). Together the observed differences in classroom learning environments and CCR-aligned mathematical engagement suggest that over time, students in different classrooms may have had meaningfully different experiences learning mathematics.

Variation in classroom learning environments and CCR-aligned mathematical engagement within teachers. Figure 3 shows instructional patterns within individual teachers. Examining the distribution of profiles within individual teachers underscores some of the differences between the teacher groups. Both the Inconsistent and Orderly teacher groups contained a sizeable number of lessons categorized as Inconsistent CLE, Infrequent CCR-M (44 and 24, respectively). In the Inconsistent teacher group these lessons are concentrated within single teachers (teachers range from 2-9 observations categorized this way). In the Orderly group, on the other hand, these lessons are distributed across teachers, such that most teachers had only one lesson characterized this way. The distribution of lesson profiles within the Inconsistent teacher group also shows that lessons sorted into the lowest profile (Turbulent CLE, Rare CCR-M) tended to cluster within particular teachers. For example, of the eight teachers

[14] The exception was Opportunities to Engage with Mathematics. This indicator captures student opportunities to do work related to mathematics, regardless of whether it is CCR-aligned. Teachers in the Inconsistent group likely have lower scores on this indicator because instructional time was spent attending to behavior rather than on mathematics content.

who had a lesson categorized as Turbulent CLE, Rare CCR-M, six had multiple lessons categorized this way.

This figure also reveals most teachers enacted at least two different mathematics lesson profiles. This within-teacher variability was largely between adjacent profiles (e.g., profiles 2 and 3). Fourteen teachers exhibited three teaching profiles (e.g., Teacher 8). Only two of these teachers were in the Orderly group, highlighting that for lessons observed in this sample, instruction in the Orderly group was most consistent. Only three out of 49 teachers exhibited just one teaching profile. Thus, while broadly speaking, individual teachers in this sample taught mathematics in consistent enough ways that teachers could be separated into three distinct groups, there was also variability in the characteristics of the classroom learning environments and opportunities for CCR-aligned mathematical engagement students had access to from lesson to lesson, even within the same teacher.

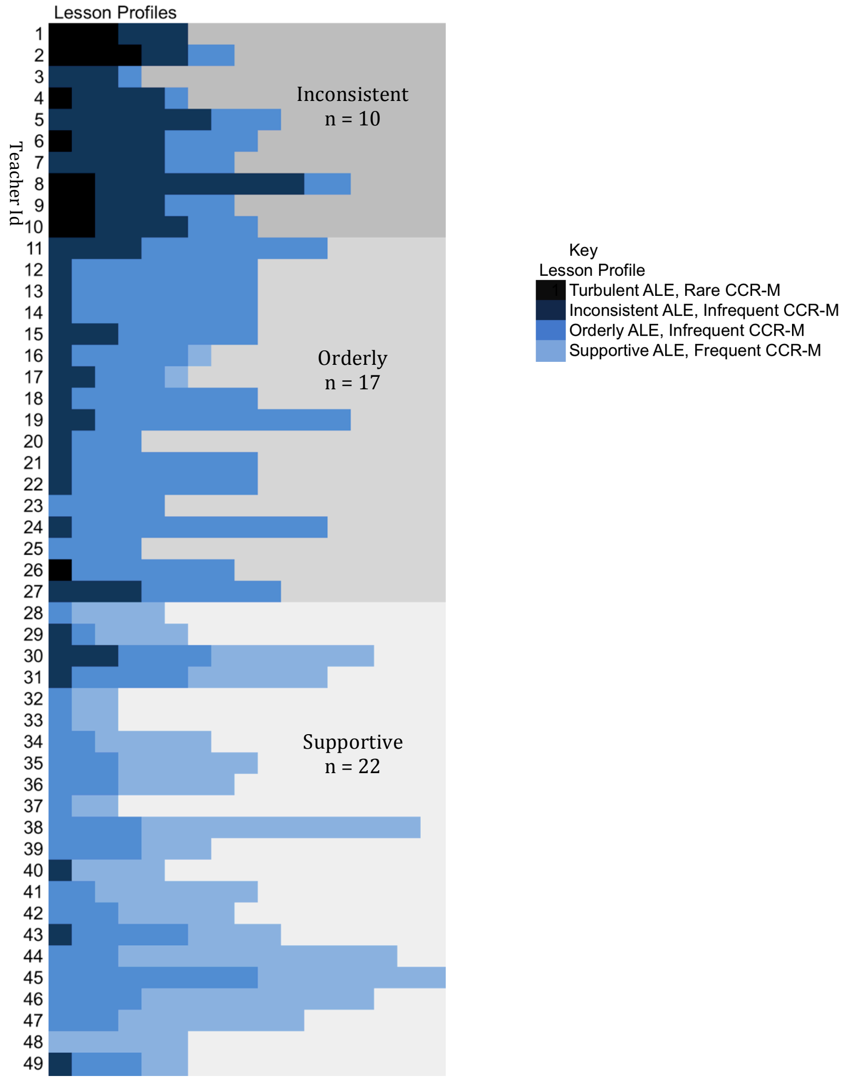


Figure 3. Distribution of lesson profiles within teachers and teacher groups.

Discussion

Interpreting Co-Occurrence of Facets of Classroom Learning Environments and CCR-Aligned Mathematical Engagement

Lesson-level findings. The first aim of this study was to identify characteristics of classroom learning environments where students do and do not engage with mathematics in the ways outlined in CCR standards. In accordance with an ecological perspective on child learning and development, these data suggest there may be more to a high-quality mathematics learning environment than whether or not it is orderly (Hamre & Pianta, 2010; Korpershoek et al., 2016, Marzano, et al., 2003; Doyle, 1977). In this sample, raters only observed consistent CCR-aligned mathematical engagement in classroom learning environments that were also safe, productive, emotionally supportive, and engaging. Raters did not observe consistent CCR-aligned mathematical engagement in classroom learning environments that were orderly, but not emotionally supportive or engaging. They also, as hypothesized, did not observe CCR-aligned mathematical engagement in turbulent or inconsistent classroom learning environments. The co-occurrence of particular characteristics of the classroom learning environment and CCR-aligned mathematical engagement provides early empirical evidence that, as others have theorized (Rimm-Kaufman & Youngs, in preparation), there may be some interplay between learning environments that support the development of social and emotional competencies and the ambitious goals for the ways students do mathematical work outlined in CCR standards.

One illustration of this potential interplay is that ratings from the IPRT-M show that, in this sample, in Turbulent, Inconsistent, and Orderly environments, students rarely shared their mathematical thinking with others. There could be a variety of reasons students chose not to share their mathematical thinking with their peers during these lessons. For example, teachers may not have provided students adequate scaffolds for discourse that would support collective mathematical reasoning (Kazemi & Stipek, 2001). Educational psychologists suggest there may also be characteristics of classroom environments—in particular affective elements that cause a student to not feel supported by their teacher or peers—that limit students' willingness to take academic risks such as sharing their ideas with others (Hawkins, 1997; Meyer & Turner, 2007; Turner & Meyer, 2004; Zins et al., 2007). It is possible similar dynamics are at play in the lessons in this sample given that the lessons profiles where students did not share mathematical thinking are characterized by lower-mid-range scores on the Positive Climate dimension of CLASS. These scores indicate there was only muted, perfunctory evidence of relationships, positive affect, positive communication, and mutual respect in these lessons.

This conjecture, that the emotional tenor of the classroom environment might be related to whether students engage in collective mathematical reasoning, may also explain differences between the Supportive CLE, Consistent CCR-M and the Orderly CLE, Infrequent CCR-M profiles. The Supportive CLE, Consistent CCR-M profile is most separated from the Orderly CLE, Infrequent CCR-M profile by its higher scores across the Emotional Support domain of CLASS and indicators of the IPRT-M that focus on students' mathematical discourse. These data cannot provide insight as to whether the classroom learning environments in this sample impacted students' willingness to share their developing mathematical thinking. These findings do, however, suggest that unpacking the co-occurrence of emotionally muted learning environments and an absence of collective mathematical reasoning is an important area for future research.

Findings from the Turbulent Environment, Rare CCR-M profile raise questions about the extent to which the classroom learning environment might be related to CCR-aligned mathematics teaching and learning in other ways. The co-occurrence of low Opportunities to Engage with Mathematics and Opportunities to Justify and Critique scores on the IPRT-M with low Behavior Management and Productivity scores on the CLASS for this profile indicates that in this sample, when students and teachers interact in ways that are chaotic or unpredictable, teachers offer limited opportunities for mathematical engagement. Prior research suggests this may be because concerns about the learning environment, particularly student engagement and behavior, *feel* pressing to teachers in a way that other instructional issues may not (Akdag & Haser, 2016; Meister & Melnick, 2003).

In this profile, there was little evidence of productive struggle (low scores on Depth, Prompting Student Thinking, and Responding to Student Misunderstanding). At the opposite end of the spectrum, in lessons sorted into the Supportive Environment, Frequent CCR-M profile, there was frequent evidence of facets of productive mathematical struggle (Granberg, 2016; Kapur, 2014; Warshauer, 2015). One interpretation of this finding is that it aligns with prior research that shows that in physically and emotionally unpredictable environments, students' resources for self-management may be diverted toward managing emotions and stress and/or controlling impulses, in general, and therefore are unlikely to be applied in service of academic goals (Maslow, 1943; Williams & Williams, 2011).

Again, these data provide no evidence that specific classroom learning environments inhibit or enable students' demonstration of social and emotional competencies in service of mathematical learning. Rather, they provide early empirical evidence of the convergence of CCR-aligned mathematical engagement and safe, productive, emotionally supportive classroom environments—a relationship that many researchers have hypothesized (Dana Center & CASEL, 2016; Rimm-Kaufman & Youngs, in preparation; Zins et al., 2007).

Teacher-level findings. The three distinct teacher groups present in this data reveal that in the lessons observed here, students in different classrooms had divergent access to positive, supportive learning environments. Researchers have found evidence of a reciprocal relationship between social and emotional learning and the classroom learning environment—a positive, supportive learning environment supports and reinforces students' social and emotional learning and vice versa (Zins et al., 2007). Zins and colleagues (2007) work would suggest that because of their divergent access to positive, supportive learning environments, students assigned to different teachers in this sample may have had meaningfully different opportunities to develop social and emotional competencies.

It is exciting that the largest group of teachers in this sample was classified as Supportive; these teachers' lessons were consistently characterized as safe, productive and often emotionally supportive and engaging. It is notable however, that for 55% of the teachers in this sample (teachers in the Inconsistent and Orderly groups), observers never or only rarely saw evidence of high levels of Emotional Support or Student Engagement. It is also noteworthy that in only two of the 125 lessons taught by teachers in the Inconsistent and Orderly groups, raters observed consistent CCR-aligned engagement. This suggests that not only did students in different classrooms have divergent opportunities to develop social and emotional competencies, they may have also had divergent opportunities to develop mathematical competencies.

These three distinct teacher groups make a case for differentiated supports for teachers. The fact that teachers in the Supportive group taught lessons that were categorized as a mixture of Orderly CLE, Infrequent CCR-M and Supportive CLE, Frequent CCR-M suggests teachers in the Supportive group may need help improving the consistency of emotional and CCR-aligned mathematical support across lessons. Given that there was almost no evidence of high levels of emotional support or CCR-aligned mathematical engagement in lessons taught by teachers in the Inconsistent and Orderly groups, these teachers may need more intensive supports to improve the quality of both general and mathematical interactions in their classrooms.

Importantly, neither these instructional nor teacher profiles could have been uncovered without the use of multilevel latent profile analysis. At the lesson level, only a lesson-centered approach would have surfaced the important distinctions between the Orderly and Inconsistent learning environments. For example, given that CCR-aligned mathematical engagement was almost identical during lessons with these two very different learning environments, a more traditional variable-centered approach that focused on correlations between observation ratings would have likely shown only weak associations between the classroom learning environment and CCR-aligned mathematical engagement. Thus, this more common variable-centered approach would not have uncovered that there were two substantively different profiles of learning environments that occurred during lessons with Infrequent CCR-aligned mathematical engagement in this sample. Methods such as multilevel latent profile analysis that identify common instructional patterns in lessons and their occurrence within and across teachers may prove especially beneficial to policy-makers and district personnel seeking to better understand a particular instructional landscape or to move beyond a one-size-fits-all model of teacher support.

These results would also not have been uncovered without the simultaneous use of a content-generic and mathematics-specific observation instrument. Indeed, had only the mathematics-specific instrument been used, the Inconsistent CLE, Infrequent CCR-M and Orderly CLE, Infrequent CCR-M profiles would have been indistinguishable. Similarly, if only a mathematics-specific tool or only a content-generic tool had been used to create instructional profiles, the exclusive co-occurrence of engaging, emotionally supportive learning environments and frequent CCR-aligned mathematical engagement in this sample would not have been uncovered. Given that researchers increasingly hypothesize that content-generic aspects of teaching—such as emotionally supportive classroom interactions—and content-specific aspects—such as those laid out in the CCR standards—are interrelated (Dana Center & CASEL, 2016; Rimm-Kaufman & Youngs, in preparation), researchers and practitioners alike may need systems of measurement and support that equally privilege both dimensions of practice.

Limitations

A major limitation of this study is the lack of outcome data linked to the lesson profiles and teacher groups found in this sample. An important next step in this line of work will be for researchers to examine the relationship between students' exposure to certain profiles of mathematics teaching and their mathematics learning.

In addition, because the school district that served as a site for this study has done extensive work to develop and retain high-quality teachers, the profiles here may not be representative of common styles of mathematics teaching in other districts. In 2009, this district implemented a unique high-stakes incentives-based evaluation system. The result of this is that some of the lowest performing teachers in the district, who might be most likely to struggle to create a supportive classroom environment, voluntarily exited or were let go (Adnot, Dee, Katz, & Wyckoff, 2017). Several groups of teachers that remained in the district substantially improved their performance (Dee & Wyckoff, 2015). In addition, since 2016, this district has enacted a district-wide intensive, CCR mathematics-focused weekly professional development and coaching program. Therefore, it is possible that high-quality learning environments and CCR-aligned mathematical engagement are overrepresented in this sample. Indeed, in this sample the largest group of teachers was the "Supportive" group (45% of the sample). It may be that in other contexts researchers would find larger numbers of teachers belonging to the "Inconsistent" profile, or find different instructional profiles altogether. This is an important area for future research.

Conclusion

Two clear findings emerged from these data. The first is that, in this sample, there was never evidence of consistent CCR-aligned mathematical engagement absent an engaging, emotionally supportive learning environment. The second is that, in the lessons observed here, students in different classrooms had substantively different opportunities to develop social, emotional, and mathematical competencies. While it is impossible in these data to parse the specific relationship between classroom learning environments and CCR-aligned mathematics teaching and learning, these findings suggest the two are related. In light of this relationship, advocates of CCR mathematics teaching and learning may want to consider the role of supportive classroom learning environments in efforts to support teachers in making the instructional shifts required for students to meet the ambitious goals outlined in CCR standards. Rather than continuing to conceptualize social and emotional learning and CCR standards as competing policy initiatives (Kochenderfer-Ladd & Ladd, 2016; Rimm-Kaufman et al., 2014), these findings suggest they may be deeply intertwined.

References

Adnot, M., Dee, T., Katz, V., & Wyckoff, J. (2017). Teacher turnover, teacher quality, and student achievement in DCPS. *Educational Evaluation and Policy Analysis*, *39*(1), 54-76.

Akdağ, Z., & Haser, Ç. (2016). Beginning early childhood education teachers' classroom management concerns. *Teachers and Teaching*, *22*(6), 700-715.

Alberti, S. (2012). Common core: Now what. *Educational Leadership*, *70*(4), 24-27.

Allen, J., Gregory, A., Mikami, A., Lun, J., Hamre, B., & Pianta, R. (2013). Observations of effective teacher–student interactions in secondary school classrooms: Predicting student achievement with the classroom assessment scoring system—secondary. *School Psychology Review*, *42*(1), 76.

Asparouhov, T., and Muthén, B.O. (2007). Multilevel mixture models. In G.R. Hancock and K.M. Samuelsen (Eds.), *Advances in Latent Variable Mixture Models*. Charlotte, NC: Information Age Publishing.

Baroody, A. E., Rimm-Kaufman, S. E., Larsen, R. A., & Curby, T. W. (2014). The link between responsive classroom training and student-teacher relationship quality in the fifth grade: a study of fidelity of implementation. *School Psychology Review*, *43*(1), 69-85.

Bausmith, J. M., & Barry, C. (2011). Revisiting professional learning communities to increase college readiness: The importance of pedagogical content knowledge. *Educational Researcher*, *40*(4), 175-178.

Becker, K. D., & Domitrovich, C. E. (2011). The conceptualization, integration, and support of evidence-based interventions in the schools. *School Psychology Review*, *40*(4), 582.

Ben-Avie, Michael, Norris M. Haynes, Jayne White, Jacque Ensign, Trudy R. Steinfeld, Loleta

D. Sartin, and David A. Squires. “Youth Development and Student Learning in Math and Science.” In *How Social and Emotional Learning Add Up*, ed. Norris M. Haynes, Michael Ben-Avie, and Jacque Ensign. New York: Teachers College Press, 2003.

Blazar, D., & Kraft, M. A. (2017). Teacher and teaching effects on students’ attitudes and behaviors. *Educational Evaluation and Policy Analysis*, *39*(1), 146-170.

Brock, L. L., Nishida, T. K., Chiong, C., Grimm, K. J., & Rimm-Kaufman, S. E. (2008). Children's perceptions of the classroom environment and social and academic performance: A longitudinal analysis of the contribution of the Responsive Classroom approach. *Journal of School Psychology*, *46*(2), 129-149.

Charney, R.S. (1993). *Teaching Children to Care: Management in the Responsive Classroom*. Greenfield, Mass.: Northeast Foundation for Children.

Cheema, J. R., & Kitsantas, A. (2014). Influences of disciplinary classroom climate on high school student self-efficacy and mathematics achievement: A look at gender and racial–ethnic differences. *International Journal of Science and Mathematics Education*, *12*(5), 1261-1279.

Cherng, H. Y. S., & Halpin, P. F. (2016). The importance of minority teachers: Student perceptions of minority versus white teachers. *Educational Researcher*, *45*(7), 407-420.

Cobb, P., & Jackson, K. (2011). Assessing the quality of the common core state standards for mathematics. *Educational Researcher*, *40*(4), 183-185.

Coburn, C. E., Hill, H. C., & Spillane, J. P. (2016). Alignment and accountability in policy design and implementation: The Common Core State Standards and implementation research. *Educational Researcher*, *45*(4), 243-251.

Cohen, J., Hutt, E., Berlin, R., & Wiseman, E. (under review). What is college and career ready teaching and can we measure it?.

Cohen, J., Ruzek, E., & Sandilos, L. (2018). Does teaching quality cross subjects? Exploring consistency in elementary teacher practice across subjects. *AERA Open*, *4*(3), 2332858418794492.

Common Core State Standard Initiative. (2010). *Standards for Mathematical Practice.* Retrieved from http://www.corestandards.org/Math/Practice/

Collaborative for Academic, Social, and Emotional Learning –CASEL (2005).What is SEL? Skills and competencies. Retrieved from http://www.casel.org/basics/skills.php

Cornelius-White, J. (2007). Learner-centered teacher-student relationships are effective: A meta-analysis. *Review of Educational Research*, *77*(1), 113-143.

Curby, T. W., Brock, L. L., & Hamre, B. K. (2013). Teachers' emotional support consistency predicts children's achievement gains and social skills. *Early Education & Development*, *24*(3), 292-309.

Curby, T. W., Grimm, K. J., & Pianta, R. C. (2010). Stability and change in early childhood classroom interactions during the first two hours of a day. *Early Childhood Research Quarterly*, *25*(3), 373-384.

Dana Center & CASEL. (2016). *Integrating social and emotional learning and the common core state standards for mathematics*. Retrieved from http://www.insidemathematics.org/assets/common-core-resources/social-emotional-learning/a__integrating_sel_and_ccssm_making_the_case.pdf

Denham, S. A., & Brown, C. (2010). “Plays nice with others”: Social–emotional learning and academic success. *Early Education and Development*, *21*(5), 652-680.

Dee, T. S., & Wyckoff, J. (2015). Incentives, selection, and teacher performance: Evidence from

IMPACT. *Journal of Policy Analysis and Management*, *34*(2), 267-297.

Dingman, S., Teuscher, D., Newton, J. A., & Kasmer, L. (2013). Common mathematics standards in the United States: A comparison of K–8 state and Common Core standards. *The Elementary School Journal*, *113*(4), 541-564.

Doyle, W. (1977). Paradigms for research on teacher effectiveness. *Review of Research in Education*, *5*(1), 163-198.

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.

Dymnicki, A., Sambolt, M., & Kidron, Y. (2013). Improving college and career readiness by incorporating social and emotional learning. Washington, DC: College & Career Readiness & Success Center at American Institutes for Research.

Egeberg, H. M., McConney, A., & Price, A. (2016). Classroom management and national professional standards for teachers: A review of the literature on theory and practice. *Australian Journal of Teacher Education*, *41*(7), 1.

Finch, W. H., & French, B. F. (2014). Multilevel latent class analysis: Parametric and nonparametric models. *The Journal of Experimental Education*, *82*(3), 307-333.

Flay, B. R., Allred, C. G., & Ordway, N. (2001). Effects of the Positive Action program on achievement and discipline: Two matched-control comparisons. *Prevention Science*, *2*(2), 71-89.

Frenzel, A. C., Pekrun, R., & Goetz, T. (2007). Perceived learning environment and students' emotional experiences: A multilevel analysis of mathematics classrooms. *Learning and Instruction*, *17*(5), 478-493.

Georgia Department of Education (2016). *Standards for Mathematical Practice.* Retrieved from https://www.georgiastandards.org/Georgia-Standards/Documents/CCSS_Standards_Math_Practice.pdf

Gest, S. D., Domitrovich, C. E., & Welsh, J. A. (2005). Peer academic reputation in elementary school: Associations with changes in self-concept and academic skills. *Journal of Educational Psychology*, *97*(3), 337.

Ginsburg-Block, M. D., Rohrbeck, C. A., & Fantuzzo, J. W. (2006). A meta-analytic review of social, self-concept, and behavioral outcomes of peer-assisted learning. *Journal of Educational Psychology*, *98*(4), 732.

Granberg, C. (2016). Discovering and addressing errors during mathematics problem-solving—A productive struggle?. *The Journal of Mathematical Behavior*, *42*, 33-48.

Good, T. L., & Grouws, D. A. (1977). Teaching effects: A process-product study in fourth-grade mathematics classrooms. *Journal of Teacher Education*, *28*(3), 49-54.

Halpin, P. F., & Kieffer, M. J. (2015). Describing profiles of instructional practice: A new approach to analyzing classroom observation data. *Educational Researcher*, *44*(5), 263-277.

Hamre, B. K., & Pianta, R. C. (2005). Can instructional and emotional support in the first-grade classroom make a difference for children at risk of school failure?. *Child Development*, *76*(5), 949-967.

Hamre, B. K., & Pianta, R. C. (2010). Classroom environments and developmental processes: Conceptualization and measurement. In J. L. Meece, & J. S. Eccles (Eds.) *Handbook of research on schools, schooling and human development* (pp. 43-59). Abingdon: Routledge.

Hamre, B. K., Pianta, R. C., Downer, J. T., DeCoster, J., Mashburn, A. J., Jones, S., . . . Hamagami, A. (2013). Teaching through interactions—Testing a developmental framework for understanding teacher effectiveness in over 4,000 U.S. early childhood and elementary classrooms. The Elementary School Journal, 113, 461–487. doi:10.1086/669616

Hannula, M. S. (2006). Motivation in mathematics: Goals reflected in emotions. *Educational studies in mathematics*, *63*(2), 165-178.

Hawkins, J. (1997) Academic performance and school success: Sources and consequences. In *Enhancing children's wellness* (pp. 278-305.), In R. Weissberg (Ed.). Sage: Thousand Oaks, CA.

Henry, K. L., & Muthén, B. (2010). Multilevel latent class analysis: An application of adolescent smoking typologies with individual and contextual predictors. *Structural Equation Modeling*, *17*(2), 193-215.

Hirsch, C. R., & Reys, B. J. (2009). Mathematics curriculum: A vehicle for school improvement. *ZDM*, *41*(6), 749-761.

Hodgkinson, H. (2001). Educational demographics: What teachers should know. *Educational Leadership*, *58*(4), 6-11.

Indiana Department of Education. (2016). *Mathematical process standards.* Retrieved from https://www.doe.in.gov/sites/default/files/standards/student-and-teacher-lookfors-administrators.pdf

Institute for Mathematics Education. (2007). *Progressions documents for the common core math standards*. Retrieved from http://ime.math.arizona.edu/progressions/

Kamin, D. C. (2016). The common core state standards for mathematics and college readiness. *The Mathematics Educator*, *25*(2).

Kapur, M. (2014). Productive failure in learning math. *Cognitive Science*, *38*(5), 1008-1022.

Kazemi, E., & Stipek, D. (2001). Promoting conceptual thinking in four upper-elementary mathematics classrooms. *Elementary School Journal*, 102(1), 59-80.

Keller, M. M., Becker, E. S., Frenzel, A. C., & Taxer, J. L. (2018). When teacher enthusiasm is authentic or inauthentic: Lesson profiles of teacher enthusiasm and relations to students' emotions. *AERA Open*, *4*(2).

Kern, L., & Clemens, N. H. (2007). Antecedent strategies to promote appropriate classroom behavior. *Psychology in the Schools*, *44*(1), 65-75.

Kiuru, N., Aunola, K., Lerkkanen, M. K., Pakarinen, E., Poskiparta, E., Ahonen, T., Poikkeus, A. M., & Nurmi. J.-E. (2015). "Positive teacher and peer relations combine to predict primary school students' academic skill development." *Developmental Psychology 51*(4), 434–446. doi: 10.1037/a0038911

Kochenderfer-Ladd, B., & Ladd, G. W. (2016). Integrating academic and social-emotional learning in classroom interactions. *Handbook of social influences in school contexts. Social-emotional, motivation and cognitive outcomes.* New York, NY: Routledge, 349-366.

Koestler, C., Felton-Koestler, M. D., Bieda, K., & Otten, S. (2013). *Connecting the NCTM process standards and the CCSSM practices*. Reston, VA: National Council of Teachers of Mathematics.

Korpershoek, H., Harms, T., de Boer, H., van Kuijk, M., & Doolaard, S. (2016). A meta-analysis of the effects of classroom management strategies and classroom management programs on students' academic, behavioral, emotional, and motivational outcomes. *Review of Educational Research*, *86*(3), 643-680.

Kunter, M., Baumert, J., & Köller, O. (2007). Effective classroom management and the development of subject-related interest. *Learning and Instruction*, *17*(5), 494-509.

Lazarides, R., Gaspard, H., & Dicke, A. L. (2019). Dynamics of classroom motivation: Teacher enthusiasm and the development of math interest and teacher support. *Learning and Instruction*, *60*, 126-137.

Lindsay, C. A., & Hart, C. M. (2017). Teacher race and school discipline: are students suspended less often when they have a teacher of the same race?. *Education Next*, *17*(1), 72-79.

Lockwood, J. R., Savitsky, T. D., & McCaffrey, D. F. (2015). Inferring constructs of effective teaching from classroom observations: An application of Bayesian exploratory factor analysis without restrictions. *The Annals of Applied Statistics*, *9*(3), 1484-1509.

Louisiana Department of Education. (2015). *K-12 Louisiana Student Standards for Mathematics*. Retrieved from http://www.louisianabelieves.com/docs/default-source/teacher-toolbox-resources/louisiana-student-standards-for-k-12-math.pdf

Magidson, J., & Vermunt, J. (2004). Latent class models. In D. Kaplan (Ed.), *Handbook of quantitative methodology for the social sciences* (pp. 175–198). Newbury Park, CA: Sage.

Mäkikangas, A., Tolvanen, A., Aunola, K., Feldt, T., Mauno, S., & Kinnunen, U. (2018). Multilevel latent profile analysis with covariates: Identifying job characteristics profiles in hierarchical data as an example. *Organizational Research Methods*, *21*(4), 931-954.

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.

Marzano, R. J., Marzano, J. S., & Pickering, D. (2003). *Classroom management that works: Research-based strategies for every teacher*. Alexandria, VA: ASCD.

Maslow, A. H. (1943). A Theory of Human Motivation. *Psychological Review*, 50(4), 370- 96.

McGrath, K. F., & Van Bergen, P. (2015). Who, when, why and to what end? Students at risk of negative student–teacher relationships and their outcomes. *Educational Research Review*, *14*, 1-17.

McLachlan, G., & Peel, D. (2000). Mixtures of factor analyzers. In *In Proceedings of the Seventeenth International Conference on Machine Learning*.

Meister, D. G., & Melnick, S. A. (2003). National new teacher study: Beginning teachers' concerns. *Action in Teacher Education*, *24*(4), 87-94.

Meyer, J. P., & Morin, A. J. (2016). A person-centered approach to commitment research: Theory, research, and methodology. *Journal of Organizational Behavior*, *37*(4), 584-612.

Meyer, D. K. & Turner, J. C. (2007). Scaffolding Emotions in Classrooms. In P. A Schutz & R Pekrun (Eds.), *Emotion in education* (pp. 235-249). San Diego: Academic Press.

Mishkind, A. (2014). Overview: State Definitions of College and Career Readiness. *College and Career Readiness and Success Center at American Institutes for Research*.

Nathan, M. J., Eilam, B., & Kim, S. (2007). To disagree, we must also agree: How intersubjectivity structures and perpetuates discourse in a mathematics classroom. *The Journal of the Learning Sciences*, *16*(4), 523-563.

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

Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2007). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study. *Structural Equation Modeling: A Multidisciplinary Journal*, *14*(4), 535-569.

Patrick, H., Anderman, L. H., & Ryan, A. M. (2002). Social motivation and the classroom social environment. In C. Midgley (Ed.), *Goals, goal structures, and patterns of adaptive learning* (pp. 85–108) Mahwah, New Jersey: Erlbaum.

Patrick, H., & Ryan, A. M. (2005). Identifying adaptive classrooms: Dimensions of the classroom social environment. In K. A. Moore & L. H. Lippman (Eds.), *What do children need to flourish? Conceptualizing and measuring indicators of positive development* (pp. 271–287). New York: Springer.

Patrick, H., Ryan, A. M., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology*, *99*(1), 83.

Pianta, R. C., Belsky, J., Houts, R., & Morrison, F. (2007). Opportunities to learn in America's elementary classrooms. *Science*, *315*(5820), 1795-1796.

Pianta, R. C., & Hamre, B. K. (2009). Conceptualization, measurement, and improvement of classroom processes: Standardized observation can leverage capacity. *Educational Researcher*, *38*(2), 109-119.

Pianta, R. C., Hamre, B. K., & Allen, J. P. (2012). Teacher-student relationships and engagement: *Conceptualizing, measuring, and improving the capacity of classroom interactions*. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), Handbook of research on student engagement (pp. 365–386). New York, NY: Springer.

Polikoff, M. S. (2015). How well aligned are textbooks to the common core standards in mathematics?. *American Educational Research Journal*, *52*(6), 1185-1211.

Porter, A., McMaken, J., Hwang, J., & Yang, R. (2011). Common core standards: The new US intended curriculum. *Educational Researcher*, *40*(3), 103-116.

Rathunde, K., & Csikszentmihalyi, M. (2005). Middle school students' motivation and quality of experience: A comparison of Montessori and traditional school environments. *American Journal of Education*, *111*(3), 341-371.

Reeve, J. (2006). Teachers as facilitators: What autonomy-supportive teachers do and why their students benefit. *The Elementary School Journal*, *106*(3), 225-236.

Reyes, M. R., Brackett, M. A., Rivers, S. E., White, M., & Salovey, P. (2012). Classroom emotional climate, student engagement, and academic achievement. *Journal of Educational Psychology*, *104*(3), 700.

Rimm-Kaufman, S. E., Fan, X., Chiu, Y. J., & You, W. (2007). The contribution of the Responsive Classroom Approach on children's academic achievement: Results from a three-year longitudinal study. *Journal of School Psychology*, *45*(4), 401-421.

Rimm-Kaufman, S. E., & Hulleman, C. S. (2015). Social and emotional learning in elementary school settings: Identifying mechanisms that matter. In J. Durlak & R. Weissberg (Eds.), *The handbook of social and emotional learning* (pp. 151–166). New York, NY: Guilford.

Rimm-Kaufman, S.E., Larsen, R.A.A, Baroody, A.E., Curby, T.W., Ko, M., Thomas, J.B., Merritt, E.G., Abry, T., DeCoster, J. (2014). Efficacy of the Responsive Classroom approach: Results from a 3-year, longitudinal randomized controlled trial. *American Educational Research Journal*, 51(3), 567-603.

Rimm-Kaufman, S., & Youngs, P. (in preparation). The need to integrate mathematics instruction and support for social-emotional learning.

Roth McDuffie, A., Drake, C., Choppin, J., Davis, J. D., Magaña, M. V., & Carson, C. (2017). Middle school mathematics teachers' perceptions of the Common Core State Standards for Mathematics and related assessment and teacher evaluation systems. *Educational Policy*, *31*(2), 139-179.

Ryan, R. M., & Deci, E. L. (2009). Promoting self-determined school engagement: Motivation, learning, and well-being. In K. R. Wentzel & A. Wigfield (Eds.), *Handbook of motivation in school* (pp. 171–196). New York: Taylor Francis.

Ryan, A. M., & Patrick, H. (2001). The classroom social environment and changes in adolescents' motivation and engagement during middle school. *American Educational Research Journal*, *38*(2), 437-460.

Sakiz, G., Pape, S. J., & Hoy, A. W. (2012). Does perceived teacher affective support matter for middle school students in mathematics classrooms?. *Journal of school Psychology*, *50*(2), 235-255.

Schleppenbach, M., Perry, M., Miller, K. F., Sims, L., & Fang, G. (2007). The answer is only the beginning: Extended discourse in Chinese and US mathematics classrooms. *Journal of Educational Psychology*, *99*(2), 380.

Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E. S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly, 18*(2), 158–176.

Schmidt, W. H., McKnight, C. C., Houang, R. T., Wang, H. C., Wiley, D. E., Cogan, L. S., et al. (2001). *Why schools matter: A cross-national comparison of curriculum and learning.* San Francisco: Jossey-Bass.

Schmidt, W. H., Wang, H. C., & McKnight, C. C. (2005). Curriculum coherence: An examination of US mathematics and science content standards from an international perspective. *Journal of Curriculum Studies*, *37*(5), 525-559.

Schoenfeld, A. H. (2015). Summative and formative assessments in mathematics supporting the goals of the common core standards. *Theory Into Practice*, *54*(3), 183-194.

Schoenfeld, A. H. (2004). The math wars. *Educational Policy*, *18*(1), 253-286.

Student Achievement Partners. (2018). *The Common Core Shifts at a Glance.* Retrieved from https://achievethecore.org/page/277/the-common-core-shifts-at-a-glance

Tainio, L., & Laine, A. (2015). Emotion work and affective stance in the mathematics classroom: the case of IRE sequences in Finnish classroom interaction. *Educational Studies in Mathematics*, *89*(1), 67-87.

Tennessee Department of Education. (2016). *Standards for Mathematical Practice.* Retrieved from https://www.tn.gov/content/dam/tn/education/standards/math/std_math_standards_mathemathema_practice.pdf

Turner, J. C., & Meyer, D. K. (2004). A classroom perspective on the principle of moderate challenge in mathematics. *The Journal of Educational Research*, *97*(6), 311-318.

Urdan, T., & Schoenfelder, E. (2006). Classroom effects on student motivation: Goal structures, social relationships, and competence beliefs. *Journal of School Psychology*, *44*(5), 331-349.

Warshauer, H. K. (2015). Productive struggle in middle school mathematics classrooms. *Journal of Mathematics Teacher Education*, *18*(4), 375-400.

Williams, K. C., & Williams, C. C. (2011). Five key ingredients for improving student motivation. *Research in Higher Education Journal*, *12*, 1.

Yu, H. T., & Park, J. (2014). Simultaneous decision on the number of latent clusters and classes for multilevel latent class models. *Multivariate Behavioral Research*, *49*(3), 232-244.

Zins, J. E., Bloodworth, M. R., Weissberg, R. P., & Walberg, H. J. (2007). The scientific base linking social and emotional learning to school success. *Journal of Educational and Psychological Consultation*, *17*(2-3), 191-210.

Zins, J. E., & Elias, M. J. (2007). Social and emotional learning: Promoting the development of all students. *Journal of Educational and Psychological Consultation*, *17*(2-3), 233-255.