ABSTRACT

KRUPA, ERIN ELIZABETH. Evaluating the Impact of Professional Development and Curricular Implementation on Student Mathematics Achievement: A Mixed Methods Study. (Under the direction of Drs. Jere Confrey and Allison McCulloch.)

In this era of high-stakes testing and accountability, curricula are viewed as catalysts to improve high school students' mathematics performances and a critical question is whether single subject or integrated curricula produce stronger student outcomes. This study was designed to investigate the effects of an integrated reform-based curriculum, *Core-Plus,* on student learning on statewide End of Course exams (EOC-Algebra I and II) and to contextualize these outcomes in a state-funded professional development program with the elements of a summer program, follow-up workshops, and monthly site based support with instructional coaches. The study was also designed to compare and contrast major sub-groups: teachers using *Core-Plus* who did or did not participate in different elements of the professional development. In addition, the study was designed to gather evidence on the variations among these groups on key implementation indices, and to use hierarchical linear modeling to investigate the role of these factors in predicting student outcomes.

Hierarchical linear modeling was used to account for the nesting of students within teachers within schools to investigate the impact of integrated mathematics and subject-specific curricular materials on student achievement across students in North Carolina. The sample was then restricted to students of teachers who participated in the state-funded professional development to relate teacher characteristics to student outcomes. The sample was further restricted to teachers who participated in different components of the professional development to analyze how curricular implementation effects student achievement and to examine factors that influenced decisions teachers made when implementing *Core-Plus.*

Findings from this study indicate that North Carolina students enrolled in integrated mathematics outperformed subject-specific students on the Algebra I End-of-Course exam, which was highly aligned with content in *Core-Plus* textbook, and performed no differently on the Algebra II exam, which was not aligned with the *Core-Plus* materials. There were favorable findings on the use of integrated mathematics in high minority, high need schools. Consistently prior achievement, student grade level and race, and classroom attendance were related to student achievement, as well as teacher content knowledge, most notably for teachers of *Core-Plus*.

This study documented large variance in teachers' implementation of *Core-Plus* based upon the professional development they received and their experiences using the materials with students. Factors that related to their implementation of the curriculum and related instructional practices included their beliefs about how students best learn mathematics, their trust of the curriculum, and systemic factors including mandatory state assessments, access to materials and technology resources, scheduling, and student transition to reform mathematics. Teachers enrolled in the summer workshops more faithfully implemented content from the textbook, but instructional coaches were an important component to facilitating change in teachers' instructional practices.

Results from this study demonstrate that teachers using *Core-Plus* need professional development designed to strengthen their mathematical content knowledge and reform-based instructional practices. Findings suggest encouraging results for the use of integrated mathematics with typically underserved student populations and among teachers who were provided with sustainable support following an authentic workshop experience.

Evaluating the Impact of Professional Development and Curricular Implementation on Student Mathematics Achievement: A Mixed Methods Study

by
Erin Elizabeth Krupa

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
requirements for the Degree of
Doctor of Philosophy

Mathematics Education

Raleigh, North Carolina

2011

APPROVED BY:

_____	_____
Jere Confrey Committee Chair	Allison McCulloch Committee Co-Chair
_____	_____
Karen Hollebrands	Ernie Stitzinger
_____	
Shevaun Neupert	

DEDICATION

To all who taught me along the way: family, friends, coaches, and teachers, especially those devoted to improving the quality of mathematics education.

BIOGRAPHY

Erin Elizabeth Krupa was born on April 18, 1980 in Greenville, North Carolina. She grew up in Greenville and graduated from D.H. Conley High School in 1998. She attended Elon University as a North Carolina Teaching Fellow and was in the Honors Program. Part of her Elon experience included studying abroad in London, England and learning the game of rugby. After graduating cum laude with a Bachelor of Science in Mathematics in 2002, Erin attended Wake Forest University in Winston-Salem, North Carolina. She earned her Master of Arts degree in Mathematics in 2004 and moved to Raleigh to begin her teaching career.

Erin taught at W.G. Enloe Magnet High School, a Gifted and Talented/International Baccalaureate Center for the Humanities, Sciences, and Arts from 2004 until 2008. At Enloe she taught Geometry, Algebra II, Pre-Calculus, and Introduction to College Mathematics. In 2008 Erin returned to graduate school to pursue a doctorate in mathematics education at North Carolina State University in Raleigh, North Carolina. She worked as a graduate research assistant for the North Carolina Integrated Mathematics Project during her three years in the program.

Erin has accepted an Assistant Professor position in the department of Mathematical Science at Montclair State University in Montclair, New Jersey. She will begin her job there in the fall of 2011.

ACKNOWLEDGEMENTS

I am sincerely grateful to my advisor and committee chair, Dr. Jere Confrey, for providing me with the opportunity to make this dream a reality and for providing timely and thoughtful feedback at every stage of this dissertation. Dr. Confrey was a wonderful mentor throughout this program; guiding me, yet giving me flexibility to grow as a researcher; answering my questions, yet encouraging me to think deeper. I am also tremendously grateful for my committee co-chair, Dr. Allison McCulloch, for her feedback, guidance, and encouragement throughout this process. I consider myself very lucky to have two wonderful faculty members work so closely with me on every facet of this study.

Thank you to my dissertation committee members, Dr. Karen Hollebrands, Dr. Ernie Stitzinger, and Dr. Shevaun Neupert, without your sound advice and prompt feedback this dissertation would not have come to fruition. I am especially thankful for the suggestions and availability of Dr. Shevaun Neupert on the design and analysis of this study. Also, for the comments Dr. Karen Hollebrands provided on earlier versions of this work and for the opportunity I had to work with her as a teaching assistant.

I would like to thank members of the NCIM project research team Dr. Jere Confrey, Dr. Alan Maloney, Dr. Shayla Thomas, and Drew Corley. I appreciate the great leadership of the NCIM project directors, Dot Doyle and Helen Compton. The workshop facilitators and instructional coaches worked tirelessly and with great enthusiasm throughout the project. Above all, I thank the teachers who participated in the study, especially those who welcomed me into their classrooms.

In addition, I am grateful for the staff and resources at the Friday Institute for Educational Innovation, especially the leadership and support of Dr. Glenn Kleiman. I also thank faculty members Dr. Karen Keene and Dr. Hollylynne Stohl Lee for motivating me to become a better student and researcher.

I am thankful for the support system and friendships created by my graduate school classmates. I thank Kenny for his tremendous resources and mentorship, Charity for always being a phone call away (with lots to talk about), Anthony for providing me with distractions when we needed a break, Richelle for her great library, Krista for always having an open ear when stopped by her desk to discuss a new idea, Cyndi for her encouragement, and Tina for sharing laughs with me throughout our coursework. I would also like to recognize the support that I received from Marrielle, Ayanna, Zuhal, and Nadia.

I am extremely grateful for the Raleigh Rugby Football Club and the three-time National Champion Raleigh Venom for giving me an outlet as I completed this degree. My teammates and coaches pushed me physically and mentally and being involved in such a wonderful organization gave me the strength and confidence I needed to push myself.

Above all, none of this would have been possible without the unconditional love and support from my family. My parents are wonderful and allowed me to follow my dreams, regardless of the sacrifices they had to make. My Mom instilled in me the importance of hard work, achieving my best, thoughtfulness, and balance. My Dad taught me the significance of utilizing hustle and humor throughout life. I appreciate the moral support from my sister, Kristin, and for the inspiration she gives to me through her work in special education.

Throughout this program I looked forward to my Wednesday morning phone call to Grandma and her words of encouragement and prayers. For Bop-bop, who passed away before the completion of this work, I promise I “returned to the original source” and “found a new topic to study.” I am thankful for the support and memories Aunt Jody gave me. I appreciate the encouragement from my extended family: Bob Henkel, Jason Burnette, and Ann and Bob Kitts. Marianne deserves the utmost appreciation for her daily devotion to me throughout this process. Her unwavering smile, positive energy, and belief in my abilities inspired and strengthened me throughout this journey.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1: INTRODUCTION

Statement of the Problem

As the nation prepares for the widespread implementation of the Common Core State Standards for Mathematics (CCSSM) (Common Core State Standards Initiative, 2010), districts across the United States must make decisions on curricular materials with consideration of their impact on student achievement. These decisions should be made as a result of evidence gathered from high quality curricular evaluation studies. However, it is an arduous task to link a curriculum directly to student learning when other confounding variables are involved in the complex setting of a classroom. Students and teachers interact with curricula in unique ways in creating the classroom milieu. These interactions are shaped by many conditions including, but not limited to, teacher and student beliefs, teachers' content and pedagogical knowledge, students' prior achievement, classroom resources, professional development experiences, and how curricula are implemented.

Curricular evaluations often seek to determine the comparative advantage between two sets of curricular materials based on student outcomes. However, outcomes measures vary and include high stakes tests (such as state or national assessments), tests specific to one curriculum, or tests with items common to both comparative curricula. Furthermore, these different forms of assessment can be formative and/or summative in nature. Even once a particular outcome measure is selected, higher outcome scores by students using a particular curriculum do not ensure that one curriculum is more effective than another. On the contrary, evaluators must also consider a number of other factors when conducting curricular

evaluations. To help researchers design and conduct evaluation studies the National Research Council (NRC) (2004) published *On Evaluating Curricular Effectiveness*, which provides a framework and recommendations for those seeking to evaluate curricula. They noted it was important to include: student characteristics (such as their beliefs about mathematics, socioeconomic status (SES), and prior achievement), measures of implementation, and teachers' exposure to professional development and institutional support.

The NRC (2004) also noted it is imperative to use effective experimental designs and appropriate statistical tests when evaluating curricula. Students cannot be treated as if they experience curricula independently because students are nested within classes, and classes are nested within schools. Further, determining causal inferences with randomized experiments is not always feasible in classroom settings (Brown, 1992). Sophisticated statistical techniques, such as hierarchical linear modeling (HLM), have allowed evaluators to appropriately analyze non-randomized nested data with the correct unit of analysis (Raudenbush & Bryk, 2002). HLM appropriately permits researchers to use classroom level data as the unit of analysis and avoids using the student as the unit (NRC, 2004).

In this era of high-stakes testing and accountability, many see mathematics curricula as catalysts to improve the mathematical skills of high school students. Determining curricular effectiveness is a complex problem, but one that is demanded by policy makers and the public. Information provided on standardized tests are widely accepted by the public (Chávez, Papick, Ross, & Grouws, 2010) and of major importance to school systems (NRC, 2004). However, there are unresolved questions about the effects curricular implementation

(M. K. Stein, Remillard, & Smith, 2007; M. K. Stein & Smith, 2010; Thompson & Senk, 2010) and professional development programs (Harwell, et al., 2007; Tarr, et al., 2010) have on curricular effectiveness. More important than picking a single curricula to implement in these complex settings, the real questions should be "what works, for whom, and under what conditions?" (Bryk, Gomez, & Grunow, 2011; Confrey & Maloney, 2011; Means & Penuel, 2005).

Purpose of Study

This research is an extension of a larger study known as the North Carolina Integrated Mathematics[1] (NCIM) project. The NCIM project was developed to create and support a community of teachers using the *Core-Plus Mathematics* (CPMP)[2] (Coxford, et al., 2001) integrated curriculum materials, particularly in high needs schools, and to study and evaluate the effectiveness of its use in these settings. The *Core-Plus* program integrates four strands (Algebra and Functions, Geometry and Trigonometry, Statistics and Probability, and Discrete Mathematics) across the first three years of high school mathematics, with an optional fourth year course, to prepare students for college mathematics (Hirsch, Fey, Hart, Schoen, & Watkins, 2008). Each year the strands are connected across units by fundamental themes to include visual thinking, recursive thinking, searching for and describing patterns, making and checking conjectures, reasoning with multiple representations, inventing mathematics, and

[1] The author is a member of the North Carolina Integrated Mathematics (NCIM) research group at North Carolina State University led by Dr. Jere Confrey and Dr. Alan Maloney. Former graduate student research assistants on the grant include Shayla Thomas and Drew Corely. The pronoun "we" will be used to acknowledge the work conducted by the NCIM research group and "our" to reference the aims of the project.
[2] CPMP will be used to refer to the Core-Plus Mathematics Project, and *Core-Plus* will be used to refer to both the first and second editions of the curriculum.

providing convincing arguments (Schoen, Cebulla, & Winsor, 2001; Schoen, Fey, Hirsch, & Coxford, 1999; Schoen, Hirsch, & Ziebarth, 1998; Schoen & Pritchett, 1998).

To prepare teachers to implement *Core-Plus*, in order to strengthen and invigorate STEM education at these schools, the NCIM project directors and evaluation team designed four components for the NCIM professional development. Each year, the four main elements of the professional development model were (1) a summer workshop providing in-depth education on use of curricular materials (one or two weeks), (2) a web-based environment supporting information exchange, (3) two face-to-face follow-up conferences, and (4) instructional coaches who visited each site monthly[3].

The researcher helped design and implement a professional development for teachers using the *Core-Plus* curriculum materials as part of her work on the (NCIM) project. Previous research found teachers' instructional practices have changed as they take part in the NCIM professional development (Krupa & Confrey, 2010, in press) and that teachers across the project provide their students with different opportunities to learn textbook material (Krupa, 2011). What is still unknown is whether there are differences in student achievement between students enrolled in integrated or subject-specific pathways and if and/or how the changes that were reported in instructional practice have an impact on student achievement. Further, we neither know nor understand how teacher participation in the NCIM professional development affects their implementation of the curricular materials. Therefore, this dissertation research had two goals: to provide insight into to the

[3] These components will be explained in more detail in the next chapter.

documentable effects of the *Core-Plus* curricular materials on student algebra achievement and to understand the impact the professional development model used in the NCIM project has had on student achievement and on teacher's implementation of *Core-Plus.*

Research Questions

Specifically, the research questions for this dissertation study are:

1. What evidence, if any, is there of significant differences in student achievement between students using integrated mathematics and subject-specific[4] curricula taught as Algebra I, Geometry, and Algebra II across the state of North Carolina?
    a. Is there evidence of differential impact on student performance on state tests between teachers using integrated mathematics curricular materials and those who use subject-specific curricula?
    b. For teachers who participated in the project workshops, is there evidence of differential impact on student performance on state tests between teachers using the *Core-Plus* curricular materials and those who use subject-specific curricula?
2. What evidence, if any, is there of significant differences on student achievement among teachers who participated in different components of the NCIM project?
    a. Among teachers using the *Core-Plus* curricular materials, what evidence of systematic differences is there on the effects of student learning based on

[4] Subject-specific mathematics curricula refer to the traditional sequence of mathematics courses taught in North Carolina: Algebra I, Geometry, Algebra II. Subject-specific and integrated mathematics are used throughout this paper, as opposed to traditional and reform mathematics, to refer to the curricular pathways instead of the instructional practices that may or may not accompany a curriculum.

varying levels of teacher participation (working with or without a instructional coach, frequency of professional development, type of professional development (content and/or pedagogy) in the NCIM project?

b. Among teachers using the *Core-Plus* curricular materials, to what extent and how can the variance in student outcomes be explained by teacher implementation?

The first question sought to determine if there were significant differences in student algebra achievement between integrated mathematics and subject-specific mathematics students, first with a wide lens of students and teachers across the state of North Carolina and followed by a more nuanced analysis from students of teachers involved in the NCIM summer workshops. Restricting the sample to teachers from the NCIM sample allowed the researcher to incorporate teacher background variables in the determination of curricular impact on student achievement. Further restricting the sample, the second research question provided a fine-grained analysis of student achievement for teachers using the *Core-Plus* curriculum. This provided a more intricate analysis of teachers' implementation and the influence their NCIM professional development experiences had on their implementation of *Core-Plus,* and subsequent student achievement.

Plan of Thesis

Following this introduction, in Chapter Two, a synthesis of literature justifying the need for this study will be presented on the effects of *Core-Plus* on student outcomes, evaluation of professional development, and evaluation of implementation fidelity. The

context of the research and the methods used to collect and analyze data will be discussed in Chapter Three. An analysis of the data and results from the two main research questions will be presented in Chapters Four and Five. Finally, Chapter Six discusses the overall findings from the study, explores the implications of the research, addresses limitations of the study, and makes recommendations for future research related to the study.

CHAPTER 2: REVIEW OF RELATED LITERATURE

This chapter summarizes literature that informed the development of the research in subsequent chapters. First, a brief background on the *Core-Plus* curriculum will be provided, followed by a detailed explanation of each component of the NCIM project. The NCIM project exists independently of the current study and the project history and development is relevant to a broader literature review. Then, a framework will be given to provide a structure for the in-depth literature review that follows.

The *Core-Plus* Curriculum

Historically students have typically received mathematics instruction directly from teachers as isolated content, unrelated to students' lived experiences (NCTM, 1989). The teacher was perceived as the sole authority in the classroom, and textbooks were used to provide explicit examples and problems for students. However, the introduction of the *Curriculum and Evaluation Standards for School Mathematics* (NCTM, 1989) led, the National Science Foundation (NSF) to issue calls for proposals to develop instructional materials to reform K-12 mathematics curricula to be more in line with the Standards (Senk & Thompson, 2003). Curricular reform efforts focused on strengthening the mathematical knowledge of all students and were guided by instructional practices that promoted problem solving, communication, reasoning, and creating mathematical connections (Senk & Thompson, 2003). In 1992 the NSF awarded five-year grants to four comprehensive high school curriculum projects (Schoen & Hirsch, 2003a; Schoen, et al., 1998), one of which is

the *Standards*-oriented curriculum *Contemporary Mathematics in Context* (Coxford, et al., 2001) developed by the *Core-Plus Mathematics Project* (CPMP).

Field-tested in 1994-1995, the CPMP program integrates Algebra and Functions, Geometry and Trigonometry, Statistics and Probability, and Discrete Mathematics into a three-year curriculum designed to meet the needs of all high school students, with an optional four year calculus preparation course (Hirsch, et al., 2008). Five design features of the curriculum were: 1) to have multiple connected strands, 2) an emphasis on mathematical modeling, 3) access for all students, 4) full use of graphics calculators, and 5) active learning experiences for students (Coxford & Hirsch, 1996). Another key tenet of the curriculum was to have students learn to regularly communicate mathematical ideas. (Schoen, Bean, & Ziebarth, 1996).

The United States Department of Education (1999) cited *Core-Plus* as an "exemplary" program and Strategic Ed Solutions (2008) reported it as a "program that works." The second edition of *Core-Plus* was released in 2008 and continued to build on the student-centered investigations of the first edition. It also incorporated a technology component with the addition of the *CPMP-Tools* software.

Context: The NCIM Project

The aim of the NCIM project was to educate teachers about the content and pedagogy of using integrated mathematics by creating a sustainable professional development model that could be replicated across the state and the nation. Spread throughout rural parts of the state, the seven partner schools were identified as low-performing (based on North Carolina

accountability measures). The average student population for the project schools for the 2009-2010 school year ranged between 110 to 292 students and the ethnic make-up consisted of: 0.80% American Indian or Asian, 4.66% Hispanic, 17.30% White, and 77.24% Black. Approximately 76.65% of students at each school qualified for free and reduced lunch.

The schools belonged to a collaboration called the North Carolina New Schools Project (NCNSP) and each had a focus on STEM (Science, Technology, Engineering and Mathematics) education. The NCNSP project is working to transform high schools across the state by implementing design principles created to ensure that all students are college, career, and life ready. The integrated system of school support services offered at these schools includes: on-site instructional coaching, leadership development for principals, professional development for teachers, and on-going support from NCNSP staff (North Carolina New Schools Project, 2011).

One aspect of the project was to assist rural schools in establishing statewide collaborations to address the challenges of isolation. Over the course of the three-year project, a total of 21 teachers participated in the NCIM project. However, due to teacher turnover, only a maximum of twelve teachers were actively involved in the project at any given time, some who were the only mathematics teacher at their school (Figure **1**).

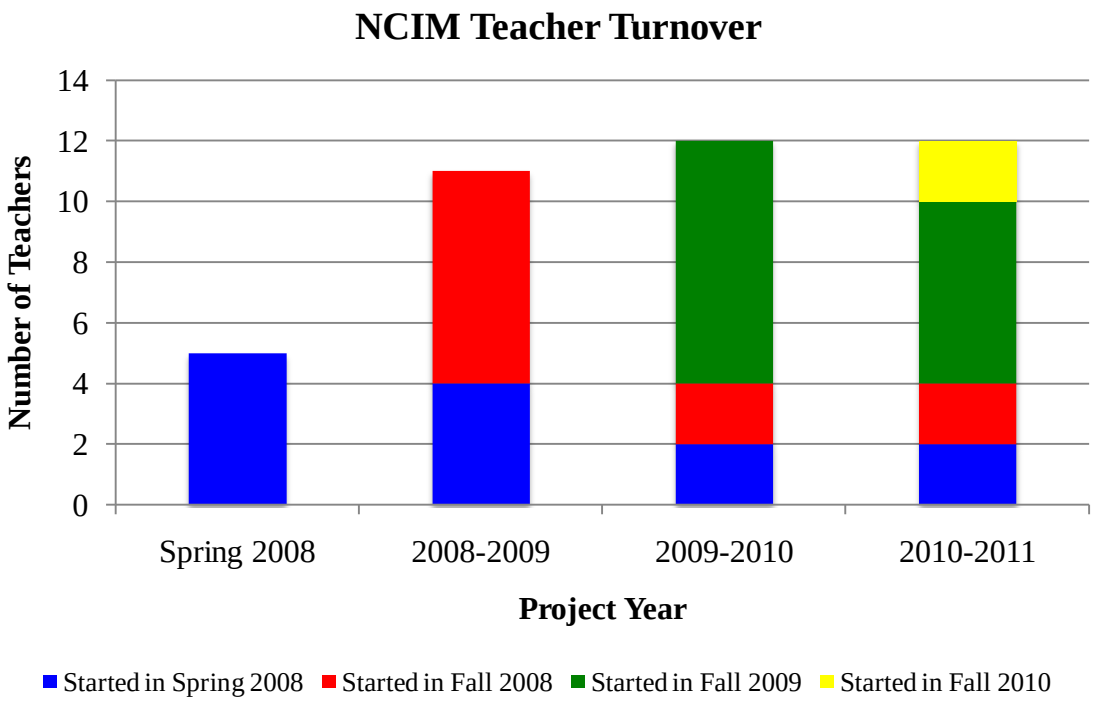


Figure 1. NCIM Teacher Participation Over the Course of Four School Years.

The goal of the overall professional development model was to educate teachers about the content and pedagogy of the *Core-Plus* curriculum and to provide them with the skills and knowledge to successfully implement the curriculum. Working together, the four main elements of the professional development model included: 1) a summer workshop providing in-depth education on use of curricular materials (one or two weeks)[5], 2) a web-based environment supporting information exchange, 3) face-to-face follow-up conferences, and 4) monthly instructional coach visits (Figure **2**).

[5] During each summer an additional fifty teachers also attended the workshop but did not receive the instructional coaches.

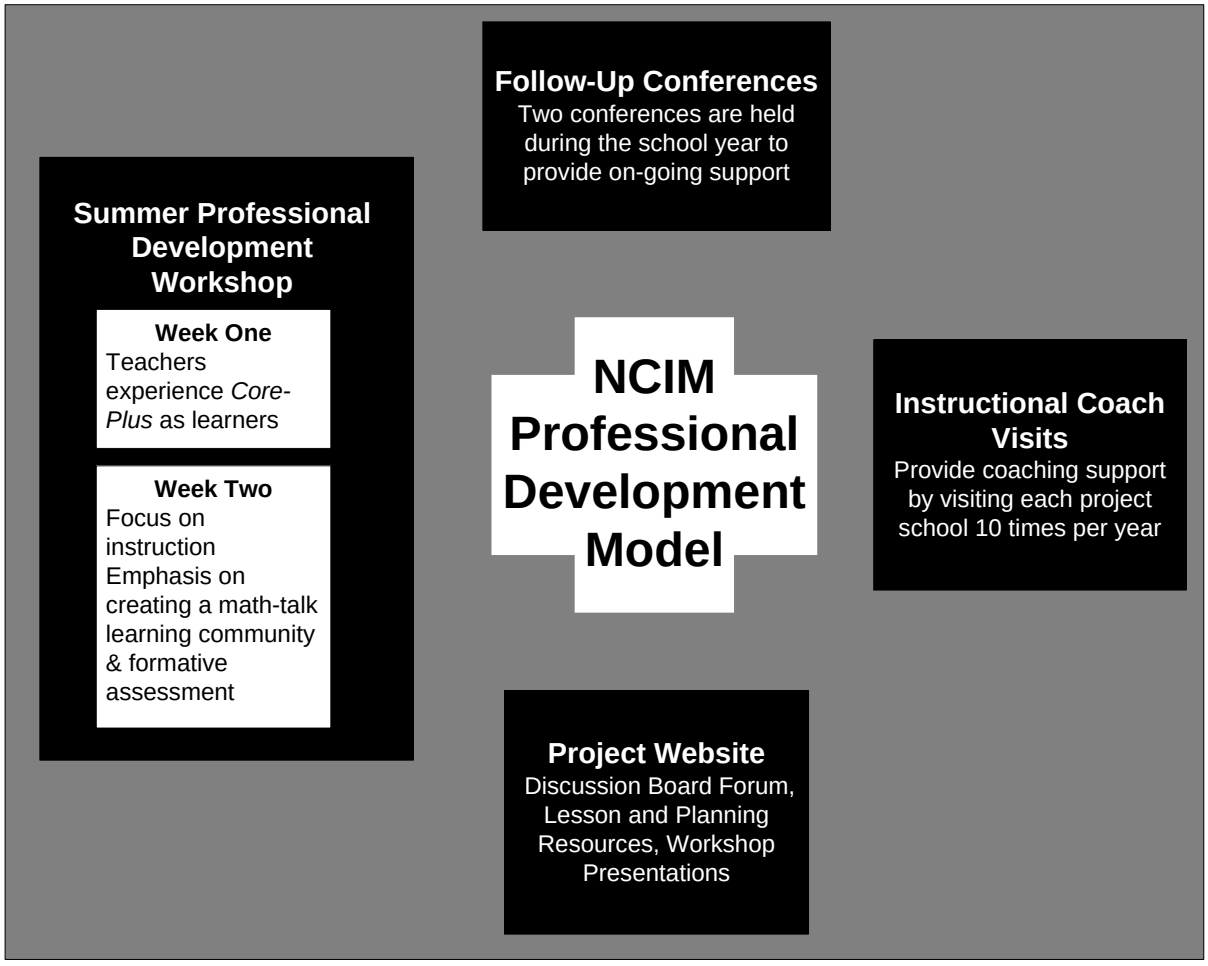


Figure 2. NCIM Professional Development Model.

Summer Workshops

One-week residential Summer Workshops were held in 2008, 2009, and 2010 and focused on using the curricular materials. At summer workshops, teachers were provided with first-hand opportunities to experience specific courses of the curriculum as students working through the investigations. These experiences enabled them to see the significance of the textbook structure and the contributions of small group work. Based on the evaluation of the 2008 workshop and reports from instructional coaches and teachers, a second week

was added, focusing on pedagogical considerations of teaching reform mathematics during the 2009 and 2010 Summer Workshop. During the academic year, 2008-9, data observation reports showed teachers were still using largely the same traditional instructional practices they experienced as students (Confrey, Maloney, & Krupa, 2008). As a result, the second week was designed to focus on how to promote discourse in the mathematics classroom.

Web-Based Learning Portal

To support communication and sharing of resources among the teachers, instructional coaches, project directors, and researchers, for the second year, a web-based learning portal was created. The web-based environment had resources for implementing *Core-Plus* and a discussion board that created a teacher communication network. This system enhanced the networking capabilities among teachers and instructional coaches.

Follow-up conferences

The third component of the model included follow-up conferences to the Summer Workshop, where teachers came back together for one-day in the fall and spring to reinforce and extend the learning that began in the summer. In total, there were four follow-up conferences. Commonly, principals attended these follow-ups in collaboration with teachers at their school; this provided principals insight on how to offer ongoing support to their teachers in reform mathematics classrooms. Topics for the follow-up conferences were selected based on the needs of the teachers, which were self-assessed by teachers and discussed among the project leadership. At the end of the Summer Workshop teachers evaluated the mathematics content of the curriculum, noted the pedagogical content areas

they wanted more exposure to, and created action plans for the upcoming year. By assessing the Summer Workshop, project directors created and improved follow-up conferences based on feedback from the teachers.

The Use of Instructional Coaches

The fourth component was comprised of monthly site visits from instructional coaches. Two experienced *Core-Plus* teachers conducted monthly site visits to identify and support teachers' needs. The NCIM coaching model, designed by the project staff, allowed the coaches to customize professional development during their visits for every NCIM teacher. The relationship between the coaches and teachers formed during the Summer Workshop when the instructional coaches acted as facilitators. When the school year started, the coaches determined and addressed the needs of each teacher. Coaches were responsible for contacting the teachers to arrange the site visits to observe and reflect with the NCIM teachers. Unlike other models, there was no prescriptive framework for these interactions, and the list of activities used to engage teachers was jointly determined by project staff and the instructional coaches (Table 1)[6].

[6] For more information on the coaching component reference: *Using Instructional Coaching to Customize Professional Development in an Integrated High School Mathematics Program* (Krupa & Confrey, 2012).

Table 1. Activities Instructional Coaches Engage in With Teachers

1. Offer feedback on class observation	2. Model teaching
3. Assist teacher in developing appropriate questioning strategies	4. Support teacher as students learn new vocabulary and read investigations
5. Define what a *Core-Plus* classroom looks like	6. Help teacher develop appropriate assessments
7. Offer guidance in developing pacing	8. Assist in lesson planning
9. Develop strategies to manage whole-class discourse	10. Determine appropriate rubrics for grading student work
11. Suggest appropriate homework for students	12. Suggest structure of daily lesson
13. Develop strategies to improve feedback to students	14. Keep in touch with teacher through email or phone conversations
15. Suggest strategies for managing groups during collaborative learning	16. Develop collaboration with other STEM teachers
17. Offer support in the use of technology	18. Encourage use of the NCIM website
19. Discuss the content or clarify content	20. Communicate needs of teacher
21. Provide resources to teacher	

Theoretical Considerations

In order to help researchers design and conduct evaluation studies, the National Research Council (NRC) (2004) published *On Evaluating Curricular Effectiveness,* which provides a framework and recommendations for those seeking to evaluate curricula. Primary components curricular evaluation should consider include program components, implementation components, and student outcomes **(**Figure **3**)[7]. It is also important to include secondary components such as systemic factors, intervention strategies, and unanticipated influences that may influence curricular implementation in curricular evaluations.

[7] Reprinted with permission from *On Evaluating Curricular Effectiveness*, by the National Academy of Sciences, Courtesy of the National Academy Press, Washington, D.C. (Appendix J), color added by author.

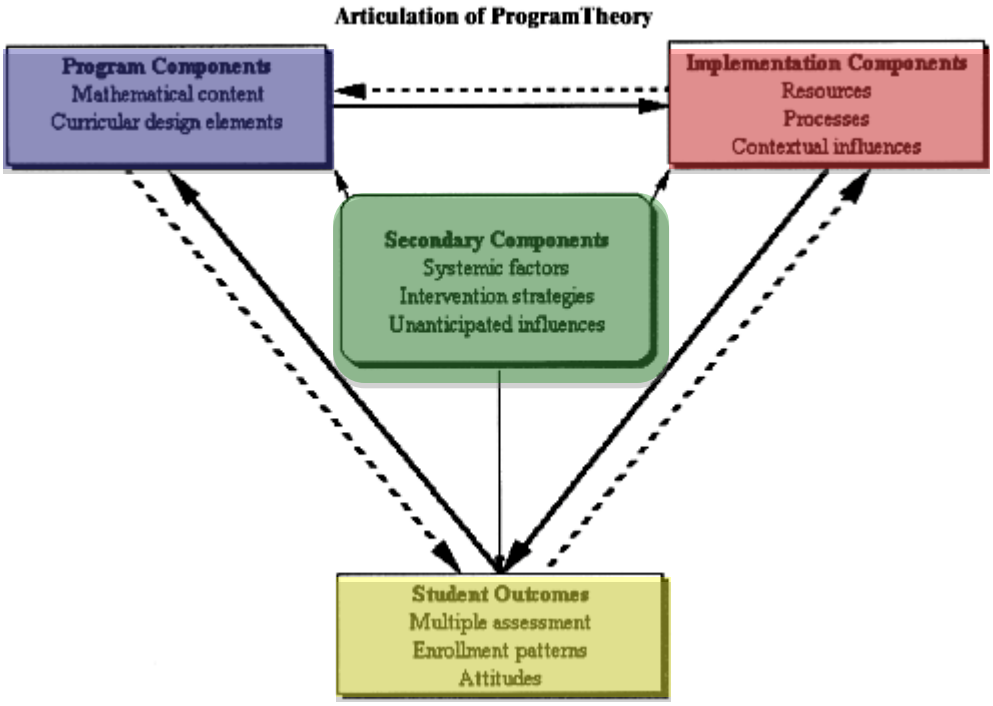


Figure 3. Primary and secondary components of mathematics curriculum evaluation (NRC, 2004, p. 40).

Program components should attend to the design elements the curriculum utilizes to present mathematics content. Often this is achieved through a content analysis of the curriculum. Implementation components include: resource variables, process variables, and community/cultural influences. Each of these implementation components influences the manner in which curricula are enacted and should be considered when making claims about curricular effectiveness. The NRC argued for multiple student outcome measures with data analysis disaggregated by subgroups to ensure equity and fairness. Tests are the primary tool used by curricula evaluators for measuring curricular effectiveness (NRC, 2004). The NRC pointed out to two types of student measures: curricular validity and curricular alignment with systemic factors. The former refers to "to the use of outcomes measures sensitive to the

curriculum's stated goals and objectives," while the later refers to "measures that facilitate student progress in the system, such as state tests, college entrance exams, and future course taking" (NRC, 2004, p. 49). It is important to document the type of outcome measures used in the evaluation.

Additionally, the NRC (2004) noted the importance of having independent evaluation teams, not affiliated with the textbook authors, conduct curricular evaluations. Further, no single study is sufficient for evaluating curricula, making it necessary to use multiple methods (NRC, 2004). As such, the NRC recommended:

> The collection of studies should use a combination of methodologies that meet these specified criteria: (1) **content analyses** by at least two qualified experts (Ph.D.-level mathematical scientist and a Ph.D.-level mathematics educator) (required); (2) **Comparative studies** using experimental or quasiexperimental designs, identifying the comparative curriculum (required); (3) one or more **case studies** to investigate the relationships among the implementation of the curricular program and the program components (highly desirable); and (4) a **final report**, to be made publicly available, should link the analysis, specify what they convey about the effectiveness of the curriculum, and stipulate the extent to which the program's effectiveness can be generalized (required). (p. 5)

Recommendations from the NRC report will be used to review results of prior analyses of the effectiveness of the *Core-Plus* curriculum to provide a foundation for the design of a next generation approach to evaluate this curriculum. Attending to the components in the NRC (2004) framework, the rest of this chapter will discuss effects of *Core-Plus* on student outcomes, evaluation of professional development (intervention strategy), and evaluation of implementation fidelity.

Evaluating the *Core-Plus* Curriculum

Numerous studies have been conducted on the effectiveness of the *Core-Plus* curriculum on student high school mathematics achievement. Typically these are comparative studies, seeking to determine if student achievement is significantly different between students using *Core-Plus* or subject-specific curricula (Huntley, Rasmussen, Villarubi, Sangtong, & Fey, 2000; Schoen, Cebulla, Finn, & Fi, 2003; Schoen, et al., 1998; Tarr, et al., 2010). The NRC (2004) presented ten elements that need to be addressed for comparative studies to determine the effectiveness of a curriculum (p. 165):

1. A better balance between experimental and quasi-experimental studies.
2. If a quasi-experimental design is selected, it is necessary to establish comparability.
3. Unit of analysis.
4. It is essential to examine the implementation components through a set of variables that include the extent to which the materials are implemented, teaching methods, professional development, teacher variables, etc.
5. Outcome data should include a variety of measures of the highest quality.
6. Statistical analysis should be conducted on the appropriate unit of analysis and should include more sophisticated methods of analysis.
7. Reports should include clear statements of the limitations to generalization of the study.
8. It is useful to report effect sizes.
9. Careful attention should also be given to the selection of samples of populations for participation.
10. The control group should use an identified comparative curriculum or curricula to avoid comparisons to unstructured instruction.

Comparative studies of *Core-Plus* date back to 1998, with the textbook developers evaluating the curriculum at field-test schools using t-tests, analysis of variance (ANOVAs), and occasionally an analysis of covariance (ANCOVA), with pretest as a covariate to compare mean student achievement (Schoen, et al., 1998). The field of curricular evaluation has developed since that time and more recent studies, conducted by external evaluators,

have increased the predictive power in modeling student achievement by including student- and classroom-level variables using Hierarchical Linear Modeling (HLM) (Harwell, et al., 2009; Post, et al., 2008; Post, et al., 2010; Tarr, et al., 2010). This section reviews *Core-Plus* evaluation studies of student achievement in algebra, focusing on how the evaluators treated the NRC (2004) recommendations for conducting comparative studies. Results will be synthesized across studies, attending to whether or not individual studies met or failed to meet the NRC recommendations, and will be presented in chronological order of their publication date[8].

In a comparative study of 11 field-test schools using the *Core-Plus* Course 1[9] materials to 20 matched comparison schools using subject-specific materials, CPMP developers and evaluators, Schoen, Hirsch, and Ziebarth (1998) administered a standardized test to *Core-Plus* and subject-specific students. The sample of subject-specific students was taken from field test CPMP sites, leading the researchers to assume comparability between the two groups. The standardized test was a subset of the *Iowa Tests of Educational Development* (ITED) test, the *Ability To Do Quantitative Thinking* (ITED-Q). The ITED-Q, a forty question multiple choice assessment, was made up of three subtests: Understanding Mathematical Concepts and Procedures, Interpreting Information, and Solving Problems, that assessed the ability to use mathematical reasoning in a variety of contexts. A comparison of means of post-test scores, using the school mean as the unit of analysis, found significant

[8] They will be presented chronologically except for one study, published in 2003, that is based on the same data as the first study, published in 1998. Hence, they will be talked about consecutively.
[9] Course 1, Course 2, and Course 3 will be used in reference to the *Core-Plus* curriculum.

differences favoring the *Core-Plus* schools over the traditionally sequenced schools on the ITED-Q for Course 1 students. Outcomes were not disaggregated by subgroups. Further, the authors failed to mention teacher professional development experience and did not measure curricular implementation. This study should be interpreted with caution due to the fact that many of the NRC (2004) recommendations were not addressed: comparability between groups was not assured, outcomes measures were not disaggregated, the statistical tests used have limited power, and there was no discussion for factors having an impact on implementation.

Five years later, in an extension of the same study, Schoen and Hirsch (2003b) disaggregated the matched comparison group into students enrolled in Prealgebra, Algebra, and Accelerated Geometry. This allowed them to match *Core-Plus* students with matched comparison students using ITED-Q pretest score, school, and gender. Establishing this comparability was an important component in the NRC (2004) recommendations. An additional outcome measure was also presented, Contextual Algebra I, Contextual Algebra II, and Procedural Algebra that were subtests of an open-ended *Core-Plus* post-test, created by the CPMP evaluation team. The tests were aligned with the *Core-Plus* curriculum and required students to explain their work. While the test was aligned with the *Core-Plus* curriculum, the authors claim it "was designed to be a test of content that both CPMP and comparison students would have had an opportunity to learn that year" (Schoen & Hirsch, 2003b, p. 9). The NRC recommends multiple outcome measures, however without collecting information on the curriculum used by comparison teachers the authors cannot ensure that

students had the opportunity to learn the assessed content. Further, even assuming comparison students were exposed to the content, there is no assurance they were exposed to instructional practices requiring them to explain their work in an open-ended format. Using this design, Course 1 students significantly outperformed all subject-specific matched comparison groups on contextual algebra problems (Table **2**.), but they were outperformed in almost every comparison involving procedural algebra skills (Schoen & Hirsch, 2003b). Entries that are blank in Table **2** did not statistically favor either student group.

Table 2. Statistically Significant Results from Course 1

Comparison	ITED-Q*	Interpret Information	Solving Problems	Context Algebra I	Context Algebra II	Procedural Algebra
Pre-Alg. vs. Course 1	Course 1	Course 1	Course 1	Course 1	Course 1	
Algebra vs. Course 1	Course 1	Course 1		Course 1	Course 1	Algebra
Accelerated Geo. vs. Course 1				Course 1	Course 1	Acc. Geo.
All Comparison vs. Course 1	Course 1	Course 1	Course 1	Course 1	Course 1	Comparison

Note. Statistically Significant Results from Course 1. * Iowa Test of Educational Development, Ability to Do Quantitative Thinking. Data compiled from Schoen and Hirsch (2003b).

Schoen and Hirsch (2003b) are currently the only researchers to analyze student achievement data from Course 2. Results from a similar matched comparison of Course 2 students to subject-specific students are shown in Table **3**. By the end of Course 2, the only matched group to outperform Course 2 students were students enrolled in Accelerated Advanced Algebra, a result that was statistically significant on procedural algebra problems. This improvement in the comparative performance of *Core-Plus* students suggests that over time they seem to be closing the gap in symbol manipulation seen in Course 1.

Table 3. Statistically Significant Results from Course 2

Comparison	ITED-Q	Interpret Information	Solving Problems	Context Algebra	Procedural Algebra
Algebra vs. Course 2	No Statistically Significant Differences			Course 2	
Geometry vs. Course 2				Course 2	
Acc. Adv. Alg. Vs. Course 2				Course 2	Acc. Adv. Alg.
All Comparison vs. Course 2				Course 2	

Note. Data compiled from Schoen and Hirsch (2003b).

Schoen and Hirsch (2003b) improved their prior study by accounting for the comparability of their groups using pre-test score and incorporating an additional outcome measure, but there are still several weaknesses with the design of their evaluation. First, they report means and standard deviations by comparison group, using t-test for basic statistical testing, but do not report effect sizes or disaggregate the sample among subgroups (gender, race, school setting). Further, they do not report other contributing factors to student achievement, such as socioeconomic status (SES), gender, or race. Finally, the authors noted that at least one *Core-Plus* teacher from each school attended a 2-week workshop prior to teaching *Core-Plus*, however this is not adequate for ensuring proper implementation (to be reported in subsequent sections). Having one teacher attend a workshop does not ensure the instructional practices and content knowledge learned at the workshop permeate to all teachers implementing *Core-Plus* within a school. Again, Schoen and Hirsch failed to quantify professional development experience or measure teachers' implementation of the materials.

In another older study, a mix of internal and external evaluators, Huntley, Rasmussen, Villarubi, Sangtong, and Fey (2000), conducted a comparative study of subject-specific students and students who completed the *Core-Plus* sequence at field-test schools. Due to the

integrated approach of *Core-Plus*, it is important to look at student achievement after students have been exposed to all three courses. They used multiple outcome measures, an open-ended contextualized problem solving assessment, a multiple-choice and constructed response context-free assessment focusing on symbol manipulation, and an open-ended contextual collaborative work assessment. A strength of their study was the use of both tests of curricular validity and curricular alignment with systemic factors, however the tests of curricular validity were solely "typical of CPMP units and other reform curricula" (Huntley, et al., 2000, p. 336) and not of the comparative curricula or mode of instruction.

Huntley, et al. (2000) used students as the unit of analysis, which is not recommended in curricular evaluations because students experience curricula in nested systems; within classrooms within schools (NRC, 2004). T-tests were used to compare means between groups. There was no assurance of comparability between the groups, though the authors "became convinced that our various strategies for assuring comparable CPMP and control-student pools at the various sites were in fact providing groups with comparable entering mathematical aptitude and achievement" (p. 340). Researchers should not have to "become convinced" of comparability, but arrive at comparability through statistical analysis. Further, prior achievement is only one mode for assuring comparability, race, gender, and SES are also important indicators for ensuring groups are comparable.

They noted a criterion for a schools' inclusion in the study was the "approximate implementation" of *Core-Plus* following the recommended conditions; though they did not state how compliance was ensured nor did they include a measure of implementation in their

quantitative data collection (Huntley, et al., 2000). They did collect interview data, with teachers of both groups, "to describe variability inherent in different implementations of the curricula" (p. 335). The addition of this data enhanced the overall study, but there were significant flaws in the design that this qualitative data could not compensate for.

Given the weaknesses in the previous studies, the summarized results (Table **4**.) from Huntley, et al. (2000) and Schoen and Hirsch (2003b) show that Course 3 students' achievement was higher on problems with context, modeling, or the use of calculators, while subject-specific students outperformed *Core-Plus* students on symbol manipulation and problems without the use of the calculator. *Core-Plus* students were able to solve the same problems in a variety of ways, many of which were calculator-based strategies.

Table 4. Results on Algebraic Problems From End of Course 3

Type of Assessment	Control/CPMP
Applied Algebra with Calculator	CPMP**
Open-ended Applied Algebra with Calculator	CPMP*
Algebraic Symbol Manipulation	Control**
Set-up Algebraic Models & Interpreting Results of Algebraic Calculation	CPMP***
Algebraic Calculations without Context or Calculator	Control**
Algebraic Calculations in Context with Calculator	Not significant

Note. Data compiled from Huntley, Rasmussen, Villarubi, Sangtong, & Fey, (2000); Schoen & Hirsch, (2003b); Schoen, Hirsch, & Ziebarth (1998).
*$p < .01$, **$p < .001$, p<.0001.

A more sophisticated study, undertaken by the Comparing Options in Secondary Mathematics: Investigating Curriculum (COSMIC) project, addressed issues with prior curricular evaluations. Tarr and colleagues (2010) recently conducted a quasi-experimental study, comparing Course 1 and Algebra I student achievement on multiple outcome measures. Participating schools offered parallel curricular paths and did not track students

into a particular pathway based on demographic data or prior achievement. In accordance with the NRC's (2004) recommendation, the COSMIC team was an independent evaluation team, not affiliated with the textbook authors.

Unique features of the COSMIC group's evaluation include using a HLM with student-level variables (prior achievement, gender, race/ethnicity, Individual Education Plan, and Limited English Proficiency) and nearly thirty teacher-level variables. Tarr and colleagues (2010) used Principal Components Analysis (PCA) to reduce the number of teacher variables into the following seven factors: standards-based instruction (SBI), implementation fidelity (Fidelity), technology and collaborative learning (T&C), opportunity to learn (OTL), knowledge of the NCTM Standards (KoS), experience (Exp), and professional development (PD). Each of these factors was used to model student achievement on multiple outcome measures.

The COSMIC team used two tests of curricular validity and one test of curricular alignment with systemic factors to measure student achievement. First, they engaged in an extensive process to create two measures of student achievement: a fair test and a reasoning and problem solving test. Chavez, et al. (2010) reported the tests they created were "*fair* tests because the tests were developed with the deliberate goal of not being biased towards either of the two curriculum programs studied" (p. 3). The Iowa Test of Educational Development (ITED): Mathematics: Concepts and Problem Solving standardized test was selected as a measure of curricular alignment with systemic factors in their evaluation of the *Core-Plus* curriculum.

Preliminary findings show the *Core-Plus* curriculum was significantly positively correlated to student mean residualized gain scores on the reasoning and fair tests, while the scores on the ITED-15 were not significantly related to curriculum type. The percentage of Free/Reduced Lunch (%FRL) was significantly negatively correlated with the fair and ITED-15 tests. OTL was significantly positively correlated to all three outcomes measures, meaning that students provided with more of an OTL had higher performance scores regardless of the assessment.

Tarr and colleagues (2010) presented additional HLM analyses controlling for different variables or interaction terms. A compilation of the correlations between curriculum type and student outcomes, when controlling for these variables is given in Table **5**. Based on the effects of partialling out different variables, the impact of curriculum type becomes more noticeable. First, note that since none of the correlation coefficients are negative, there were no perceived benefits to the subject-specific curriculum over the integrated curriculum. In other words, students of the *Core-Plus* curriculum performed as well as or better than the comparative students on each of these student outcome measures. When controlling for the %FRL there was a significant relationship between curriculum type and student achievement on all three assessments in favor of the integrated curriculum. This suggests that when %FRL was held constant across all schools, integrated mathematics students have higher achievement scores on fair, reasoning, and standardized assessments. Further, controlling for %FRL substantially reduced the magnitude of the OTL effect, signifying that %FRL and OTL seem to be closely related. When controlling for OTL, the magnitude of each

correlation between curriculum type and student outcomes increased and became statistically significant. In other words, when OTL was held constant across classrooms, students of integrated curriculum significantly outperformed students of comparative curriculum on all three assessments. Further, the effect was positive suggesting the more OTL a student was provided with, the higher their performance on these three outcome measures. The close relationship between %FRL and OTL suggests that students are provided with different opportunities to learn mathematics content based on the classroom %FRL population. These are important measures to include when measuring the impact of curriculum on student learning.

Table 5. Correlations Between Student Outcomes and Curricular Type Based on Different Analyses

		All Variables Included	Effects When Controlling for:				
			%FNL	T&C	OTL	Interaction Curr & T&C	Interaction Curr & %FRL
		Curr Type	Curr Type	Curr Type	Curr Type	Curr Type	Curr Type
Fair Test	*r*	.304	.459	.196	.411	.202	.399
	p	.038*	.001**	.192	.005**	.178	.006**
Reasoning	*r*	.518	.646	.398	.638	.438	.604
	p	.000***	.000***	.006**	.000***	.002**	.000***
ITED-15	*r*	.264	.404	.053	.335	.149	.320
	p	.073	.005**	.726	.023*	.323	.030*

*$p < .05$. **$p < .01$. *Note*. Compiled from Tarr, et al. (2010).

Collectively, data from the previous sections suggest that *Core-Plus* students outperform most subject-specific students in quantitative thinking, contextual algebraic problems, and interpreting information, highlighting its effectiveness as a reform curriculum. However, caution should be used when generalizing these results because student

populations contained low percentages (less than 7%) of minority students (Schoen & Hirsch, 2003b; Tarr, et al., 2010) or did not adequately discuss student demographics (Huntley, et al., 2000; Schoen, et al., 1998). Further, only the COSMIC group considered measures of implementation fidelity and professional development experience as classroom-level predictors for student achievement.

Using the work conducted by Tarr, et al. (2010) as a guide, studies are needed that add to the field of curricular evaluation by linking student achievement to teacher and student variables while controlling for student characteristics. Also, only few studies examined student achievement for students exposed to the second edition of the text. Based on this review, studies are needed which involve non field-test sites, serve a high percentage of minority students, and are conducted by external evaluators. It is important to study non field-test sites to represent the impact of the curricula in schools that do not have access to the expertise of the textbook authors and the CPMP professional development. As presented in the next two sections, curricular evaluations should also incorporate measures for professional development and implementation into the analyses, which are currently lacking in most curricular studies.

Evaluating Professional Development

It is difficult to discuss the impact of a curriculum when teachers are not provided with necessary support. We cannot expect curricula to be tossed into the hands of teachers without a structure for supporting them in their use of the materials. The *Principles and Standards for School Mathematics* (NCTM, 2000) recommend teachers take part in

"ongoing, sustained professional development" (p. 369). While studies of curricular evaluation may report on professional development experiences, and the NRC (2004) argues for the importance of gathering such data, few provide adequate detail to position those treatments in light of research on effective professional development. It is important to both understand what the field knows about effective professional development and how this subsequently impacts student achievement. As a result, this section is split into research on components of effective professional development and on studies that have used professional development as a predictor for student achievement.

Components of effective professional development

Teachers typically use the same instructional practices used by their teachers (Ball, 1988; Tyack & Cuban, 1995), and in order for them to change their instructional practices to inquiry-based or reform instruction they need ongoing and sustainable support (Ball & Cohen, 1999; Loucks-Horsley, Love, Stiles, Mundry, & Hewson, 2003; Putnam & Borko, 1997; S. M. Wilson & Berne, 1999). Numerous case studies show that when teachers using the *Core-Plus* materials are provided with support, their beliefs change as they transition from teaching traditional mathematics to *Core-Plus*, however the transition takes time (Arbaugh, Lannin, Jones, & Park-Rogers, 2006; Lloyd, 1999, 2002; M. S. Wilson & Lloyd, 2000). The NRC (2004) contends teachers need adequate professional development *before implementing* new curricular materials, continued support *while implementing*, and time for reflection *during and after implementation* (p. 46). The rest of this section identifies effective

professional development practices in each of those three areas: workshop offerings (before), ongoing support and coaching (while), and collaboration and resources (during and after).

Workshops. Teachers implementing reform materials for the first time need first-hand opportunities to experience the curriculum as students working through the investigations (Confrey, et al., 2008; Darling-Hammond & McLaughlin, 1995; Lloyd, 1999; Schoen & Hirsch, 2003b; Ziebarth, 2001). These experiences enable them to see the significance of the textbook structure and the contributions of small group work. These opportunities also improve and strengthen teachers' mathematical content and pedagogical knowledge (Loucks-Horsley, Stiles, Mundry, Love, & Hewson, 2010). Workshops can also be designed to focus on how to promote discourse in the mathematics classroom (Herbel-Eisenmann, 2009; Hufferd-Ackles, Fuson, & Sherin, 2004; Nathan & Knuth, 2003; National Council of Teachers of Mathematics, 1991; Sherin, 2002; C. C. Stein, 2007; White, 2003).

On-going support and coaching. Ball (1996) states, "the most effective professional development model is thought to involve follow-up activities, usually in the form of long-term support, coaching in teachers' classrooms, or ongoing interactions with colleagues" (pp. 501-502). This customized mentoring is responsive to individual teacher needs (Darling-Hammond, 1997). Loucks-Horsley's et al. (2010) provided five key elements of coaching: *(1)* teachers focus on learning or improvement; *(2)* a climate of trust, collegiality, and continuous growth is cultivated; *(3)* coaches are well prepared with in-depth content knowledge; *(4)* mechanisms for observing practice and providing feedback are critical; and *(5)* opportunities for interaction are provided.

Through an analysis of research on different types of teacher training, Joyce and Showers (2002) found that training programs which incorporate a coaching component are more effective in increasing teachers' knowledge, skills, and transfer to practice than trainings without such a component. In particular they advocate for peer coaching or trainer coaching, but contend that hiring trainers is not practical in most settings and that peer coaching is aimed at school-wide implementation of innovations through school-level collaborations. Table 6 displays the percentage of workshop participants who attained an adequate understanding of a new concept based on different training combinations: study of theory, demonstration, practice, and coaching. Gains in knowledge acquisition were determined by giving pre- and post-tests to participants in a workshop setting. Skill percentages were measured by having workshop participants demonstrate the intended skill set during simulations. The table shows that when teachers were only provided theory, demonstration, or practice alone, they did not transfer the knowledge or skill into their practices. When peer coaching was incorporated with teacher training, the percentage of teachers transferring the new knowledge and/or skill into practice skyrocketed to 95%.

Table 6. Training Components and Attainment of Outcomes in Terms of Percent of Participants

Component	Knowledge	Skill	Transfer
Study of Theory	10	5	0
Demonstration	30	20	0
Practice	60	60	5
Peer Coaching	95	95	95

Note. Adapted from Joyce and Showers (2002).

One type of coaching model implemented with success is cognitive coaching (Costa & Garmston, 1994). This type of professional development consists of pre- and post-observations between the coach and teacher to discuss their initial concerns about a particular lesson, observation of the lesson with specific data collection, and analysis of the data by the teacher and coach in a post-observation reflection. Olson and Barret (2004) expanded cognitive coaching with additional strategies to improve teachers' instructional practices using additional activities: rich mathematical tasks (M. K. Stein & Smith, 1998), model instruction (Becker, 2001), co-teaching (Showers & Joyce, 1996), and reflection (Schon, 1987). While these activities have been utilized by coaches to improve teachers' instruction, evidence on the effectiveness of these activities is thin (Cobb, Box, & Jackson, under review) and is not often linked to student performance data.

Collaboration and resources. Encouraging professional communication can increase the motivation teachers have to continually improve their instructional practices (Garet, Porter, Desimone, Birman, & Yoon, 2001; Lieberman & McLaughlin, 1992) and foster collegiality among colleagues (Loucks-Horsley, et al., 2010). In a study of teachers experiences implementing *Core-Plus* following participation in the NCIM project, Thomas (2010) argued for the importance of teacher collaboration with both experienced and less-experienced teachers. She noted, "They should seek out more experienced teachers for assistance, but also maintain communication with less-experienced teachers to share frustrations and successes, so that they realize their experiences are not unique to them" (p. 183). Thomas also discussed the benefits of having access to an on-line community, where teachers can go throughout the year to share strategies and resources.

While it is apparent that ongoing support, before, during and after implementation, can help teachers change their practice, it is yet to be determined how this type of support has an impact on the implementation of the curricular materials, and subsequently student achievement.

Linking professional development to student achievement

The NRC (2004) recommends that curricular evaluations include data on professional development duration, timing, and type received. They found only 1.5% of the 63 studies that were deemed *minimally methodologically adequate*[10] reported and adjusted for a measure of professional development in their curricular evaluations, and 71.5% did not mention professional development at all. As previously mentioned, it is unusual for *Core-Plus* curricula evaluators to include a measure of professional development when modeling student achievement. Often this results in researchers making broad claims about the professional development teachers received, without quantifying it in for inclusion in models of student achievement. For example, Schoen and Hirsch (2003b) claim teachers at the field-test schools received professional development prior to teaching *Core-Plus*, but this neither ensured the professional development had an impact on their instruction nor provides evidence about which aspects of the professional development were beneficial to the teacher.

For those studies that did include a measure for professional development, there were mixed results when determining the impact of the professional development on student

[10] To be classified as minimally methodologically adequate curricular evaluations had to include a quantifiable outcome measure, judge the comparability of groups, and included at least one of the following: report of implementation or professional development, disaggregated results by content strands or subgroups, or use multiple outcome measures (NRC, 2004, p. 102).

achievement in *Core-Plus* classrooms. Schoen, Finn, Griffin, and Fi (2003) found that Course 1 teachers with the greatest growth in student achievement, as measured by ITED-Q pre and post-tests, were more likely to have attended a *Core-Plus* professional development than teachers with lower growth (90% vs. 30%). They concluded that participation in a workshop designed to expose teachers to Course 1 material, both as learners and teachers, was positively and significantly related to student achievement, though no predictor for professional development was included in statistical models of student achievement. Contradictory results were presented by Tarr, et al. (2010), as they found professional development, as measured by hours of professional development received and included in HLM models of student achievement, had no effect on the reasoning or fair assessments and was negatively correlated (though non-significant) with student achievement on the ITED-15. In a follow-up study, Taylan and Tarr (2011) found teachers reported little to no impact of professional development activities on teaching practices, in part, because they perceived the activities merely confirmed what they were already doing in their classroom instruction.

Though literature linking predictors for professional development to student achievement is limited from earlier curricular evaluations of *Core-Plus*, more recent studies of other NSF-curricula have included professional development measures into their multi-level models. Two separate examinations of *Standards*-based mathematics on standardized achievement test, one conducted by Harwell and colleagues (2007) on secondary students and another by Post and colleagues (2008) on middle-grades students, found professional development hours were not a significant predictor of student achievement. Teachers in each

study participated in three components of the same professional development: a 2-week summer training, sessions during the school year, and newly implementing teachers received a mentor. All of the teachers in these studies took part in an average of 144 hours (Harwell, et al., 2007) or 162 hours (Post, et al., 2008) of professional development over a three-year period. There was no attempt to determine significant differences among teachers' exposure to the three components of the workshop.

These discrepant findings point to the need for additional studies beyond correlational studies. It is important for researchers and designers of professional development to understand why the professional development is effective, ineffective, or non-significant, as perceived by these measures. Therefore, additional research needs to be conducted to analyze the effectiveness of teachers' exposure to different components of professional development and how these professional development experiences impact their instruction, implementation of the materials, and student achievement.

Evaluating Implementation Fidelity

Regardless of the professional development experiences, it can neither be assumed that access to teachers' guides and curriculum materials will ensure curricula are being properly implemented (Scott, 1994), nor that teachers within the same school will implement curricula materials in the same way (Bowzer, 2008; Chávez, 2003; M. K. Stein, et al., 2007). Thus, it is important that claims made about curricular effectiveness include measures to judge the adequacy of implementation (NRC, 2004; Senk & Thompson, 2003). In more detail

the NRC (2004) noted the importance of documenting the faithfulness of the implementation and recommended:

> Evaluations should present evidence that provides reliable and valid indicators of the extent, quality, and type of the implementation of the materials. At a minimum, there should be documentation of the extent of coverage of curricular material (what some investigators referred to as "opportunity to learn") and of the extent and type of professional development provided. (p. 194)

Implementation fidelity measures the extent to which textbook materials are used for instruction and are not indicative of the quality of teaching (McNaught, Tarr, & Grouws, 2008; National Research Council, 2004).

Implementation measures and student achievement

In a longitudinal mixed methods study of the effectiveness of reform curricula on heterogeneous groups of students, Boaler and Staples (2008) found the most significant factor in explaining differences in students' Algebra achievement was the variance among teachers using the same curriculum. Variance in teachers' implementation of textbook material has been reported elsewhere (Huntley, et al., 2000; McNaught, Tarr, & Sears, 2010; Swafford, 2003).

In a study of high poverty middle schools, Balfanz, et al. (2006) examined the impact of implementation on student achievement for teachers using a middle school reform mathematics textbook, exposed to a comprehensive professional development around the use of the curricular materials, and proved with in-class coaching. They created an

implementation index that included textbook coverage, number of years teaching, use of pedagogy, curriculum coverage, and professional development attendance. Textbook coverage was defined as the number of textbook units the teacher had taught by the end of the year. Use of pedagogy was measured by curriculum coaches on teachers' use of nine recommended instructional practices and the suggested frequencies of those practices. Incorporating the number of years taught and the number of hours a teacher attended professional development were unique variables to include in an implementation index. Using a three-level HLM model, they found that higher implementation of the reform curricula produced higher student outcomes on a state exam.

Tarr and colleagues (2010) used the percent of the textbook taught, the extent to which the textbook was used, and seating arrangement for their implementation measure. Seating arrangement was quantified as the relative frequency that students were in groups during an observation. Implementation was found to be significantly related to the three different outcome measures, such that higher the implementation correlated to higher student achievement.

Evidence shows "that students do not learn content to which they are not exposed" (M. K. Stein, et al., 2007, p. 327), therefore studies of curricular effectiveness should account for the opportunities students are given to learn content that will likely be on their assessments (Tarr, et al., 2010). The COSMIC team has made great strides in the field of curricular evaluation in creating innovative ways to measure implementation fidelity. They have created indices for opportunity to learn (OTL), extent of textbook implementation

(ETI), textbook content taught (TCT), content fidelity, and presentation fidelity. The content and presentation fidelity ratings are based on classroom observations; helping researchers understand *how* the content is being taught.

Implementation indices

The COSMIC team measured the OTL, ETI, and TCT[11] through Table of Contents Records (TOC-logs), which were self reported by the teachers and customized for the textbook they were using (McNaught, et al., 2008). The TOC-logs have each lesson of the textbook in a table, and teachers check if they taught the content primarily from the textbook, with some supplementation, primarily from an alternative source, or not at all. The OTL index measures the percentage of textbook content that was taught, either solely from the textbook or through supplemental materials. COSMIC researchers found the mean OTL index across teachers using *Core-Plus* was 60.81, with standard deviation 19.98, and across teachers using subject-specific texts was 76.63, with standard deviation 17.02 (McNaught, et al., 2010). These data indicate on average less than about 60% of the content in the *Core-Plus* textbooks was covered, and just over 75% of the subject-specific content was covered, though there was substantial variance in teachers' OTL indices.

The ETI weights the options in the TOC-log, giving the content taught primarily from the textbook the most weight, a weight of one, and decreasing the weights so content not taught was weighted as zero. The weights were then summed and divided by the total number of lessons contained in the textbook. This measures the degree to which the textbook was

[11] The methods section in Chapter 3 will provide more detail on the calculations of these indices.

used directly to teach the content. The mean ETI index for integrated teachers from the COSMIC project was 50.73, with a standard deviation of 20.20 and for subject-specific teachers it was 57.15 with standard deviation 18.94 (McNaught, et al., 2010).

Similarly, the TCT used the same weighted sum but divided by the total number of lessons taught through any means. This is a measure of how the textbook was used in lessons that were taught and ignores the topics students were not taught. McNaught, et al. (2010) reported the total TCT index for COSMIC teachers using *Core-Plus* was 81.96 with standard deviation 14.50 and 74.93, with standard deviation 18.29 for subject-specific teachers. On average, teachers of *Core-Plus* used the textbook for teaching content more frequently than subject-specific teachers. Each of these three indices was measured at the course level.

Content and presentation fidelity

Content fidelity refers to *what* was taught, while the presentation fidelity refers to *how* the content was taught. The content and presentation fidelity measures are independently calculated using a Classroom Visit Protocol (CVP). This was created after interviewing the authors of each textbook to determine their expectations for the implementation of their curricular materials. Researchers rated each lesson on a 1 (lowest) to 5 (highest) scale for both content and presentation. These two data sources were at the lesson level.

McNaught, et al. (2010) found, across 109 *Core-Plus* and subject-specific teachers in 326 lessons, the mean content fidelity rating was 3.67, standard deviation 1.02, and presentation fidelity was 3.11, standard deviation 0.96. The presentation ratings were significantly lower than the content ratings, suggesting it is important to analyze both

measures. They also reported that *Core-Plus* teachers had significantly lower presentation ratings than subject-specific teachers. In other words, classroom observation data showed that subject-specific teachers used instructional practices consistent with their textbook authors' intentions more frequently than *Core-Plus* teachers. Taken together data from the previous two sections shows that *how* curricula are taught is an important factor in student outcomes and the implementation can be drastically different from one classroom to the next. The next section will present research of factors influencing teachers' implementation.

Teachers' implementation

Implementing reform mathematics curricula represents a challenging transition for many teachers (Ziebarth, 2003), especially for those whose perceptions of mathematics education are grounded in traditional views of teaching mathematics. Although many view the textbook as the most important catalyst for changing what occurs in mathematics classrooms, the adoption of the curriculum alone will not likely transform teachers' instructional practices (Arbaugh, et al., 2006; M. S. Wilson & Lloyd, 2000). Teachers' beliefs and backgrounds greatly influence how they implement curricula, potentially causing the implemented curricula to be significantly different from the developers' intended curricula (Ball & Cohen, 1996; Remillard, 2000) and from a teacher's own intentions (M. K. Stein, et al., 2007; M. K. Stein & Smith, 2010).

Wilson and Lloyd (2000) discovered *Core-Plus* teachers with strong content knowledge were able to more easily implement CPMP's suggested instructional methods because they did not have to learn new math content, thus providing them with more time to

concentrate on changing their instructional practices. They also found that without adequate professional development teachers implementing *Core-Plus* for the first time were more likely to spend a disproportionate amount of time on whole-class activities. While there is no prescriptive amount of time to spend on each facet of a *Core-Plus* lesson, Schoen, Finn, Griffin, and Fi (2001) found positive results for teachers who aligned their teaching practices with the textbook materials. Course 1 teachers with the greatest growth in student achievement, as measured by ITED-Q pre and post-tests, used a variety of strategies not displayed by the teachers with the lowest growth. The greatest growth teachers tended to spend more time on investigations (64.2% vs. 50.7%), spend less time on homework exercises (11.4% vs. 22.2%) and less time on whole-class lecture during investigations (19.6% vs. 45.6%).

Implementation and professional development

The NRC (2004) first argued that professional development could be thought of separately as a program element or as part of the implementation process, but subsequently decided when coding the studies it was important to separate professional development and implementation. In an analysis of postsecondary mathematics performance for students who were enrolled in *Standards*-based high school curricula, Post, et al. (2010) claim, "teachers whose students were represented in this study received relatively extensive targeted professional development…It therefore seems reasonable to posit a moderate degree of fidelity of implementation of these three NSF curricula" (p. 281). However, due to the variance in implementation across teachers and schools reported earlier, researchers cannot

make the assumption that because teachers attended a professional development they were able to effectively implement reform curricula. Before such claims can be made the field needs to research the impact professional development has on the implementation of curricular materials. The field needs research designed to examine the interaction of professional development experience and teachers' fidelity of implementation. More specifically, research should focus on determining how participation in various components of high quality professional development have an impact on teachers' implementation of *Core-Plus*.

Setting up the Study

As indicated in this review, there is a need for curricular evaluations that link student achievement to teacher and student variables using more sophisticated techniques than previous *Core-Plus* evaluations. In addition, curricular evaluations should be guided by the assumption that exposure to professional development does not imply effective implementation of curricular materials. While it is apparent that ongoing support can help teachers change their instructional practices, it is yet to be determined *how* this type of support impacts implementation of curricular materials and subsequently student achievement. Rarely does research link teacher professional development or implementation fidelity to student outcomes.

The design for this empirical study was guided by the gaps in previous research. The aim was to conduct a curricular evaluation on *Core-Plus,* contextualized within the NCIM professional development population of students and teachers, using recommendations from

the NRC (2004). Further, to understand the impact of different components of the NCIM professional development, this study analyzed the impact of different components of the NCIM professional development model on teachers' instruction, implementation, and subsequently on student achievement. The context of the NCIM project, non-field test sites and ones with a high percentage of minority students, supports research in areas that have not been well researched.

CHAPTER 3: METHODOLOGY

Overview

This research was an extension of a larger study known as the North Carolina Integrated Mathematics (NCIM) project. Using data collection tools developed from the larger study and ones created for this study, the impact of the *Core-Plus Mathematics* curriculum on student achievement was examined. In addition, this study examined the influences the NCIM professional development had on teachers' implementation of the curricular materials and how implementation had an impact on student achievement.

This chapter begins with the research questions and methods. Next, the sample population is described, followed by a discussion of the data sources and data analysis that were conducted. The data sources included student achievement, teacher content knowledge, teacher belief and background surveys, Table of Content Records, classroom observations, and teacher interviews. Each instrument is described in this chapter, followed by a discussion of how the data were analyzed.

Research Questions

The purpose of this research was two-fold, to provide insight into to the documentable effects of the *Core-Plus* curricular materials on student achievement across North Carolina and to analyze the impact the professional development model used in the NCIM project had on teacher's implementation of *Core-Plus* and subsequent student achievement. Specifically, the research questions for this dissertation study were:

1. What evidence, if any, is there of significant differences in student achievement between students using integrated mathematics and subject-specific curricula taught as Algebra I, Geometry, and Algebra II across the state of North Carolina?
    a. Is there evidence of differential impact on student performance on state tests between teachers using integrated mathematics curricular materials and those who use subject-specific curricula?
    b. For teachers who participated in the project workshops, is there evidence of differential impact on student performance on state tests between teachers using the *Core-Plus* curricular materials and those who use subject-specific curricula?
2. What evidence, if any, is there of significant differences on student achievement among teachers who participated in different components of the NCIM project?
    a. Among teachers using the *Core-Plus* curricular materials, what evidence of systematic differences is there on the effects of student learning based on varying levels of teacher participation (working with or without a instructional coach, frequency of professional development, type of professional development (content and/or pedagogy) in the NCIM project?
    b. Among teachers using the *Core-Plus* curricular materials, to what extent and how can the variance in student outcomes be explained by teacher implementation?

Methods

To address the research questions, this study used a mixed methods design. A mixed approach was used to incorporate "supplemental research strategies to collect data that would not otherwise be obtainable by using the main method" (Morse, 2003, p. 191). In particular a variant of the embedded design called "an embedded correlational model" (Creswell & Clark, 2007) (Figure **4**) was used. This type of design is represented with the notation: QUAN(qual). As Creswell and Clark explain this model, "quantitative data are used to answer the primary question in a correlational design and qualitative data are embedded within the correlational design with the intent of explaining the mechanisms that relate to the predictor and outcome variables" (p. 85). The rationale for using this design was based on the relative importance of the quantitative findings and the ability of the qualitative data analysis to refine the statistical results by exploring teachers' implementation in greater detail.

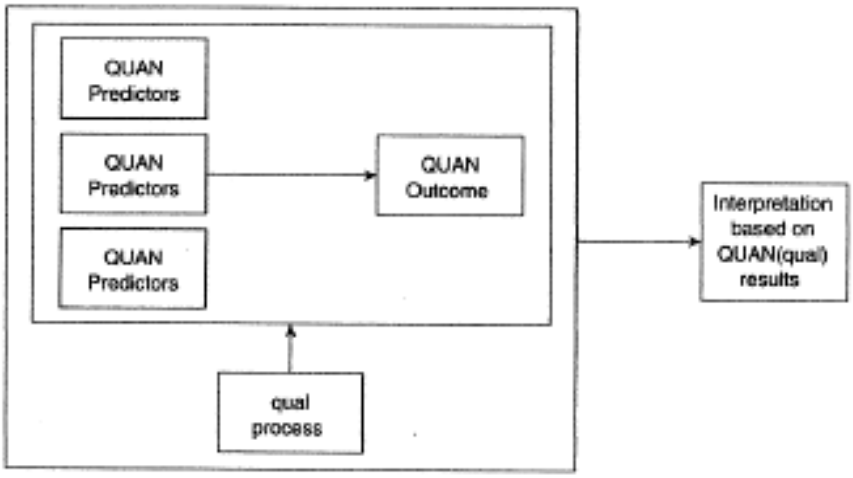


Figure 4. Embedded Correlational Model. Taken from Creswell and Clark (2007).

The first question was answered using a quasi-experimental matched group design, while the second question was analyzed utilizing both quantitative and qualitative methods. The qualitative methods assisted in the interpretation of the findings from the quantitative analysis and in supporting the results (Morse, 2003). 2009-2010 student test data from the seven STEM schools implementing the *Core-Plus* curriculum was contrasted with test data from comparison schools (selected based on the presence and level of implementation of the *Core-Plus* curriculum). Following recommendations from the NRC (2004), students were not used as the sole unit of analysis because they are nested within groups and do not experience the curricula as individuals. Hierarchical linear modeling (Raudenbush & Bryk, 2002) was utilized to model variation in student achievement since randomization of students was impractical. Using student, teacher, and school level data, three-level models were used with students nested within teachers and teachers nested within schools.

Sample

The sample included six separate target populations (Table **7**). First were teachers from the NCIM project schools (Group A) who participated in the professional development and who received instructional coach visits each month. The summer conferences accommodated an additional 150 participants. These workshop participants had not received an instructional coach and were either in the process of implementing *Core-Plus* (Group B) or were trying to integrate reform practices into a subject-specific curriculum (Group C). Two other groups of teachers were selected from a comparative group of non-workshop

participants. They either used an integrated mathematics curriculum (Group D) or used a subject-specific curriculum (Group E).

A final group emerged during the study. Initially, teachers in Group A were classified as NCIM teachers because they received follow-up instructional coach support throughout the school year following the summer workshops and worked at schools affiliated with the NCIM project. However, after analyzing quantitative measures (to be described in this chapter) of Group A teachers' implementation of *Core-Plus*, commonalities were identified among the teachers with outlier measures. There were two teachers in Group A with lower implementation measures than the rest of the group, and upon further analysis it was discovered they had not attended a summer workshop prior to the 2009-2010 school year. So while they had received an instructional coach throughout the school year, it was decided they could not be classified with the rest of the Group A teachers due to lack of workshop attendance. Hence, a new group, F, was formed with these two teachers.

Table 7. Groups of Teacher Participants in Large Study

Group	Sample Size	Curricular Materials Used	NCIM Professional Development Experience
A	12	*Core-Plus*	Workshop participants Received an instructional coach
B	10	*Core-Plus*	Workshop participants
C	74	Subject-Specific	Workshop participants
D	13	Integrated Mathematics	None
E	304	Subject-Specific	None
F	2	*Core-Plus*	Received an instructional coach

The second research question involved recruiting a subset of twenty-one teachers that used the *Core-Plus* curricular materials, seven from Group A, the two from Group F, and six each from Groups B and D. Each teacher was given a pseudonym. Group A teachers were: Chloe, Kathleen, Leigh, Luke, Maria, Missy, and Nicole. The two Group F teachers were Heather and Nora. The Group B teachers were: Ben, Beth, Ellie, Jenni, Steven, and Taylor. The Group D teachers were: Calley, Delores, Josie, Kasey, Kristin, and Shauna. To be considered for this focus study, a teacher must have taught a *Core-Plus* course during the 2009-2010 school year that culminated with students taking the North Carolina Algebra I or Algebra II End-of-Course (EOC) test. These assessments typically follow Course 2 and 3 respectively. Teachers in Group A and F were involved in the NCIM project for at least a year and were already identified for participation in this extended study. Teachers in Group B were recruited based on their participation in the Summer 2009 workshop, ensuring teachers had student EOC scores.

Six Group D participants were selected from across the state of North Carolina so that they meet the following criteria: were exposed to a minimal amount of professional development on implementing *Core-Plus*, taught *Core-Plus* Course 2 or 3 in the Spring of 2010, and were using the *Core-Plus* materials during the Fall of 2010. It was rare to find a teacher that received no support before implementing the curriculum and difficult to find a district willing to allow the researcher into their schools to examine teachers' implementation. These six teachers ended up being recruited from one district in the western part of the state that had been using the first edition *Core-Plus* materials for ten years. During

the Fall of 2010 they transitioned to using the second edition textbook. Each teacher in Group D had a one- or two-day local workshop facilitated by the math coordinator of the district and two of them had been to *Core-Plus* training facilitated by the curriculum developers. The six participants from Group D received a modest stipend for their participation, dispensed after they completed the data collection. The stipend was not offered to the teachers in Group A, F, and B because they were familiar with the researcher, were willing to work with researcher, and had completed a significant portion of the data collection at the summer 2010 workshop. Further, Group A and B teachers received a stipend from the NCIM project following their participation in the summer workshops.

Data Sources

Data were collected from many sources. For a visual reference of the data sources and methods that were used to answer each specific question refer to Appendix A. In the next section, student level data will be explained, followed by teacher and school level data.

Student level

Tests are the primary tool used by curricula evaluators for measuring curricular effectiveness (NRC, 2004). This study involved a test of curricular alignment with systemic factors, which refers to "measures that facilitate student progress in the system, such as state tests, college entrance exams, and future course taking" (NRC, 2004, p. 49). The researcher received access to student level achievement data from the 2009-2010 school year for all of the study participants in Groups A, B, C, D, E, and F from the North Carolina Department of Public Instruction (NC-DPI). More specifically, if a teacher participated in any facet of the

NCIM professional development, student achievement data was provided for every teacher in the participating teacher's school.

The dependent variable was student achievement on the North Carolina Algebra I and Algebra II EOC standardized assessments. The Geometry EOC assessment was excluded from this analysis because the NCIM partner schools had been granted a waiver for the Geometry assessment since 2008 from the NC-DPI, and in 2009-2010 the state discontinued administering the Geometry EOC. Independent student level variables from the DPI reports included: gender, race/ethnicity, number of days present for class, number of days absent, EOC Algebra I level, and EOC Algebra II level.

Teacher level

Instruments at the teacher level included Table of Contents Records (TOC-log), content knowledge assessment, background questionnaire, beliefs survey, instructional coach reports, semi-structured interviews, and classroom observations. Each of these will be explained in more detail below.

Table of Contents Records. To address the second research question, all teachers in Groups A, F, B, and D completed TOC-logs (McNaught, et al., 2008) to measure the implementation fidelity of using the *Core-Plus* materials. These logs were replicas of the table of contents in the Core-Plus textbooks so that teachers could readily recall the material they taught (Appendix B). For each lesson of the textbook, teachers indicated if they taught the content (a) primarily from the textbook, (b) primarily from the textbook with some supplementation, (c) primarily from an alternative source, or (d) not at all (McNaught, et al.,

2008). These logs corresponded to the textbook edition they used during the semester students took the EOC. It was important that the implementation indices corresponded to the same semester the student data were collected. All of the teachers in Groups A, F, and B and one teacher in Group D completed TOC-logs for the second edition of the textbook, the other five teachers in Group D completed them for the first edition of the textbook.

Content Knowledge Assessment. Each workshop participant took a paper-and-pencil pre-test on day one and an alternative form post-test at the end of day five of the NCIM summer workshop (Appendix C). Non-workshop participants completed this assessment after they were recruited in Fall 2010. This test was solely a measure of teachers' mathematical content knowledge and not pedagogical content knowledge. The researcher designed this assessment for the Summer 2009 Workshop to align with the new North Carolina mathematics Essential Standards[12]. The items were selected from a draft of the Essential Standards created by the North Carolina Department of Public Instruction (NC-DPI) and then agreed upon by the project evaluation team. Each member of the team individually took the assessment and then aligned each question to objectives in the Essential Standards. The assessment was designed to be comprehensive across the first three years of high school mathematics, regardless of curricular pathway, and participants were encouraged to use a graphing calculator. To reduce a potential practice effect in these data, three forms of the assessment were given so that each participant had a different pre- and post-test form. For this study the teachers' most recent content knowledge assessment was used.

[12] This took place prior to the release of the Common Core State Standards.

Background Questionnaire and Beliefs Survey. The teacher background instrument was created to collect basic demographic data, including years of experience, degrees obtained, use of technology, NCIM workshops attended, teacher certification status, and number of students taught during the school year (Appendix D). The teacher belief survey was created to solicit responses to questions about teachers' preparedness to teach different populations of students, beliefs about effective instructional strategies, and perceptions of mathematics (Appendix E). This instrument was adapted from instruments created by Horizon Research (2000) and McCaffrey, et al. (2001) and includes 7 four-point and 22 five-point Likert Scale items. Teachers self-reported their answers for both of these instruments.

Instructional Coach Reports. To address question 2b, more extensive measures were collected on teachers in Groups A, F, B, and D. Throughout the NCIM project, instructional coaches completed a report after each of their site visits with teachers in Groups A and F to document their observations and activities with each teacher (Appendix F). The reports provided information on the teachers' use of time and pacing, instructional behaviors, use of collaborative learning, use of technology, use of formative assessment, and classroom management. Since the inception of the project in 2008, there have been approximately 140 reports for teachers in Groups A and F. This report structure was used to collect data on the additional twelve teachers from Groups B and D. The classroom observations provided additional information regarding an external evaluation of the teacher's implementation of the materials. Since the observations did not occur until Fall 2010 and because NC-DPI provided released student test scores from Spring 2010, the reports were not used for indices

in the HLM analysis. Further, because of the student data set structure, it was impossible to link a student's class period to the class period the instructional coach observed. The observation reports were used to qualitatively explain differences in teachers' implementation.

One of the two NCIM instructional coaches was hired to conduct the additional twelve classroom observations to ensure there was consistency between the current reports, for Group A and F teachers, and the reports from Groups B and D. She used the same instrument for the new groups and observed each Group B and D teacher once. There was no attempt made to measure change in teachers' instructional practices, so as an added benefit to teachers' participation in this study, the instructional coach was asked to interact with the teachers she observed as if she were their instructional coach, providing feedback at the conclusion of the observation and following-up with teachers via email.

Teacher Interviews. In addition to the instructional coach reports, semi-structured interviews and classroom observations were conducted by the researcher with the participants in Groups A, F, B, and D. The interview protocol (Appendix G) was the same that was used over the duration of the project. Questions on the protocol were informed by initial classroom observations of NCIM project teachers. The purpose of the interviews was to collect information on the teachers' background, beliefs about reform mathematics, and decisions they make with respect to implementing the curriculum. Also, for the Group A and B participants that may have been exposed to professional development during the summer of 2010, the interviews served as a way to explain changes the teachers made to their

instructional practices. Specifically, the interviewer asked the teacher if there were things they were doing in the fall, from going to the workshop, that they did not do last spring. The interviews were audio recorded and then transcribed.

Classroom Observations. It was important to understand the dynamics in each teacher's classroom. Therefore, the researcher conducted classroom observations on Group A, F, B, and D teachers. Group B and D teachers were observed before the instructional coach visit. Altheide (1987) found that using only structured protocols can mask significant concerns of a studied group and suggested using field-notes to support main assertions. During the observations the researcher took notes on what was observed in the class, first with a description of the setting and population of students in the room, followed by an account of the instruction and interactions that took place during the class period. After each observation detailed field-notes (Emerson, Fretz, Shaw, & Thompson, 1995) were written on the researchers experiences in the classroom. The purpose of the observations was to provide the researcher with experience in each teacher's classroom to make it possible to triangulate observations with the instructional coach reports and interviews that were conducted.

School Level

The NC-DPI Accountability Department reported that a Free and Reduced Lunch (FRL) tag was not allowed when compiling data. Further, they were not able to identify students as FRL because it is private information. Due to the importance of including a measure for socioeconomic status, and since the NC-DPI neither collects this information at

the student nor teacher level, school level variables were collected to have a measure for socioeconomic status to use in three-level HLMs.

School level variables included average daily membership (ADM), number of reduced lunch applications, number of free lunch applications, needy percent for the school, Title One status, number of males and females in each ethnic category, and total students enrolled at the school. NC-DPI defines a school's needy percent as the sum of the reduced and free lunch applicants divided by the average daily membership, this will be referred to as %FRL. The data for each of these variables was collected from Excel spreadsheets downloaded at http://www.dpi.state.nc.us/data/reports/.

In addition, the teachers in the subset of Groups A, F, B, and D had school level measures from their work with the North Carolina New Schools Project (NCNSP) which included: number of teachers, Algebra I EOC achievement, Algebra II EOC achievement, Geometry EOC achievement, average SAT score, and graduation rate. Since several of these schools had opened recently, measures for SAT and graduation rate were not available for all schools.

Assessing the Curricular Alignment and Outcome Measure

The outcome measures used in this study were curricularly aligned with systemic factors (NC state standards), and were not tests aligned with a particular curriculum. Therefore, it was necessary to assess the content of the *Core-Plus* text to determine how well textbook content aligned with questions on the Algebra I and II exams. This provided a measure of the percent of EOC topics that were taught in *Core-Plus* Courses 1, 2, or 3. Four

people were recruited and hired to assist with assessing the curricular alignment to the exam, three of which were experienced *Core-Plus* teachers in North Carolina and instructors of *Core-Plus* professional development workshops. The other was affiliated with the CPMP and had extensive experience developing and using the curricular materials.

Assessing the alignment between the textbook and the exam included four separate analyses to match both of the *Core-Plus* textbook editions to both of the Algebra EOCs. Two coders independently assessed the alignment for each textbook edition. Released Algebra I and Algebra II EOCs, published on the NC-DPI website, were used to complete the analysis. The coders were given an EXCEL spreadsheet with a tab for the Algebra I exam and one for the Algebra II exam, along with the released exams. The Algebra I tab contained a column with the title of each unit, lesson, and investigation in the *Core-Plus* Course 1 and 2 textbooks, whereas Algebra II tab contained the same information for Courses 1, 2, and 3 (Figure **5**). Coders were instructed to read the questions in the released EOC and put a "1" in the columns for each *Core-Plus* investigation that addressed the content of the test item. Each question has its own row in the table and can have multiple entries in the row if it is taught in more than one *Core-Plus* investigation. There was also a column labeled "not taught in CPMP" for content the reviewer did not believe was included in the text.

		2nd Edition Course 1							
Unit		Unit 1 Patterns of Change							
Lesson		Lesson 1 Cause and Effect			Lesson 2 Change Over		Lesson 3 Tools for Studying Patterns		
Invest.	Not in CPMP	*Physics and Business at Five Start Amusement Park*	*Taking Chances*	*Trying to Get Rick Quick*	*Predicting Population Change*	*Tracking Change with Spreadsheets*	*Physics and Business at Five Start Amusement Park*	*Taking Chances*	*Trying to Get Rick Quick*
EOC Item									
1									
2									
3									
4									
5									
6									

Figure 5. Screen shot of instrument used to assess curricular alignment to the EOCs.

Data Analysis

It was important to use prior achievement to establish the equivalence or control for the non-equivalence of treatment groups (Tarr, et al., 2010). Based on the data set provided by NC-DPI, when modeling Algebra II achievement, it was possible to control for a student's prior achievement using their Algebra I score. However, there was no score to use when modeling Algebra I achievement. NC-DPI would not provide access to each student's 8th Grade Mathematics End-of-Grade (EOG) assessment score, so it was not possible to control for prior achievement when using Algebra I achievement as the dependent variable. Another option considered was using students' mathematics grades to control for their prior achievement. However, grades are not standardized among teachers and collecting such data was not feasible due to the large number of students and teachers in the sample. The use of HLM helped control for the non-equivalence of student groups based on student gender, race, attendance, and grade level. Further, teacher and school level variables were also used to control for the possible non-equivalence of treatment groups.

Hierarchical linear modeling was selected as an analysis technique for the quantitative data because the observations were nested within larger groups, neither balanced nor complete data was required, and HLM partitioned the variability within-teacher, between-teacher, and between-schools (Raudenbush & Bryk, 2002). In this case, students were nested within teachers and teachers were nested within schools. Students were used at level one in this study, with teachers at level two, and schools at level three. The assumptions of HLM allow researchers to run analyses regardless of having the same number of schools or teachers in each comparative group or the same number of students in each class.

Finally, Raudenbush and Bryk (2002) claim that it is "essential that the variables under study have precise meaning so that statistical results can be related to the theoretical concerns that motivate the research" (p. 31). They suggest paying attention to the metric of all level-1 predictors and possibly transforming the location of the predictors to make the results interpretable. As such, in this study each independent variable was grand mean centered so that the intercept is the expected student achievement score when the value of a predictor is set to its grand mean. Centering was also important to reduce the non-essential multicollinearity between variables and was essential for testing for cross-level interactions.

Research Question 1a

Is there evidence of differential impact on student performance on state tests between teachers using integrated mathematics curricular materials and those who use subject-specific curricula? To determine the effect of being taught using integrated or subject-specific curricular materials, student test scores from Groups C and E were compared to

Groups A, F, B, and D. While controlling for student gender, race, prior achievement (in the case of Algebra II), and school ADM, %FRL, and Title One status, student achievement was modeled using a three level HLM model (Appendix H) (Raudenbush & Bryk, 2002). Singer's method (1998) was used to determine which model was a better fit, one with the slopes allowed to vary randomly or one with constrained slopes.

Since there was student data for both Algebra I and Algebra II EOC scores, separate models were run using each of those as dependent variables. Teachers were coded (Appendix I) as either teaching integrated mathematics (0 for Groups A, F, B, and D) or subject-specific mathematics (1 for Groups C and E) based on the state course number associated with the data set.

Research Question 1b

For teachers who participated in the project workshops, is there evidence of differential impact on student performance on state tests between teachers using the Core-Plus curricular materials and those who use subject-specific curricula? Student test scores from Groups A, F, B were compared to Group C; essentially this was to determine the effect of being taught using integrated or subject-specific curricular materials while considering teacher characteristics as gathered during the summer workshops. Student achievement was modeled using a three level HLM while controlling for student gender, race, prior achievement, and school ADM, %FRL, Title One status, and NCNSP status, (Raudenbush & Bryk, 2002). Again, separate models were run for Algebra I and Algebra II achievement. Teachers were coded as either teaching integrated mathematics (Groups A, F, B) or subject-

specific mathematics (Groups C) based on the state course number associated with the data set. Additional teacher level predictors included teacher background characteristics and content knowledge. Teacher background variables used include years of experience, gender, degrees obtained, use of technology, NCIM workshops attended, and teacher certification status. The teacher background variables not providing any explanatory power were be removed from the analysis (McCaffrey, et al., 2001).

The content knowledge assessment was also used as a predictor during this phase of the study. Two members of the NCIM evaluation team graded the assessments using a rubric-scoring guide. Each of the ten items was independently scored on a 0-5 point scale. To ensure consistency among raters, Pearson's ***r*** was calculated after each scorer graded one form of the assessment. The correlation between the ratings of each scorer on the assessment was 0.98. Due to the high correlation, one of the scorers graded the rest of the assessments.

Assessing the Curricular Alignment and Outcome Measure

Two individual alignment studies were completed for each edition of the textbook, and for each Algebra EOC, to ensure reliability between the two coders. This was a simple alignment study to determine the percent of content on the Algebra I and II EOC exams that was represented in the textbook. Eight alignments were conducted, two each for: Algebra I to first edition *Core-Plus,* Algebra I to second edition *Core-Plus,* Algebra II to first edition *Core-Plus,* and Algebra II to second edition *Core-Plus.*

This analysis only examined whether or not content was included in the textbook for each EOC question and not the particular investigation that content appeared in the textbook.

The alignments were analyzed by determining the percent agreement and Cohen's κ (1960) for the number of agreements between the coders on "content in text" and "content not in text" for each exam question. Landis and Koch (1977) suggested the following measures of Cohen's κ strength of agreement: less than zero are poor, between 0.00 and 0.20 are slight, 0.21 to 0.40 are fair, 0.41 to 0.60 are moderate, 0.61 to 0.80 are substantial, and 0.81 to 1.00 are almost perfect. This study used values greater than 0.41 to determine sufficient reliability between the two coders to warrant an alignment measure. If there was not sufficient agreement, then the researcher settled disputes between the two coders in a third external and independent review. The percent of the *Core-Plus* text that was aligned with the EOC was calculated for each coder. Results are presented at the end of Chapter 4.

Research question 2a

Among teachers using the Core-Plus curricular materials, what evidence of systematic differences is there on the effects of student learning based on varying levels of teacher participation in the NCIM project? To determine what factors influenced student learning in *Core-Plus* classrooms based on teacher participation in the NCIM project, dummy coding was used to compare the four groups of teachers using the *Core-Plus* materials (Groups A, F, B, and D). The dummy codes allowed for the determination of differences in student achievement between any one group and the other three. For example, if Group A was the referent group, the HLM analysis would produce estimates for the significance between Group A and F, Group A and B, and Group A and D. The teacher level variables become more complex during this phase of the study, the focus study, because data

include: teacher mathematics content knowledge, teacher background variables, teacher beliefs, and the implementation indices.

Teacher beliefs survey. Similar to approaches by Tarr, et al. (2010) and McCaffrey, et al. (2001) exploratory factor analysis was used to reduce the 29 teacher beliefs into factors used in the HLM analysis. Due to the small size of teachers in the focus study, teacher beliefs for teachers who attended the Summer 2010 workshop were combined with the focus teachers for the data reduction. Principal component analysis was used on data collected from these 75 teachers to extract five factors. There are different recommendations for the number of participants necessary (Field, 2009). The Kaiser-Meyer-Olkin (KMO) measure of sampling adequacy for this sample was 0.613. Kaiser (1974) suggests values less than 0.5 are unacceptable, whereas factors in the 0.60's are mediocre. Bartlett's test of sphericity was significant ($\chi^2 = 1014.08, df = 406, p < 0.0001$), suggesting the appropriateness of factor analysis.

To determine the degree to which the variables load onto the five factors, a varimax rotation was conducted. A varimax rotation was chosen because "it attempts to maximize the dispersion of loadings within factors" (Field, 2009, p. 636), which results in variables loading highly onto factors for ease in the interpretation of the factors. The varimax rotation was conducted in SPSS with a Kaiser Normalization and converged in eleven iterations. Table 8 provides the beliefs and loadings for the five factors.

Table 8. Factors Related to Teachers Beliefs About Mathematics

Factor 1: Reform Based Practices	Loading
Belief 9: Encourage students to use manipulative in solving a problem	.562
Belief 10: Ask students to look for alternative methods for solving a problem	.456
Belief 11: Require students to explain their reasoning when giving an answer	.659
Belief 13: Review student notebooks/journals	.565
Belief 14: Give tests requiring open-ended response	.425
Belief 19: Work on solving real-world problems	.573
Belief 20: Record, represent, and/or analyze data	.697
Belief 21: Collaborate with other students in solving a problem	.562
Belief 22: Write a description of a plan, procedure, or problem-solving process	.669
Belief 25: Students write reflections about something they learned	.604
Belief 28: Read from a mathematics textbook in class	.588
Factor 2: Collaborative Learning	**Loading**
Belief 2: Have students work in cooperative learning groups	.752
Belief 4: Teach groups that are heterogeneous in ability	.454
Belief 8: Arrange seating to facilitate small group work	.736
Belief 16: Have students complete performance tasks in groups	.801
Belief 19: Work on solving real-world problems	.477
Factor 3: Diversity and Encouragement	**Loading**
Belief 1: Take students' prior understanding into account when planning	.488
Belief 4: Teach groups that are heterogeneous in ability	.571
Belief 6: Encourage participation of females in mathematics	.888
Belief 7: Encourage participation of minorities in mathematics	.906
Factor 4: Explanations	**Loading**
Belief 5: Teach students who have limited English proficiency	.568
Belief 17: Explain new mathematics content to the whole class	-.578
Belief 23: Work on extended mathematics investigations or projects	.627
Belief 24: Make formal presentations about their projects to the rest of the class	.472
Belief 27: Answer textbook/worksheet questions	-.633
Factor 5: Traditional Mathematics Practices	**Loading**
Belief 15: Have students complete performance tasks individually	.588
Belief 18: Give short answer tests	.736
Belief 29: Practice mathematics facts, rules, or formulas	.662

Once the beliefs were rotated into the five factors, a composite score was calculated to estimate each teacher's score on each factor. The Anderson-Rubin method was used to determine the factor scores because it produces scores that are uncorrelated and standardized (Field, 2009). These factor scores for each belief factor were incorporated into the HLM analyses for the focus study teachers.

Implementation indices. The OTL, TCT, and ETI indices were computed using the COSMIC approach (McNaught, et al., 2008). Recall, for each lesson of the textbook, teachers indicated if they taught the content (a) primarily from the textbook, (b) primarily from the textbook with some supplementation, (c) primarily from an alternative source, or (d) not at all (McNaught, et al., 2008). Figure **6** provides an example of a completed TOC-log, which will be used to explain how the implementation indices were calculated.

NCIM Project
***Core-Plus, Course 3* Table of Contents Record**

Unit 1 **Multiple-Variable Models**		Taught primarily from *Core-Plus* textbook	Taught from *Core-Plus* textbook with some supplementation	Taught primarily from alternative(s) to *Core-Plus*	Did not teach content
Lesson 1 *Linked Variables*					
Inv 1	*Stressed to the Breaking Point*	☑	☐	☐	☐
Inv 2	*Go with the Flow*	☐	☑	☐	☐
Inv 3	*Combining Rates and Times*	☐	☐	☑	☐
Inv 4	*Taking Algebraic X-Rays*	☐	☑	☐	☐
Lesson 2 *Algebra, Geometry, and Trigonometry*					
Inv 1	*Triangulation*	☐	☐	☐	☑

Figure 6. Example Table of Contents Record Completed.

The OTL index measured the percentage of textbook content that was taught, either solely from the textbook or through supplemental materials. In the case above (Figure **6**), content was taught in four of the five lessons. The OTL index would be 80, calculated as $\frac{4}{5}*100$.

The ETI index weighted the options in the TOC-logs, giving the content taught primarily from the textbook the most weight (a weight of one), content with some supplementation a weight of two-thirds, content mostly from alternative sources a weight one-third, and content not taught a weight of zero. The weights were then summed and divided by the total number of lessons contained in the textbook. This measured the degree to which the textbook was used directly to teach the content. The ETI for the example TOC-log (Figure **6**) was calculated as $\frac{2+\frac{2}{3}*2+\frac{1}{3}+0}{5}*100$, or 73.3.

Similarly, the TCT index used the same weighted sum as the ETI index but was divided by the total number of lessons taught through any means. This is a measure of how the textbook was used to teach content in the textbook and ignores the topics students were not taught. The TCT for the example TOC-log (Figure **6**) is calculated as $\frac{2+\frac{2}{3}*2+\frac{1}{3}+0}{4}*100$. The resulting TCT index is 91.67. Each of these three indices was measured at the course level.

It is important to note that teachers completed one TOC-log for each unique course they taught. For example if they taught two sections of Course 2 and one section of Course 3

then they would have one measure of OTL, ETI, and TCT for Course 2, with both sections having the same measures, and one measure of OTL, ETI, and TCT for Course 3. Implementation indices from Course 2 were used in models of student Algebra I achievement, whereas indices from Course 3 were used in models of student Algebra II achievement.

Since the dependent measure was Algebra student achievement, it was important to measure opportunity to learn Algebra content. The researcher identified which units in the *Core-Plus* textbook corresponded to the four different content strands, to determine the differences in implementation across major topics. Recall the four strands are: Algebra and Function, Probability and Statistics, Geometry and Trigonometry, and Discrete Mathematics. Based on these strands, the OTL, ETI, and TCT were recalculated for each teacher for each content strand. The Algebra OTL (OTLAlg) was used as a predictor representing students' opportunity to learn the Algebra topics from the textbook.

Due to the sample size, the Kruskal-Wallis distribution free test (Hollander & Wolf, 1999) was used to determine quantitative differences in teachers' implementation across courses and content strands in the textbook. The Jonckheere post-hoc test (Hollander & Wolfe, 1999) was used to test for ordered alternatives to see if there was a decrease in the indices across the course sequence or content strands. Finally, Dunn's Test (Elliott & Hynan, 2010), a nonparametric multiple comparisons test based on pair-wise rankings, was used to determine differences among specific groups.

Since there were at least 16 teacher variables excluding the factors from the beliefs survey, an attempt was made to reduce the number of teacher variables using Principal Components Analysis (PCA). However, due to the small size of teachers in the focus group and large number of variables, it was not possible to reduce the data set.

Research question 2b

Among teachers using the Core-Plus curricular materials, to what extent and how can the variance in student outcomes be explained by teacher implementation? This question helped paint a picture of how teacher experience with (or without) the NCIM professional development had an impact on their implementation of *Core-Plus*, following findings from the previous question regarding student achievement. Here field-notes from the researcher's observations, the teacher interviews, and instructional coach reports combine with the aforementioned data to provide a detailed account of the teachers' implementation of *Core-Plus*.

A prior study has provided evidence that teachers in Group A are using the textbook as their main source to teach textbook content (Krupa, 2011), however this does not imply their instruction is consistent with the written text. The interviews and observations helped the researcher understand *how* the content was taught. The instructional coach observation reports were organized and combined into one spreadsheet and the interviews were transcribed.

To analyze the qualitative data from the two observations and interviews, the researcher found emergent patterns and themes using a grounded theory (Glasser & Strauss,

1967) approach. Grounded theory consists of initially making data collection the main focus of the research and coding it in as many ways as possible. This approach included open coding (Strauss & Corbin, 1990), in as many ways as possible, each teacher's interview transcript, observation report, and field-note. The researcher was able to triangulate these three data sources, which allowed for cross-checking of the multiple data sources (O'Donoghue & Punch, 2003) and provided corroborating evidence for more detailed account of the setting under study (Creswell, 1998). After the initial coding, data from the three sources were reviewed to detail common and contradictory themes and characteristics for each teacher. Next, the open codes were used to identify common themes and characteristics for each teacher and were reviewed to extract broad categories for each of the focus groups. Then, emergent themes among teachers in each group were determined for the broad categories (Table **9**). These broad categories were used to determine differences among the groups and to embed the qualitative findings with the quantitative results.

Table 9. Themes Influencing Teachers' Implementation of *Core-Plus*

Theme	Subcategories
Teacher beliefs about how students learn mathematics	Influenced how teachers facilitated collaborative groups and presented content to their students
Trust of the curriculum	Influenced teachers' attitudes about how the curriculum is structured and how they handled the gradual refinement of key concepts overtime
Systemic factors	State assessments, access to materials for instruction, school scheduling, and student transition to reform mathematics influenced curricular implementation

CHAPTER 4: RESULTS QUESTION ONE

This chapter summarizes the quantitative results of the impact of curricular materials on student mathematics performance. The results are organized into four sections. The first two sections report results from the entire sample. First, results from the Algebra I End-of-Course (EOC) exam will be presented, followed by results from the Algebra II EOC. The third and fourth sections contain results from the sample of workshop participants using the same two outcome measures. The chapter will conclude with a summary of the results surrounding the first research question.

Modeling Student Algebra I Achievement: Entire Sample

The North Carolina Department of Public Instruction (NC-DPI) provided access to student achievement data for schools in which there was at least one teacher who participated in the NCIM professional development. For the Algebra I EOC assessment this resulted in 10,399 student scores from 75 schools. This accounted for 9% of the total student Algebra I test scores for North Carolina during the 2009-2010 school year.

Background and Demographics

Two different curricular pathways for mathematics were offered at the 75 schools, 15 offered integrated mathematics and 66 offered subject-specific mathematics (6 offered both options). On average, schools offering integrated and subject-specific pathways had similar percentages of students qualified for Free and Reduced Lunch (%FRL) (Table **10**). Schools that offered integrated mathematics tended to have a smaller student population, as measured

by the average daily membership (ADM) and were more likely to be classified as a Title One school.

Table 10. School Demographics for Algebra I Sample

	Average %FRL	Average ADM	Percent Title One
Integrated Mathematics	50.09	551.86	20.00
Subject-Specific	51.61	748.25	6.06
Total	51.45	713.45	8.00

Students take the Algebra I EOC exam after completing the second course in the integrated mathematics sequence or Algebra I on the subject-specific pathway. As a result, there were some differences in the percentage of students taking the Algebra I EOC at each grade level (Table 11). Several teachers participating in the NCIM workshops taught at middle schools, so there was some representation of middle school students who took this exam after completing Algebra I. There were a higher percentage of female students enrolled in Course 2 who took the Algebra I EOC. The percentage of Asian and White students was higher for Course 2, there were higher percentages of Black students in Algebra I, and the other races were similar between the two groups.

Table 11**.** Percentages of Student Within Each Category for Algebra I Sample

			Grade			**Gender**	
	6-8th	9th	10th	11th	12th	Male	Female
Integrated Math	0.00	56.70	39.56	3.26	0.48	43.49	56.51
Subject-Specific	5.11	69.03	19.10	5.60	1.15	51.18	48.82
Total	4.60	67.80	21.15	5.37	1.09	50.40	49.60
				Race			
	AmIn	Asian	Hispanic	Black	White	Multi	
Integrated Math	1.34	4.98	10.34	18.59	62.26	2.49	
Subject-Specific	1.74	1.56	10.63	33.03	50.29	2.75	
Total	1.70	1.90	10.60	31.58	51.49	2.72	

Distribution and Summary Statistics

The distribution of scores for this sample is presented in Figure 7. The possible range of scores on this assessment was 118-181. The vertical red bar represented the minimum score a student must earn to receive a passing grade on the assessment, 148. The average score among all North Carolina students taking the assessment during the 2009-2010 school year was 153.4 with a standard deviation of 10.1; the average for this sample was 151.77 with standard deviation 9.20.

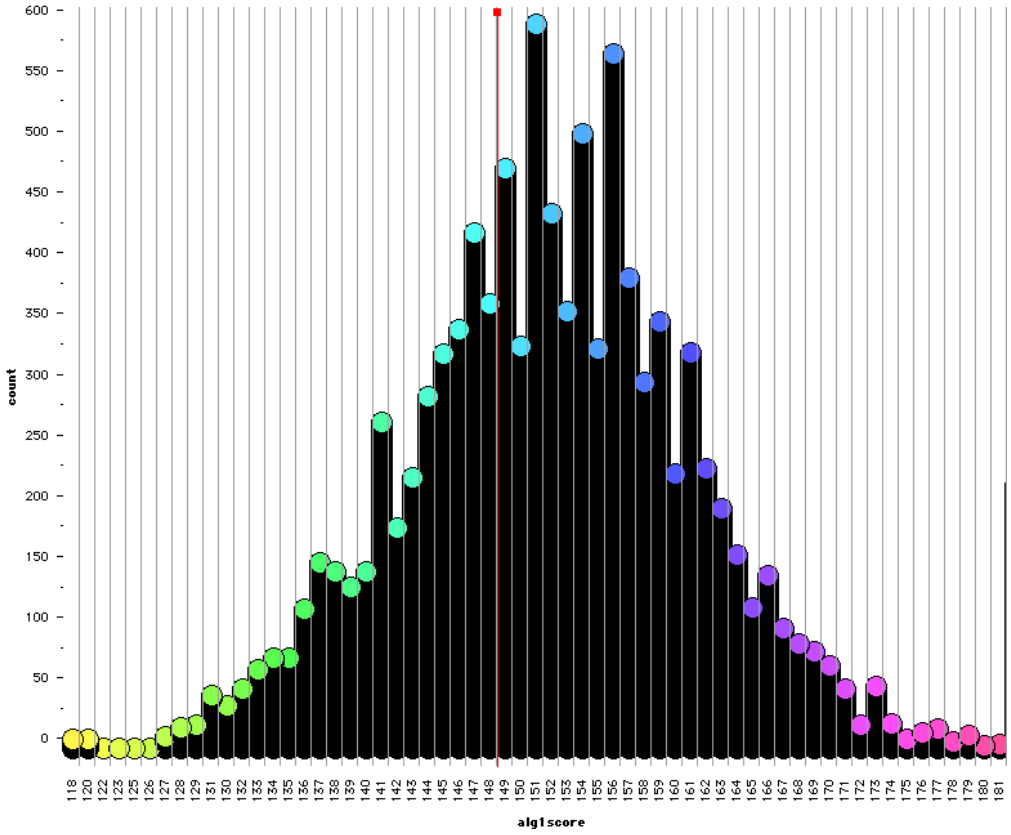


Figure 7. Distribution of Algebra I EOC scores for Research Question 1A.

For the sample of students taking the Algebra I EOC, the mean for Course 2 students was 156.87 and for Algebra I students it was 151.19, the passing rates were 83.72% and

66.98% respectively (Table 12). Students enrolled in integrated mathematics had higher means and passing rates than subject-specific students in all analyzed subgroups. Across both curricular pathways, students' mean Algebra I exam score and passing rates decreased as student grade level increased ($\mu_{9th} = 152.98, \mu_{10th} = 148.09, \mu_{11th} = 145.56, \mu_{12th} = 144.93$). This suggests students at higher grade levels tended to have weaker performance scores on the Algebra I EOC. Regardless of curriculum, females performed slightly higher than males and there was a decrease in mean achievement for students with more than twenty absences ($\mu_{Females} = 152.25, \mu_{Males} = 151.31; \mu_{DaysAbs \leq 20} = 152.40, \mu_{DaysAbs > 20} = 145.34$).

Table 12. Algebra I EOC Means and Passing Rates for Students Within Each Subgroup

		Grand Mean	Grade Level					Gender	
			6-8th	9th	10th	11th	12th	Male	Female
IM	Sample Size (n)	1038	-	588	411	34	5	451	587
	Mean	156.87	-	159.78	153.55	148.29	146.40	156.04	157.51
	Standard Dev.	9.62	-	8.87	9.14	8.67	10.69	10.07	9.22
	Passing Rate	83.72%	-	92.35%	74.45%	52.94%	40%	80.93%	85.86%
SS	Sample Size (n)	9139	476	6360	1711	498	94	4669	4470
	Mean	151.19	158.92	152.35	146.77	145.37	144.85	150.85	151.56
	Standard Dev.	8.97	9.81	8.56	7.90	6.74	6.90	9.28	8.63
	Passing Rate	66.98%	87.81%	73.33%	47.45%	39.16%	34.04%	66.16%	67.83%
Entire Sample	Mean	151.77	158.92	152.98	148.09	145.56	144.93	151.31	152.25
	Standard Dev.	9.20	9.81	8.82	8.58	6.91	7.06	9.46	8.91

		Race						Attendance	
		AmIn	Asian	Hispanic	Black	White	Multi	Absent ≤ 20 Days	Absent > 20 Days
IM	Sample Size (n)	14	51	107	191	650	26	987	51
	Mean	154.21	165.51	157.02	154.22	157.04	156.04	157.31	148.33
	Standard Dev.	8.70	7.78	10.0	8.55	9.66	156.04	9.46	8.70
	Passing Rate	78.57%	98.04%	85.98%	77.49%	84.15%	80.77%	85.31%	52.94%
SS	Sample Size (n)	160	145	964	3012	4607	251	8288	851
	Mean	147.54	155.60	149.76	148.42	153.25	152.02	151.81	145.16
	Standard Dev.	7.71	9.87	8.98	8.31	8.80	9.15	8.76	8.82
	Passing Rate	51.25%	80.69%	61.20%	55.05%	75.88%	70.92%	69.55%	41.95%
Entire Sample	Mean	148.08	158.18	150.49	148.77	153.72	152.40	152.40	145.34
	Standard Dev.	7.97	10.32	9.34	8.43	8.99	9.16	9.00	8.84

Fully Unconditional Model

To determine if there was sufficient variability in student achievement at each level for further analysis, it is recommended that a preliminary model, a fully unconditional model, be conducted (Raudenbush & Bryk, 2002). From this preliminary analysis, results indicated that 25.50% of the variance in student test score was between schools $(\tau_\beta = 25.52, Z = 4.27, p < 0.0001)$, 16.87% was between teachers $(\tau_\pi = 16.89, Z = 8.20, p < 0.0001)$, and 57.63% was within-teachers $(\sigma^2 = 57.69, Z = 70.20, p < 0.0001)$. Therefore, the fully unconditional model signified there was sufficient variability in each level to warrant further analysis. The grand mean for the sample was 151.64 $(t = 225.34, p < 0.0001)$.

Incorporating Student- and School-level Predictors

A second model (Appendix H) was used to determine, while controlling for student and school characteristics, if there were teacher level differences in student achievement associated with the curriculum a teacher used. The model with the slopes constrained to be equal across classes and schools resulted in a better fit than the model with the slopes free to vary $(\chi^2_{18} = 648.3, p < .05)$. Therefore, the slopes were constrained in subsequent analyses.

Controlling for student and school characteristics, there were significant curriculum type differences on the average Algebra I EOC score between teachers (Table 13). On average, integrated mathematics teachers had students who scored 4.64 points higher on the Algebra I EOC than subject-specific teachers. There were significant negative correlations between student achievement and grade level, days absent, all non-Asian minorities, and

school %FRL. There was a large effect of grade level, suggesting that for each increase in grade level there was a 2.86-point decrease in Algebra I EOC score. Each day a student was absent resulted in a slightly lower test score, as did the %FRL in a school.

Table 13. HLM 2 for Entire Algebra I Sample

Fixed Effects	Model 2		
	Coefficients	Std. Errors	t-Value
Algebra I Score, β_{00k}			
Intercept γ_{000}	151.48	0.49	307.43 ***
Curriculum Type γ_{010}	4.64	1.07	4.34 ***
Grade Level γ_{100}	-2.86	0.14	-20.75 ***
Days Absent γ_{200}	-0.17	0.008	-21.03 ***
Gender γ_{300}	-0.13	0.15	-0.90
American Indian γ_{400}	-2.01	0.59	-3.43 ***
Asian γ_{500}	0.31	0.55	0.55
Hispanic γ_{600}	-2.32	0.26	-9.01 ***
Black γ_{700}	-3.51	0.19	-18.49 ***
Multi-racial γ_{800}	-1.21	0.46	-2.63 **
ADM γ_{001}	-0.0005	0.0007	-0.65
%FRL γ_{002}	-0.062	0.031	-2.02 *
Title One γ_{003}	0.66	1.85	0.36
Random Effects	Estimate	Std. Errors	Z-Value
Within-teacher, σ^2	51.19	0.73	69.67 ***
Between-teacher, τ_π	7.94	1.06	7.48 ***
Between-school, τ_β	9.25	2.41	3.84 ***

Note: $*p<.05, **p<.01, ***p<.001$

There were significant negative effects on student achievement for non-Asian minorities. On average, Black students scored 3.51 points lower than White students, Hispanics 2.32 points lower than Whites, American Indians 2.01 points lower than Whites, and multi-racial students scored 1.21 points lower than Whites. The difference between achievement scores for Whites and Asians was non-significant. This model accounted for

11.25% of within-teacher variability, 52.98% of the between-teacher variability, and 63.75% of the between-school variability in student Algebra I EOC score.

Interactions of Curriculum Type

A third model was conducted to test for significant interactions of curriculum type. To determine if the relationship between student grade level and test score depended on curriculum type or if the relationship between school %FRL and test score depended on curriculum type, predictors were added for Curriculum x Grade Level and for Curriculum x %FRL. Further, due to the significance of grade level and %FRL in the previous analysis, the third model tested for a possible three-way interaction between curriculum type, %FRL, and grade level. Also note, there were no significant interactions between curriculum type and race, so they were removed from this model.

The fixed effects of this model were similar to the second model (Table 14). Integrated mathematics teachers tended to have students with higher test scores than subject-specific teachers. There were still significant negative correlations between grade level, days absent, and non-Asian minorities. The interaction between curriculum type and %FRL was non-significant. However, there were two significant cross-level interactions: Curriculum x Grade level and Grade level x %FRL.

Table 14. HLM 3 for Algebra I Entire Sample

Fixed Effects	Model 3		
	Coefficients	Std. Errors	t-Value
Algebra I Score, β_{100}			
Intercept γ_{000}	151.64	0.54	279.29 ***
Curriculum Type γ_{010}	4.77	1.14	4.19 ***
Grade Level γ_{100}	-2.87	0.14	-20.67 ***
Days Absent γ_{200}	-0.17	0.008	-21.00 ***
Gender γ_{300}	-0.11	0.15	-0.72
American Indian γ_{400}	-1.96	0.59	-3.34 ***
Asian γ_{500}	0.29	0.55	0.53
Hispanic γ_{600}	-2.24	0.26	-8.69 ***
Black γ_{700}	-3.48	0.19	-18.40 ***
Multi-racial γ_{800}	-1.12	0.46	-2.44 *
ADM γ_{001}	-0.0003	0.0008	-0.33
%FRL γ_{002}	-0.057	0.033	-1.70
Title One γ_{003}	0.61	2.10	0.29
Curriculum*%FRL γ_{012}	0.018	0.066	0.27
Curriculum*Grade Level γ_{110}	-1.33	0.54	-2.48 *
Grade Level*%FRL γ_{102}	0.064	0.01	6.18 ***
Curriculum*Grade*%FRL γ_{112}	-0.10	0.038	-2.73 **
Random Effects	Estimate	Std. Errors	Z-Value
Within-teacher, σ^2	50.86	0.73	69.64 ***
Between-teacher, τ_π	8.06	1.08	7.48 ***
Between-school, τ_β	11.59	2.98	3.89 ***

Note: $*p<.05, **p<.01, ***p<.001$

First, the relationship between grade level and student achievement was dependent on curriculum type. Integrated mathematics students enrolled in low grades levels (operationalized as one standard deviation below the average grade level of the sample) scored on average 1.76 points higher than subject-specific students enrolled in low grade

levels ($p<.0001$). Integrated mathematics students at higher grade levels (operationalized as one standard deviation above the average grade level of the sample) tended to score 1.19 points higher than their subject-specific peers ($p=0.001$). This suggests, for the Algebra I EOC, integrated mathematics students scored higher than subject-specific students regardless of the grade level of the student. Further, integrated mathematics and subject-specific mathematics students at lower grade levels had higher achievement scores than students at higher grade levels ($p<.0001$).

Secondly, the relationship between grade level and student achievement was dependent on the school %FRL. Students at lower grade levels had higher achievement. For students at lower grade levels, each one unit increase in a school's %FRL resulted in a 1.32 point decrease in Algebra I EOC achievement ($p=0.004$). However, students at high grade levels performed similarly, regardless of school need.

Although there was neither a significant relationship between %FRL and student achievement, nor a significant interaction between curriculum type and %FRL, there was a significant three-way interaction between curriculum type, grade, and %FRL. The relationship between grade level and student achievement significantly depends on %FRL and curriculum type (Figure **8**). In order to interpret the three-way cross-level interaction, two additional models were conducted; one for students enrolled in a lower grade level and one for those in a higher grade level. The slopes in school %FRL (low %FRL was operationalized as one standard deviation below the average %FRL of the sample and high %FRL as one standard deviation above the average %FRL of the sample) from the two

additional models were then plotted for each curricular pathway. The difference in contrasts for high %FRL schools was significant for both integrated and subject-specific mathematics, such that teachers of integrated mathematics with students at low grade levels had a 1.50 point ($p<0.001$) higher performance on the Algebra I EOC than subject-specific teachers, and at high grade levels integrated mathematics outperformed subject-specific by 2.00 points ($p<0.0001$). There was no difference between the curriculum types for students in low grade levels at low %FRL schools, however there was a difference for students at higher grade levels. Integrated mathematics teachers of students at higher grades in low %FRL schools had student achievement scores that averaged 2.00 points ($p<.0001$) higher than students of subject-specific teachers. Thus, with the exception of curriculum not impacting low %FRL students at lower grade levels, teachers of integrated mathematics saw significantly higher results at higher grade levels and in high need schools.

The slope of the integrated mathematics lines (red) were not significant for either low or high grade levels. This suggests that integrated mathematics had similar effects on student achievement at both low and high %FRL schools. The moderate significance for the negative slope of the subject-specific lines (blue) indicated that for each one unit increase in the school need there was a 1.18 point decrease on the Algebra I EOC for teachers using subject-specific curricula at low grade levels ($p=0.075$) and a 0.76 point decrease at high grade levels ($p=0.10$).

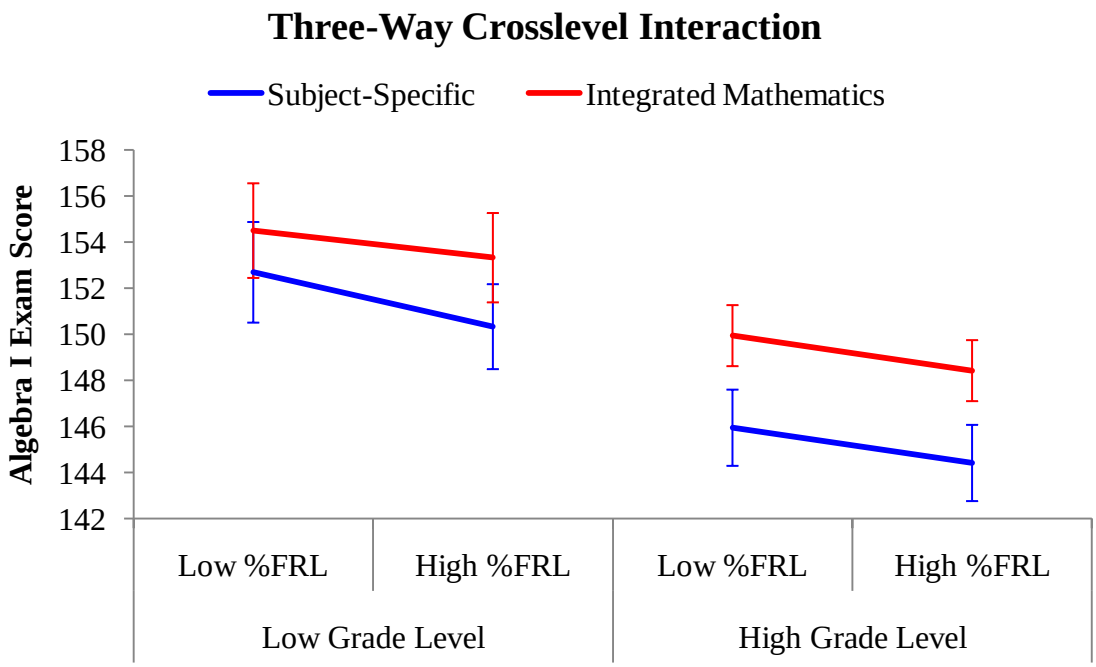


Figure 8. Three-Way Crosslevel Interaction: Curriculum, %FRL, Grade Level

The third model accounted for 11.82% of within-teacher variability, 52.27% of the between-teacher variability, and 54.58% of the between-school variability in student Algebra I EOC score.

Data in this section suggest that integrated mathematics teachers had higher student performance on the Algebra I exam than subject-specific teachers, regardless of student grade level. Individually both curricular pathways had higher achievement for students at lower grades, but integrated mathematics teachers tended to have students at high grade levels perform better than subject-specific students at low grades. Thus, integrated mathematics teachers had more consistent results across the grade levels. In addition, integrated mathematics appeared to be more effective at schools with high %FRL student populations. Another trend was the significantly lower results for non-Asian minorities, as

compared to White students. These negative results were apparent in students for both pathways.

Modeling Student Algebra II Achievement: Entire Sample

Background and Demographics

Algebra II EOC data was provided for 67 schools. Seven schools offered integrated mathematics and 63 offered subject-specific mathematics (three offered both). The sample was comprised of 676 students enrolled in integrated mathematics Course 3 and 8,641 in Algebra II (subject-specific). This sample accounted for 12% of the total population of students taking the Algebra II EOC in NC during the 2009-2010 school year. Schools with students in Course 3 had a higher %FRL, lower average daily membership, and were more frequently classified as Title One than the schools with Algebra II students (Table **15**).

Table 15. School Demographics for Algebra II Sample

	Average %FRL	Average ADM	Percent Title One
Integrated Mathematics	61.56	395.29	42.86
Subject-Specific	50.39	682.98	4.76
Total	50.99	653.28	7.46

Demographics for students in the integrated and subject-specific mathematics samples were similar in gender and race[13] (Table **16**). There was a higher percentage of students taking Algebra II in 9th grade than taking Course 3. Students taking the Algebra II exam in 9th grade are typically advanced for their grade level. A higher percentage of Course

[13] Overall, while there were more White students included in this sample than other minority populations, the other minorities were similarly represented in the Algebra II sample as the Algebra I sample. In other words, it does not appear that the representation of minorities in Course 3 or Algebra II had decreased after the first high stakes exam. Keep in mind, this was not longitudinal data, both Algebra I and II exams were given during the 2009-2010 school year on a different sample of students.

3 students took the Algebra II EOC in tenth or eleventh grade compared to Algebra II students, though more Algebra II students took the exam their senior year. Students in Course 3 averaged 6.75 days absent, while Algebra II students averaged 7.11 days.

Table 16. Percentages of Students Within Each Category for Algebra II Sample

			Grade			**Gender**	
	6-8th	9th	10th	11th	12th	Male	Female
Integrated Math	0.00	2.37	47.93	46.60	3.11	43.05	56.95
Subject-Specific	0.01	10.52	44.40	38.61	6.45	46.64	53.36
Total	0.01	9.93	44.66	39.20	6.20	53.62	46.38
				Race			
	AmIn	Asian	Hispanic	Black	White	Multi	
Integrated Math	0.30	1.63	6.36	31.07	58.14	2.51	
Subject-Specific	1.56	2.84	6.86	28.04	58.22	2.48	
Total	1.47	2.75	6.83	28.26	58.22	2.48	

Distribution and Summary Statistics

The distribution of scores for this sample is presented in Figure **9**. The possible range of scores on this assessment was 129-178. The vertical red bar represented the minimum score a student must earn to receive a passing grade on the assessment, 148. The average score among all North Carolina students taking the assessment during the 2009-2010 school year was 153.1 with standard deviation 8.90, the average for this sample was 152.54 with standard deviation 8.73.

For this sample, the percent of students passing the Algebra II exam were similar between Course 3 and Algebra II students at 72.26% and 72.65%, respectively (Table **17**). Despite integrated mathematics students having an almost two-point higher mean prior achievement score $(\mu_{IM} = 158.02, \sigma_{IM} = 8.51; \mu_{S-S} = 156.07, \sigma_{S-S} = 8.88)$, as measured by the

Algebra I exam, subject-specific students had a slightly higher mean Algebra II score at 152.59, compared to a mean of 151.93 for integrated mathematics students. Subject-specific students had higher mean achievement on all subgroups except for students classified as American Indian or Black and for 11th and 12th grade students. Similar patterns persisted on the Algebra II exam, there was a decrease in mean achievement as grade level increased ($\mu_{9th} = 158.12, \mu_{10th} = 154.61, \mu_{11th} = 149.73, \mu_{12th} = 146.29$), females had slightly higher results than males ($\mu_{Female} = 152.63, \mu_{Male} = 152.45$), and there was a decrease in performance for students with more than 20 absences ($\mu_{DaysAbs \leq 20} = 152.86, \mu_{DaysAbs > 20} = 146.39$).

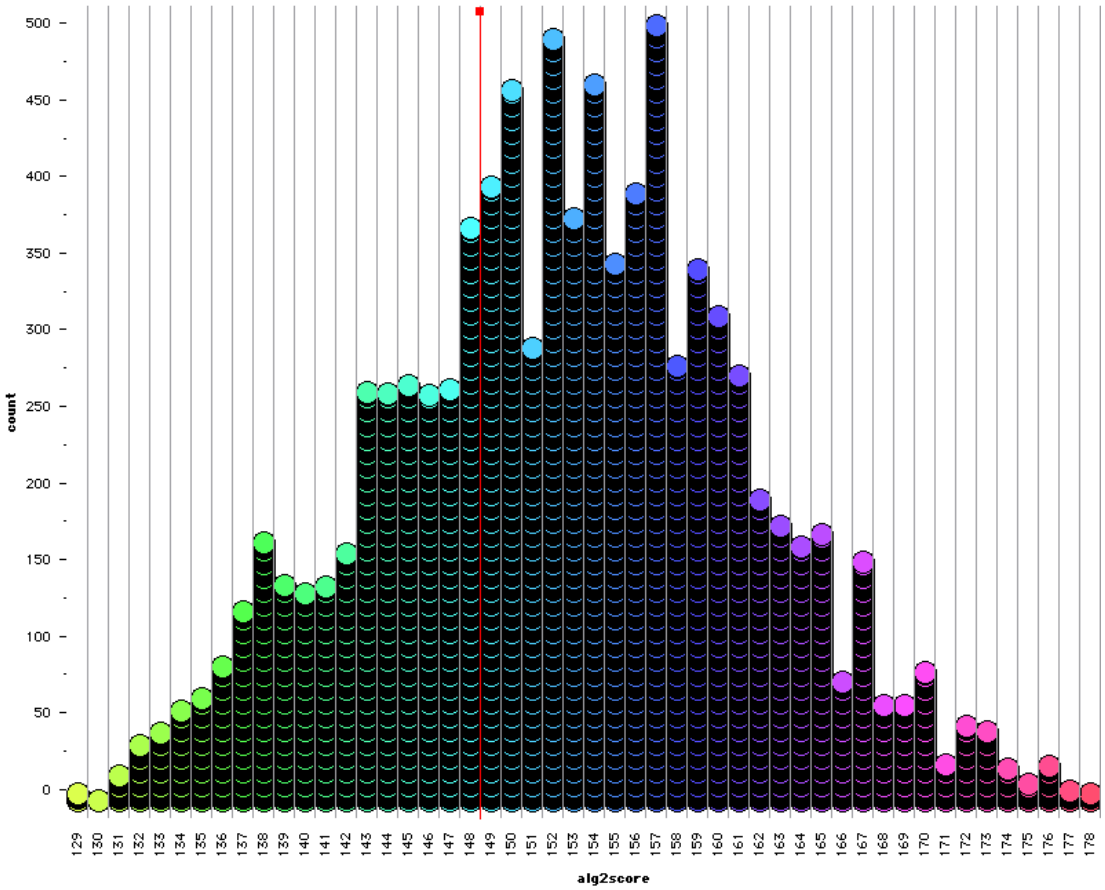


Figure 9. Distribution of Algebra II EOC scores for Research Question 1A.

Table 17. Algebra II EOC Means and Passing Rates for Students Within Each Subgroup

		Grand Mean	Grade Level					Gender	
			8th	9th	10th	11th	12th	Male	Female
IM	Sample Size (n)	674	-	16	324	313	21	290	384
	Mean	151.93	-	153.81	153.69	150.35	146.90	151.79	152.04
	Standard Dev.	7.90	-	5.95	7.89	7.52	8.44	7.64	8.10
	Passing Rate	72.26%	-	87.50%	79.63%	64.54%	61.90%	72.07%	72.40%
SS	Sample Size (n)	8566	1	908	3818	3291	548	3984	4582
	Mean	152.59	157.00	158.20	154.68	149.67	146.26	152.49	152.68
	Standard Dev.	8.79	-	8.66	8.65	7.53	6.99	8.98	8.62
	Passing Rate	72.65%	100.00%	89.32%	80.57%	63.69%	43.61%	71.74%	73.44%
Entire Sample	Mean	152.54	157.00	158.12	154.61	149.73	146.29	152.45	152.63
	Standard Dev.	8.73	-	8.63	8.59	7.53	7.04	8.90	8.59

		Race						Attendance	
		AmIn	Asian	Hispanic	Black	White	Multi	Absent ≤ 20 Days	Absent > 20 Days
IM	Sample Size (n)	2	11	43	210	391	17	651	23
	Mean	154.00	156.91	150.40	149.65	153.13	152.94	152.15	145.83
	Standard Dev.	0	6.01	8.94	7.29	7.65	12.03	7.85	7.19
	Passing Rate	100.00%	100.00%	65.12%	63.33%	77.24%	64.71%	73.27%	43.48%
SS	Sample Size (n)	135	242	587	2406	4984	212	8143	423
	Mean	149.59	158.61	151.76	148.75	154.30	153.35	152.91	146.42
	Standard Dev.	8.72	8.61	7.94	7.90	8.63	8.30	8.71	8.05
	Passing Rate	60.74%	90.91%	72.23%	57.27%	79.45%	75.00%	74.09%	44.92%
Entire Sample	Mean	149.65	158.54	151.66	148.82	154.22	153.32	152.86	146.39
	Standard Dev.	8.68	8.51	8.01	7.86	8.57	8.60	8.65	8.00

Fully Unconditional Model

The fully unconditional model indicated that 10.03% of the variance in student test score was between schools ($\tau_\beta = 8.24, Z = 2.61, p < 0.01$), 26.73% was between teachers ($\tau_\pi = 21.97, Z = 7.12, p < 0.0001$), and 63.24% was within-teachers ($\sigma^2 = 51.98, Z = 67.23, p < 0.0001$). There was sufficient variability in each level to warrant further analysis. The grand mean for the sample was 151.12 ($t = 284.91, p < 0.0001$).

Incorporating Student- and School-level Predictors

A second model was used to determine, while controlling for student and school characteristics, if there were teacher level differences in student achievement associated with the curriculum a teacher used. Controlling for student and school characteristics, there were no significant curriculum type differences on the average Algebra II EOC score between teachers (Table **18**).

Prior achievement was a positive significant predictor for Algebra II achievement. On average, for each 1-point increase in a student's Algebra I EOC score there was a 0.49-point increase in their Algebra II EOC score. On average, Asian students outperformed White students by 1.68 points and White students outperformed Black students by 0.77 points. White students neither outperformed American Indian, Hispanic, nor Multi-racial students on the Algebra II exam. Female students significantly outperformed male students on the Algebra II EOC. There were significant student level negative effects on Algebra II achievement for grade level and days absent.

Table 18. HLM 2 for Entire Algebra II Sample

Fixed Effects	Model 2		
	Coefficient	Std. Error	t-Value
Algebra II Score, β_{0i}			
Intercept γ_{000}	152.29	0.35	441.30 ***
Curriculum Type γ_{010}	-0.97	1.19	-0.81
Algebra1score γ_{100}	0.49	0.0086	56.58 ***
Grade Level γ_{200}	-1.65	0.11	-15.23 ***
Days Absent γ_{300}	-0.14	0.0086	-16.08 ***
Gender γ_{400}	-0.65	0.12	-5.26 ***
American Indian γ_{500}	-0.66	0.52	-1.26
Asian γ_{600}	1.68	0.40	4.21 ***
Hispanic γ_{700}	-0.016	0.26	-0.06
Black γ_{800}	-0.77	0.17	-4.53 ***
Multi-racial γ_{900}	-0.54	0.40	-1.33
ADM γ_{001}	0.0011	0.0005	2.28 *
%FRL γ_{002}	-0.054	0.022	-2.68 **
Title One γ_{003}	3.32	1.56	2.14 *
Random Effects	Estimate	Std. Error	Z-Value
Within-teacher, σ^2	30.19	0.47	63.73 ***
Between-teacher, τ_{π}	6.79	1.10	6.20 ***
Between-school, τ_{β}	2.66	1.24	2.15 *

Note: $(*p < .05, ** p < .01, ***p < .001)$

Each of the school level predictors had a significant effect on student Algebra II achievement. There were positive effects for ADM and Title One status and a negative school effect for %FRL. Schools that were classified as Title One tended to have scores 3.32 points higher on the Algebra II EOC than non-Title One schools. Recall, the majority of Title One schools in this sample teach using the integrated mathematics pathway. Considering Title One classification means a school is in need of significant academic improvement, this was an encouraging result for these schools. However, upon further analysis, students at Title One schools had a mean achievement of 149.58 and non-Title One schools had a mean

achievement of 152.67. Therefore, the significant positive effect of Title One schools is a result of holding student and school level characteristics constant. In other words, if all students and schools had similar traits then Title One schools would outperform non-Title One schools on the Algebra II EOC. Model 2 accounted for 41.91% of within-teacher variability, 69.10% of the between-teacher variability, and 67.76% of the between-school variability in student Algebra II EOC score.

Interactions of Curriculum Type

A third model was conducted to test for significant interactions of curriculum type. This model tested for a possible three-way cross-level interaction between curriculum type, prior achievement, and grade level to determine if the relationship between curriculum type and Algebra II score depended on prior achievement and student grade level. Despite a significant effect of Title One classification, there was no interaction between Title One status and curriculum type. In addition, despite the significance of each of the school-level variables (ADM, %FRL, and Title One) there was no interaction between the three and student Algebra II achievement. This suggests each of the school-level factors individually accounts for significant variance in student achievement.

The fixed effects of the interaction model were similar to the second model (Table 19). Curriculum type was not associated with Algebra II achievement. Neither American Indian, Hispanic, nor Multi-racial ethnicities were significantly different than Whites on student Algebra II achievement. There were still significant negative associations between grade level and Algebra II achievement and days absent and average Algebra II score. On

average, females scored higher than males on the Algebra II exam. The average daily membership in a school was positively associated with Algebra II achievement, and Title One schools tended to perform 3.3 points higher than non-Title One schools. The school's %FRL was negatively correlated to Algebra II achievement. Each of the interaction terms were significant. The relationship between prior achievement and Algebra II score depended on curriculum, the relationship between grade level and Algebra II achievement depended on curriculum, and the relationship between prior achievement and Algebra achievement depended on grade level. Furthermore, the relationship between curriculum type and Algebra II score significantly depended on prior achievement and student grade level. These interactions will be explained in more detail below. The final model accounted for 62.68% of the between-school, 71.12% of the between-teacher, and 41.91 within-teacher variance in Algebra II score.

Table 19. HLM 3 for Algebra II Entire Sample

Fixed Effects	Model 3		
	Coeff.	Std. Error	t-Value
Algebra II Score, B_{0i}			
Intercept γ_{000}	151.74	0.36	424.47 ***
Curriculum Type γ_{010}	-0.77	1.21	-0.64
Algebra1score γ_{100}	0.52	0.0084	61.30 ***
Grade Level γ_{200}	-1.77	0.11	-16.81 ***
Days Absent γ_{300}	-0.13	0.0083	-16.00 ***
Gender γ_{400}	-0.63	0.12	-5.30 ***
American Indian γ_{500}	-0.66	0.50	-1.32
Asian γ_{600}	1.50	0.38	3.90 ***
Hispanic γ_{700}	-0.018	0.25	-0.07
Black γ_{800}	-0.83	0.16	-5.04 ***
Multi-racial γ_{900}	-0.58	0.39	-1.48
ADM γ_{001}	-0.001	0.0005	2.06 *
%FRL γ_{002}	-0.054	0.122	-2.43 *
Title One γ_{003}	3.30	1.57	2.11 *
Curriculum*Algebra1score γ_{110}	0.12	0.031	3.97 ***
Curriculum*Grade Level γ_{210}	1.23	0.43	2.85 **
Grade Level*Algebra1score $\gamma_{(10)00}$	-0.18	0.0083	-22.21 ***
Curriculum*Grade*Algebra1score $\gamma_{(10)10}$	0.17	0.045	3.70 ***
Random Effects	Estimate	Std. Error	Z-Value
Within-teacher, σ^2	28.18	0.44	63.70 ***
Between-teacher, τ_{π}	6.35	1.06	5.96 ***
Between-school, τ_{β}	3.08	1.32	2.33 **

Note: $^{*}p < .05, ^{**}p < .01, ^{***}p < .001$

In order to interpret the three-way cross-level interaction, two additional models were conducted; one for students enrolled in a low grade level and one for those in a high grade level. The slopes in prior achievement from the two additional models were then plotted for each curricular pathway (Figure **10**). The left side of the graph shows results for students in low grade levels when taking the Algebra II EOC. For students at low grade levels, the slopes

for both subject-specific and integrated mathematics were positive and significant. For each one point increase in Algebra I prior achievement, there is a 5.60 point increase in Algebra II achievement for students taking integrated mathematics at low grade levels ($p<0.0001$) and a 5.34 point increase in Algebra II achievement for students taking subject-specific mathematics at low grade levels ($p<0.0001$). Integrated mathematics students in low grade levels with low prior achievement score 0.85 points lower on the Algebra II EOC than their subject-specific peers ($p=0.0086$). Integrated mathematics students enrolled in high grade levels tend to perform 0.59 points lower on the Algebra II EOC than their subject-specific peers ($p=0.05$). This suggests that teachers using subject-specific mathematics have higher Algebra II performance than integrated mathematics teachers for students at lower grade levels, regardless of students' prior achievement.

Results for students at higher grade levels are presented on the right of Figure **10**. For each one point increase in Algebra I achievement, there was a 3.79 point increase in Algebra II achievement for students taking integrated mathematics at high grade levels ($p<0.0001$) and a 2.54 point increase in Algebra II achievement for students taking subject-specific mathematics at high grade levels ($p<0.0001$). Differences in contrasts between integrated and subject-specific mathematics achievement on the Algebra II exam for students at high grade levels with low prior achievement favors subject-specific teachers, marginally significant ($p=0.086$). The difference in student Algebra II achievement for students at high grade levels with high prior achievement favors integrated mathematics students by 0.66 points, but was only marginally significant ($p=0.06$). These results suggest at higher grade levels there were

no significant performance differences on the Algebra II exam, but that integrated mathematics teachers had a greater increase in student Algebra II achievement score given a student's prior achievement.

Further, looking across the left and right sides of the graph in Figure **10** shows that integrated mathematics teachers had consistent results in Algebra II achievement regardless of the grade level of the student. Conversely, between subject-specific teachers with students that had high Algebra I scores, there was an almost 5 point difference in Algebra II score between low and high grade level students.

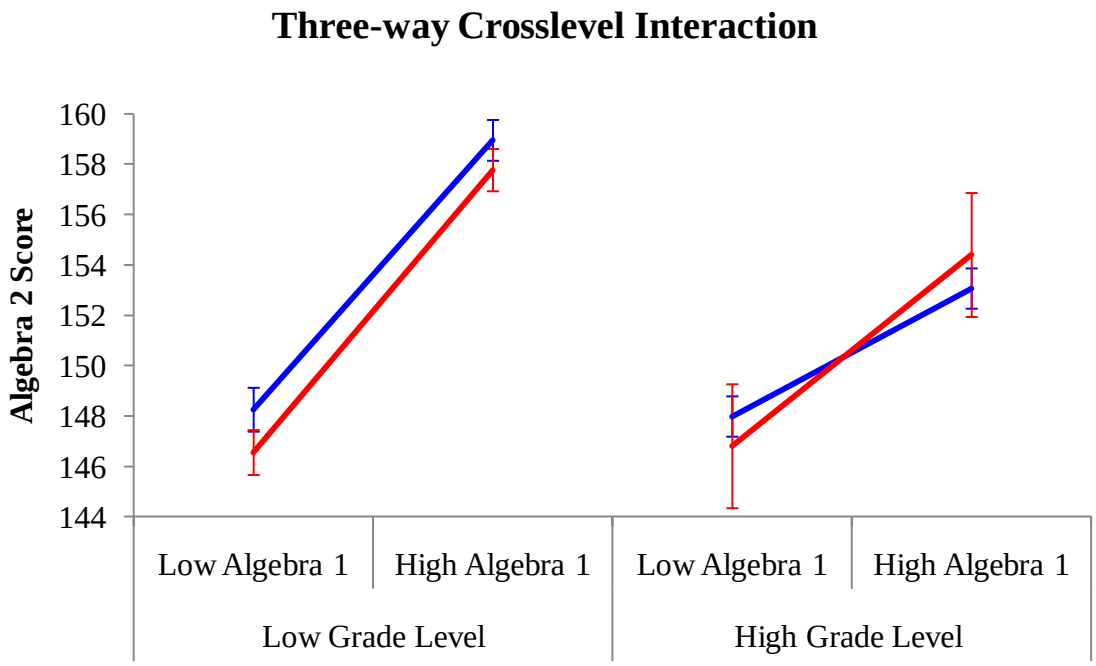


Figure 10. Three-Way Crosslevel Interaction Between Curriculum Type, Grade Level, and Prior Achievement

Results from the Algebra II EOC showed no significant differences between teachers of the two curricular pathways. The inclusion of students' prior achievement in models of

Algebra II achievement increased the amount of within-teacher variance that was explained. Neither pathway helped students with low prior achievement close the achievement gap with students of high prior achievement. Similar to the Algebra I results, integrated mathematics teachers had consistent results for students across grade levels, whereas subject-specific teachers had drops in student achievement for students at higher grade levels.

The magnitude of the significant negative results for non-Asian minorities on the Algebra I EOC decreased on the Algebra II EOC, suggesting the achievement gap may close after students are exposed to content through Course 3 or Algebra II. However, due to the correlation between race and prior achievement, the inclusion of prior achievement in models of Algebra II achievement could suppress the effect of race. To explore this further, the interaction between race, prior achievement, curriculum, and Algebra II achievement was tested to determine if the relationship between curriculum type and Algebra II score depended on prior achievement and student race. This three-way crosslevel interaction was significant for Black, Hispanic, and American Indian students (Appendix K) and non-significant for Asian and Multi-racial students. Despite the significant interaction, there were no significant differences in contrasts for any race based on curricular pathway and the only significant slopes were between White and Black students. White students using subject-specific mathematics with low prior achievement outperformed Black students using subject-specific mathematics with low prior achievement by 0.59 points on the Algebra II exam ($p<.0001$). Also, White students using integrated mathematics with high prior achievement outperformed Black students using integrated mathematics with high prior achievement by

0.29 points on the Algebra II exam (p=0.05). These results suggest, on average, the achievement gap dissipated by the end of Course 3, but there were still some differences between student achievement for White and Black students based on curriculum type.

An encouraging result was the strong positive correlation between Title One schools and Algebra II achievement. In the sample, 43% of schools that taught integrated mathematics were Title One and 5% of subject-specific schools were Title One. The strong effect between Title One status and the outcome measure could speak favorably about the use of integrated mathematics in Title One schools.

Modeling Student Algebra I Achievement: Workshop Sample

This section models student Algebra I EOC achievement for the teachers that participated in one or more phases of the NCIM project. The workshop sample of teachers allowed for the introduction of additional predictors at level 2. In order for the correlations of each of these predictors to be analyzed this section is split into four separate pieces; following a description of the sample. First, the unconditional model will be presented. Second, controlling for curriculum type, level 1, and level 3 predictors, each teacher level predictor was individually added into the model to test for correlations between that specific predictor and Algebra I score. Third, a model simultaneously incorporating various teacher level predictors was analyzed. Finally, significant interactions between curriculum type and significant individual teacher level predictors were tested.

Background and Demographics

The subset of Algebra I EOC data for workshop participants was provided for 48 schools. Twelve schools offered integrated mathematics and 37 offered subject-specific mathematics (one offered both). The sample was comprised of 572 students enrolled in integrated mathematics Course 2 and 1,986 in Algebra I (subject-specific). Schools offering integrated mathematics had a slightly higher %FRL, lower average daily membership, and were more frequently classified as Title One or as a North Carolina New Schools Project School (NCNSP) school than the schools with Algebra I students (Table **20**).

Table 20. School Demographics for Workshop Teachers on Algebra I EOC

	Average %FRL	Average ADM	Percent Title One	Percent NCNSP
Integrated Mathematics	52.07	499.64	25.00	75.00
Subject-Specific	49.12	694.43	2.70	35.14
Total	49.80	646.81	8.16	44.90

There were 15 integrated mathematics and 51 subject-specific mathematics teachers included in this sample (three teachers taught both curricula). Teachers scored about 74 points on the content knowledge assessment (CKA), regardless of the curriculum they taught (Table 21). Integrated mathematics teachers were exposed to about 45 additional hours of NCIM professional development than subject-specific teachers, while subject-specific teachers tended to have 5 more years of teaching experience.

Table 21. Teacher Demographics for Workshop Teachers on Algebra I EOC

	Average CKA Score	Average Professional Development Hours	Average Number Years Taught
Integrated Mathematics	73.93	72.25	7.80
Subject-specific	73.74	26.71	13.18

Demographics for students in the integrated and subject-specific mathematics samples were not as similar as the larger data set (Table **22**). There was a higher percentage of students taking Algebra I in 9th grade than taking Course 2, and a higher percentage of Course 2 students in 10th grade than Algebra I students. There were about the same number of males as females for subject-specific students, but there were a much higher percentage of females enrolled in integrated mathematics than males. There was a higher percentage of Asian and Hispanic students represented in the integrated mathematics sample. There were no American Indian students in this sample, so the American Indian predictor has been removed from further Algebra I EOC analyses. Students in Course 2 averaged 6.24 days absent, while Algebra I students averaged 8.31 days.

Table 22. Percentages of Students Within Each Category for Workshop, Algebra I Sample

			Grade			**Gender**	
	6-8th	9th	10th	11th	12th	Male	Female
Integrated Math	0.00	65.73	31.99	2.10	0.17	38.99	61.01
Subject-Specific	3.12	74.92	15.96	4.83	1.16	49.40	50.60
Total	2.42	72.87	19.55	4.22	0.94	47.07	52.93
				Race			
	AmIn	Asian	Hispanic	Black	White	Multi	
Integrated Math	0.00	7.34	11.89	30.07	45.80	3.15	
Subject-Specific	0.00	1.21	8.26	36.86	48.48	2.52	
Total	0.00	2.58	9.07	35.34	47.89	2.66	

Means and Passing Rates

For this sample, integrated mathematics students had a mean of 157.80, 5.22 points higher than the mean for subject-specific students $\left(\mu_{S-S} = 151.80\right)$. Integrated mathematics students had a mean of at least four points higher than subject-specific students for every subgroup except for 11th and 12th grade students. The largest mean difference, in favor of integrated mathematics, was for Asian students at 13.80 points. There were much higher passing rates for integrated mathematics students overall (86.29% vs. 69.00% for subject-specific students), for all subgroups except 12th grade (there was only one integrated mathematics student in the 12th grade sample) (Table 23).

Table 23. Algebra I EOC Means and Passing Rates for Students Within Each Subgroup

		Grand Mean	Grade Level					Gender	
			8th	9th	10th	11th	12th	Male	Female
IM	Sample Size (n)	569	-	375	181	12	1	222	347
	Mean	157.80	-	159.69	154.86	145.42	135.00	157.15	158.22
	Standard Dev.	9.63	-	9.10	9.49	5.18	0.00	10.13	9.30
	Passing Rate	86.29%	-	91.73%	77.90%	50.00%	0.00%	84.68%	87.32%
SS	Sample Size (n)	1942	61	1463	306	89	23	960	982
	Mean	151.58	151.77	153.03	146.68	145.70	146.61	151.25	151.90
	Standard Dev.	8.53	9.00	8.29	7.53	6.24	7.31	8.98	8.06
	Passing Rate	69.00%	70.49%	75.73%	47.06%	38.20%	47.83%	67.81%	70.16%
Entire Sample	Mean	152.99	151.77	154.39	149.72	145.66	146.13	152.36	153.55
	Standard Dev.	9.17	9.00	8.87	9.20	6.10	7.53	9.49	8.84

		Race						Attendance	
		AmIn	Asian	Hispanic	Black	White	Multi	Absent ≤ 20 Days	Absent > 20 Days
IM	Sample Size (n)	10	41	68	171	261	18	548	21
	Mean	154.20	165.93	157.97	154.19	159.10	156.22	158.13	149.43
	Standard Dev.	8.42	7.42	10.44	8.65	9.50	7.19	9.54	8.27
	Passing Rate	80.00%	100.00%	88.24%	77.19%	90.04%	83.33%	87.23%	61.90%
SS	Sample Size (n)	51	24	157	717	944	49	1800	142
	Mean	147.73	152.13	150.68	148.97	153.87	152.24	152.06	145.53
	Standard Dev.	7.92	8.38	7.92	8.55	7.95	9.15	8.37	8.33
	Passing Rate	56.86%	79.17%	65.61%	56.49%	79.34%	71.43%	70.89%	45.07%
Entire Sample	Mean	148.79	160.83	152.88	149.98	155.00	153.31	153.47	146.03
	Standard Dev.	8.29	10.23	9.36	8.81	8.58	8.79	9.03	8.40

Similarities and Differences Between Two Samples: Larger and Workshop

To situate the workshop sample within the larger sample, a brief comparison of the two will be given. Integrated mathematics students in the workshop Algebra I sample averaged one point higher on the EOC than integrated mathematics students in the larger sample (Table 12, Table 23). Similarly, subject-specific students in the workshop sample scored higher on the EOC, by an average of 0.40 points. The combined means for the integrated mathematics and subject-specific workshop sample was higher for all subgroups except for 6-8th grade students (which were only represented by SS sample). The workshop sample had a higher percentage of integrated and subject-specific mathematics students taking the exam in the 9th grade, so these students were taking Algebra I earlier than students in the larger sample. There were higher percentages of female, Black, and Asians students in the workshop sample. The percentage of Black students in workshop sample was over 10% more than larger sample, as a result the percentage of White students decreased for integrated mathematics workshop sample. School %FRL was two percentage points higher at the workshop schools, ADM was about 50 students lower, and the percent of Title One was much higher for the workshop sample of schools (Table **10**, Table **20**). Taken together the workshop sample of students had higher achievement on the Algebra I exam $(\mu_{Lar.} = 151.77, \mu_{S-S} = 152.99)$, contained a more diverse student population.

Fully Unconditional Model

The fully unconditional model indicated that 12.26% of the variance in student test score was between schools $(\tau_\beta = 10.54, Z = 8.17, p = 0.098)$, 21.16% was between teachers

($\tau_{\pi} = 18.20, Z = 7.08, p = 0.005$), and 66.58% was within-teachers ($\tau_{\beta} = 57.27, Z = 1.64, p < 0.0001$). The variance between schools was moderately significant at $\alpha = 0.10$, therefore the fully unconditional model signified there was sufficient variability in each level to warrant further analysis. The grand mean for the sample was 152.62 ($t = 203.95, p < 0.0001$).

Individual Teacher Level Predictors

The next models were used to determine, while controlling for curriculum type and student and school characteristics, if there were individual teacher level differences in student achievement associated with individual teacher level predictors. These models were the same as the one presented in Appendix H with the addition of $\beta_{02k}(Level2\,\text{var}. - \overline{X})$ at level 2. In seven separate analyses, each of the teacher level variables was individually tested to determine its correlation with Algebra I achievement (Table **24**).

Table 24. Correlations Among Individual Level-2 Predictors and Algebra I Score, When Controlling for Curriculum Type

Level-2 Predictor	Coefficients	Standard Error	t-Value
Curriculum Type γ_{010}	3.96	1.24	3.19 ***
License γ_{020}	1.81	1.82	0.99
Content Knowledge γ_{020}	0.07	0.03	2.18 *
Years Taught γ_{020}	-0.04	0.04	-0.66
Teacher Gender γ_{020}	-1.52	1.09	-1.40
College Major γ_{020}	0.04	1.06	0.04
NCIM Professional Development γ_{020}	-0.02	0.02	-1.32

Note. ($*p < 0.05, ** p < 0.01, ***p < 0.001$)

Without including any other teacher level predictors, curriculum type was significantly related to Algebra I EOC score. On average, students of teachers of integrated

mathematics scored almost four points higher than students of subject-specific teachers on the Algebra I EOC. Teacher level differences in Algebra I EOC score were also associated with teacher content knowledge. The other five teacher-level predictors had non-significant correlations with Algebra I achievement when controlling for curriculum type. Despite teacher gender and NCIM professional development hours being non-significant, the magnitude of the effect of curriculum type jumped when controlling for each of these individually. In other words, when controlling for teacher gender, the effect of curriculum type rose to 4.28 ($t = 3.45, p = 0.0006$), and when controlling for NCIM professional development hours the effect of curriculum rose even higher to 4.43 $t = 3.24, p = 0.0012$). Taken together, these findings suggest that while teacher gender and NCIM professional development hours were not directly correlated to student achievement, these factors interacted with curriculum type in unique ways. The impact of these two factors, along with other teacher-level factors, is the focus of the next two sections.

Combination of Teacher Level Predictors

Accounting for student, teacher, and school characteristics, a third set of models were run to determine if there was evidence of differential impact on student Algebra I performance between the two curricular types. Including all seven teacher level predictors was problematic for several reasons. First, including each variable conflated the missing data and resulted in losing 45% of the original data. Further, there was not enough variance in teacher licensure causing no estimate for this predictor. As a result, licensure was removed from this analysis.

Model 3a, 3b, and 3c incorporated the same student and school level variables as well as various teacher level variables. Model 3a excluded professional development hours, 3b excluded teacher content knowledge, and 3c excluded both professional development hours and teacher content knowledge. Each of the three models has similar effects on student level predictors, but there were differences in the magnitude and significance of the effects at the teacher and school levels (Table **25**). Furthermore, there were significant negative associations of grade level, days absent, Hispanic, and Black on Algebra I achievement for each model and non-significant effects for student gender and for Asian and Multi-racial students.

In addition to controlling for teacher gender, college major, and number of years taught, Model 3a included a predictor for teacher content knowledge. There were significant teacher level differences in Algebra I EOC score associated with teacher content knowledge. Further, students of teachers of integrated mathematics scored on average 3.36 points higher than students of subject-specific teachers on the outcome measure. None of the other teacher levels variables were correlated to Algebra I score. None of the school level effects were significantly associated with Algebra I achievement, neither ADM, %FRL, nor Title One status. This model accounted for 82.83% of the between-school, 53.46% of the between-teacher, and 11.18% of the within-teacher variance in Algebra I EOC score.

Model 3b was the same as the previous model except that teacher content knowledge was replaced by hours of NCIM professional development. There was a non-significant negative relationship between professional development and Algebra I score, suggesting that

more hours of NCIM professional development did not directly strengthen student outcomes on the Algebra I EOC[14]. However, teacher gender was significantly related to the outcome variable. On average, female teachers had student Algebra I scores that were two points higher than male teachers. The only other significant level-2 predictor was curriculum type. Of these three models, the magnitude of the fixed effect for curriculum was the highest for Model 3b, at 5.06 in favor of integrated mathematics. Model 3b resulted in significant school effects for Title One status and average daily membership. The direction of the Title One effect suggests that schools classified as Title One tended to score 3.25 points higher than non-Title One schools on the Algebra I EOC. Average daily membership was negatively correlated to Algebra I achievement. Model 3b accounted for all of the between-school, 42.20% of the between-teacher, and 11.96% of the within-teacher level variance in Algebra I EOC score.

The final model in the third set excluded both content knowledge and professional development. Aside from the significant level-1 predictors mentioned earlier, there were only two significant effects on Algebra I achievement, curriculum type and %FRL. The effect of curriculum favored integrated mathematics teachers by 4.24 points. This model accounted for 59.69% of between-school, 51.60% of between-teacher, and 12.29% of within-teacher variance in Algebra I achievement.

These three models showed that regardless of which level-2 predictors were included in the models, it was important to consider student grade level, attendance, and ethnicity.

[14] Teachers with the most NCIM professional development experience taught in the highest need schools. These teachers will be separately analyzed in Chapter 5.

Further, curriculum type was significant in favor of integrated mathematics regardless of which level-2 predictors were added. Each model resulted in no significant between-school variability in the level of Algebra I score, which suggests that the majority of variance had been accounted. Finally, interesting effects occur when including either content knowledge assessment or professional development. The fourth set of models will continue to unravel the intricacies involved in determining the impact curriculum type had on Algebra I achievement.

Table 25. Third Set of HLM for Algebra I Workshop Sample

Fixed Effects	Model 3a (No PD)			Model 3b (No CKA)			Model 3c (Neither PD nor CKA)		
	Coeff.	Std. Error	t-Value	Coeff.	Std. Error	t-Value	Coeff.	Std. Error	t-Value
Algebra I Score, β_{01k}									
Intercept γ_{000}	1.52.32	0.55	274.61 ***	152.70	0.48	320.03 ***	152.41	0.53	285.06 ***
Curriculum Type γ_{010}	3.36	1.33	2.52 **	5.06	1.38	3.66 ***	4.24	1.29	3.29 ***
Grade Level γ_{100}	-3.25	0.38	-8.61 ***	-2.60	0.35	-7.51 ***	-2.84	0.32	-9.00 ***
Days Abs. γ_{200}	-0.22	0.024	-9.41 ***	-0.22	0.021	-10.81 ***	-0.22	0.019	-11.28 ***
Gender γ_{300}	0.55	0.36	1.53	0.12	0.32	0.39	0.19	-.30	0.65
Asian γ_{400}	0.028	1.20	0.03	0.30	1.07	0.28	0.19	1.03	0.19
Hispanic γ_{500}	-2.33	0.70	-3.32 ***	-2.12	0.57	-3.70 ***	-2.30	0.55	-4.21 ***
Black γ_{600}	-3.61	0.45	-8.05 ***	-3.14	0.37	-8.40 ***	-3.38	0.36	-9.50 ***
Multi-racial γ_{700}	-0.82	1.12	-0.74	-1.17	1.00	-1.16	-0.83	0.96	-.87
CKA γ_{020}	0.071	0.031	2.28 *						
NCIM PD Hours γ_{030}				-0.024	0.018	-1.37			
Yrs. Taught γ_{040}	-0.031	0.056	-0.56	0.021	0.053	0.39	-0.029	0.055	-0.53
T. Gender γ_{050}	-2.00	1.20	-1.67	-2.08	1.03	-2.03 *	-1.79	1.12	-1.59

Table **25**. Continued

College Major γ_{060}	0.96	1.19	0.80	-0.74	1.10	-0.67	0.0061	1.09	0.01
ADM γ_{001}	-0.0005	0.0009	-0.50	-0.079	0.035	-2.25 *	-0.00087	0.001	-0.93
%FRL γ_{002}	-0.0003	0.044	0.01	-0.001	0.001	-0.83	-0.084	0.040	-2.10 *
Title One γ_{003}	-0.25	2.31	-0.11	3.25	1.61	2.01 *	1.65	1.91	0.86
Random Effects	Estimate	Standard Errors	Z-Value	Estimate	Standard Errors	Z-Value	Estimate	Standard Errors	Z-Value
Within-teacher, σ^2	50.87	1.78	28.54 ***	50.42	1.57	32.18 ***	50.23	1.47	34.26 ***
Between-teacher, τ_{π}	8.47	4.85	1.75 *	10.52	2.52	4.17 ***	8.81	4.48	1.97 *
Between-school, τ_{β}	1.81	5.05	0.36	0.00			4.25	5.25	0.81

Note. $(*p < 0.05, ** p < 0.01, *** p < 0.001)$.

Interaction of Curriculum Type When Including Teacher Level Predictors

The fourth set of models were identical to the ones used in the previous section, with the addition of interaction terms to determine if the relationship between Algebra I score and various independent variables depended on curriculum type. There were non-significant interactions between curriculum type and content knowledge and between curriculum type and professional development. As a result, these were removed from the analysis. Based on the increased effect sizes on curriculum when including teacher gender, each model analyzed the interaction between curriculum type and teacher gender. Further, due to the strong negative correlation between students whose race was identified as Black and Algebra I achievement, an interaction term was added to test if the relationship between Algebra I achievement and being classified as Black depended on curriculum type. Note, there were similar negative significance between the Hispanic race and Algebra I achievement, but the interaction between Hispanic students and curriculum were non-significant and removed from the analyses.

These three models had similar results for level-1 predictors as the previous analyses (Table **26**). Compared to their respective models in section three, the magnitude of the effect of curriculum increased and was significant for each model, reaching the highest for the model that included professional development hours at 6.11 in favor of integrated mathematics. Not only were there significant interactions in each model, but also the fixed effect for curriculum was affected by the inclusion of the interaction terms.

Table 26. Interaction Terms for HLM for Algebra I Workshop Sample

Fixed Effects	Model 4a (No PD)			Model 4b (No CKA)			Model 4c (No PD nor CKA)		
	Coeff.	Std. Error	t-Value	Coeff.	Std. Error	t-Value	Coeff.	Std. Error	t-Value
Algebra I Score, β_{0jk}									
Intercept γ_{000}	152.41	0.52	295.40 ***	152.77	0.42	365.54 ***	152.41	0.52	294.44 ***
Curriculum Type γ_{010}	4.45	1.25	3.56 ***	6.11	1.25	4.88 ***	5.49	1.24	4.43 ***
Grade Level γ_{100}	-3.34	0.37	-9.05 ***	-2.58	0.34	-7.57 ***	-2.89	0.31	-9.35 ***
Days Absent γ_{200}	-0.23	0.023	-9.61 ***	-0.22	0.020	-10.94 ***	-0.22	0.019	-11.41 ***
Gender γ_{300}	0.58	0.36	1.63	0.16	0.32	0.49	0.21	0.30	0.71
Asian γ_{400}	-0.22	1.10	-0.21	0.039	1.06	0.04	-0.011	1.03	-0.01
Hispanic γ_{500}	-2.40	0.70	-3.43 ***	-2.16	0.57	-3.79 ***	-2.36	0.55	-4.31 ***
Black γ_{600}	-3.58	0.45	-7.98 ***	-3.20	0.37	-8.59 ***	-3.47	0.36	-9.76 ***
Multi-racial γ_{700}	-0.96	1.11	-0.87	-1.27	1.00	-1.27	-0.91	0.96	-0.95
CKA γ_{020}	0.052	0.027	1.93 *						
NCIM PD Hours γ_{030}				-0.031	0.016	-2.01 *			

Table **26**. Continued

Yrs. Taught γ_{040}	-0.049	0.048	-1.01	-.039	0.047	0.82	-0.027	0.049	-0.55
Teacher Gender γ_{050}	-1.20	1.08	-1.11	-1.68	0.91	-1.85	-1.31	0.98	-1.34
College Major γ_{060}	0.49	0.99	0.49	-0.98	0.97	-1.00	-0.19	0.91	-0.21
ADM γ_{001}	0.0001	0.0009	0.17	-0.00001	0.0008	-0.02	-0.0003	0.0009	-0.32
%FRL γ_{002}	0.0038	0.042	0.09	-0.070	0.031	-2.27 *	-0.082	-.039	-2.11 *
Title One γ_{003}	-1.43	2.23	-0.64	2.68	1.42	1.89	0.67	1.94	0.35
Title*Black γ_{610}	-2.40	0.94	-2.54 **	-1.98	0.85	-2.33 *	-1.52	0.84	-1.81
Title*Teacher Gender γ_{070}	-8.09	2.14	-3.79 ***	-7.89	2.05	-3.85 ***	-9.11	2.13	-4.28 ***
Random Effects	Estimate	Standard Errors	Z-Value	Estimate	Standard Errors	Z-Value	Estimate	Standard Errors	Z-Value
Within-teacher, σ^2	50.72	1.78	28.52 ***	50.34	1.56	32.17 ***	50.19	1.47	34.25 ***
Between-teacher, τ_{π}	3.53	2.59	1.36	7.66	1.94	3.94 ***	3.73	2.04	1.83 *
Between-school, τ_{β}	4.30	3.46	1.25	0.00			7.24	3.30	2.19 *

Note. $(*p<0.05,** p<0.01,***p<0.001)$.

Model 4a, controlling for content knowledge assessment, led to non-significant between-school and between-teacher variability in the level of Algebra I score. As a result, discussion will revolve around this model in interpreting the interaction terms. Similar to model 3a, there were significant correlations between Algebra I achievement and the following student level variables: grade level, days absent, and designations of race as Hispanic and Black. There were no significant predictors at the school level. On average, integrated mathematics teachers had student results 4.45 points higher than subject-specific teachers on the Algebra I EOC exam. The only other significant level-2 predictor was content knowledge, such that each one point increase in teacher content knowledge resulted in a 0.05 point increase in student Algebra I achievement. Both the interaction terms, curriculum type X Black and curriculum type X teacher gender, were significant. These interactions will be presented below. Model 4a accounted for 59.22% of between-school, 77.05% of between-teacher, and 11.44% of within-teacher variance in Algebra I achievement.

The relationship between Algebra I score and the designation of Black significantly depended on curriculum type (Figure **11**). White students in integrated mathematics tended to score 2.33 points higher than their subject-specific peers on the Algebra I EOC ($p<0.0001$). Similarly, Black integrated mathematics students outperformed their subject-specific peers by 1.38 points on the Algebra I EOC ($p=0.014$). The negative slope of the integrated mathematics line suggests that on average, White students outperform Black students by 2.19 points ($p<0.0001$). The negative slope of the subject-specific line suggests that on average, White students outperform Black students by 1.24 points ($p<0.0001$). The significant slopes

showed that White students outperform Black students regardless of curricular pathway, however there is an even greater advantage to White and Black students using *Core-Plus* over subject-specific curricula.

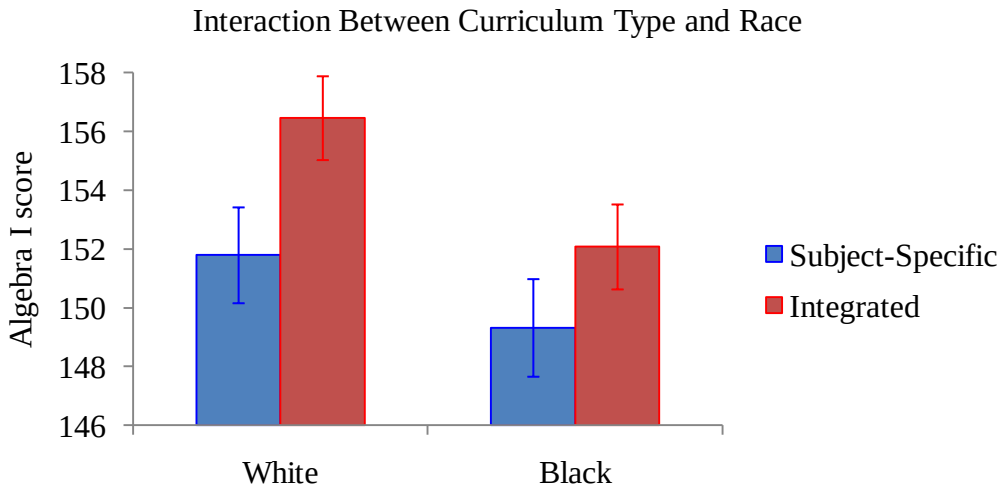


Figure 11. Interaction Between Curriculum Type and Race.

There was a significant interaction between curriculum type and teacher gender on Algebra I achievement (Figure **12**). Students of female integrated mathematics teachers scored 2.11 points higher results on the Algebra I EOC than students the male integrated mathematics teachers ($p<0.0001$)[15]. There were no significant differences in the outcome measure based on teacher gender for subject-specific teachers. Students of integrated mathematics female teachers had results on the Algebra I EOC 3.41 points higher than subject-specific female teachers ($p<0.001$). There were non-significant differences between

[15] Note, there was a relatively small number of male teachers $\left(n_{IM}=5, n_{S-S}=15\right)$ compared to female teachers $\left(n_{IM}=10, n_{S-S}=35\right)$.

curriculum type for male teachers on the Algebra I EOC. These findings suggest that students of female teachers using integrated mathematics had stronger performance results on the Algebra I EOC than students of male teachers or female teachers using subject-specific mathematics. To explore this result further, a test was conducted to determine if there was a significant interaction between teacher gender and student gender, given the high percentage of female students in integrated mathematics. This interaction was non-significant suggesting the effect of teacher gender occurs regardless of the gender of the student. In other words, it is not the case female integrated mathematics students have higher student achievement as a result of the gender of their students.

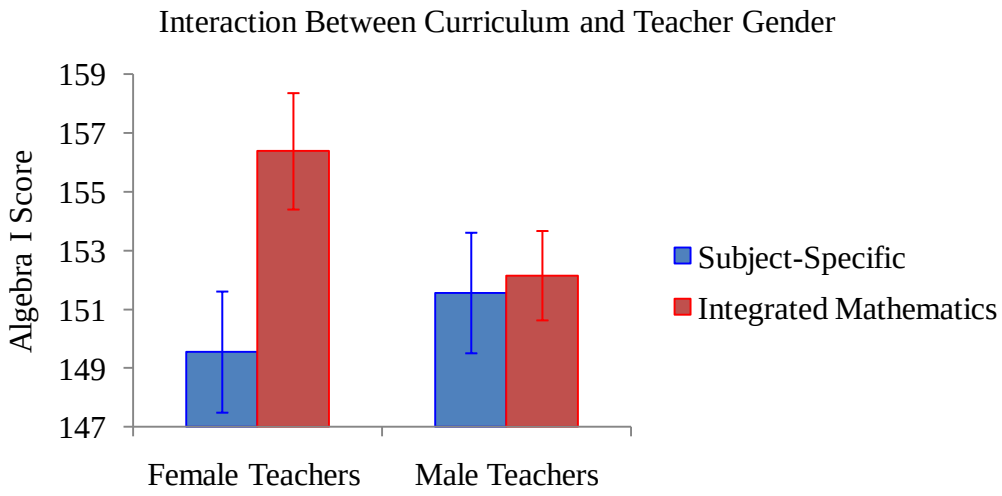


Figure 12. Cross-level Interaction Between Curriculum Type and Teacher Gender.

Incorporating teacher-level predictors with the workshop sample increased the variance that was explained at level-2 over the first analysis of Algebra I models with the entire sample. Student Algebra I scores for workshop teachers favored integrated mathematics teachers by almost four points. When holding teacher level predictors constant,

most notably professional development, the effect for curriculum increased to over 6 points in favor of integrated mathematics.

The impact of teacher characteristics became more apparent at this stage of the analysis. Teacher content knowledge had a significant impact on Algebra I achievement and teacher gender interacted with curriculum type in a unique manner. Female integrated mathematics teachers had higher student performance than any other, female subject-specific and all male teachers. Subject-specific teachers had consistent results across male and female teachers.

Modeling Student Algebra II Achievement: Workshop Sample

This section models student Algebra II EOC achievement for the teachers that participated in one or more phases of the NCIM project. Findings will be presented in a similar fashion to the previous section.

Background and Demographics

The subset of Algebra II EOC data for workshop participants was provided for 35 schools. Five schools offered integrated mathematics and 32 offered subject-specific mathematics (two offered both). The sample was comprised of 290 students enrolled in integrated mathematics Course 3 and 1,689 in Algebra II (subject-specific). Schools offering integrated mathematics had a higher %FRL, lower average daily membership, and were more frequently classified as Title One than the schools with Algebra II students (Table **27**). All of the schools offering Course 3 were involved in the NCNSP.

Table 27. School Demographics for Workshop Teachers on Algebra II EOC

	Average %FRL	Average ADM	Percent Title One	Percent NCNSP
Integrated Mathematics	72.85	275.25	60.00	100.00
Subject-Specific	49.81	605.93	3.13	56.25
Total	52.52	567.03	10.81	62.16

There were eight integrated mathematics and 36 subject-specific mathematics teachers included in this sample (one teacher taught both curricula). Subject-specific teachers scored about 16 points higher than integrated mathematics teachers on the content knowledge assessment (Table **28**). Integrated mathematics teachers were exposed to about 54 additional hours of NCIM professional development than subject-specific teachers, while subject-specific teachers tended to have 6.3 more years of teaching experience. All of the teachers in this sample held a secondary mathematics teaching licensure.

Table 28. Teacher Demographics for Workshop Teachers on Algebra I EOC

	Average CKA Score	Average Professional Development Hours	Average Number Years Taught
Integrated Mathematics	61.88	88.42	9.25
Subject-specific	77.96	34.85	15.54

There was a higher percentage of subject-specific students taking the Algebra II EOC as high school freshman, which suggests students on this pathway were tracked in middle school (Table **29**). There were similar percentages of males and females for each pathway, with a higher overall percentage of female students. The two pathways had a similar percentage of each ethnicity, except for White and Black percentages. The majority of students in subject-specific mathematics were White, while the majority for integrated

mathematics was Black. Students in Course 3 averaged 6.87 days absent, while Algebra II students averaged 7.20 days.

Table 29. Percentages of Students Within Each Category for Workshop, Algebra II Sample

			Grade		**Gender**	
	9th	10th	11th	12th	Male	Female
Integrated Math	4.14	48.97	44.83	2.07	42.41	57.59
Subject-Specific	16.38	44.44	34.98	4.20	43.14	56.86
Total	14.52	45.13	36.48	3.88	43.03	56.97
			Race			
	AmIn	Asian	Hispanic	Black	White	Multi
Integrated Math	0.00	1.03	7.24	61.38	26.90	3.45
Subject-Specific	3.03	1.92	5.25	29.36	58.10	2.35
Total	2.57	1.78	5.56	34.22	53.35	2.52

Means and Passing Rates

Integrated and subject-specific mathematics students who took the Algebra II EOC had remarkably similar prior achievement scores on the Algebra I EOC, 155.35 and 155.59 respectively. Subject-specific students had a 2.34-point higher mean achievement on the Algebra II exam than integrated mathematics students. Further, subject-specific students had a higher mean achievement than integrated mathematics students for all subgroups except for Black students and a higher passing rate for all subgroups except for Black and Asian students (All Asian students in both pathways passed the exam) (Table 30).

Table 30. Algebra II EOC Means and Passing Rates for Students Within Each Subgroup

		Grand Mean	Grade Level					Gender	
			8th	9th	10th	11th	12th	Male	Female
IM	Sample Size (n)	290	-	12	142	130	6	123	167
	Mean	148.83	-	152.67	150.64	146.94	139.33	148.77	148.87
	Standard Dev.	7.12	-	4.58	6.91	6.77	5.43	7.04	7.20
	Passing Rate	58.28%	-	83.33%	69.72%	45.38%	16.67%	56.91%	59.28%
SS	Sample Size (n)	1612	-	265	719	560	68	695	917
	Mean	151.17	-	154.74	152.59	148.25	146.31	150.71	151.52
	Standard Dev.	8.44	-	7.91	8.58	7.45	6.71	8.29	8.54
	Passing Rate	67.80%	-	83.40%	73.85%	55.89%	41.18%	65.76%	69.36%
Entire Sample	Mean	150.81		154.65	152.27	148.00	145.74	150.42	151.11
	Standard Dev.	8.29	-	7.80	8.35	7.34	6.86	8.14	8.40

		Race						Attendance	
		AmIn	Asian	Hispanic	Black	White	Multi	Absent ≤ 20 Days	Absent > 20 Days
IM	Sample Size (n)	-	3	21	178	78	10	280	10
	Mean	-	153.33	147.48	148.60	149.88	146.30	149.05	142.60
	Standard Dev.	-	3.79	7.90	6.85	7.27	9.07	7.03	7.37
	Passing Rate	-	100.00%	47.62%	57.87%	62.82%	40.00%	59.64%	20.00%
SS	Sample Size (n)	49	31	85	474	935	38	1521	91
	Mean	150.80	157.87	151.13	147.23	152.92	152.34	151.59	144.19
	Standard Dev.	8.77	5.21	7.92	8.03	8.07	7.11	8.23	8.74
	Passing Rate	67.35%	100.00%	70.59%	50.84%	74.65%	78.95%	70.09%	29.67%
Entire Sample	Mean	150.80	157.47	150.41	147.60	152.69	151.08	151.20	144.03
	Standard Dev.	8.77	5.22	8.01	7.74	8.05	7.85	8.11	8.60

Similarities and Differences Between Two Samples: Larger and Workshop

The workshop sample of Algebra II students was a subset of the larger sample presented earlier. The larger sample of students had higher mean achievement on the Algebra II exam ($\mu_{IM} = 158.02, \mu_{S-S} = 156.07$) than the workshop sample of students ($\mu_{IM} = 155.35, \mu_{S-S} = 155.59$). Every subgroup in the larger sample had a higher mean Algebra II score for both pathways, integrated mathematics and subject-specific, except for American Indian students (only represented in the subject-specific sample) and students in the 12th grade in subject-specific mathematics (a 0.05-point difference) (Table **17**, Table 30). The workshop Algebra II sample had a much higher percentage of Black integrated mathematics students (over 30% more) and a much lower percentage of White integrated mathematics students (31.24% less). The percentages of subject-specific students within each race were similar between the two samples (29.36% and 28.04% Black and 58.10% and 58.22% White for the workshop and larger samples respectively). The %FRL for schools using integrated mathematics in the workshop sample was over 10 percentage points higher than larger sample, while the workshop sample average %FRL for subject-specific was slightly lower (0.58 percentage points less) for workshop sample (Table **15**, Table **27**). There was a higher ADM for larger sample (653.28 vs. 567.03 for workshop sample) and a much higher percentage of Title One schools in the workshop sample (over 3 percent higher, integrated mathematics was 18 percent higher in workshop sample), regardless of curricular pathway. While the mean prior achievement and Algebra II achievement was higher for the larger sample (prior achievement $\mu_{IM} = 158.02, \mu_{S-S} = 156.07$ and Algebra II achievement

$\mu = 152.54$) than the workshop sample (prior achievement $\mu_{IM} = 155.35, \mu_{S-S} = 155.59$ and Algebra II achievement $\mu = 150.81$), the workshop sample included a higher percentage of students enrolled in integrated mathematics from typically underrepresented groups (students identified as Black and higher %FRL schools more likely to be classified as Title One). Further, considering teachers' experience with the NCIM professional development, the workshop offered Course 3 training to teachers for the first time during the summer of 2009. As a result, most of the integrated mathematics teachers in the workshop sample were teaching Course 3 for the first time during the 2009-2010 school year, the year corresponding to the sample of Algebra II exam scores. Collectively, the subset of Algebra II scores for the workshop sample included integrated mathematics students from underserved populations and teachers with limited experience teaching Course 3.

Fully Unconditional Model

The fully unconditional model indicated that 22.87% of the variance in student test score was between schools ($\tau_\beta = 16.39, Z = 2.71, p = 0.0033$), 9.79% was between-teachers ($\tau_\pi = 7.02, Z = 2.12, p = 0.017$), and 67.34% was within-teachers ($\sigma^2 = 48.27, Z = 30.47, p < 0.0001$). Therefore the fully unconditional model signified there was sufficient variability in each level to warrant further analysis. The grand mean for the sample was 150.40 ($t = 181.31, p < 0.0001$).

Individual Teacher Level Predictors

The next models were used to determine, while controlling for curriculum type and student and school characteristics, if there were teacher level differences in student

achievement associated with individual teacher level predictors. In six separate analyses, each of the teacher level variables was tested to determine its correlation with Algebra II achievement (Table **31**).

Table 31. Correlations Among Individual Level-2 Predictors and Algebra II Score, When Controlling for Curriculum Type

Level-2 Predictor	Coefficients	Standard Error	t-Value
Curriculum Type	-2.22	1.94	-1.14
Content Knowledge	0.066	0.028	2.34 *
Years Taught	0.033	0.047	0.69
Teacher Gender	-2.89	0.99	-2.93 **
College Major	-0.67	1.31	-0.52
NCIM Professional Development	-0.0033	0.019	-0.18

Note. ($p < 0.05$*, $p < 0.001$**, $p < 0.001$***).

Without including any other teacher level predictors, curriculum type not was significantly related to Algebra II EOC score. While controlling for curriculum type, teacher level differences in Algebra II EOC score were also associated with teacher content knowledge assessment and gender. The other three teacher-level predictors had non-significant correlations with Algebra II achievement. The magnitude of the curriculum effect remained non-significant when including the individual teacher level predictors. Thus, even when controlling for differences in teachers gender, content knowledge, college major, years taught, or professional development exposure, there were still no differences in Algebra II achievement based on curriculum type.

Combination of Teacher Level Predictors

Accounting for student, teacher, and school characteristics, the next two models were run to determine if there was evidence of differential impact on student Algebra II

performance between the two curricular types. Controlling for all six teacher level predictors continued to have no significant effect on the relationship between curriculum type and the outcome measure. However, the magnitude of the effect of curriculum transitioned from being negative in the previous analyses of Algebra II score, to being positive in favor of integrated mathematics, though still non-significant.

Algebra II achievement was modeled in Model 2 using all six teacher level predictors in addition to the student and school level predictors that were previously used (Table **32**). Student's prior achievement was significantly related to current achievement, such that for each one-point increase in Algebra I score there was a 0.50 point increase in Algebra II score. Student grade level, days absent, student gender, and designation as Black were all significant and negatively correlated to Algebra II EOC score. Teacher content knowledge was significant and positive and professional development hours were significant and negatively related to Algebra II achievement. There were no significant school level effects. This model accounted for 38.85% of the between-school, 100% of the between-teacher, and 69.46% of the within-teacher variance in Algebra II score.

Table 32. HLM 2 and 3 for Algebra II Workshop Sample

Fixed Effects	Model 2			Model 3		
	Coefficients	Std. Error	t-Value	Coefficients	Std. Error	t-Value
Algebra II Score, β_{0jk}						
Intercept γ_{000}	150.88	0.72	209.07 ***	151.06	0.74	205.21 ***
Curriculum Type γ_{010}	1.26	1.80	0.70	2.01	1.83	1.10
Algebra1score γ_{100}	0.50	0.02	23.11 ***	0.50	0.021	23.07 ***
Grade Level γ_{200}	-1.35	0.30	-4.48 ***	-1.41	0.29	-4.83 ***
Days Absent γ_{300}	-0.14	0.021	-6.3 ***	-0.14	0.02	-6.62 ***
Gender γ_{400}	-1.01	0.32	-3.12 **	-0.98	0.31	-3.12 **
American Indian γ_{500}	-1.55	1.29	-1.2	-1.44	1.25	-1.15
Asian γ_{600}	1.85	1.24	1.49	1.68	1.20	1.40
Hispanic γ_{700}	0.80	0.71	1.13	0.82	0.69	1.19
Black γ_{800}	-1.05	0.46	-2.3 *	-0.92	0.44	-2.08 *
Multi-racial γ_{900}	-1.19	1.05	-1.13	-1.14	1.02	-1.11
Teacher Gender γ_{020}	-3.13	1.77	-1.77	-3.13	1.78	-1.76
CKA γ_{030}	0.11	0.027	4.02 ***	0.060	0.043	1.37
PD Hours γ_{040}	-0.04	0.01	-2.7 **	-0.035	0.014	-2.56 **
College Major γ_{050}	-1.66	1.48	-1.11	-1.68	1.46	-1.15
Years Taught γ_{060}	0.044	0.065	0.68	0.048	0.068	0.71
%FRL γ_{001}	0.0024	0.044	0.05	-0.033	0.046	-0.71
ADM γ_{002}	0.0014	0.0016	0.88	0.0018	0.0016	1.12
Title One γ_{003}	2.28	2.72	0.84	2.64	2.73	0.97

Table **32**. Continued

Curriculum*Alg1score γ_{110}				0.042	0.053	0.81
Curriculum*CKA γ_{070}				0.085	0.058	1.48
CKA*Algebra1score γ_{130}				0.0078	0.0011	7.28 ***
Curriculum*CKA*Alg1score γ_{170}				-0.013	0.0028	-4.72 ***
Random Effects	Estimate	Standard Errors	Z-Value	Estimate	Standard Errors	Z-Value
Within-teacher, σ^2	28.49	1.20	23.75 ***	26.75	1.13	23.73 ***
Between-teacher, τ_{π}	0			0		
Between-school, τ_{β}	10.02	4.01	2.50 **	10.07	3.99	2.52 **

Note. $(p < 0.05^{*}, p < 0.001^{**}, p < 0.001^{***})$.

Interaction of Curriculum Type When Including Teacher Level Predictors

The third model considered interactions between curriculum type and other predictors on student Algebra II achievement. There were no significant level-2 interactions with curriculum type, and unlike the Algebra I analysis there were no significant interactions between curriculum type and race. Due to the emergence of a significant negative effect of NCIM professional development hours when including multiple teacher-level predictors, interactions were tested between NCIM hours and all other teacher-level predictors to try to determine for what teacher group this negative result was most prevalent. There were no significant interactions between NCIM professional development hours and any other teacher-level predictor. However, Model 3 showed a significant three-way cross-level interaction between prior achievement, teacher content knowledge, curriculum type, and Algebra II achievement (Table **32**). The student level fixed effects remained almost identical to Model 2, as did the fixed effect for professional development. While content knowledge was no longer significant, there was a significant content knowledge X Algebra I score interaction and the significant three-way cross-level interaction, curriculum X content knowledge X Algebra I score.

To interpret the three-way interaction, two additional models were conducted, one for low prior achievement (centered Algebra I scores less than or equal to 0) and one for high prior achievement (centered Algebra I scores greater than 0) (Figure **13**). For students with low Algebra I scores, teachers of integrated mathematics with high content knowledge had Algebra II EOC results that were 0.99 points higher than teachers with low content

knowledge, but this result was only moderately significant (p=0.10). For students with high prior achievement, teachers of integrated mathematics with high content knowledge had Algebra II EOC results that were 2.54 points higher than teachers with low content knowledge (p=0.0047). There were no significant differences in the slopes of the lines for subject-specific teachers, nor between any of the contrasts between curriculum types. These results suggest that teachers of Course 3 with higher content knowledge have more successful student results on the Algebra II exam than teachers with lower content knowledge. As students progress to Course 3 the textbook content gets more difficult, and these results showed that teacher knowledge was important to increased student achievement.

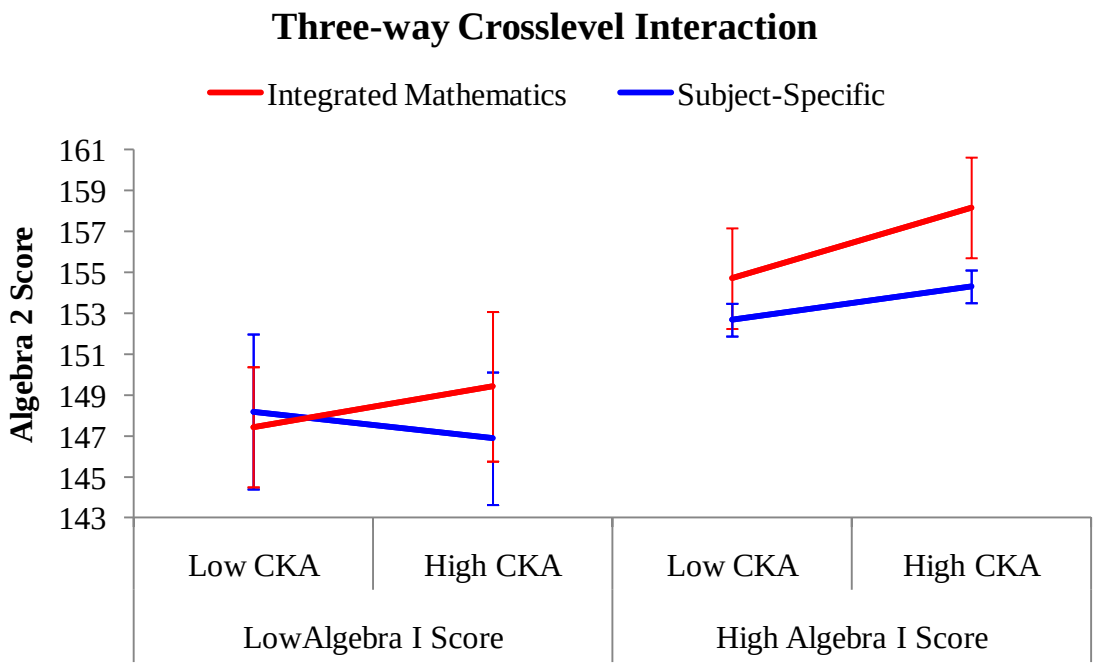


Figure 13. Three-way Crosslevel Interaction Between Curriculum, Prior Achievement, and Teacher Content Knowledge.

Similar to the previous analysis of Algebra I achievement for the workshop sample, modeling Algebra II achievement with teacher level predictors was influenced by teacher content knowledge, gender, and professional development. There were no significant differences between curricular pathways and the Algebra II exam scores, although the effect transitioned from negative to positive, in favor of integrated mathematics, as teacher variables are included in the models.

Though the effect size was small, there was a negative relationship between NCIM professional development hours and Algebra II score when holding multiple teacher level variables constant. This significance was prevalent when only including curriculum type and professional development at level-2. This suggests that as differences among teachers was held constant, the effect of professional development became more prevalent.

Teacher content knowledge was positive and significantly correlated with Algebra II achievement throughout the analysis of Algebra II scores. It was evident from the interaction between prior achievement, curriculum type, and Algebra II achievement that integrated mathematics teachers with high content knowledge had students that performed better than integrated teachers with low content knowledge. Interestingly, there were no significant differences in student achievement for subject-specific teachers based on their content knowledge, suggesting that teacher content knowledge had a smaller an impact on student achievement in subject-specific classrooms. Integrated mathematics teachers of students with high prior achievement performed better than subject-specific students regardless of teacher content knowledge.

Results From Assessing the Curricular Alignment and Outcome Measures

Recall an analysis was conducted to determine how much of the content on the Algebra EOCs were included in the *Core-Plus* textbooks (both first and second editions). This section is devoted to presenting the results for the percentage of Algebra I and II exam questions that were included in the *Core-Plus* textbook content.

First Edition Algebra I

The percent agreement between the two coders was 84.37%, with a Cohen's κ of 0.35. The fair Cohen's κ resulted in the researcher conducting an external analysis to settle disputes between the first two coders. Despite the initial coders both believing that content on the EOC was included in the text 85.94% of the time, they only agreed on the exact question in 50 out of 64 cases, or 78.13%. Incorporating the third coder's results increased Cohen's κ to 0.78. The third coder analyzed the items the initial coders disagreed on and determined there was a 90.63% alignment between the exam and the textbook. Therefore, the *Core-Plus* textbook is aligned with the released EOC between 85.94 and 90.63%.

Second Edition Algebra I

The percent agreement between the two coders was 98.43% and the Cohen's κ for the two coders of the second edition text and the Algebra I EOC was 0.66. One coder believed that 98.44% of the EOC topics were included in the Course 1 and 2 textbooks, while the other coder believed there was a 96.88% alignment. In a question-by-question analysis, these coders only disagreed that content was taught for one question, leading to a 96.88% agreement for content being taught. The released Algebra I EOC is aligned to Courses 1 and

2 materials between 96.88 and 98.44% of the time. Therefore, it is safe to assume a high degree of alignment between the Algebra I EOC and the first two courses in the *Core-Plus* series.

First Edition Algebra II

The percent agreement between the two coders of the first edition textbook alignment to the released Algebra II EOC was 80.93%, with a Cohen's κ of 0.562. Coders agreed content was in the textbook for 28 of the 51 released EOC questions (54.90%), agreed that content was not in the textbook on 13 of the 51 questions (25.49%), and disagreed on the other 10 questions (19.61%). One coder believed there was a 64.71% alignment between the text and the exam, while the other coder rated the alignment as 60.78%. This was considered a low alignment.

Second Edition Algebra II

Coders for the second edition textbook on the Algebra II EOC had a 82.34% agreement and a Cohen's κ of 0.61. They agreed content on the EOC was included in the text in 30 out of 51 questions (58.82%), though one review felt nine additional questions were also included in the text for a 76.47% alignment. Due to the high level of agreement, it is concluded the released Algebra II EOC is aligned to the second edition between 58.8 and 76.5%. Again, there was a low alignment between the Algebra II EOC and the *Core-Plus* textbook series.

Discussion and Conclusions From Chapter

A discussion of results surrounding the two parts of the first research question will be provided below. Then, the chapter will conclude with overall conclusions from both analyses and set the stage for the focus study in the following chapter.

Research Question 1a

Impact on student performance between teachers using different curricula across N.C.

Students exposed to integrated mathematics had higher Algebra I achievement than students on the subject-specific pathway, however there were no differences between the two curricular pathways on the Algebra II EOC. Recall, the most prominent integrated mathematics textbook used in the state, *Core-Plus*, was highly aligned to the Algebra I EOC, but not well aligned to the Algebra II EOC. It is encouraging that students of teachers using integrated mathematics materials performed no different on the Algebra II EOC than teachers who use subject-specific text, given this lack of alignment.

There were big effects of students' grade level on the two outcome measures, such that students at lower grade levels perform higher on EOC mathematics assessments than students at higher grade levels. It is no surprise that more advanced students progress through course sequences at a faster rate, resulting in them taking state assessments at lower grade levels. However, the disparities between scores for students at low and high grade levels is a disheartening trend for the achievement gap. Teachers of integrated mathematics had consistent results across grade levels for both outcome measures, while there was a drop in subject-specific teacher scores for students at high grade levels.

School %FRL was negatively correlated to student achievement on both mathematics EOC assessments, suggesting the importance of including such a measure into models for student achievement. On the Algebra I EOC, teachers of integrated mathematics had consistent results for both low and high %FRL schools, while subject-specific teachers had a drop in student achievement at high %FRL schools. One very encouraging result was that when controlling for student and school characteristics Title One schools scored significantly higher on the Algebra II exam than non-Title One schools. Due to the low number of Title One schools in the sample using the subject-specific pathway, and given the %FRL results, this speaks highly of the use of integrated mathematics in typically low performing schools with high a %FRL population.

It was possible to include student prior achievement in models of Algebra II achievement, which greatly increased the amount of explained student-level variance. This is an important predictor to consider collecting when conducting curricular evaluations. Overall, students in this sample were not disadvantaged from being exposed to integrated mathematics. They performed as well, if not higher, than students on the subject-specific pathway on NC standardized mathematics assessments.

Research Question 1b

Impact on student performance between workshop teachers using different curricula

Findings from research question 1b supported results on the effects of curriculum from the much larger data set of question 1a, and allowed for integration of teacher level predictors. Incorporating these into models of student achievement increased the amount of

school and teacher level variance in the outcomes measures. This was important for determining what teacher characteristics have the largest effects on student achievement. Similar to the previous analysis, *Core-Plus* teachers had higher scores on the Algebra I EOC, while there was no difference between curricular pathways on the Algebra II EOC.

Teacher content knowledge was an important predictor of student achievement for both outcome measures, both individually and when controlling for other teacher level variables. In models of student achievement content knowledge was positively correlated to student achievement, with effect sizes ranging from 0.05 to 0.11. There were no differences in teacher content knowledge among teachers in the Algebra I sample, but subject-specific Algebra II teachers scored 16 points higher than *Core-Plus* Course 3 teachers on the teacher content knowledge assessment. The impact of teacher content knowledge was most prevalent when modeling Algebra II achievement incorporating interaction terms. Regardless of student prior achievement, *Core-Plus* teachers with high content knowledge had higher student Algebra II achievement than *Core-Plus* teachers with low content knowledge. There was no difference in student Algebra II achievement based on subject-specific teachers' content knowledge. These results suggest the importance of having Course 3 teachers with a strong mathematical background.

There were also unique differences in student achievement based on teacher gender, though not consistent among the more intricate models. Female teachers of *Core-Plus* had higher student Algebra I achievement than all subject-specific teachers and male *Core-Plus* teachers. Similar to the content knowledge predictor, there were no significant differences on

student Algebra I achievement based on teacher gender for subject-specific teachers. Further, in a model of student Algebra II achievement when only including curriculum type and teacher gender at level-two, students of female teachers scored almost three points higher than students of male teachers. The effect of teacher gender was not prevalent in models of Algebra II achievement that incorporated multiple teacher level predictors. These results point to the impact teacher gender on student achievement, most notably for *Core-Plus* teachers.

For both the Algebra I and II workshop samples, there was a significant negative correlation between student achievement and the number of hours a teacher spent in the NCIM professional development. NCIM professional development hours was only significant in the Algebra I sample when including interaction terms and in the Algebra II sample when including multiple teacher-level predictors or interaction terms. The effect sizes ranged between -0.031 and -0.04. These findings suggest that when holding teachers' background constant, using numerous teacher-level traits, students of teachers exposed to increased amounts of NCIM professional development had lower performance on the Algebra I and II EOCs. Recall, teachers with the most NCIM professional development experience came from the highest need schools. Chapter 5 will provide an in-depth of analysis of groups of teachers involved in the NCIM professional development project, including teachers that received instructional coaches throughout the year (teachers with the highest amount of professional development hours in high poverty settings), to understand this result more fully.

Overall Conclusions

As demonstrated across multiple forms of analyses, students of teachers of integrated mathematics had higher Algebra I achievement than subject-specific teachers, with or without including teacher-level predictors. There was a positive effect on the use of the integrated materials on student Algebra I EOC, based on the use of an outcome measure that was previously found to be well-aligned with the *Core-Plus* curriculum. The strong positive effect of using integrated materials was prevalent in all models that were analyzed, regardless of what predictors were added or removed.

There were no significant differences between teachers of subject-specific and integrated curricula on the Algebra II exam, even after controlling for different teacher characteristics. Given the fact that the released Algebra II EOC was not strongly aligned to the *Core-Plus* integrated materials, it was encouraging that integrated mathematics teachers experienced the same level of success as subject-specific teachers.

Across each of the prior models there were significant negative effects on student achievement for grade level, days absent, designation as Black, and %FRL. For each one point increase in student grade level, there was at least a 2.86 point decrease in Algebra I score and at least a 1.35 point decrease in Algebra II score. Students taking these exams in ninth grade were more advanced and tended to have higher EOC achievement. However, students enrolled in integrated mathematics had more consistent results across the grade levels than subject-specific students (Figure **8**, Figure **10**). Black students performed significantly lower than Whites on both outcome measures for both samples; up to 3.61

points lower on the Algebra I EOC and up to 1.50 points lower on the Algebra II EOC. Black students enrolled in integrated mathematics had higher mean and passing rates than their subject-specific peers on both outcome measures, in both samples. Further, when controlling for student-, teacher-, and school-level variables, curriculum type significantly interacted with the classification as Black on the Algebra II exam. This suggests that while Black students were outperformed by White students, that there was a comparative advantage to Black students using integrated mathematics over subject-specific. The Algebra I models had additional negative results for several races: American Indian, Hispanic, and Multi-racial. However, the strong negative effect on the Algebra I exam for non-Asian ethnicities dissipated on the Algebra II exam when adding prior achievement as a predictor, suggesting prior achievement measures may suppress effects of race. Closer analysis of this result showed that the only significant differences between race and Algebra II achievement were between White and Black students based on their prior achievement. White students with low prior achievement had significantly higher Algebra II scores than Black students when using subject-specific mathematics, whereas White students with high prior achievement had significantly higher Algebra II scores than Black students when using integrated mathematics.

Finally, when modeling Algebra II achievement there were significant positive results in all models for prior achievement. On average, for each one-point increase in student prior achievement, there was around a 0.50-point increase in Algebra II exam score. Further,

including a measure for prior achievement helped explain more of the student-level variance than models without prior achievement measures.

Overall results from these four analyses imply that students at lower grade levels with high attendance in low %FRL schools performed better on state assessments. These data did not detail how teachers implemented the curricula and there are unresolved questions about the influence the NCIM professional development had on student achievement. The next chapter restricts the sample under investigation even further to document teachers' implementation of *Core-Plus* and how their implementation and experiences in the NCIM professional development affect student achievement.

CHAPTER FIVE: RESULTS QUESTION TWO

This chapter details the influence of the NCIM professional development project on teacher's implementation of the *Core-Plus* mathematics curricular materials and factors that are associated with student learning. Quantitative results will be presented on teachers' implementation of the *Core-Plus* materials and the effect these materials have on student mathematics performance. Qualitative results will be presented to describe how the variance in student outcomes might be explained by teacher implementation of the materials. Following a brief description of the participants' background and beliefs, the results presented are organized into four sections. The first section examines teachers' implementation of *Core-Plus* materials based on varying levels of participation in the NCIM project, as quantified by the Table of Contents Log (TOC). Second, the implementation indices and belief factors will be included into HLM models of student Algebra achievement. Third, teacher interviews, classroom observations, and instructional coach reports will be triangulated to report factors that affect teachers' implementation of *Core-Plus.* The final section will link the first three sections to discuss how teachers' implementation of *Core-Plus* explains variance in student achievement among teachers with varying levels of NCIM professional development experience. The chapter will conclude with a summary of the results surrounding the second research question regarding evidence of significant differences on student achievement among teachers who participated in different components of the NCIM project.

Background and Beliefs of Participants in Each Group

Recall, the groups of teachers used in the focus group analysis were A (participants in the NCIM project), F (instructional coach only), B (participants in the workshop only), and D (not involved in any aspect of NCIM project). Initially, teachers in Group A were classified as NCIM teachers because they received follow-up instructional coach support throughout the school year following the summer workshops and worked at schools affiliated with the NCIM project. However, after calculating and analyzing the Group A teachers' implementation indices, two teachers with outlier indices were identified as profiling quite differently from the rest. There were two teachers in Group A with lower indices than the rest of the group, and upon further analysis it was discovered they had not attended a summer workshop prior to the 2009-2010 school year. So while they had received an instructional coach throughout the school year, it was decided they could not be classified with the rest of the Group A teachers due to lack of workshop attendance. Hence, a new group, F, was formed with these two teachers.

Teachers from ten different schools were represented in the focus group sample. Teachers receiving instructional coaches (Groups A and F) taught at schools with a much smaller student population $(\mu_A = 275.25, \mu_F = 152.00, \mu_B = 781.40, \mu_D = 726.00)$ and with a much higher %FRL $(\mu_A = 72.85, \mu_F = 96.00, \mu_B = 39.03, \mu_D = 42.29)$ than the other two schools (Table 33). Further, three out of the four Group A schools were classified as Title One, while none of the Group B or D schools had Title One status.

Table 33. School Characteristics Based on Groups

	Average ADM	Average %FRL	Percent Title One
Group A	275.25	72.85	75.00
Group F	152.00	96.00	100.00
Group B	781.40	39.03	0.00
Group D	726.00	42.29	0.00

There were some distinguishable differences among the participant groups involved in the focus group analysis (Table **34**). The average content knowledge assessment (CKA) score of teachers varied among each group, with a range of 37 points. All of the teachers in Group D either majored in mathematics or mathematics education in college, while neither of the Group F teachers had a mathematics background. Teachers in Group D taught *Core-Plus* for an average of 7.17 years, whereas Groups F and B had just begun teaching *Core-Plus.* However, on average, teachers in Groups F and B had more overall teaching experience, and Group A had the least.

Table 34. Teacher Background Demographics Among Groups

	Average NCIM PD Hours	Average CKA Score	Average Yrs. Taught *Core-Plus*	Average Years Taught	Percent Non-Math Majors	Percent Male
Group A (n=7)	111.98	69.71	3.71	5.57	28.57	14.29
Group F (n=2)	23.93	48.00	1.00	12.00	100.00	0.00
Group B (n=6)	55.50	77.80	1.33	10.33	33.33	50.00
Group D (n=6)	0.00	85.00	7.17	9.67	0.00	0.00

Recall a teacher beliefs survey was given to all workshop and focus group teachers and Principal Components Analysis was used to reduce the beliefs into five factors (Table 8). The standardized average for each belief factor varied among groups (Table **35**). Standardized averages closest to positive one suggest teachers feel strongly in favor of that belief.

Table 35. Average Teacher Beliefs by Belief Factors Among Groups

	Belief 1 Reform	Belief 2 Diversity	Belief 3 Collaborative	Belief 4 Explanations	Belief 5 Traditional
Group A	0.761	0.751	-0.364	0.497	-0.066
Group F	-0.705	0.640	-0.940	0.190	-0.165
Group B	0.523	0.350	-0.192	0.128	-0.015
Group D	0.682	0.950	0.335	-0.113	-0.662

Group A had the highest average for their beliefs about reform practices, 0.761, suggesting that Group A teachers tended to agree with statements about reform-oriented instructional practices more often than the other groups. The high negative value for Group F teachers in beliefs about reform practices, -0.705, provides empirical evidence supporting the claim that these teachers do not believe in reform practices to the same degree as the other groups. Inversely, Group F had the highest values for beliefs about traditional practices, -0.165, whereas Group D had the lowest values for traditional practices, -0.662. Group D was the only group with a positive value for beliefs about collaborative learning, 0.335, though Group A (-0.364) and B's (-0.192) values were not as strongly opposed to collaborative learning as Group F (-0.940). Group D's negative value for belief about explanations, -0.113, suggests that they tended to believe the student should explain their mathematical ideas,

whereas positive values for this belief suggests that teachers tended to agree with statements pointed towards teachers explaining the content to the students.

Teacher Implementation Indices Disaggregated by NCIM Participation

The analysis of Table of Content Logs (TOC-logs) was based on the 41 logs that were completed by teachers in Groups A, F, B, and D from the 2009-2010 school year. As presented in Chapter 3, the TOC-logs were analyzed by computing implementation indices for Opportunity to Learn (OTL), Extent of Textbook Implementation (ETI), and Textbook Content Taught (TCT) (Table 36). Recall, the OTL index measured the percentage of textbook content that was taught, either solely from the textbook or through supplemental materials. The ETI index weighted the options in the TOC-logs to measure the degree to which the textbook was used directly to teach the content. Finally, the TCT used the same weighted measure as the ETI but only included content that was taught through any means. This provided a measure of how the textbook was used to teach content in the textbook and ignored the topics students were not taught.

Table 36. Mean (and Standard Deviations) of the Implementation Indices

	OTL	ETI	TCT
Group A	56.19 (14.31)	51.21 (15.13)	90.65 (8.32)
Group F	33.80 (6.30)	26.18 (6.54)	77.98 (17.14)
Group B	58.79 (9.90)	57.69 (8.44)	98.48 (3.83)
Group D	46.67 (9.74)	33.85 (14.90)	70.58 (17.99)
Entire Sample	52.30 (13.93)	46.51 (16.73)	87.29 (14.94)

Opportunity to Learn

The mean OTL index across all *Core-Plus* teachers was 52.30 with a standard deviation of 13.93 (Figure **14**). These data indicate that on average just over half of the

content in the CPMP textbook was covered, though there was considerable variation among teacher's OTL indices, ranging from 27.69 to 81.71. Teachers who participated in the project workshops have higher OTL indices than non-workshop participants. An ANOVA for unbalanced data showed significant differences in the mean OTL across the four groups $(F = 6.52, p = 0.0012)$. Scheffe's post hoc test found significant differences in the mean OTL between Groups B and F and Groups A and F. These data suggest the importance of workshop attendance on textbook implementation. The two teachers who were provided with an instructional coach, absent workshop attendance, had significantly lower textbook OTL.

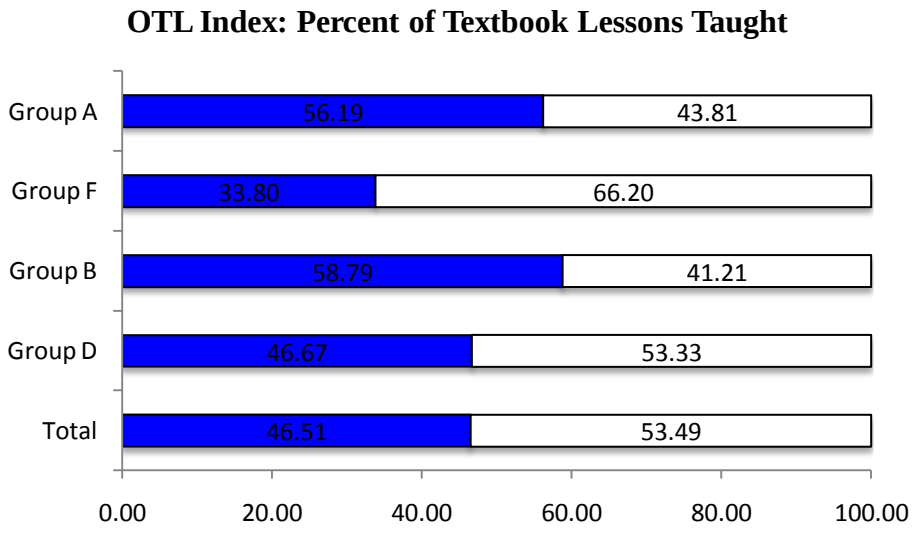


Figure 14. Summary of OTL indices.
Note. Blue shading indicates content taught primarily from textbook, with some supplements, or primarily from an alternative source. White space indicates content not taught.

Extent of Textbook Implementation

Recall the ETI index is a weighted measure describing the degree that the textbook, rather than other materials, was used to teach the content. The mean ETI for this sample was

46.51 with a standard deviation of 16.73 (Figure **15**). There was a significant difference in the mean ETI across teachers in all four groups $(F = 10.31, p < 0.0001)$. Scheffe's post hoc analysis determined differences in Groups B and D, B and F, A and D, and A and F $(\alpha = 0.05)$. These data indicated that participation in the workshop significantly increased teachers' ETI indices.

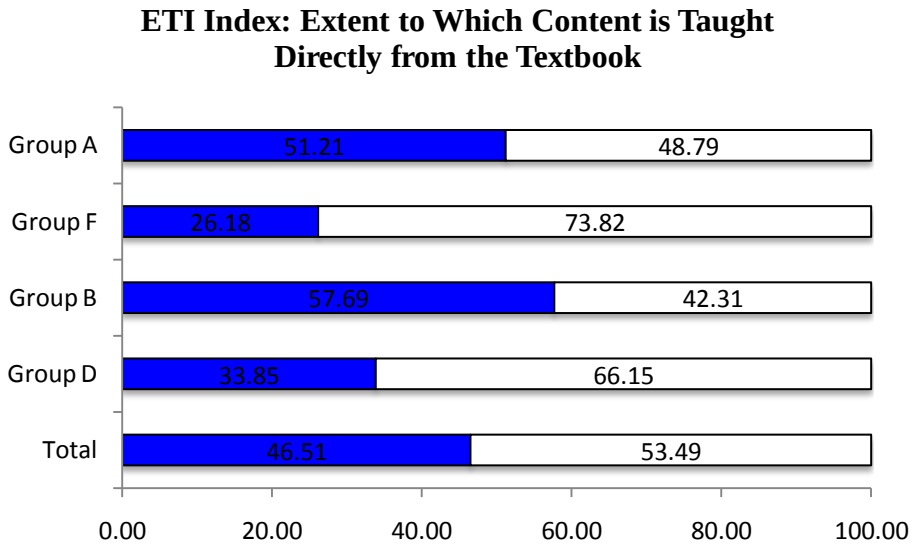


Figure 15. Summary of ETI Index Across Groups.

It was also helpful to look at the ETI index disaggregated by the weighted options: content taught primarily from the textbook, with some supplementation, primarily from alternative sources, or not at all (Figure **16**). Teachers involved in the NCIM professional development supplemented the textbook less frequently and rarely used alternative sources. Groups B and D supplemented textbook with almost the same frequency that they primarily used the textbook for instruction. Group D teachers used alternative sources more frequently than others groups.

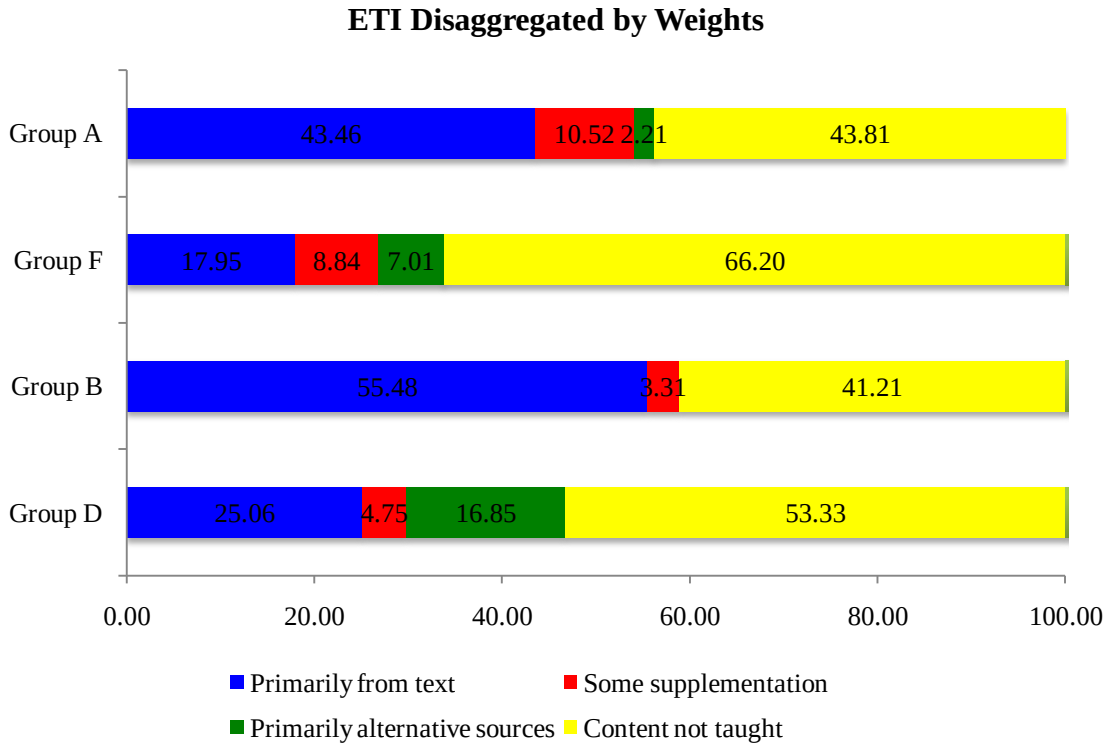


Figure 16. ETI Disaggregated by Weights

Textbook Content Taught

Recall that the TCT index restricted the ETI to consider only the *Core-Plus* content that was taught. The average TCT index for this sample was 87.29 with standard deviation 14.94 (Figure **17**). These data were not normally distributed. The nonparametric Kruskal-Wallis test showed differences in location for these groups ($\chi^2 = 17.02, p = 0.0007$). To determine which groups had significantly different TCT scores a nonparametric post hoc test for multiple comparisons was utilized (Dunn, 1964). Dunn's test showed that the TCT indices were significantly different for Groups B and D and Groups B and F ($\alpha = 0.05$).

When Group B teachers taught content in the textbook, they were directly using the textbook for their instruction instead of supplements.

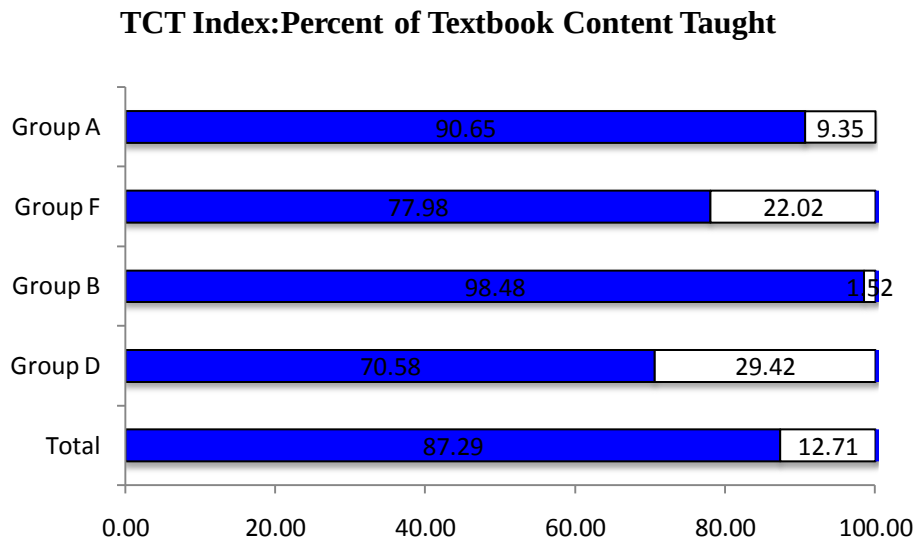


Figure 17. TCT Index Across the Groups.

Differences in the TCT indices became more prevalent when considering the TCT indices for content taught primarily from the textbook, with supplementation, and from alternative sources (Figure **18**). Group D teachers used alternative sources for instruction more frequently than teachers who took part facets of the NCIM professional development. Groups with teachers attending the summer workshops rarely used alternative sources for instruction and utilized the textbook as the primary resource in their instruction.

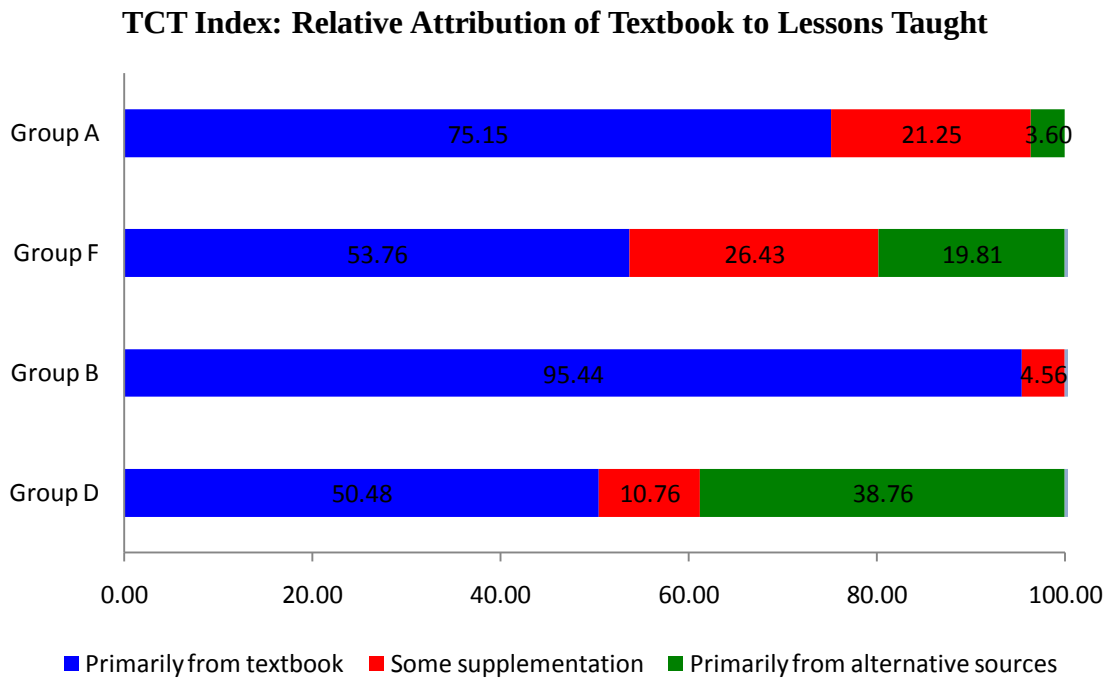


Figure 18. TCT Index Disaggregated by Weights.

The TOC-logs provided evidence of the variance in teachers' implementation of textbook content among teachers with varying levels of NCIM professional development experience. It was clear that teachers who participated in the NCIM summer workshops utilized the textbook for teaching content in the *Core-Plus* curriculum more frequently than teachers who did not attend the workshops. Reasons for the differences among groups will be described in a later section.

Content strands

To further explore textbook implementation, differences in implementation across major topics were analyzed. Units in *Core-Plus* corresponding to four different content strands were identified: Algebra and Function, Probability and Statistics, Geometry and

Trigonometry, and Discrete Mathematics. Each unit of each course was aligned to one (or sometimes two) content strands. The implementation indices were recalculated for each teacher for each of the four content strands (Table **37**). The initial question was: *Are there differences in the implementation indices across the four content strands?* The small size of the data set led to the use of a nonparametric equivalent to the one-way ANOVA, the Kruskal-Wallis distribution free test (Hollander & Wolf, 1999). Results of the Kruskal-Wallis tests showed significant differences in textbook use across teachers and content strands for each of the three implementation indices: OTL $(H' = 89.45, p < 0.0001)$, ETI $(H' = 75.75, p < 0.0001)$, and TCT $(H' = 49.97, p < 0.0001)$.

Table 37. Mean Implementation Indices Across Content Strands

	OTL	ETI	TCT
Algebra and Function	85.53	73.75	86.93
Probability and Statistics	24.58	23.00	65.09
Geometry and Trigonometry	39.71	34.58	69.68
Discrete Mathematics	13.59	14.47	11.83

The Jonckheere post-hoc test (Hollander & Wolfe, 1999) was used to test for ordered alternatives to see if there was a decrease in the indices across the ordered strands: Algebra and Function, Geometry and Trigonometry, Probability and Statistics, and Discrete Mathematics. Results from the distribution-free Jonckheere test showed evidence that all three indices decreased across the ordered content strands: OTL $(J = 2375.50, p < 0.0001)$, ETI $(J = 2586.50, p < 0.0001)$, TCT $(J = 3382.50, p < 0.0001)$. This would imply that the average ETI for the Algebra and Function content strand was greater than the other three, yet

the ETI for the Geometry and Probability and Statistics was greater than the Discrete Mathematics strand ETI.

Dunn's Test, a nonparametric multiple comparisons test based on pair-wise rankings, was used to determine which content strands differed in location when $\alpha = 0.05$. Results are presented in Table **38**; blank entries were not tested.

Table 38. Two-sided Multiple Comparisons Based on Pair-wise Rankings

	Q		
Comparison	OTL	ETI	TCT
Algebra to Statistics	73.8*	67.07*	6.33
Algebra to Geometry	54.3*	50.01*	
Algebra to Discrete	93.86*	87.06*	57.36*
Statistics to Geometry	19.5	17.06	
Statistics to Discrete	20.06	19.99	56.96*
Geometry to Discrete	39.56*	37.05*	57.36*

*Significant at $\alpha = 0.05$

Using the OTL index, there were significant differences between the Algebra $(\bar{x} = 85.53)$ strand and all other strands; Statistics $(\bar{x} = 24.58)$, Geometry $(\bar{x} = 39.71)$, and Discrete $(\bar{x} = 13.59)$. These results were favorable to Algebra, suggesting students were exposed to more Algebra content than the other strands. In addition, the Geometry strand had a significantly higher OTL than the Discrete strand. The Geometry strand had a significantly higher OTL and ETI than the Discrete strand. Results from the ETI index were similar, there were significant differences between the Algebra strand $(\bar{x} = 73.75)$ and all others; Statistics $(\bar{x} = 23.00)$, Geometry $(\bar{x} = 34.58)$, and Discrete $(\bar{x} = 14.47)$. The results from the TCT index show differences between the Discrete strand $(\bar{x} = 11.83)$ and all other strands;

Algebra $(\bar{x} = 86.93)$ Statistics $(\bar{x} = 65.09)$ and Geometry $(\bar{x} = 69.68)$. These findings suggest students had more of an opportunity to learn Algebra than the other topics and that the Discrete topics were not a focus of instruction during the 2009-2010 school year. This is not surprising as the discrete math topics were not covered on the state assessment and as Thomas (2011) reported this material is often unfamiliar to teachers.

Taken together these findings point to the wide variance in implementation across the four main content strands in the textbook. This also provides empirical evidence to suggest teachers were skipping major units in the textbook. Given the strong emphasis on the Algebra content across the teachers, a separate analysis of Opportunity to Learn Algebra (OTLAlg) was conducted across the four groups. The mean OTLAlg for teachers in Group A was 87.41, Group F was 60.21, Group B 94.42, and Group D 87.95. A Kruskal-Wallis test showed differences among these groups $(\chi^2 = 16.83, p = 0.0008)$. Group F had a significantly different OTLAlg index than the other three groups $(\alpha = 0.05)$.

Summary of Implementation Indices

There was wide variance among the groups in their textbook implementation. Teachers with NCIM workshop attendance (Groups A and B) more faithfully implemented textbook content, as self-reported by the teachers. Group D teachers, without exposure to the NCIM professional development, more frequently supplemented textbook content. Qualitative explanations for these results are provided in the fourth section of this chapter.

Due to the wide variance of content taught across the four content strands, and because the student outcome measure was algebra achievement, it was important to calculate

and use the OTLAlg index as a predictor in the HLM analyses. This index provided a measure for exposure to algebra topics in the *Core-Plus* curriculum, which helps determine the claims that can be made about the effectiveness of the *Core-Plus* curriculum on student outcomes. The next section investigates how these and other predictors impact student achievement.

Modeling Student Algebra I Achievement: Focus Group

This section models student Algebra I achievement to determine if there were differences on the effects of student learning based on varying levels of teacher participation in the NCIM professional development. In addition, the indices presented in the previous section were incorporated into the HLMs to determine the impact of teachers' textbook implementation on student outcomes.

Recall, teachers in the focus group were required to have either Algebra I or Algebra II EOC data from the 2009-2010 school year. There was no requirement that each teacher have student scores from both the Algebra I and Algebra II exams. As a result, five of the seven Group A, one of the two Group F, all six of the Group B, and four of the six Group D teachers were represented in the focus group Algebra I data set.

Background and Demographics

The majority of students took the Algebra I EOC in 9^{th} (59.73%) or 10^{th} grades (37.37%), with little representation from upperclassmen **(**Table **39**). Each of the four groups in the sample had more female students taking the exam than males. The groups served different ethnic populations of students. The most prevalent ethnicity in Groups B and D

were White students, 76.22% and 76.63% respectively, while Groups A and F had higher percentages of Black students, 40.54% and 65.71% respectively.

Table 39. Percentages of Students Within Each Category for Integrated Algebra I Sample

	Grade				**Gender**	
	9th	10th	11th	12th	Male	Female
Group A	64.32	33.51	2.16	0.00	43.24	56.76
Group F	82.86	14.29	2.86	0.00	37.14	62.86
Group B	67.65	30.39	1.96	0.00	36.76	63.24
Group D	39.51	55.56	3.70	1.23	43.21	56.79
Total	59.73	37.37	2.56	0.34	40.61	59.39
	Race					
	AmIn	Asian	Hispanic	Black	White	Multi
Group A	0.00	1.08	16.22	40.54	37.30	4.86
Group F	2.86	5.71	0.00	65.71	25.71	0.00
Group B	0.98	9.80	9.31	1.23	76.22	2.45
Group D	0.62	0.62	6.79	12.25	76.63	3.09
Total	0.68	5.97	10.24	21.33	58.54	3.24

Subgroup Means and Passing Rates for Each Group

Student Algebra I achievement was very similar among the groups. The mean for each group was within 0.77 points of the sample mean (156.88) and all groups had passing rates of 85%. There was a large decrease in achievement based on the grade level during which a student enrolled in Algebra I for all groups $(\mu_{9th} = 159.97, \mu_{10th} = 152.57, \mu_{11th} = 147.87)$. Ninth grade students from each group outperformed 10th grade students from the same group by 8.90 to 11 points. The mean for 9th grade students in Group B (161.04) was higher than the mean for 9th grade students in all other groups, by at least 1.10 points. Group F had the lowest mean for 9th grade students (157.03). Tenth grade students in Group D (155.07) had a higher mean than all other groups

by at least 3 points. Due to the small sample sizes for 11th and 12th grade students, it was difficult to draw conclusions across the groups. Female students had a higher overall mean than male students ($\mu_{Female} = 157.12, \mu_{Male} = 156.53$), with the largest differences between genders for Group F, favoring males by 1.65 points, and for Group B, favoring females by 2.83 points. Across all groups Asian students had a 100% passing rate and the highest mean scores for any subgroup. Hispanic students in Group A had a higher mean (158.37) than the Hispanic students in Groups B (154.47) and D (155.91). Black students had similar passing rates among Groups A, F, and B, ranging from 75% to 78.26%. Group D had a small number of Black students (n=2), and they both passed the Algebra I exam. White students in Group A had the highest mean achievement (160.03) and white students in Group D had the lowest mean achievement (156.17). There was a wide variance in Multi-racial students' mean and passing rates due to low sample sizes for this subgroup.

Table 40. Algebra I EOC Means and Passing Rates for Students Within Each Subgroup

		Grand Mean	Grade Level				Gender	
			9th	10th	11th	12th	Male	Female
Group A	Sample Size (n)	185	119	62	4	-	80	105
	Mean	156.62	159.94	151.02	144.50	-	156.19	156.94
	Standard Dev.	8.64	7.71	6.77	5.26	-	9.17	8.25
	Passing Rate	85.95%	96.64%	69.35%	25.00%	-	86.25%	85.71%
Group F	Sample Size (n)	35	29	5	1	-	13	22
	Mean	156.11	157.03	151.80	151.00	-	157.15	155.50
	Standard Dev.	8.09	8.04	8.23	-	-	8.96	7.68
	Passing Rate	85.71%	86.21%	80.00%	100.00%	-	84.62%	86.36%
Group B	Sample Size (n)	202	138	60	4	-	74	128
	Mean	157.59	161.04	150.50	145.25	-	156.85	158.02
	Standard Dev.	10.79	9.97	8.67	6.18	-	11.34	10.49
	Passing Rate	85.15%	92.03%	70.00%	75.00%	-	82.43%	86.72%
Group D	Sample Size (n)	162	64	90	6	2	70	92
	Mean	156.45	159.08	155.07	151.33	150.00	156.47	156.43
	Standard Dev.	8.68	8.37	7.92	14.15	16.97	9.04	8.45
	Passing Rate	86.42%	93.75%	83.33%	66.67%	50.00%	84.29%	88.04%
Entire Sample	Mean	156.88	159.97	152.57	147.87	150.00	156.53	157.12
	Standard Dev.	9.42	8.85	8.07	9.83	16.97	9.80	9.15

Table 40. Continued

		Race						Attendance	
		AmIn	Asian	Hispanic	Black	White	Multi	Absent ≤ 20 Days	Absent > 20 Days
Group A	Sample Size (n)	-	2	30	75	69	9	180	5
	Mean	-	168.50	158.37	152.76	160.03	154.11	156.82	149.20
	Standard Dev.	-	6.36	8.83	7.95	7.56	8.18	8.60	7.53
	Passing Rate	-	100.00%	93.33%	76.00%	95.65%	66.67%	86.11%	80.00%
Group F	Sample Size (n)	1	2	-	23	9	-	34	1
	Mean	161.00	161.00	-	154.39	158.89	-	156.59	140.00
	Standard Dev.	-	16.97	-	8.47	4.31	-	7.70	-
	Passing Rate	100.00%	100.00%	-	78.26%	100.00%	-	88.24%	0.00%
Group B	Sample Size (n)	2	30	19	24	122	5	192	10
	Mean	146.50	165.30	154.47	152.29	157.31	160.00	158.06	148.70
	Standard Dev.	6.36	7.28	13.98	8.88	10.57	4.00	10.77	6.93
	Passing Rate	50.00%	100.00%	73.68%	75.00%	85.25%	100.00%	86.46%	60.00%
Group D	Sample Size (n)	1	1	11	2	142	5	153	9
	Mean	141.00	177.00	155.91	162.50	156.17	162.20	156.91	148.67
	Standard Dev.	-	-	7.16	12.02	8.54	8.35	8.38	10.52
	Passing Rate	0.00%	100.00%	90.91%	100.00%	85.92%	100.00%	88.89%	44.44%
Entire Sample	Mean	148.75	165.57	156.68	153.13	157.43	157.79	157.26	148.44
	Standard Dev.	9.32	7.78	10.47	8.29	9.15	7.89	9.30	8.22

Similarities and Differences Between Three Samples: Larger, Workshop, and Focus Group

This section will briefly situate the population of students in the focus group sample with the larger and workshop samples of integrated mathematics teachers. The focus group sample is a subset of the larger sample of integrated mathematics teachers, which also contains integrated mathematics teachers from the workshop sample. The percentage of focus group students from each grade level (9^{th}: 59.73%, 10^{th}: 37.37%, 11^{th}: 2.56%, 12^{th}: 0.34%) was similar to the larger sample of integrated mathematics students (9^{th}: 56.70%, 10^{th}: 39.56%, 11^{th}: 3.26%, 12^{th}: 0.48%). The percentage of female students in the Algebra I focus group sample was at least 6.46% higher than the other two samples. Groups A and F in the focus group had a higher percentage of Black (40.54% and 65.71% respectively) students than the larger and workshop sample, and Groups B and F had very few Black students taking the Algebra I exam (1.23% and 12.25% respectively). Similarly, Groups A and F had a lower ADM (275.25 and152), higher %FRL average (72.85% and 96%), and were more likely to be classified as Title One (75% and 100%) than the workshop sample, but Groups B and D had a much higher ADM (781.4 and 726), and much lower %FRL (39.03% and 42.29%) and Title One classification (0% and 0%) versus the workshop sample. The average Algebra I achievement score for the focus group ($\mu = 156.88$) was the same as the mean for the larger integrated mathematics sample ($\mu = 156.87$) and one point lower than the workshop integrated mathematics sample ($\mu = 157.80$). 9^{th} grade students had similar means across the three samples ($\mu_L = 159.78, \mu_{WS} = 159.69, \mu_{Focus} = 159.94$), 10^{th} grade students in the focus group had a mean one to two points lower than the other two samples

($\mu_L = 153.55, \mu_{WS} = 154.86, \mu_{Focus} = 152.57$), 11th grade students in the focus group had a mean higher than the workshop sample but lower than the larger sample, ($\mu_L = 148.29, \mu_{WS} = 145.42, \mu_{Focus} = 144.50$) and for 12th grade integrated mathematics students the sample size was too small to make generalizations among the three samples. The mean for Black students in the workshop and larger samples was at least 0.90 points higher than the focus group sample. White students in the workshop sample had the highest mean achievement (159.1), followed by the focus group teachers (157.43), and the larger sample of teachers (157.04). Taken together, these three samples have a lot of similarities in terms of the population of students they serve and in the mean achievement for subgroups of those students. Groups A and F were higher poverty settings than the larger or workshop samples, whereas Groups B and D were much lower.

Fully Unconditional Model

Before running sophisticated models of student Algebra I achievement, it was necessary to determine if there was sufficient variability in student achievement at each level for further analysis. The preliminary analysis showed there was insufficient variability at the school-level when running the three-level fully unconditional model for the Algebra I sample ($\pi_\beta = 2.92, Z = 0.34, p = 0.37$). As a result, the school-level variables were not included in these analyses. The two-level HLM model for student Algebra I EOC score, showed 24.30% of the variance in student Algebra I achievement was between-teacher ($\tau_{00} = 21.95, Z = 2.44, p = 0.0073$), and 75.70% was within-teacher

($\sigma^2 = 68.37, Z = 16.85, p < 0.0001$). The grand mean for the sample was 156.01 ($t = 126.49, p < 0.0001$), 4.37 points higher than the grand mean for the entire sample.

Differences in Algebra I Achievement Among Groups

To determine if there were differences among the four groups when controlling for student background characteristics, the next model only incorporated the Groups at level-2, with Group A as the referent. The effects of grade level and the number of days a student was absent were both significantly negative predictors of Algebra I achievement (Table **42**). American Indian, Hispanic, and Black ethnicities each performed significantly lower on the Algebra I exam than White students. Students of teachers in Groups B and D scored lower than students of teachers in Group A on the Algebra I exam, but this result was non-significant. Students of Group F teachers had a higher, non-significant score on the Algebra I EOC than students of teachers in Group A. This model accounted for 52.94% of the between-teacher variability in test score and 15.59% within-teacher.

Including the same student-level predictors and changing the referent groups resulted in similar results. There were no differences between groups B and D, B and F, or D and F on the Algebra I EOC. Regardless of the referent group, there were no statistically significant differences in student Algebra I achievement for teachers with varying levels of participation in the NCIM professional development. Recall the differences in the student population among the groups, Groups A and F coming from high need schools, and that neither prior achievement nor school-level factors, such as %FRL, Title One status, or ADM, could be included in these models. It is encouraging that students in each of these groups has similar

Algebra I performance given the lack of control for the distinct differences in these populations. In other words, students at high-needs schools with large numbers of underrepresented populations performed similarly to students at low-need schools reporting student populations of greater than 76% White students, in HLM models that did not attend to differences in prior achievement or school need.

Table 41. Differences Among Groups Modeling Algebra I Achievement

Level-2 Predictor	Coefficient	Standard Error	t-Value
Algebra I Score, β_{0i}			
Intercept γ_{00}	156.28	0.87	177.76 ***
Gender γ_{10}	0.79	0.66	1.20
Grade Level γ_{20}	-4.52	0.73	-6.18 ***
Days Absent γ_{30}	-0.28	0.050	-5.68 ***
American Indian γ_{40}	-10.43	3.91	-2.67 **
Asian γ_{50}	2.10	1.51	1.39
Hispanic γ_{60}	-2.22	1.14	-1.96 *
Black γ_{70}	-5.33	0.97	-5.52 ***
Multi γ_{80}	-0.040	1.84	-0.02
Group F γ_{01}	1.01	3.82	0.27
Group B γ_{02}	-1.06	2.16	-0.49
Group D γ_{03}	-0.62	2.40	-0.26
Random Effects	Estimate	Standard Errors	Z-Value
Within-teacher, σ^2	57.71	3.45	16.72 ***
Between-teacher, τ_{00}	10.33	5.31	1.95 *

Note. $(p < 0.05^{*}, p < 0.001^{**}, p < 0.001^{***})$.

Individual Level-2 Predictors When Modeling Algebra I Achievement

To determine the impact of each individual teacher-level variable on student achievement, when controlling for student demographics and participant groups, thirteen additional individual models were run. Due to the numerous teacher-level data collected for this sample, it was not informative to simultaneously include them all in models of student

achievement. Hence, each one was individually tested to determine its impact on student achievement. Teacher gender and belief about traditional practices were the only significant predictors of Algebra I achievement (Table **42**). On average, students of female teachers had Algebra I EOC performance results almost six points higher than students of male teachers[16]. For each one-point increase in teachers' traditional beliefs (Belief Factor 5) about mathematical practices there tended to be a two-point decrease student in Algebra I score.

Table 42. Correlations Among Individual Level-2 Predictors and Algebra I Score, When Controlling for Participant Group and Student-Level Variables

Level-2 Predictor	Coefficient	Standard Error	t-Value
Teacher Gender γ_{03}	-5.97	1.91	-3.12 **
Content Knowledge γ_{03}	0.024	0.075	0.32
Years Taught CPMP γ_{03}	-0.012	0.31	-0.04
Years Taught γ_{03}	0.22	0.13	1.64
OTL γ_{03}	0.041	0.083	0.50
ETI γ_{03}	0.030	0.074	0.40
TCT γ_{03}	-0.020	0.095	-0.21
OTLAlg γ_{03}	0.092	0.15	0.62
Belief Factor 1 γ_{03}	2.59	1.62	1.61
Belief Factor 2 γ_{03}	0.59	1.54	0.38
Belief Factor 3 γ_{03}	-0.10	1.51	-0.07
Belief Factor 4 γ_{03}	1.58	1.01	1.56
Belief Factor 5 γ_{03}	-2.03	1.03	-1.97 *

Note. $(p < 0.05^{*}, p < 0.001^{**}, p < 0.001^{***})$.

Incorporating Multiple Level-2 Predictors in a Model of Algebra I Achievement

The next model was used to explain more of the level-2 variance and to determine if there were differences between Groups F, B, and D to Group A when controlling for multiple teacher-level variables. The five teacher-level variables with the lowest p-values in the

[16] The Algebra I focus group contained students from 13 female and three male teachers.

previous analysis were included with the student background and group variables. Holding teacher characteristics constant, there were still significant negative differences in Algebra I score based on grade level and days a student was absent (Table **43**). On average, Asian students scored significantly higher than Whites, and Black students scored significantly lower than Whites. The only significant level-2 predictor remained teacher gender, with a strong effect size of -5.17. This result indicated that students of female teachers scored over five points higher than students of male teachers. There were no significant differences between Group A and the other groups. This model accounted for 67.29% of the between-teacher variability in Algebra I score and 14.40% of the within-teacher variability.

Table 43. Multiple Level-2 Predictors Modeling Algebra I Achievement

Level-2 Predictor	Coefficient	Standard Error	t-Value
Algebra I Score, β_{0i}			
Intercept γ_{00}	156.15	0.93	168.60 ***
Gender γ_{10}	0.44	0.69	0.64
Grade Level γ_{20}	-3.96	0.77	-5.15 ***
Days Absent γ_{30}	-0.28	0.052	-5.32 ***
American Indian γ_{40}	5.04	2.17	2.32 *
Asian γ_{50}	-2.25	1.23	-1.83
Hispanic γ_{60}	-4.55	1.00	-4.53 ***
Black γ_{70}	0.23	1.95	0.12
Multi γ_{80}	1.68	6.69	0.25
Group F γ_{01}	0.31	3.02	0.10
Group B γ_{02}	-0.65	3.39	-0.19
Teacher Gender γ_{01}	-5.17	2.67	-1.94*
CKA γ_{01}	0.0067	0.14	0.05
OTLAlg γ_{01}	-0.028	0.20	-0.14
Years Taught CPMP γ_{01}	0.038	0.30	0.13
Belief Factor 1 γ_{01}	1.70	2.43	0.70
Belief Factor 5 γ_{01}	0.68	1.75	0.39

Table **43**. Continued

Random Effects	Estimate	Standard Error	Z-Value
Within-teacher, σ^2	58.52	6.13	1.17
Between-teacher, τ_{00}	7.18	3.70	15.82 ***

Note. $(p < 0.05^{*}, p < 0.001^{**}, p < 0.001^{***})$.

Two models were run with the same student- and teacher-level predictors, but with changes to the referent group. Similarly, these models showed there were no statistically significant results between groups B and D, B and F, or D and F on the Algebra I EOC. Therefore, there were no statistically significant differences in student Algebra I achievement for teachers with varying levels of participation in the NCIM professional development.

Summary of Algebra I EOC Results

Similar to the Algebra II results for the workshop sample in Chapter 4, there was an effect of teacher gender. Though the number of male teachers in the sample was limited, female integrated mathematics teachers had higher student Algebra I results than male integrated mathematics teachers. Teacher beliefs about traditional mathematics practices were negatively correlated to student achievement, but the effect diminished when controlling for other teacher characteristics. Due to the wide variance in the indices generated from the TOC-logs, it was surprising there were no correlations between the implementation indices and student Algebra I achievement.

Controlling for student and teacher traits, student achievement was the same for this sample of *Core-Plus* teachers, regardless their level and type of NCIM professional development exposure. While there was no statistical difference in these groups, considering the differences in student population for each group, with varying school need that could not

be accounted for, results are encouraging for the use of integrated mathematics by the NCIM teachers in Groups A and F.

Modeling Student Algebra II Achievement: Focus Groups

Similar to the previous section, the goal of this section is to determine if there were differences on the effects of student learning, as measured by the Algebra II EOC, based on varying levels of teacher participation in the NCIM professional development. Group B schools were implementing integrated mathematics for the first time during the 2009-2010 school year. As a result, there were no Group B teachers with students taking the Algebra II exam and they are removed from this analysis. Five of the seven Group A teachers, both of the Group F teachers, and three of the six Group D teachers had student Algebra II achievement data and were included in this sample.

Background and Demographics

Students in this sample were evenly distributed between 10^{th} and 11^{th} grades (46.89% and 47.10% respectively), with few students taking the Algebra II EOC as freshmen (3.32%) or seniors (2.70%) (Table **44**). The percentage of male and female students in each group were generally the same, except for Group F, which had a much higher percentage of female students (69.39%). There were no American Indians in this sample and very few Asians (0.83%). Groups were not similar in their percentages of each ethnicity; Group D had over 90% White students, while Group F had over 87% Black students.

Table 44. Percentages of Students Within Each Category, Integrated Algebra II Sample

	Grade				Gender	
	9th	10th	11th	12th	Male	Female
Group A	5.09	35.19	57.41	2.31	48.15	51.85
Group F	2.04	89.80	8.16	0.00	30.61	69.39
Group D	1.84	48.85	45.62	3.69	47.93	52.07
Total	3.32	46.89	47.10	2.70	46.27	53.73
	Race					
	AmIn	Asian	Hispanic	Black	White	Multi
Group A	0.00	0.00	7.87	56.48	31.94	3.70
Group F	0.00	2.04	2.04	87.76	4.08	4.08
Group D	0.00	1.38	6.45	0.92	90.32	0.92
Total	0.00	0.83	6.64	34.65	55.39	2.49

Subgroup Means and Passing Rates for Each Group

The average prior achievement scores among groups differed by as much as four points on the Algebra I EOC; Group A 155.84 (8.79), Group F 153.69 (8.20), and Group D 157.97 (7.50). Students of teachers in Groups F and D had similar mean Algebra II achievement $(\mu_F = 152.22, \mu_D = 152.49)$, but students of Group A teachers had a mean achievement four points lower than the other two groups $(\mu_A = 148.29)$ (Table 45). Group F had the highest overall passing rate at 81.63% and Group A had the lowest at just over half of the students passing the Algebra II exam. Group A students had the lowest means across all subgroups with at least five students in the sample. Group D had the highest mean Algebra II achievement for ten of the 14 subgroups reported. Though the sample only included two Black students taking the Algebra II exam from Group D, this subgroup had one of the lowest mean achievement scores at 139.5. The only average lower than that subgroup was for 10^{th} grade students in Group A, with an average of 139.4 points.

Table 45. Algebra II EOC Means and Passing Rates for Students Within Each Subgroup

		Grand Mean	Grade Level				Gender	
			9th	10th	11th	12th	Male	Female
Group A	Sample Size (n)	216	11	76	124	5	104	112
	Mean	148.29	153.27	150.33	146.96	139.40	148.25	148.33
	Standard Dev.	7.30	4.27	7.46	6.83	6.07	7.28	7.34
	Passing Rate	54.63%	90.91%	65.79%	45.97%	20.00%	52.88%	56.25%
Group F	Sample Size (n)	49	1	44	4	-	15	34
	Mean	152.22	146.00	152.66	149.00	-	153.27	151.76
	Standard Dev.	6.15		6.22	4.69	-	3.24	7.06
	Passing Rate	81.63%	0.00%	86.36%	50.00%	-	93.33%	76.47%
Group D	Sample Size (n)	215	4	106	97	8	103	112
	Mean	152.49	157.25	153.66	151.01	152.50	153.17	151.87
	Standard Dev.	7.27	8.88	7.63	6.45	8.75	6.51	7.89
	Passing Rate	76.28%	100.00%	80.19%	70.10%	87.50%	81.55%	71.43%
Entire Sample	Mean	150.57	153.81	152.35	148.74	147.46	150.87	150.32
	Standard Dev.	7.46	5.95	7.44	6.91	10.05	7.14	7.72

Table 45. Continued

		Race					Attendance	
		Asian	Hispanic	Black	White	Multi	Absent ≤ 20 Days	Absent > 20 Days
Group A	Sample Size (n)	-	17	122	69	8	208	8
	Mean	-	145.82	147.79	150.12	145.50	148.44	144.38
	Standard Dev.	-	5.99	7.04	7.42	10.07	7.28	7.19
	Passing Rate	-	41.18%	54.10%	62.32%	25.00%	55.77%	25.00%
Group F	Sample Size (n)	1	1	43	2	2	48	1
	Mean	156.00	172.00	151.74	153.50	149.50	152.58	135.00
	Standard Dev.	-	-	5.60	7.78	2.12	5.68	-
	Passing Rate	100.00%	100.00%	79.07%	100.00%	100.00%	83.33%	0.00%
Group D	Sample Size (n)	3	14	2	194	2	205	10
	Mean	161.33	151.29	139.50	152.55	154.50	152.61	149.90
	Standard Dev.	9.45	6.90	4.95	7.16	0.71	7.35	5.17
	Passing Rate	100.00%	85.71%	0.00%	75.77%	100.00%	76.59%	70.00%
Entire Sample	Mean	160.00	149.03	148.71	151.92	147.67	150.73	146.79
	Standard Dev.	8.16	7.97	6.94	7.29	8.80	7.44	7.01

Similarities and Differences Between Three Samples: Larger, Workshop, and Focus Group

This section will situate the sample of students in the focus group within the larger sample of integrated mathematics students, to which the focus group is a subset, and to the workshop sample of integrated mathematics students, to which the focus group intersects but is not contained within. Overall, the focus group had a smaller percentage of 10^{th} grade students (46.89%) and a larger percentage of 11^{th} (47.1%) grade students than either the larger or workshop samples. Thus, on average students in the focus group took Course 3 later in high school than students in the other samples. There was a larger percent of male students in the focus group sample (46.27%) than the other two, however there were still a higher percentage of female students in all three samples (A: 51.85%, F:69.39%, D: 52.07%). There were similar percentages for non-Black minority students for each of the three samples. On average the workshop sample of integrated mathematics students had the highest percentage of Black students(61.38%), but Group F from the focus group had the highest individual percentage at 87.76%. The larger sample had the highest percentage of White students at 58.14%, but the individual percent for White students for Group D was 90.32%. The average %FRL for the larger sample was 61.56% and for the workshop sample was 72.85%. Groups A and F had a higher %FRL (72.85% and 96%) than those two, whereas the %FRL for Groups B and D (39.03% and 42.29%) were lower. Similarly, Groups A and F were more likely, than the larger or workshop samples, to be classified as Title One (75% and 100%), whereas neither Group B nor D included Title One schools. The grand mean for the focus group was 1.74 points higher than the mean for the workshop sample and 1.36 points lower

than the larger sample. The average Algebra II achievement for 9th grade students in the focus sample (153.81) was the same as the larger sample (153.81) and higher than the workshop sample (152.67), 10th and 11th grade students in the focus sample had a mean Algebra II achievement (10th: 152.35, 11th: 148.74) in the median of the workshop (lowest mean) (10th: 150.64 and 11th: 146.94) and larger (highest mean) (10th: 153.69 and 11th: 150.35) samples. 12th grade students in the focus group had a higher mean score than the larger and workshop samples by 0.50 and 8.13 points respectively. Students in the larger sample had higher mean achievement for each ethnicity represented in the sample. To conclude, students from Groups A and F came from schools with higher need and higher minority populations than the larger or workshop samples. Conversely, students from Group B and D came from larger schools with lower need and a smaller percent minority population. On average, the focus sample had higher mean achievement across the subgroups than the workshop sample, but slightly lower than the larger sample.

Fully Unconditional Model

There was insufficient variability at the school level when running the three-level fully unconditional model for the Algebra I sample ($\pi_\beta = 9.31, Z = 0.89, p = 0.19$). As a result, the school level variables were not included in these analyses. The two-level HLM model for student Algebra II EOC score, showed 20.51% of the variance in student Algebra II achievement was between-teacher ($\tau_{00} = 12.30, Z = 1.84, p = 0.032$), and 79.49% was within-teacher ($\sigma^2 = 47.66, Z = 15.32, p < 0.0001$). The grand mean for the sample was 149.82 ($t = 128.18, p < 0.0001$), 1.82 points lower than the grand mean for the entire sample.

Differences in Algebra II Achievement Among Groups

To determine if there were differences among the three groups' student Algebra II EOC scores, when controlling for student background characteristics, the next model only incorporated the Groups at level-2, with Group A as the referent. Prior achievement and the number of days a student was absent were the only significant predictors of Algebra II achievement (Table **46**). There were no significant differences in student achievement based on race. There were neither significant differences between Group F nor D with Group A. This model accounted for 43.94% of the between-teacher variability in Algebra II EOC score and 53.33% of the within-teacher variability.

Table 46. Differences Among Groups Modeling Algebra II Achievement

Level-2 Predictor	Coefficient	Standard Error	t-Value
Algebra I Score, β_{0i}			
Intercept γ_{00}	150.46	0.92	164.32 ***
Algebra I Score γ_{10}	0.63	0.032	20.03 ***
Gender γ_{20}	-0.085	0.45	-0.19
Grade Level γ_{30}	-0.39	0.43	-0.91
Days Absent γ_{40}	-0.095	0.037	-2.56 *
Asian γ_{50}	0.94	2.78	0.34
Hispanic γ_{60}	1.53	0.96	1.60
Black γ_{70}	-0.063	0.72	-0.09
Multi γ_{80}	-0.59	1.47	-0.40
Group F γ_{01}	3.78	2.37	1.60
Group D γ_{02}	2.16	2.03	1.06
Random Effects	Estimate	Standard Errors	Z-Value
Within-teacher, σ^2	22.24	1.48	15.03 ***
Between-teacher, τ_{00}	6.90	4.17	1.65 *

Note. $(p < 0.05^{*}, p < 0.001^{**}, p < 0.001^{***})$.

Including the same student-level predictors and changing the referent group resulted in similar results. There were no differences between groups D and F on the Algebra II EOC. Thus, there were no statistically significant differences in student Algebra II achievement for teachers with varying levels of participation in the NCIM professional development.

Individual Level-2 Predictors When Modeling Algebra II Achievement

Due to the numerous teacher-level data collected for this sample, not all could be included in models of student achievement. To determine the impact of each individual teacher-level variable on student achievement, when controlling for student demographics and participant groups, 13 additional individual models were run (Table 47). The only variable with a significant impact on Algebra II achievement was teacher content knowledge, such that for each 1-point increase in teacher content knowledge, there tended to be a 0.13-point increase in Algebra II score. Incorporating individual teacher-level variables did not change the insignificance between Groups F and D to Group A, except when adding content knowledge. When including teacher content knowledge into the model, there was a significant difference between Group F and Group A. When keeping teacher content knowledge constant among teachers, Group F teachers had Algebra II scores that were, on average, 6.46 points higher than Group A teachers $(t = 5.86, p < 0.0001)$. Possible explanations for this result will be provided in the next section.

Table 47. Correlations Among Individual Level-2 Predictors and Algebra II Score, When Controlling for Curriculum Type and Student-Level Variables

Level-2 Predictor	Coefficient	Standard Error	t-Value
Teacher Gender γ_{03}	1.93	3.24	0.60
Content Knowledge γ_{03}	0.13	0.023	5.75 ***
Years Taught CPMP γ_{03}	0.22	0.29	0.76
Years Taught γ_{03}	-0.098	0.14	-0.71
OTL γ_{03}	-0.042	0.082	-0.51
ETI γ_{03}	-0.027	0.077	-0.35
TCT γ_{03}	-0.025	0.078	-0.32
OTLAlg γ_{03}	0.036	0.14	0.25
Belief Factor 1 γ_{03}	4.13	2.71	1.53
Belief Factor 2 γ_{03}	0.30	1.64	0.18
Belief Factor 3 γ_{03}	-0.34	1.38	-0.25
Belief Factor 4 γ_{03}	-0.039	1.03	-0.04
Belief Factor 5 γ_{03}	0.65	1.19	0.54

Note. $(p < 0.05^{*}, p < 0.001^{**}, p < 0.001^{***})$.

Incorporating Multiple Level-2 Predictors in a Model of Algebra II Achievement

The next model was used to explain more of the level-2 variance and to determine if there were differences between Groups F and D to Group A when controlling for multiple teacher-level variables. The five teacher-level variables with the lowest p-values in the previous analysis were included with student background and the group variables. Prior achievement and days a student was absent were still significant predictors of Algebra II achievement (Table **48**). In addition, content knowledge, OTLAlg, and the number of years a teacher had taught using *Core-Plus* were all significant predictors of the outcome variable. Content knowledge and the number of years teaching *Core-Plus* both had positive fixed effects. OTLAlg was negatively correlated to Algebra II score. Controlling for these variables resulted in Groups F and D becoming significantly different than Group A. When

holding the five teacher-level variables and student background characteristics constant, Group F scored 5.22 points higher than teachers in Group A and Group D scored 3.22 points lower than teachers in Group A. This model accounted for all of the between-teacher variability and 53.41% of the within-teacher variability in Algebra II achievement.

Table 48. Multiple Level-2 Predictors Modeling Algebra II Achievement

Level-2 Predictor	Coefficient	Standard Error	t-Value
Algebra I Score, β_{0i}			
Intercept γ_{00}	150.49	0.22	691.65 ***
Algebra I Score γ_{10}	0.65	0.031	20.84 ***
Gender γ_{20}	-0.14	0.44	-0.32
Grade Level γ_{30}	-0.43	0.43	-1.00
Days Absent γ_{40}	-0.10	-.037	-2.75 **
Asian γ_{50}	0.99	2.77	0.36
Hispanic γ_{60}	1.46	0.95	1.53
Black γ_{70}	-0.15	0.72	-0.20
Multi γ_{80}	-0.76	1.28	-0.52
Group F γ_{01}	5.22	2.18	2.39 *
Group D γ_{02}	-3.22	1.61	-2.00 *
CKA γ_{03}	0.14	0.026	5.22 ***
OTLAlg γ_{04}	-0.12	0.048	-2.30 *
Years Taught CPMP γ_{05}	0.36	0.12	2.93 **
Belief Factor 1 γ_{06}	0.20	1.28	0.16
Belief Factor 5 γ_{07}	-0.24	0.41	-0.60
Random Effects	Estimate	Standard Errors	Z-Value
Within-teacher, σ^2	22.20	1.47	15.07 ***
Between-teacher, τ_{00}	0.00		

Note. $(p < 0.05^{*}, p < 0.001^{**}, p < 0.001^{***})$.

Including the same student- and teacher-level predictors and changing the referent group resulted in an additional difference between groups. Group F outperformed Group D by 8.44 points $(t = 5.06, p < 0.0001)$. Therefore, each of the three groups had significantly

different student Algebra II achievement when controlling for student and teacher-level variables. Possible reasons for these differences will be provided at the end of the chapter.

Summary of Algebra II EOC Results

There were no differences between Group A and F teachers or Group A and D teachers on the Algebra II EOC when the groups were the only level-2 predictors. However, when controlling for additional teacher characteristics, significant differences among these groups emerged. Further, teacher-level predictors for OTLAlg and teaching experience with *Core-Plus,* that were non-significant in individual models, became significant when incorporating multiple teacher-level traits. Findings showed that experience teaching *Core-Plus* had a positive impact on student achievement, while OTLAlg was negatively related to Algebra II scores. Possible reasons for this will be provided in the next section.

Teacher content knowledge was positive and significant throughout the analysis of Algebra II results. The content in the Course 3 textbook is difficult, and due to the correlation between teacher content knowledge and Algebra II achievement, these results point to the need for highly knowledgeable Course 3 teachers. Recall, the Group F teachers had the lowest content knowledge assessment scores. When controlling for differences in content knowledge, there was evidence of teachers in Group F having a significantly higher student achievement than Group A teachers, however there was no difference among groups when content knowledge was not included as a predictor.

Factors Influencing Teacher Implementation

Across the focus groups, quantitative data in the previous three sections have begun to provide detail on differences in implementation and student achievement for teachers with varying levels of NCIM professional development experience, yet the full story is still far from complete. Thus far, the wide variance in teachers' implementation of *Core-Plus* among the four groups has been documented. Teachers with NCIM workshop experience had much higher textbook implementation indices as measured by the mean OTL, ETI, and TCT indices, which were at least 10, 17, and 13 points higher, respectively, than teachers without NCIM workshop experience (Table 36). Difference in OTLAlg between Group F and the others were significant, suggesting Group F teachers provided their students with less opportunity to learn Algebra content. Using these indices in HLM models of student Algebra I achievement showed no correlation between student outcomes and implementation indices and non-significant differences in student achievement among the groups in the focus study. Then, in models of Algebra II achievement, there was a significant negative correlation between OTLAlg and student achievement. Further, when including multiple teacher-level predictors there were differences in student achievement among teachers in each group, these differences ranged from 3.22 to 8.44 points. A lot has been learned about teachers' implementation of *Core-Plus*, the influence of the NCIM professional development on teachers' implementation of *Core-Plus,* and how implementation and NCIM experience influenced student achievement, however these quantitative results raise many questions that were unable to be answered with these data alone. Using data gathered from teacher

interviews, classroom observations, and the instructional coach reports, this section will provide more insight into the quantitative findings by exploring decisions and actions teachers made when implementing *Core-Plus*.

Chapter 3 outlined that this study utilized a mixed methods approach, more specifically an embedded correlational model, to explain "the mechanisms that relate to the predictor and outcome variables" (Creswell & Clark, 2007, p. 85). Using mixed methods allowed for the analysis of the qualitative data and then interpreting them with the major quantitative results to form an integrated explanation, in this case, for teachers' implementation of *Core-Plus*. Recall, the emergent themes from the qualitative data came from a sequence of analyses of the teacher interviews, classroom observations, and the instructional coach reports; first for each individual teacher, then for each group of teachers, and finally across the four groups of teachers. This section is devoted to presenting the emergent themes from the qualitative data and then weaving these findings into the quantitative results that have already been presented in order to provide a more detailed account of teachers' implementation of *Core-Plus* based on their professional development experiences. First, evidence explaining the wide variance in teachers' implementation will be provided. Next, though student-level differences in achievement were not always prevalent, using information provided on teachers' implementation the variance, or lack thereof, on student outcomes will be explained. This chapter will conclude by linking evidence from each section with evidence on how varying levels of NCIM participation had an impact on teacher implementation and student outcomes.

Variance among the four groups' implementation indices has been well documented. Three broad themes were identified, providing factors that influenced decisions teachers made about textbook implementation among the four groups of teachers. The factors that had an impact on teacher implementation were: teacher beliefs about how students learn mathematics, teachers' trust of the curriculum, and systemic factors. Below, each category is organized to provide a brief overview of the theme, followed by an explanation of commonalities within each group, connecting to how this influenced teachers' implementation of *Core-Plus*, and concludes with summarizing major differences among the groups.

Teacher Beliefs About How Students Learn

Teachers' beliefs about how students learn mathematics influenced how they implemented the curriculum, including how they facilitated collaborative groups and presented content to their students. Recall, the belief factors (Table **35**) presented at the beginning of the chapter. The collaborative learning factor provides a lens to detail how teachers' beliefs about how students learn mathematics influenced their instruction. This section will first provide an overview for teachers' beliefs about collaborative learning and then examines how their beliefs about how students learn connected to their implementation of *Core-Plus.*

Averages for teachers' beliefs about collaborative learning $\left(\mu_A = -0.36, \mu_F = -0.94, \mu_B = -0.19, \mu_D = 0.34\right)$ showed that Group D teachers were the most open to a collaborative learning environment, whereas Group F had beliefs strongly opposed

to collaborative learning. Group A and B teachers had averages below the mean, but they were not as severe as Group F. Supporting evidence from instructional coach reports and the researcher's observations confirmed that Group F teachers did not often utilize groups in their instruction. Both Group F teachers had their desks arranged in rows instead of in groups at the start of each class. In an instructional coach report, Nora's instructional coach noted, "She appears to be a teacher who has taught from the front of the classroom by giving notes…She is trying to incorporate group work but I don't think she is very comfortable with it."

All Group A teachers, except for Maria, had their desks arranged in groups of four. Despite many attempts from her instructional coach to save instructional time by having desks in groups, the desks in Maria's classroom were arranged in rows with students sitting near their groups. The majority of Group B teachers had tables facing the front of the classroom and students worked in groups of two to four students. Group D teachers had a mix between desks in groups of four (3 teachers), rows (2 teachers), and pairs facing the front of the classroom (1 teacher). The arrangement of the desks was important because teachers with desks in rows or with tables facing the front of the classroom tended to provide direct instruction to students prior to the launch of the daily investigation. In several of these instances groups were only used after a teacher lectured or note taking took place. It appears that teachers who choose to arrange their desks in groups were able to facilitate the collaborative environment recommended by *Core-Plus* authors. The rest of this section will

detail how each group of teachers believes students learn mathematics, drawing on teachers use of collaborative groups.

Teachers in Group A tended to believe that students learn best by "seeing connections" and determining where the mathematics comes from. They think students learn through collaboration and noted the importance of having students verbalize their mathematical ideas so they have a deeper understanding of mathematical topics. Several teachers mentioned using multiple learning styles (i.e., visualizing, hearing, representing, and practicing) help students learn mathematics. This group of teachers was also concerned about the mathematics making sense to students and showing connections by relating to students' out-of-school experiences. For example, Maria was observed connecting coordinate methods to students' interest in video games and Luke was anxious to use Clicker technology in his class to get students more involved. Missy explained the importance of relating to students' experiences and having them work in groups:

> I am a proponent for integrated math because what worked for me when I was in school is just not going to work for these kids, because we are in this 21st essential standard and all this technology. I think if kids can get up and move and talk and discover the math themselves and then when they hit little snags I can be more of a facilitator and help them more when it is discovery based. I think that is the way they learn because a lot of time they are just parroting what the teacher has said and it still doesn't make any sense to them. So I think in groups, small groups, small setting, hands on.

The belief that students learn in groups and can discover the mathematics had an impact on the design of teachers' classrooms (seating arrangements) and their willingness to let students encounter roadblocks. During observations, it was common to see a student in Group A teachers' classes get out of their seat, ask another group a question and/or assist

another group with their question, before returning to his or her seat without disrupting the rest of the class. In the case of Luke's class, one student was repetitively called over by members of other groups because they had such confidence in his abilities to explain the mathematics they were discovering. Similarly, during Missy's classes, students engaged in mathematics together, and students were observed collaborating outside of their own group. Missy stated during her interview that to create a math community she encourages "them to get up and go ask her what she just did and come back and report to me. Or you go explain that to them." Students were getting out of their seats and discussing mathematics with other groups without the teacher's permission. She explained the importance of allowing students to collaborate across the room during her interview:

> If you noticed that in some of my classes some of my kids got up and moved to another group. They know they can do that because I am a firm believer that if you can explain something, then you understand it. Sometimes (because) their language they can say the same thing I am saying but because it is coming from one of their peers they get it a whole lot faster than they do from me. So I am more of a facilitator, and then if I interject, I try to interject with higher order thinking questions instead of just trying to give them the answer I let them struggle.

Missy believed that students learn best by explaining mathematics to other students. Her belief that collaborative groups help students articulate their ideas to understand mathematics, enabled her become a facilitator instead of an instructor. Also, because students are able to rely on their group members, Missy believed she could let students struggle with mathematical content instead of providing them with direct answers when they have questions. In practice, Group A teachers had varying levels of success in allowing students to verbalize their ideas. Similar to Missy, Maria created a classroom environment where

students were comfortable exchanging ideas. She was frequently observed stopping by groups of students, who were engaged in conversation or debating about a mathematical topic, just to listen to their conversations. Her ability not to "butt-in" allowed students to freely express their ideas and gave them ownership of the mathematics they were creating. On the other hand, two teachers were observed providing quick and direct answers to students as they walked by groups. In one case, the teacher answered student questions without providing an explanation and only addressed individuals within a group, ignoring the group as a whole. Allowing students to productively collaborate with other students around the room, outside of their own groups, was a unique feature observed only in Group A's classes.

Both of the Group F teachers believed that students learn mathematics best by receiving content directly from the teacher and undertaking extensive practice. In her interview, Heather said, "I don't like it (*Core-Plus*) overall so I don't feel really super comfortable teaching it. I don't think it is like the best method always for them to learn." Following her belief that students learn from the teacher, she was observed extracting the discovery from the curriculum. She provided extensive notes to her class, seated in rows, before they began working in their groups. It was important for her to provide students with direct instruction *before* they began to work through problems in the textbook, thus leaving nothing for students to discover. Similarly, Nora noted, "students best learn math if they understand the concept and then there is manipulation." By this, she expressed her belief that the students should fully understand the concept before attempting to practice the concept.

Nora concluded that, "with concept correct practice makes perfect." As a result, she lectured to her class before giving them practice worksheets (not *Core-Plus* materials). Her instructional coach noted her content delivery was a "traditional classroom with practice problems." Heather and Nora's beliefs about how students learn mathematics had an impact on the manner in which they implemented content from the textbook, typically through direct instruction followed by practice worksheets. Group F teachers' beliefs about learning mathematics did not align with the *Core-Plus* authors' vision, causing them to implement the curriculum with low fidelity.

As stated in their interviews, Group B teachers tended to believe that students learn mathematics best when it is tied to real world problems, helping them see the connections between mathematical relationships and their applications. For example, Ellie used a video clip of the creation of a Pixar film to motivate the need to learn about coordinate geometry. She tied the students' interests in movies to the creation of the computer animations on a coordinate axis. However, this was the only example witnessed in classroom observations from the Group B teachers. Their statements in the interviews seemed more of a belief about ideal ways students learn than what was actually observed in their practice.

Similar to the other Group B teachers, One teacher, Taylor, believed students learn best through problem-solving where they learn to reason about the mathematics, but he added the caveat, while "I think *Core-Plus* is best for more students, some students, I think, would be better off with traditional maybe. Maybe they just need that straight-forward, 'this is what you do (instruction)'." While he valued problem-solving and applications as a way to learn

mathematics, he also felt that some students learn better by direct instruction. With the exception of Taylor who admittedly provided "straight lecture" to students between days of note-taking and investigations, the Group B teachers were observed using the textbook during instruction.

Though the textbook was used to teach content, this did not ensure teachers' implementation was consistent with the authors' intentions. Teachers' beliefs about how students learn did not always correlate to their instructional practices. The next examples show the ways in which teacher practice contradicted their beliefs about learning mathematics. One Group B teacher, Beth, taught the vocabulary that was intended to be discovered by students in an investigation before the students had even started the investigation. In another class, Ben's Course 2 students were observed working on an investigation to understand how to solve systems of equations by elimination. Instead of presenting the problem students were going to investigate from the text, the teacher used a worksheet that simply had the system of equations listed. For example, Question 1a in the Course 2 text read, "What do the variables p and q represent in the given system, and why do the given equations accurately represent the problem conditions?" (Hirsch, et al., 2008, p. 54). However, the corresponding question on the worksheet read, "What does each equation represent?" Later, instead of having students derive equations to match a contextual situation, the worksheet provided the equations and merely had the students solve to get the answer. In my field notes, I observed:

> Students begin working through the packet. Their department has created these packets to help students work through the problems in the book… I noticed that

> hardly any students did the reading…He did encourage them to read, but the packet, in my opinion, took their thinking out of the work and made the steps to elimination more explicit than the book.

This packet decontextualized the content from the textbook and watered-down the intended curriculum. This example showed that supplemental material does not always consist of extra practice worksheets and can become entwined with instructional practices to drastically change the investigative approach of the curriculum. Examples from Beth and Ben highlight, that while teachers of Group B believed students learn best when they engage in application, they also stripped the discovery nature of the written curriculum. In other words, teachers used the context of the problems in the textbook, but they did not allow students to experience the discovery of mathematical ideas like the textbook authors intended. However, their beliefs about how students learn mathematics made them more willing to use the textbook than teachers in Group F.

Five of the six Group B teachers spoke to the instructional coach, Margaret, about difficulties they have with collaborative grouping. Their discussions ranged from concerns about the number of students in each group to problems facilitating productive collaborations among group members. In one case, Margaret pointed out a teacher's difficulty with collaboration, though he did not realize it was an area for improvement:

> Teacher stated that he does not need any assistance…From my observations, however, this teacher needs support on collaborative group work and classroom management. The students do not demonstrate that they have been trained to work productively in groups. The teacher circulates among the groups as they work but does not use this time to help the students figure things out but rather to tell them how to do it. When the class pulls together for a whole class discussion, student input is limited and the teacher is by far the biggest source of content and explanation.

Group B teachers were utilizing collaborative techniques, but still show some discomfort integrating this into their instruction.

All of the Group D teachers said that students best learn mathematics by doing it and figuring out for themselves. As an added caveat, there was one teacher who also suggested that some students needed to be shown the traditional way. However, as a whole Group D teachers felt similar to Kasey, who believed students learn best "from doing it and talking about it. Not from watching it and trying to recreate it." They felt that hands-on investigations were a great way for students to figure out the mathematics for themselves. Five of the six implemented the first edition of the textbook during the 2009-2010 school year, and believed the second edition provided more hands-on activities and practice problems. Several of these teachers reported supplementing hands-on activities into the first edition to keep the students engaged in class. Josie discussed the importance of the hands-on activities for student learning, "They don't have a whole lot of hands-on investigations and that is not good for them. They need some more hands on stuff. But I think hands on and technology, anyway at all to help them learn."

Kasey facilitated discovery and communication in her classroom by having daily process goals to encourage collaboration among students within each group. Margaret noted, "Teacher (Kasey) circulates to encourage and ask probing questions but does not serve as a source of information," which provided students with the opportunity to discover the mathematics with their groups. Delores created an environment where students created mathematical ideas by constantly encouraging alternative approaches to the same problem

and by having students consider incorrect answers. To create this environment, she frequently asked for alternative solutions and had students brainstorm to link similarities and differences among different ways of solving a problem. Then, if a student provided an incorrect answer, she would neither dismiss the answer, nor tell the student they were incorrect. Rather, she would have students debate possible answer choices until they agreed on the solution, or she would prompt them with another problem, asking them to utilize their method, to test the validity of their approach for a different circumstance.

In interviews, Group D teachers noted the importance of allowing students to struggle through the investigations, and this belief had an impact on when and how they intervened to assist students. With the exception of the two youngest teachers, Calley and Shauna, in practice Group D teachers were observed providing students with time to struggle. Shauna was observed teaching a lesson using *CPMP-Tools* involving special segments of a triangle. It was frequently documented that when a student would call her over for help, she would take their computer mouse and perform the desired construction without questioning the student on their conceptions. She quickly intervened to take control of their computers without providing them time to conceive of the mathematical relationships displayed on the computer. Similarly, when a student or group ran into difficulties, Holly was observed imposing her way of solving on students. This was documented in my field-notes with my initial reactions in italics:

> Calley walks by group 6 and student #1 asks, "Can you put that in the calculator to find the center?" Calley replied, "you guys actually know the center already, go back and look at number 2…" *She totally gave them directions for where to look instead of using the answer the student gave as the starting point for the discussion.* Calley

> asked them, "So how do you know the center?" *which they obviously don't yet.* #1 replies, "use the distance formula." Calley responds, "No you should be able to tell by the equation." *Still trying to impose her way when they don't see what she would like them to see.*

In their interviews, Calley and Shauna discussed the importance of providing students with time to develop their ideas. However, observations reports showed they still need practice developing this skill.

Group D's most experienced teachers showed more patience than any other group in providing time for students to struggle through the investigations. For example, students called Delores over because they did not understand a problem about deriving the midpoint formula; she reminded them to think back to the previous class. One student gave a possible (but incorrect) answer and she merely said, "there you go, think along those lines" as she walked away. This sparked a heated debate among the three students in this group. Delores allowed them to struggle through the derivation, which led to their creation of two separate agreed upon midpoint formulas[17], $\frac{(A,B)+(C,D)}{2}$ or $\frac{(A,B)-(C,D)}{2}+(C,D)$. Similarly, Kristin often addressed a group's question by rephrasing the question and then giving them more time to ponder possible solutions. Only if she noticed the majority of groups were stuck on the same problem did she bring the whole class back together to have a student lead discussion about their confusion. In practice, experienced Group D teachers' beliefs, that

---

[17] They did not empirically prove this in class, but a possible proof is: $\frac{(A,B)-(C,D)}{2}+(C,D)=$
$\frac{(A,B)-(C,D)}{2}+\frac{2(C,D)}{2}=\frac{(A,B)-(C,D)+2(C,D)}{2}=\frac{(A,B)+(C,D)}{2}$

students learn mathematics by doing and figuring it out, was observed in their ability to provide students with time to develop ideas before resorting to whole class discussions or prompting them with a directed answer.

Group A and B teachers had similar beliefs about how students learn mathematics, but observations and instructional coach reports pointed to differences in how their beliefs shaped how the curriculum was enacted in their classrooms. Group F teachers' beliefs had the least alignment with other groups and contributed to their poor implementation of the curricular materials. Teachers in Group D had beliefs about how students learn mathematics that were aligned to the nature of the *Core-Plus* textbook.

Teachers' Trust of the Curriculum

Teachers held a variety of attitudes about how the *Core-Plus* curriculum is structured, sequenced, and organized and influenced their implementation of the textbook. This section will discuss how teachers' trust of the curriculum shaped their implementation of *Core-Plus.* Recall, content in *Core-Plus* maybe be introduced in one unit and then refined later in an entirely different chapter. Teachers' trust of *Core-Plus* would often determine how they handled the gradual refinements of key concepts overtime rather than to rush in and overteach a concept on its first introduction.

Group A's overall positive trust in the curriculum led to their high implementation of the textbook content, despite a few examples of low presentation fidelity. Missy, the most experienced *Core-Plus* teacher in Group A, understood the looping nature of the curriculum:

> I think it is because there is so much within the *Core-Plus* course, how they go back and loop to something they have done before. And it just keeps going back spiraling

> and looping things and brining it back, so that if they didn't get it the first time you go, "okay I am going to give it to you one more time in a different world problem." So they can see it, they can apply it and stuff.

However, Chloe's (Group A) instructional coach noted the difficulties she was initially having trusting the curriculum in one of her reports:

> She is still trying to figure out the pacing and trying to make herself trust the curriculum. I think maybe we should spend some time going over the scope and sequence in the summer workshops so teachers will realize that "it" will be taught at a later time

Even after workshop exposure, teachers did not often understand the spiraling nature of the curriculum until after they taught it. Group A teacher, Chloe, reflected on her first year experiences with the curriculum:

> Everything just builds on itself. That is one thing I learned last year is, not going too far into it, and I didn't realize it was going to pick it back up in the next course. But it is just neat the way it builds and even in the investigations everything, questions you can't answer question 4 without doing 1, 2, and 3.

It was not until she was able to see the entire *Core-Plus* sequence, and understand that concepts are built over time, that she was able to trust the curriculum enough not to supplement extra content the first time material was presented. Similarly, Nicole's instructional coach reported, "Teacher does not trust the investigations to teach the material, rather she tried to pre-teach the material before assigning the investigation." Other Group A teachers noted the importance of trusting the curriculum and not "over teaching" it, to preserve the discovery nature of the written curriculum. The workshops increased their tendency to trust the curriculum, but experience teaching with *Core-Plus* was also critical to shaping their beliefs.

It was obvious through the interviews, instructional coach reports, and observations that the Group F teachers, Heather and Nora, did not like, and were not comfortable with, the curriculum. As a result they rarely provided students with the opportunity to learn content from the textbook and typically presented content in the textbook with low fidelity. Heather's instructional coached noted "she still needs reinforcement that Core Plus works."

Without exposure to the curriculum before implementing it, Heather and Nora did not understand the reoccurrence of ideas in the textbook or how mathematical notation is built throughout the sequence. For example, Heather noted her frustration with the curriculum:

> It (*Core-Plus* textbook) does everything kind of backwards like for them discovering. And it gets really confusing and I don't always understand what their objectives are because they put it in Core Plus terms instead of math terms. I mean like it's great once you really know the curriculum, I guess, but it is hard to really know…I first try to understand what it is doing.

Since Heather was not comfortable with the way ideas in the textbook were developed she often used supplemental material. Her instructional coach reported, "it drives me crazy that she doesn't use the book enough." To ensure a topic was fully taught the first time it was introduced in the textbook, Heather often resorted to replacing the textbook with teacher notes or practice worksheets. A consequence of Group F teachers' negative beliefs towards the curriculum was their low presentation fidelity of textbook content.

Similar to Group A, Group B teachers also had high implementation indices and share many of the same beliefs about the curriculum as Group A teachers. Overall, Group B teachers trusted the curriculum to provide students with a rich foundation for the interconnected topics in mathematics. They believe *Core-Plus* is a curriculum for all

students, is rich in application, and allows students to apply the knowledge they have to new situations. Taylor trusted the textbook provided rich problems for student discovery:

> I think it's a framework and it's better than any of the other textbooks. It's a lot easier for me to sit here and come up with 10 one-step equations then it is to sit here and think of some step-by-step way of getting a kid to realize what they are doing…It's a lot easier to supplement traditional stuff than it is to do the discovery based stuff.

Steve, saw the benefit of trusting the curriculum to provide students with a strong mathematical foundation, regardless of student ability. He believed, "you can go way down in it and come out with a great understanding of how these concepts have developed." Similarly, Beth, who attributed her trust for the curriculum to the NCIM summer workshops, believed in the curriculum because it helps students build relationships among topics. She stated:

> Nothing is ever done in isolation and by teaching Algebra and Geometry separately the kids never really got the relationships, and now, I have kids writing proofs that are really elegant and when I taught Geometry it was pulling teeth to get them to write proofs.

Further, Ellie trusts *Core-Plus* creates mathematically proficient students by facilitating the use of multiple representations. Ellie stated one of the benefits to using *Core-Plus* is that her class can "look at the graph the table the word problem the symbolic rule all at the same time and it is not in chucks. We look at the whole thing all together so they can see the whole relationship." In practice, some teachers were more comfortable integrating multiple representations into their lessons than others. Margaret witnessed Steve relying too heavily on the written equation for learning about exponentials and suggested "it is really helpful if you can help them see what an exponential should look like in the table, graph AND

equation, not just the equation. The opening activity would be a great place to emphasize all three."

Ben stated the NCIM trainings helped him "see how it works" so that he could understand the curriculum. However, as Margaret pointed out Ben's classes often began with a "general introduction to topic given by teacher, including pre-teaching of the content before book is opened." He was also observed expanding on concepts after the investigation had been completed. His actions show a lack of recognition, confidence, or support that the concept will be fully developed throughout the sequence. Recall from the previous section, that pre-teaching content was observed in several of the Group B teachers' classes. Ellie admitted, "last year I would anticipate that they were going to have trouble with questions, and I would, um, explain and explain and explain before they started. This year, I think, I am more willing to let them flounder." Experience at the NCIM summer workshop and her experience teaching *Core-Plus* the previous year, combined to give her more comfort with using the curriculum. This year, she was not observed pre-teaching content or over explaining ideas before students started an investigation. These examples highlighted that Group B teachers have strong positive beliefs about the use of the *Core-Plus* curriculum. However, even among teachers who attended the summer workshops, observations highlighted variance in how teachers' trust of the curriculum had an impact on their implementation.

Teaching *Core-Plus* an average of 7.17 years, Group D teachers were very supportive of the curriculum. Several teachers noted they find using these curricular materials harder to

teach than subject-specific texts because it takes more time to plan, more time for student ideas to develop, and they have to be comfortable giving the authority to the students, which may result in them heading in the wrong mathematical direction. Previous examples showed Group D teachers provided students with authority for their learning and were more patient allowing students to develop their own ideas.

Interviews depict high levels of Group D teachers' trust of the curriculum. They realized the care that must be taken when implementing the curriculum so that they do not go beyond the current development of a concept. Kristin summarized:

> Its [*Core-Plus* text] very rich and that is something you have to be careful of when you are teaching it and it can go beyond where you really want your students to go. I mean there is a point where "Okay, you've got enough right now. You need to kinda let that settle for a little while." It is very rich in mathematics and there's a lot there.

Since she was able to distinguish the refinements of concepts overtime, she does not overexpose students to concepts when they are first introduced.

Teachers in Group D also recognized this level of trust is difficult for teachers who have not been exposed to the curriculum. Excerpts from teacher interviews with Josie and Kristin highlighted the importance of teachers being exposed to content in the textbook prior to teaching it so that they learn to trust the spiraling nature of the curriculum. Josie stated,

> It takes getting used to, to go through the material. And it is hard to explain to somebody who hasn't taught it before that yes they are just covering slope right now but they will come back to it later. That is probably the biggest thing for people to get over is you didn't spend a month teaching slope and minor sets, but they do come back to it later on and they revisit it again and again and again. I think that is a good thing.

Kristin stated,

> One of the difficulties you have is if you skip an investigation because the next investigation may depend on that. You have to be very aware of the whole scope of the course. And that is the reason, I fully, I tell every teacher they need to go to a workshop where you cover the book from page one to the last page. You know, you may not work every problem but you see where things are, make connections, because the spiral curriculum.

The trust Josie and Kristin discussed was similar to what Chloe experienced during her first year of teaching and what the Group F teachers had not yet discovered. Even Delores, from Group D, admitted to showing students certain formulas her first year implementing *Core-Plus.* She did not believe ideas were formalized for students and wanted to move along quickly. Overtime she has begun to trust the curriculum and provides her students with time for concepts to develop. Similarly, Kasey explained her hesitation her first year implementing:

> The first time we taught a unit we would look through it and wonder "are kids really gonna get this?" I mean we see where they [textbook] are going but is it really going to happen this way. And at the end of each unit, it never failed, we were like "that was brilliant." Kids really got it and I am so glad that I stuck to it. Because there were several times when we were like "we could knocked this out in a day if we just taught it to them." and then we were like "well, maybe if we trust the curriculum and do it." And every time we trusted the curriculum it worked to our benefit.

Teachers' experiences with the curriculum helped them see the spiraling nature of the curriculum, which improved their trust of the textbook materials and subsequently how they implemented the materials. Group D teachers showed the most trust of the curriculum. Group A teachers' trust of the curriculum developed as a result of the workshop and their experiences implementing the curriculum. Teachers implementing for the first time during the 2009-2010 school year supplemented more content than teachers who had prior experiences using the curriculum for instruction. There was no clear consensus for Group B

teachers' trust of the curriculum based on interview data, yet it was clear that Group F teachers neither understood nor trusted the spiraling nature of the textbook. This resulted in Group F teachers over-teaching content and supplementing content more frequently than other groups.

Systemic Factors

In addition to teacher beliefs about how students learn mathematics and their trust for the curriculum, systemic school and/or district wide policy mandates affordances and constraints that had an impact on the manner in which teachers in the focus group implemented textbook content. Four factors that contributed to variance in teachers' implementation were: (1) End-of-Course assessments and countywide benchmarking, (2) access to materials for instruction, (3) scheduling, and (4) student transition to reform mathematics. Each of these will be discussed below.

End-Of-Course Assessments and Benchmarking. Across all four groups, teachers consistently skipped the discrete mathematics and probability and statistics units of the textbook. Interview data determined they omitted these sections because of the constraints imposed by the fact that these topics were not tested on the NC EOC assessments. A typical statement about skipping these units was noted in an interview with Josie when she said, "We know we gotta give the Algebra II test at the end of the year in Integrated 3. So we are skipping some of the discrete and statistics." In addition, when asked how she decides which investigations to teach and which ones to skip, Ellie stated, "I'll be honest, it is based on the

Algebra I Standard Course of Study." Teachers seemed frustrated that testing influenced the content they felt they had to omit. Missy pointed out:

> And I wish I could teach a *Core-Plus* book from the beginning of the book of the end of the book. Instead of chopping it up and dicing it in order to get them ready for the state test. I wish I could teach every chapter in 1 then teach every chapter in 2…we have to skip this chapter because your kids are not being taught statistics on the algebra 1 so that is what the state said.

Teachers frequently supplemented materials to prepare their students for mandated assessments. Luke used worksheets because he believed, "EOC review is just drill and kill." In addition, they supplemented to get students familiar with the format of the EOC. Missy pointed this out when she said that she supplements to "get them used to that format because if you noticed they rarely have multiple choice test in intergraded math. So I gotta get them used to that multiple choice format." Teachers in each group admittedly supplemented content from the textbook with traditional practice worksheets. For example, Delores pointed out she used worksheets for extra practice, "Maybe they are still having trouble with the symbol manipulations so I'll use um, we have a worksheet generator to just generate some symbol manipulation stuff." There were differences, however, among groups in the manner supplemental content was used to prepare students for these assessments.

Group A justified supplementing in worksheets by declaring this was to prepare students for the EOC assessment and typically occurred at the end of the year or as an out-of-class review assignment. Missy from Group A noted, "I never bring anything in to help them grasp the concept I just bring in stuff for repetition or to go over it again to make sure you really have it." Group F teachers supplemented content in the *Core-Plus* textbook with

material from the subject-specific Glencoe textbook and worksheets almost daily. The main difference between Group F's supplements and the other groups was that Group F supplemented content, whereas the other groups supplemented for practice.

Group B teachers also used worksheets for practice but the majority indicated it was necessary because of county benchmarking requirements or due to mandatory tutoring sessions built into the school day to prepare students for the EOC. This did not hinder these teachers' implementation of the textbook, but it is important to note that it did allow their students more time to practice traditional problems and multiple-choice assignments; practice aligned to the outcome measure available for use in this study. These teachers supplemented practice more regularly throughout the year, at times due to administrative decisions. For example, two teachers had time built into their school day for tutoring or mandatory practice. Steve's schedule included lengthening the school day to remove Friday as an instructional day. As a result, students were required to spend an extra 30-minutes a day on review for EOC subjects. When Steve was asked how he used his review time, he responded:

> We go back to straight Algebra 1 EOC where we work, review, do all of this. I mean the county says we have to benchmark and do some of that…In class, nothing. On the outside yea I do stuff to get ready, only for the Algebra 1 EOC because it is real, it is a different from free response to multiple choice and all of that.

Similar, Ellie's school built in tutoring time everyday. Two other teachers noted they supplemented with practice, one throughout the year in class and one used traditional mathematics problems for homework. The exception to supplementing traditional practice problems was Beth who admitted her tests were harder than the EOC. During her observation, she told her students, "You have learned to read and think and take things step-

by-step and work your way through it. As long as you do what you have been doing, (the EOC is a) walk in the park." In the case of Steven and Ellie, the structure at the school necessitated traditional practice problems to prepare students for the EOC, and as a whole Group B teachers supplemented de-contextualized practice problems throughout the school year outside of regular instruction in order to prepare students for the EOC.

Similar to the other groups, Group D supplemented with worksheets designed to improve students symbol manipulation. The difference was in the frequency of the practice problems during the 2009-2010 school year. Teachers using the first edition noted frequent use of worksheets and quiz generators throughout the school year. Group A typically supplemented at the end of the year and Group B tended to supplement outside of instructional time. However, Group D teachers supplemented practice worksheets into their daily schedules, as did Group F but their supplements were for content as well as practice. The Group D teacher using the second edition materials supplemented traditional practice worksheets weekly during the support class she taught to all students enrolled in integrated mathematics.

Access to Materials for Instruction. Teachers expressed concerns over the lack of materials provided for instruction. The investigative nature of *Core-Plus* requires that teachers have access to materials (e.g. balls, ramps, tape measures, stop watches, coins, etc.) to fully implement the textbook investigations. In addition, one of the instructional features of *Core-Plus* is the integration of technology throughout the sequence. *CPMP-Tools*, a computer software program that can be downloaded free of charge, and graphing calculators

are frequently incorporated into the *Core-Plus* textbook. Some of the Group A and B teachers used *CPMP-Tools*, however Groups F and D had not, for reasons provided below. Typically supplements act in conjunction with, or in addition to, textbook material, however teachers were observed replacing investigations and technology lessons with supplemental material. In other words, there were instances when textbook content was not used for instruction and teachers replaced these units altogether with instruction that did not require the use of materials or technology. There was little consensus within each group for the frequency with which individual teachers used computer technology.

Teachers in Groups A and F had the most technology built into their classrooms. Each of these teachers, except one, had classrooms equipped with either a Promethean or Smart Board attached to a computer. Instructional coaches reported Group A teachers frequently used these to project the textbook, display timers for keeping students on pace, and announce daily objectives and assignments. Principals of schools in the New Schools project ensured that teachers had access to technology in their classrooms, though these were often used more for a classroom management tool rather than an instructional tool to present mathematical content.

Group A frequently asked their instructional coaches for technology assistance, both on the computer and with the calculator. Group A teachers did not report any issues securing non-technology materials for investigations (ramps, balls, coins, etc). However two of them expressed concerns for the lack of student access to computer carts or labs at their schools, while the other teachers had easily accessible computer labs or worked in a 1:1 environment.

Teachers in the latter group did not have to supplement content during the lessons that incorporate technology, whereas the former teachers resorted to skipping these lessons altogether or presenting the content to the class using one teacher demonstration computer. Teachers reported using applets in *CPMP-Tools* and the slider functions of the CAS software during teacher-led demonstrations. During the only observation I witnessed involving *CPMP-Tools*, Maria's students used individual computers in their groups in a state-of-the-art computer lab to learn about coordinate geometry. It was clear that she was comfortable implementing the technology components of the textbook and that the students were highly engaged in the lesson. The accessibility Group A teachers had to technology was unparalleled by any other group. This enhanced their instruction regardless of whether it was used as a teacher demonstration, student investigation, or as a classroom management tool.

In their first year of teaching *Core-Plus,* without attending the workshops, the two Group F teachers had not used *CPMP-Tools* in their instruction. The instructional coach frequently observed them using calculators for instruction, and towards the end of their first year teaching *Core-Plus* they had both begun using a computer connected to a data projector, mainly to display a timer to keep the class on pace throughout the class period, but had not used technology as an instructional tool.

Only two of the six Group B teachers, Ellie and Jenni, had attempted to use *CPMP-Tools* in their instruction. Other teachers in this group mentioned they could not get access to the technology when they needed it or they had not explored the technology enough to feel comfortable with it. Of the two that had used *CPMP-Tools* in their instruction, Ellie works at

a school that provides each student with a computer and calculator and one that requires her to show proof she uses technology. I observed her taking pictures of the students using *CPMP-Tools* so that she could use it for her professional portfolio. The other teacher, Jenni, mentioned that it was difficult to get access to enough computers for her students to use. They often resorted to one teacher demonstration computer or groups of four students using one computer. Margaret noted in her observation report that Jenni "stated that she was still learning with technology, but it was clear that her students use it regularly and are comfortable with it." Margaret also pointed out that the use of *CPMP-Tools* in a lesson on regression provided Jenni's students with a better understanding of influence points and outliers because they were able to visualize how things change when data points were removed from the data set. Ellie and Jenni were fortunate to have access to a computer or a class set of computers for their instruction, allowing them to properly implement investigations that incorporated *CPMP-Tools.*

Group B teachers that had not used *CPMP-Tools* before this year were beginning to phase it into their lessons despite the accessibility constraints. During the 2009-2010 school year they had a difficult time reserving computers or had not practiced enough on their own. Beth believes the most difficult lessons to teach were the ones that required access to computers; she felt she had to supplement these lessons. Beth said, "I really want some more experience with *CPMP-Tools,* and I just need the time to sit down and play with it…a lot of times there's one problem in an entire investigation and getting a lab in here to do one problem is hard." Similarly, Ben mentioned the most difficult lessons to implement were the

ones that used applets created for *CPMP-Tools.* He has begun to use the software, but was uncomfortable last year because he had never used it with a class before. Taylor has not used it because of access issues. After meeting with Taylor, Margaret reported, "Teacher stated he needs more technology. He does not have a sufficient supply in his room of calculators and has to share with another teacher. Teacher stated that having a more permanent set up of computers would be beneficial." I observed Taylor running between several other teachers' classes at the beginning of his class to find enough calculators for his students to use. Overall, Group B access to computers and calculators has hindered the manner in which the technology units in the text are implemented, if taught at all.

The five Group D teachers using the first edition of the textbook during the 2009-2010 school year had not used *CPMP-Tools* because it was not incorporated into the first edition textbook. They did report supplementing in technology such as Geometer's Sketchpad or Excel on occasion. The other Group D teacher, using the 2nd edition textbook, noted that it was a big ordeal to get the laptop carts with the correct software for them to use. This constraint forced her to either utilize a teacher-led demonstration of *CPMP-Tools* or skip the technology aspect altogether.

The difficulties teachers faced getting access to technology and their comfort with the tools impeded their ability to faithfully implement the technology units in the textbook. This was especially true for Group F who did not attempt to use *CPMP-Tools* during their instruction. Group A and B teachers had some training on the software program at the NCIM

summer workshops but had a difficult time implementing these lessons because of access issues and the lack of practice with the tool.

Scheduling. Several focus Groups had unique scheduling opportunities for students. Three scheduling types were identified that influenced the time teachers had to implement the curriculum: (1) addition of a mathematics course, (2) Early College campuses, and (3) student transition to reform mathematics. These schedules may explain some of the variance in implementation indices between certain groups of teachers. Teachers within each group with unique scheduling (Table **49**) will be described, connecting to how that influenced their textbook implementation.

Table 49. Number of Teachers in Each Group with Unique Scheduling

	Additional Mathematics Course	Early College Campus	Reform Mathematics in Middle School
Group A	1 of 7	0 of 7	0 of 7
Group F	2 of 2	0 of 2	0 of 2
Group B	0 of 6	4 of 6	1 of 6
Group D	0 of 6	1 of 6	6 of 6

Addition of a mathematics course. One of the schools began a schedule that allowed students to take four semesters of mathematics before taking the Algebra II EOC exam. This provided teachers with more time to cover the content in the textbook before the exam, leading to a more comprehensive implementation of textbook content. Both teachers from Group F and one teacher from Group A worked at a school that added a discrete mathematics course between Courses 2 and 3. This provided students with four semesters of mathematics before taking the Algebra II EOC, whereas all other teachers had students taking the Algebra II EOC after exposure to the first three *Core-Plus* courses. Maria's instructional coach

documented her concerns, as well as Maria's, and the benefits of adding this additional course:

> I am worried that she is feeling pressure from her principal about her students being successful on the EOC. She is also very concerned that she does not have time to finish all of the units that she needs to. I think having the discrete course next semester will help alleviate some of this stress.

By incorporating this additional course, these teachers could cover more of the textbook content in the series and could spread the topics in Courses 2 and 3 out over three semesters. This could explain the low average OTL index for Group F teachers (33.80), but the low TCT indices and classroom observations still indicated these teachers were not using the textbook for the primary source of instruction. Upon further analysis of the TOC-logs, these three teachers had Course 3 OTL indices at 39.70, 29.41, and 36.76, remarkably lower than the means for Groups A, B, and D (Table 36). The Group A teacher, Maria, had the lowest OTL index for all of the Group A teachers and this was her lowest OTL index for Courses 1 thru 3 over the past two years. This provides evidence that the lower implementation indices for these three teachers could be a result of the additional mathematics course. However, the Group F teachers had low OTL indices for the other courses they taught as well: Course 1- 41.43, Course 2- 27.69 and 30.77, suggesting they have a low average OTL index regardless of the course they teach. Keep in mind Group F teachers were implementing for the first time, so it is difficult to state the causal effect of their low implementation indices.

Early College campuses. Several of the teachers taught at Early Colleges, on local community college campuses. This allowed these schools to be more flexible with students'

scheduling. Five of the teachers taught at Early Colleges, four of these were from Group B and one from Group D. Recall from earlier, that three teachers worked at schools with time built in daily for EOC exam review and extra tutoring. Each of these three teachers worked at an Early College high school. Early College schedules allowed teachers to provide targeted EOC practice to students outside of regular instruction to prepare them for the state assessments.

The other two (of the five) teachers worked at a school that had all gender-specific freshman year classes. As a result, all freshmen take Courses 1 and 2 in classrooms with either all females or all males. Beth discussed the benefit, especially for females, of having freshmen in gender-specific courses:

> And all the sudden we've got girls asking, girls arguing with boys, when before if we had put them together as freshman I don't think the girls would have stepped up as much. I don't think there would be the level of engagement that there is because the girls confidence level is a whole lot higher because they had that year to get the foundation and its really been an added bonus to brining *Core-Plus* in.

Jenni attributed this schedule to female students taking a more active role in her classroom. While gender-specific classes did not directly impact teachers' implementation of the materials, it helped the teachers more easily transition students to a reform-oriented environment. During the classroom observation female students were seen questioning one another, as documented in the researchers field-notes, "The students are questioning one another and using each other for the source of ideas." Both teachers believed this unique scheduling helped female students become more confident and argumentative in mathematics.

Student transition to reform mathematics. Struggles transforming to reform mathematics curricula are not unique to teachers, there is also an adjustment period for students as well. The curricula students were exposed to in middle school influences the length of time it takes students to adjust to *Core-Plus.* Teachers indicated students have to adjust to the reading in mathematics, working in a collaborative environment, and arriving at answers multiple ways when they are first exposed to *Core-Plus*. Each of these adjustments affect decisions teachers made about how to implement the curriculum. The adjustment period hinders teachers' abilities to implement the curriculum at the beginning of the year. Group A and B teachers believe the adjustment period only occurs during Course 1 and lasts between four to eight weeks.

One school district offered reform mathematics to middle school students, decreasing the training necessary to transition students to a *Core-Plus* classroom. According to teacher interviews, students exposed to reform mathematics curricula in middle school have an easier transition to *Core-Plus* than students from traditional mathematics backgrounds. Across their district, students of teachers in Group D were exposed to reform mathematics in middle school. According to teachers, this exposure decreases the adjustment period for students as they transition into *Core-Plus* classrooms. Rarely did students from the other groups have experiences with reform mathematics prior to high school, however in one district it was possible for advanced middle school students to take *Core-Plus* in middle school. Missy explained the differences in the transition period for students based on their middle school experiences:

> You can tell the kids that did integrated 1 in the 8th grade when they come to high school to do integrated 2 it is just a piece of cake. You know, but the kids who come into high school who have never had integrated and they have to go into it, they are looking for worksheets, that regurgitation where Mrs. Missy you do it and I do 15 like that. And it's not like that. I try to make my parents understand you have to give them at least 6-8 weeks to get used to it. You have to because the units and stuff are just different and they are like a fish out of the water you got to give them time.

One Group D teacher pointed out that the transition as become easier since the feeder middle school transitioned to reform mathematics:

> Its not as terrible as it used to be because our middle school has been *Connected Math* and has really been on top of that the past year or two but like the first year or two there was like a huge transitions they were like "what do you mean I have to think about it?" " Just tell me what to do" was constantly the thing.

As a result students in Group D come in with more experience with reform mathematics practices. Highlighted below, teachers in other groups have to devote time to training students to become comfortable with the new instructional practices.

When asked how her students responded to *Core-Plus,* Group F teacher Heather stated, "What scares me the most is that they don't really respond. I mean they don't even talk that it is different or they like it better or that they like it worse. They just kind of go." The bulk of the transition lies in Course 1, and one plausible explanation for Heather not having a difficulty with students' transition to *Core-Plus* could be due to her not teaching Course 1. As her instructional coach pointed out, "She has pretty good kids that Maria has trained...Hopefully the best kids will step up and set a good example for the rest of the class." Her instructional coach believed that the Course 1 teacher, Maria, had successfully used reform instructional practices to transition students to *Core-Plus.* An alternative reason could be that Heather utilized traditional practices, which the students were used to prior to Course

1. Neither Group F teacher taught Course 1 so they will be removed from the discussion about students' transition to *Core-Plus.*

Group A and B teachers believed training students to read and work in collaborative groups were the most difficult transitions to reform mathematics. Of the transitions, training students to read in mathematics had the most significant impact on the manner in which teachers' implemented textbook content. Stated simply, Ben noted, "A lot of them (students) have to get used to the idea of actually reading a math book." Teachers dealt with these reading difficulties in various ways. I observed several of them taking the authority of reading away from the students. Instead, teachers would read to students, rephrasing the questions to make the content easier to interpret. One Group B teacher stated he did this to "save time." Margaret noted the difficulties several Group B teachers had with reading levels in her reports:

- Teacher (Taylor) stated that students have a hard time reading the investigations, especially his IEP kids. They also have issues with organization and staying on task. These are all issues that make working with CORE-Plus difficult.
- Teacher (Ellie) stated that she needed more reading support…The challenge the teacher is having the hardest time with is that the students don't understand what the questions in the investigations are asking. Constantly fighting the students on this is exhausting. We discussed how the first year the students are in this curriculum is especially challenging but that if you put in the time it gets easier.

- Teacher (Steve) and his colleagues struggle with where to place students coming from non-integrated math places and dealing with poor reading skills of students

Margaret's documentation of Group B teachers' difficulties with students' reading abilities highlights a unique challenge, not observed in the classes of other groups. Though Group A teachers mentioned there was an adjustment period for students learning to read in mathematics, the instructional coaches did not record this as challenge area for any Group A teacher. Group A teachers utilized strategies to ease the transition students faced reading in mathematics. I observed Missy interacting with groups that changed readers in the middle of the class because they thought the reading was too hard. She would not allow group roles to change due to the reading level. Instead, she enforced the group roles and would make the same person read for the entire period. If the reader had a difficult time, she would have him or her read a sentence, stop, and interpret what they had just read. She would also periodically stop the reader to ensure that everyone in the group understood the passage. This tactic was one way that she scaffolds the reading for students while not losing the context of the mathematics. Chloe pointed to the tough time students have reading and interpreting the textbook, but she constantly encouraged them to use their groups and to not get caught up in finding the exact answer. Similarly, I observed some of Maria's students struggle with reading aloud to the class. She offered words of encouragement to get them to continue reading.

The distinction between how Group A and B teachers handled student difficulties reading textbook content was subtle, yet changed in the manner in which textbook content

was presented and analyzed by students. Group B teachers lower the presentation fidelity of the text by reading for students or interpreting what a question was asking before providing students time to think about the problem. Group A teachers experienced some of the same difficulties transitioning students to reading in mathematics, but have a wider range of instructional tools they employed to keep the authority of reading on the students.

Summary of Differences in Implementation Among Groups

Teachers' implementation of textbook content was influenced by their beliefs about how students learn mathematics, trust of the curriculum, and systemic factors of the educational system. Variance among the groups' implementation indices, reported in the first section, can be attributed to the factors discussed above. Teachers supplemented textbook material for a host of different reasons, most notably due to their beliefs about how students learn, their comfort with the curriculum, and the confines of the EOC assessments.

Groups A and B had higher textbook implementation indices than groups that did not have teachers at the NCIM summer workshops. This indicates that NCIM workshop participants used the textbook as the primary source of instruction more frequently than teachers not involved in the workshops. However, this did not ensure teachers' instructional practices were aligned with reform-oriented practices, merely that teachers in Groups A and B used the textbook, rather than supplements, as the main source for their instruction. Through the follow-up support of their instructional coaches, qualitative data showed that Group A teachers were becoming more comfortable with reform-based practices and utilizing these practices more frequently than Groups F and B. Group B teachers could

benefit from additional individual support. Explanations for the low indices for Group F teachers include that it was the first time either teacher was implementing *Core-Plus* and that the addition of the discrete mathematics course provided teachers with more time to teach the content, lowering their implementation. However, qualitative data highlighted the deliberate use of traditional instructional practices these two teachers used in their classrooms. Group D teachers, whose beliefs were highly aligned with reform-oriented practices, also had lower textbook implementation indices than Groups A and B. One plausible explanation for this was their use of the first edition of the textbook, which did not incorporate much technology and lacked practice problems. Further, they have taught using CPMP materials for the longest period of time, which has given them more autonomy to use alternative sources. Typically, the teachers using the first edition supplemented technology and hands-on activities into their instruction. It should be noted that unlike other groups, they supplemented activities that were consistent with reform-based practices. Qualitative data indicated that Group D teachers were very comfortable with the curriculum and they frequently employed reform based instructional practices.

Discussion of Variance in Student Outcomes as Related to Teacher Implementation

The previous section documented differences in how and why teachers implemented *Core-Plus.* The current focus draws on these differences to explain the variance, or lack of, in student achievement among the four groups, Group A who attended the NCIM workshop and had monthly coaching assistance (n=7), Group F who had monthly coaching assistance (n=2), Group B who attended the summer workshops but had no coaching support (n=6), and

Group D who were not affiliated with the NCIM project in any capacity (n=6). First, based on the analysis of the surveys, observations, test analyses, interviews and instructional coach report, possible explanations for the lack of significance among the groups on the Algebra I EOC will be examined and discussed. This will be followed by a similar discussion to consider how teacher implementation can help explain the differences in student achievement on the Algebra II EOC between each of the groups.

Variance in Algebra I Achievement

There were no differences in student achievement between any two groups on the Algebra I EOC. Given the population of students at Group A and F schools (small ADM, high minority, high %FRL, and high percentage Title One) and the other schools (Groups B and D) these results represent encouraging results for typically disadvantaged students. These results are encouraging because these groups of teachers had similar student achievement in the models without controlling for differences in school need and prior student achievement. Further, looking only at mean achievement for each group, Black students in Groups A and F outperformed Group B students (Group F's Black population only had an n=2), and White students in Groups A and F had a higher mean Algebra I achievement than either Group B or F. These are positive results for the use of integrated mathematics in these high poverty settings. To provide further evidence in support of using *Core-Plus* in high poverty school, future research should include school-level predictors and prior achievement measures into the HLM models or use a comparable population of subject-specific students from high need schools.

What remains perplexing is the lack of evidence of differences in student performance among Group A and other groups given their extensive and ongoing professional development. Perhaps the lack of significance between the groups is due to their participation in the NCIM workshops and without these experiences they would have had significantly lower results. This conjecture could not be validated by the data in this study because there was not a comparable population of integrated mathematics teachers, with limited experience teaching *Core-Plus,* in high poverty settings to compare to student achievement from Group A teachers. Another plausible explanation would be that the workshop and continued coaching assistance was ineffective and did not translate into classroom practice. However, qualitative data showed that Group A teachers were presenting the content in the textbook with more fidelity than other groups with NCIM project experiences. Further, it has been reported in other studies of the NCIM project that Group A teachers were able to transform their practice based on their workshop experiences and follow-up support but that these changes were gradual and took time to develop (Krupa & Confrey, in press; Thomas, 2011). It is possible that changes in instructional practice do have a significant impact on student achievement but the effect is not being accurately measured. Given the constraints of the outcome measure and the extra tutoring time built into the day for 4 of the 6 Group B teachers and Group F and D teachers frequent supplements for practice, it is difficult to measure the influence of instructional practices based on a standardized multiple-choice assessment of student achievement. The results from teachers'

implementation of *Core-Plus* suggest additional reasons there were no significant results between groups.

The Algebra I EOC released exam was highly aligned to the second edition of *Core-Plus* Courses 1 and 2 and moderately aligned to the first edition. The groups with high implementation indices and low supplementation, Groups A and B, used textbook content for their instruction that was aligned with the state assessment. Though there were variations in the manner in which content was presented between these two groups, they both utilized content in the textbook for their main mode of instruction. As a result, there were no differences in student outcome measures between these two groups. Based on this analysis, the most compelling explanation for the insignificant difference between these two groups seems to be due to the fact that they covered approximately the same amount of textbook content with little supplementation.

Groups F and D supplemented content more frequently than the other two groups. They supplemented based on their perceptions about mathematics teaching and to provide extra practice to prepare their students for the exam. In the case of Group F, the instruction was rooted in traditional mathematics instruction of teacher lecture and student note taking followed by considerable practice on each concept. Unsurprisingly, students exposed to instruction targeted at passing the Algebra I EOC performed no different than students exposed to more reform oriented instruction, given the state assessment is a test of curricular alignment with systemic factors (NRC, 2004). In other words, the test measured students'

progress within high school mathematics and was not sensitive to the content, goals, or objectives of the *Core-Plus* curriculum.

Group D's results are harder to explain given the teachers' experience with the *Core-Plus* curriculum and beliefs about mathematics, in favor of reform practices. Group D teachers utilized reform instructional practices frequently but tended to supplement content more than any other group. Their supplementation is likely due, in large part, to the use of the first edition by almost all Group D teachers. They also supplemented content from other sources and used practice worksheets throughout the school year to prepare students for the Algebra I EOC. There were limited practice problems and hands-on activities in the first editions, so teachers felt they had to bring in additional materials. The combination of their use of the text with extra practice could have attributed to the lack of significant differences between them and other groups. On the other hand, given the population of students in Group D schools, low minority and lower school %FRL, and the teachers' experience teaching *Core-Plus,* the lack of differences could suggest positive results for teachers who participated in any facet of the NCIM project. Stated alternatively, given the fact there were no differences between Group D and the other groups, perhaps the support the NCIM project gave to teachers in Group A, F, and B closed the gap in achievement between students at these schools and students at Group D schools.

Though teachers attending the summer workshops had a higher fidelity of content implementation than non-workshop teachers, this result did not permeate to influence student outcome measures. Workshop participation, along with teaching experience with the

materials, made teachers more comfortable with trusting the curriculum, but this did not have an impact on student achievement.

Variance in Algebra II Achievement

The focus group included student Algebra II achievement data for teachers representing Groups A, F, and D. Without controlling for teacher characteristics, there were no significant differences in student Algebra II achievement between Groups A and F, Groups A and D, or Groups F and D. However, when controlling for teacher content knowledge we began to see differences between any two groups. When content knowledge, OTLAlg, years taught *Core-Plus*, beliefs about reform mathematics, and beliefs about traditional mathematics, along with student background, were controlled for there were significant differences between each group of teachers. Group F significantly outperformed Group A by 5.22 points and Group D by 8.44 points, while Group A outperformed D by 3.22 points on the Algebra II EOC. Considering the mean achievement for each Group, $\left(\mu_A = 148.29, \mu_F = 152.22, \mu_D = 152.49\right)$, showed that students in Group A had a mean achievement approximately four points lower than either of the other two groups. Together, these results highlight the importance of teacher content knowledge, experience with the curriculum, textbook implementation, and beliefs about mathematics. Holding these constant sheds light on differences not apparent without incorporating them into models of student achievement.

Data have pointed to the traditional practices of Group F teachers and their weak implementation of textbook content and instructional practices. Their significantly higher

student achievement on the Algebra II EOC, controlling for their weak content knowledge, could be attributed to their desire to "teach to the test" or to their weak content knowledge. First, consider content knowledge. There were no differences between Group F and the other two groups when the CKA predictor was not included, once including this predictor there were large fixed effects for the differences between Group F and the other groups. It appears their weak content knowledge hindered their students from being even more successful on the assessment. However, Group F students still had an average higher than the sample average, suggesting that if Group F teachers improved their content knowledge their students would have even greater achievement. Secondly, Group F teachers utilized traditional instruction of directing learning from the front of the classroom, rarely used collaborative groups, and frequently supplemented practice worksheets into instruction. These teachers were providing instruction devoted to their students doing well on the Algebra II EOC throughout the year. Alternatively, these significant results could be due to the addition of the discrete mathematics course between Courses 2 and 3, which allowed students in Group F's classes to have an additional mathematics course before taking the Algebra II exam. This provided these students with more time to learn content and more time to mathematically mature. Additionally, the majority of these students had Maria, from Group A, for Courses 1, 2, and discrete mathematics. This complicates the interpretation we can make on these students' achievement, as it would be fair to assume they had positive exposure to reform instruction prior to entering into a Group F teacher's classroom.

Briefly recall that five of the six teachers in Group D used the first edition of the textbook during the 2009-2010 school year, which was poorly aligned to the Algebra II EOC. Similarly, the Course 1, 2, and 3 books of the second edition, used by Group A and F teachers, had a weak alignment to the released Algebra II EOC. Variance in student achievement between Group A and D teachers could be attributed to Group D teachers supplementation of the text. Teachers in Group D supplemented textbook content to handle the weak alignment between the textbook and the exam, whereas Group A teachers remained faithful to implementing textbook content, despite the weak alignment. Alternatively, the variance could be attributed to the manner in which textbook content was presented to students. Though, Group D's supplementation aside, these two groups were the most similar in their abilities to provide reform instruction to students. The differences between these two groups were only apparent when controlling for teacher background and beliefs. Group D teachers had higher content knowledge, more experience teaching with *Core-Plus*, were more likely to hold a college degree in mathematics or mathematics education, and had significantly lower student achievement when incorporating these characteristics. However, students in Group D had a mean achievement over four points higher than Group A students. Thus, the variance in student achievement, only apparent when teacher traits were held constant among teachers, could be attributed to Group A's weak content knowledge and lack of experiences teaching Course 3.

Discussion and Conclusions From Chapter

This chapter highlighted significant differences in teachers' implementation of *Core-Plus* textbook content, factors that correlated to student Algebra I and II achievement among groups with varying levels of NCIM professional development experience, examined themes that influenced teachers' instructional practices, and drew on multiple forms of analysis to detail how these various data sources and results combine to provide a rich picture for teachers implementation of *Core-Plus.* First, an analysis of textbook implementation showed significant differences between teachers that attended the NCIM workshop and those that had not. This suggests that workshop attendance can increase teachers' trust for the curriculum or knowledge about how to implement it with students.

Next, these implementation indices were used to model student Algebra achievement. These models underscored the importance that teacher content knowledge and experience teaching a curriculum had on student achievement, as well as support findings from the previous analysis on the impact student background variables had on student outcomes. Teacher content knowledge and years teaching *Core-Plus* were significant positive predictors for student Algebra II achievement. However, these were non-significant for Algebra I achievement, suggesting that as the content gets more difficult in Course 3, students need experienced teachers with high levels of mathematics content knowledge. Further, differences in student achievement among the focus groups was only prevalent on the Algebra II exam when controlling for teacher content knowledge, years experience teaching *Core-Plus,* OTLAlg, and teacher beliefs about traditional and reform mathematics.

There were similarities and differences in the resources used and mode of instruction among the four groups of teachers using the *Core-Plus* materials. Qualitative data indicated that teachers' beliefs about how students learn mathematics, their trust for the curriculum, and systemic factors influenced decisions teachers made about textbook implementation. For example, their beliefs about how students learn mathematics had an impact on teachers' use of collaborative groups, supplements they provided to students, and their ability to let students struggle through mathematics.

Teachers' trust for the curriculum was a barometer for how frequently they utilized reform or traditional instructional practices. Teachers with low trust tended to provide students with lots of standard practice problems following direct instruction (notes at the board). It took teachers time to understand the spiraling nature of content throughout the *Core-Plus* curriculum. Participation at the NCIM summer workshop helped Group A and B teachers feel more confident and trusting of the *Core-Plus* curriculum and instructional practices, and the instructional coaches supported these ideas throughout the year for Group A teachers. Teachers need an understanding for the spiraling nature of content throughout the *Core-Plus* curriculum materials before implementing it with students.

Systemic factors within a school or system also had an impact on teachers' implementation of *Core-Plus.* There is no denying the EOC drives mathematical instruction across the state. The reason many of the focus group schools began implementing *Core-Plus* was to increase student achievement at these typically isolated low performing sites. The constraints and affordances of the system (access to materials, scheduling, and student

adjustment to *Core-Plus*) contributed to the variance among teachers' implementation of the textbook materials.

Finally, using a mixed methods approach provided a deeper understanding for differences in the implementation indices, for the lack of significant differences among groups in the focus study when modeling Algebra I achievement, and for the varying degree of differences among groups in the Algebra II sample. Analyzing teacher interviews, instructional coach reports, and field-notes from classroom observations helped to detail instructional practices teachers used in their classrooms and decisions they made about supplementing the textbook. The factors that had an impact on teachers' implementation (beliefs about how students learn, trust for the curriculum, and systemic factors) provided evidence that helped explain the quantitative results. Further, the factors provide a foundation for recommendations, to be presented in the next chapter, which should be considered to successfully implement *Core-Plus.*

CHAPTER SIX: DISCUSSION AND CONCLUSIONS

Fifteen year trends of North Carolina passing rates for the five core courses (Algebra I, Civics and Economics, Biology, U.S. History, and English I), show that student Algebra I passing rates fell twenty percentage points by the school year 2006-2007 and have since consistently maintained the lowest passing rates out of the core EOC exams (Public Schools of North Carolina, State Board of Education, & Department of Public Instruction, 2011). To improve student achievement on the Algebra I exam, several districts and/or schools adopted an integrated mathematics curriculum. However, the impact of such curricula on student achievement across North Carolina has not been well researched. One goal of this study was to determine if there was evidence of differences in student achievement for students who took integrated mathematics and those who took subject-specific mathematics.

Research has established convincingly that implementing reform curricula is a difficult and uneven process (Arbaugh, et al., 2006; Chval, Chavez, Reys, & Tarr, 2008; McNaught, et al., 2008; McNaught, et al., 2010; Swafford, 2003; M. S. Wilson & Lloyd, 2000; Ziebarth, 2003). Sufficient professional development is critical to ensure teachers are equipped with the knowledge and skills to transition to implementing reform-based instructional practices (Ball & Cohen, 1999; Loucks-Horsley, et al., 2003; Putnam & Borko, 1997; S. M. Wilson & Berne, 1999). Consequently, an additional goal of this study was to provide an account of the influence of a variety of components of a professional development program on teachers' implementation of the *Core-Plus* materials and their subsequent impact

on student learning in *Core-Plus* classrooms. Specifically, the two research questions addressed were:

1. What evidence, if any, is there of significant differences in student achievement between students using integrated mathematics and subject-specific curricula taught as Algebra I, Geometry, and Algebra II across the state of North Carolina?
    a. Is there evidence of differential impact on student performance on state tests between teachers using integrated mathematics curricular materials and those who use subject-specific curricula?
    b. For teachers who participated in the project workshops, is there evidence of differential impact on student performance on state tests between teachers using the *Core-Plus* curricular materials and those who use subject-specific curricula?
2. What evidence, if any, is there of significant differences on student achievement among teachers who participated in different components of the NCIM project?
    a. Among teachers using the *Core-Plus* curricular materials, what evidence of systematic differences is there on the effects of student learning based on varying levels of teacher participation (working with or without a instructional coach, frequency of professional development, type of professional development (content and/or pedagogy) in the NCIM project?

b. Among teachers using the *Core-Plus* curricular materials, to what extent and how can the variance in student outcomes be explained by teacher implementation?

This chapter will present a brief overview of the study, discuss key findings, discuss implications of the findings for various audiences, detail limitations of this study, offer recommendations for future research, and close with final remarks on the significance of this research.

Overview of Study

This section summarizes the research that was conducted as part of this study and discusses the significance of this work. This study analyzed student Algebra I and II EOC data, first with a larger sample of integrated mathematics and subject-specific students, and then with a restricted sample of students whose teacher had attended one of the NCIM Summer Workshops. The workshop sample examined the impact of the curricular materials with more diverse, high need populations and afforded to the researcher opportunities to include teacher characteristics into the analyses. Compared to the population of NC students who took the exams in 2009-2010, the sample for this larger study was representative of students with lower passing rates and mean performance on both Algebra exams, and therefore, caution should be considered when generalizing the results in this portion of the overall study to the broader set of students across North Carolina.

Quantitative data provided insight into differences in student achievement between the curricular pathways and effects of teacher-level predictors on student achievement, but

the data collected in those two analyses yielded limited explanations for their findings and could not explain teachers' implementation of the *Core-Plus* materials, a popular integrated mathematics textbook used across North Carolina. While it could be determined that curriculum choice made a difference with certain samples, understanding what specific conditions seemed to make a difference was equally important. Thus, to answers these questions, it was necessary to further restrict the sample, and so 21 teachers were chosen for a focus study detailing their implementation of the *Core-Plus* textbook and instructional practices. Teachers were selected from four different groups (Figure **19**), Group A teachers participated the NCIM Summer Workshop and received an instructional coach (n=7), Group F teachers received an instructional coach only (n=2), Group B teachers participated in the Summer Workshop (n=6) with no coach, and Group D teachers did not participate in any facet of the NCIM project. Groups A and F contained a majority of students from high poverty settings, whereas students from Groups B and D had a significantly lower needy percentage than the larger sample of students. Restricting the larger sample to students of focus group teachers provided an opportunity to examine teacher implementation of *Core-Plus* and on other teacher-level factors that had an effect on integrated mathematics students' Algebra performance.

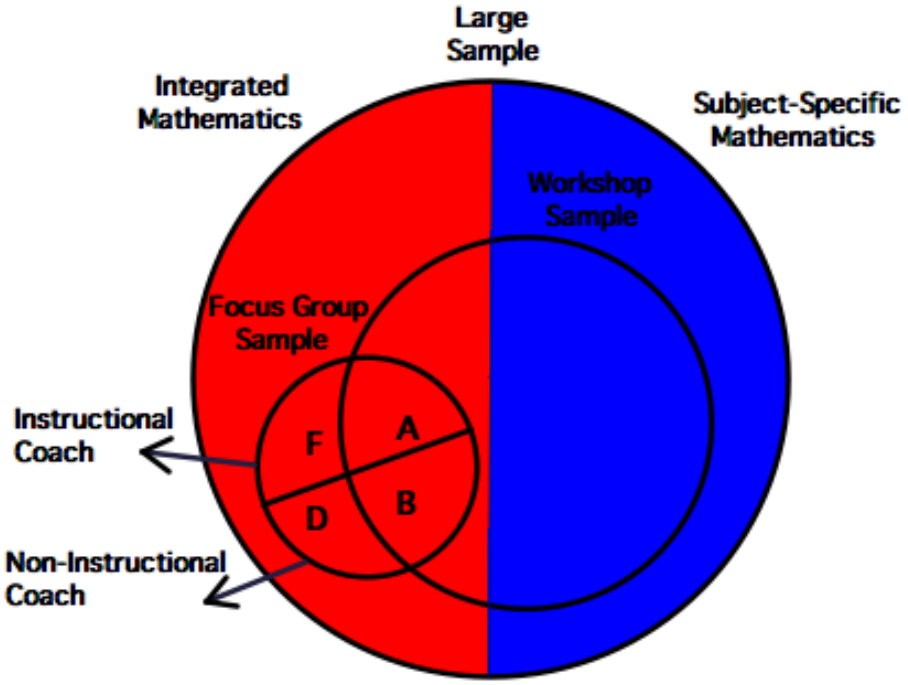


Figure 19. Sample of participants used in this study.

Discussion of Key Findings

Algebra I

Across multiple analyses with the larger and workshop student samples, students enrolled in integrated mathematics consistently had higher student achievement than subject-specific students on the Algebra I EOC. Not only did integrated mathematics students have a higher mean achievement on the Algebra I EOC, but after controlling for student and school characteristics, HLM analyses indicated integrated mathematics students had Algebra I scores ranging from 3.36 to 5.06 points higher than subject-specific students. A content analysis conducted prior to the analysis of the results, showed content in *Core-Plus,* was highly aligned to items on the Algebra I EOC.

Some might question the significance of the differences between the two curricular pathways due to the fact that integrated mathematics students took the exam after Course 2, or their second high school mathematics course, whereas subject-specific students took the exam after Algebra I, which is often their first high school mathematics course. The percentage of students from each pathway who took the Algebra I EOC in 9th grade would confirm this as a possibility. However, analysis of the mean achievement showed that not only did integrated mathematics students outperform subject-specific students from the same grade level, they also outperformed subject-specific students in the grade levels below and above them. That is, integrated mathematics students had a higher mean achievement than subject-specific students who were exposed to more than one high school mathematics course prior to taking the Algebra I exam, though typically these are the weaker subject-specific students. Also, they had higher achievement than their more advanced subject-specific peers who took Algebra I in earlier grade levels. For example, integrated mathematics students who took the Algebra I exam in 9th grade had a higher mean achievement than subject-specific middle school students, who were on an accelerated track, 9th grade students, who were on a standard track, and 10th grade students, who were exposed to at least two mathematics courses prior to the Algebra I exam.

Analysis of the data from the focus group also provided possible alternative reasons that integrated mathematics students could have had higher Algebra I achievement. First, North Carolina students take the Algebra I EOC after Course 2 because that is when the state determined teachers would have covered content on the Algebra I exam due to the additional

content strands that are included in the *Core-Plus* textbook. However, it was clear from an analysis of the logs and surveys of teachers that teachers rarely introduced non-Algebra content into their instruction and would consequently have more time to devote to students' development of algebraic concepts. Further, due to the documented time it takes to transition students to integrated mathematics, both adjusting to the collaborative nature of the curriculum and the reading required, students and teachers needed the additional course to provide students with a strong mathematical foundation.

It is also important to consider the population of students that benefited or were disadvantaged by using a certain curriculum, and it appears that integrated mathematics was more effective than subject-specific in high needs settings. Keeping in mind that students who are more strongly prepared tend to take advanced math courses earlier in high school, one way to examine who is being most successfully served is to consider the grade level at which the EOC was taken. No differences in student learning were evident for students at lower grade levels (9^{th} vs. 11^{th}) in low %FRL schools between integrated mathematics and subject-specific mathematics students. However, integrated mathematics students from higher grade levels and in high needs schools had higher Algebra I scores than subject-specific students from the same demographic. In addition, teachers of integrated mathematics had more consistent student results regardless of the school %FRL and across grade levels, whereas subject-specific teachers had a decrease in student achievement as school %FRL increased and for students at higher grade levels.

In considering patterns of performance by race across all three samples, White students had higher Algebra I achievement than all non-Asian minorities. Further, White students in integrated mathematics revealed an advantage over White subject-specific students. Similar trends were not observed for Black students, as there were no significant curricular differences in favor of Black students on the Algebra I exam. This should neither be interpreted as integrated mathematics only works for White students, nor that subject-specific was ineffective for White students. Rather, in this sample no perceived benefits of using one curriculum over the other with Black students was found, and White students had a higher mean performance when enrolled in integrated mathematics.

Algebra II

The arguments presented to explain the Algebra 1EOC scores point to the need to analyze student achievement after students had been exposed to the entire integrated mathematics sequence. In both curricular options, students took the Algebra II exam after three high school courses, after Course 3 on the integrated mathematics pathway and after Algebra II on the subject-specific pathway. There were no significant differences between students of teachers of subject-specific and integrated curricula on the Algebra II exam, even after controlling for different teacher characteristics.

Some might suggest this result is, in part, favorable to subject-specific students given that integrated mathematics students had higher prior achievement. These readers might suggest that since integrated mathematics students began with higher prior achievement, the lack of differences on the Algebra II exam indicates that integrated mathematics students did

not get any further increase in achievement, or alternatively, that subject-specific students showed more growth on the Algebra II exam. Analysis of mean achievement in the larger sample could support such a theory, integrated mathematics students were found to have a higher mean prior achievement, and yet in this larger sample they scored no differently than subject-specific students on the Algebra II exam. Furthermore, integrated mathematics and subject-specific students from the workshop sample had the same prior achievement and still no significant differences between the two student groups were in evidence. These inconsistent findings may suggest a lack of alignment or curricular sensitivity of the outcome measures, which was confirmed by the content alignment study showing that at least for the released Algebra II items, the EOC was not well-aligned to the *Core-Plus* integrated materials. A possible further disadvantage of the outcome measure for students in integrated mathematics was that they were often given tests requiring them to explain their work in contextual situations and less accustomed to the multiple choice format. For some readers, given the lack of content and format alignment between the integrated materials and the exam, it could be argued that it is, in fact, encouraging that integrated mathematics teachers experienced the same level of success as subject-specific teachers.

Another explanation that merits careful and further examination is that the lack of differences in student learning could be due to the inexperience and content weaknesses of the integrated mathematics teachers. The data showed that the majority of teachers in the workshop sample teaching Course 3 were doing so for the first time during the 2009-2010 school year. It also showed that teachers of integrated mathematics with higher content

knowledge produced students with higher Algebra II achievement. Perhaps with more experience teaching Course 3 and improved content knowledge, integrated mathematics teachers would obtain higher student achievement, a hypothesis worthy of further investigation.

Student Effects

Findings from this study also provide insight into student characteristics that influence student achievement. Based on the HLM analyses, there were consistent differences in student learning based on several student characteristics, most notably grade level and racial status. Both will be briefly discussed below.

Grade level. Consistently students at lower grades had higher student achievement than students at higher-grade levels. The one exception to this finding was in the focus group sample, which contained only integrated mathematics students. There were no differences in student Algebra II achievement among students at different grade levels. Focus group schools equitably served students at all grade levels by the end of the *Core-Plus* sequence, regardless of the year students enrolled in the curricular sequence. This finding supports prior analysis that integrated mathematics students had consistent results across the grade levels, while subject-specific students saw a decrease in achievement for each increase in grade level. Subject-specific students that were not tracked into Algebra at early grade levels were disadvantaged in the knowledge they obtained in high school, whereas student tracking on the integrated pathway did not effect the knowledge they obtained by the end of the curricular sequence.

Race. After discouraging differences in Algebra I achievement based on student race, especially for non-Asian minorities, the differences seemed to dissipate over time with the inclusion of prior achievement in models of Algebra II achievement, with Black students as the only racial group that continued to have lower mean performance than that of White students. Even more encouraging, the focus group, from high need schools, had no racial differences in student learning on the Algebra II exam, when controlling for Algebra I achievement. When removing prior achievement as a predictor there were racial differences in student achievement in all samples, with smaller effect sizes in the focus group sample. This suggests both the importance of including this type of measure in models of student achievement and its ability to suppress the effects of race. One possible explanation for these positive outcomes at focus group schools may be due to how they used *Core-Plus* to engage their students in mathematics and implemented unique scheduling to provide their students with extra support. These scheduling decisions included the addition of a mathematics course to provide students with more time to develop deep content knowledge or time built into the daily schedules for extra tutoring on mathematics content. Due to the unique structure of the schools in the focus group, it appears these schools were designed to improve the success of all students.

Teacher Effects

This study found several teacher traits that influenced student achievement, some were anticipated and others more surprising, but each warrants further discussion. The

sections below will discuss the effect teacher content knowledge, teacher gender, OTLAlg, and NCIM professional development experience had on student achievement.

Teacher content knowledge. Throughout multiple forms of analysis on both outcome measures, teacher content knowledge was positively correlated with student achievement. The effects were the strongest for integrated mathematics teachers, especially on the Algebra II exam, suggesting that teachers of Course 3 need a strong mathematical background. Integrated mathematics teachers in Course 3 had very weak content knowledge, so much so that teachers in Group F were only comfortable when they were the source of mathematical knowledge in the classroom. On the other hand, the integrated mathematics teachers with the highest content knowledge, Group D, were more comfortable facilitating collaborative learning and also had more autonomy in supplementing textbook content. This result speaks to the essential need for substantial opportunities for teachers implementing reform approaches to have sustained professional development specifically targeted at the topics and related instructional practices in the advanced courses.

Opportunity to Learn Algebra Content. Based on analysis of the focus group sample, an increase in a teachers' OTLAlg index resulted in a decrease in student Algebra II achievement, but only when holding differences in teacher characteristics constant. This result must be considered in light of qualitative analysis of the focus group teachers. Group F teachers, who received no workshop training but had an instructional coach, were implementing *Core-Plus* for the first time during the 2009-2010 school year, had weaker content knowledge than the other groups of teachers, had a significantly lower mean OTLAlg

index than other groups, and had beliefs about instructional practices aligned with traditional views of teaching mathematics. Only when controlling for these differences did Group F teachers emerge as producing higher student achievement than teachers who participated in different components of the NCIM professional development. It seems likely that their lower OTLAlg indices account for the negative correlation between OTLAlg and Algebra II achievement. Since these results were only documented when controlling for large differences between teachers in the sample, this outcome, which warrants further investigation, should not be interpreted to indicate that lower OTLAlg indices produce higher student achievement scores. Further caution is warranted because Group F teachers' low OTLAlg indices could be attributed to the addition of a Discrete Mathematics course at their school, which allowed them to distribute the Course 3 content up over two semesters. So, while their OTLAlg indices may appear low, students may have been exposed to more Algebra content over the course of the two years than the TOC-log calculated. The surprising results do, however, illustrate why quantitative analyses of student outcomes need to be carefully examined with consideration of multiple methods providing information on implementation, teacher variables, and content monitoring.

Professional Development. Increases in the number of hours workshop teachers spent in the NCIM professional development decreased student Algebra I and II achievement. This surprising outcome should not be interpreted as the workshop was ineffective in its goals to educate teachers about the content and pedagogy of using integrated mathematics. On the contrary, this study showed that integrated mathematics teachers attending the workshop had

higher textbook implementation than non-workshop teachers and that workshop experiences helped teachers build trust for the curriculum. Additionally, it is important to note that teachers with the most professional development, NCIM teachers, were employed at schools serving underrepresented students, with more challenging levels of student need and preparation, and were located in historically low-performing school systems. Perhaps the professional development helped increase student achievement at these schools over time, however data for this study were not collected to explore this hypothesis further and needs further study. The finding could be a result of a ceiling effect. The NCIM teachers received an extensive amount of professional development hours, with an average of 112 hours compared to the rest of the workshop sample of 33 hours, there could be a point at which increasing the hours of professional development do not further improve student achievement. That does not mean that the professional development does not continue to improve teachers' instruction with *Core-Plus*.

Teacher Gender. Across the workshop and focus group samples, students of female teachers had higher achievement than students of male teachers. Results indicated that female integrated mathematics teachers had higher student Algebra I achievement than any other teacher and that on average, students of female teachers had Algebra II scores three points higher than students of male teachers. However, with only four male teachers in the focus group, it was difficult to make generalizations of male teachers' implementation of *Core-Plus.*

Teacher Implementation

The perceived benefit using integrated mathematics with high poverty, diverse populations is encouraging. To understand these results and suggest directions for further elaboration, researchers should engage in discussions about the use of the *Core-Plus* curriculum with these groups of students and should include factors that relate to the materials high implementation fidelity and the role of professional development in those processes. The focus group allowed the researcher to identify what factors seemed to relate to *Core-Plus* being implemented successfully in these low performing settings among teachers with varying levels of NCIM professional development experience. Factors that influenced teachers' implementation of *Core-Plus* were: teacher beliefs about how students best learn mathematics, their trust for the curriculum, and systemic factors.

Teacher beliefs about how students learn mathematics influenced their use of collaborative groups, the manner in which they presented content to their students, and the frequency in which they supplemented textbook content with additional materials. Based on teacher beliefs about how students learn mathematics and NCIM professional development experiences, teachers implemented the curriculum differently. Teachers from Groups A and B had NCIM workshop experience and similar beliefs about how students learn mathematics, however teachers in Group B were often observed pre-teaching content or decontextulizing problems in the textbook. While Group B teachers whole-heartedly tried to implement *Core-Plus* with high fidelity, there were still gaps in their comfort for and trust in using the textbook to address student learning needs. Teachers in Group A experienced many of the

same struggles when they first began implementing the curriculum, but the support of their instructional coach facilitated a much easier transition than Group B experienced, emphasizing the importance of site-based support, either from a coach or possibly a more-experienced peer. Group F teachers believed students needed direct instruction, and despite receiving an instructional coach they made little effort at utilizing reform based instruction consistent with the *Core-Plus* textbook.

Teachers' trust for the curriculum influenced profoundly how they handled the gradual refinement of key concepts in the *Core-Plus* curriculum. Regardless of professional development, inexperienced teachers were uncomfortable with the spiraling nature of the curriculum and did not understand the reoccurrence of content at deeper levels, often pre-teaching or over-explaining concepts before they were fully developed in the course sequence. Though instructional coaches made teachers aware of instances when they explained too much, teachers' trust for the curriculum seemed to increase gradually more as a result of using the materials directly with students.

Finally, the affordances and constraints of the school system influenced how teachers implemented the curriculum. Student achievement on standardized tests often drives education decisions across the country, and can be a deterrent to an emphasis on student understanding. This was one of the main reasons teachers across the study reported using traditional practice worksheets and why they felt the need to supplement with other materials that appeared to weaken the organizational structure of the *Core-Plus* sequence. Additionally, too often, they reported needing to supplement materials due to the lack of

sufficient access to materials for the investigations. Limited access to manipulatives and technology recourses constrained teachers from implementing the curriculum as intended. A positive example of a systemic factor that seemed to affect student outcomes positively was affordances in scheduling that provided students with more time to learn mathematics content through extra tutoring built into the school day or with the addition of a mathematics course between Courses 2 and 3. The use of reform mathematics in middle school eased students' transition to *Core-Plus*, providing teachers more time to focus on implementing the curriculum rather than on transitioning students to work in collaborative groups and on reading in mathematics.

These three themes, views of student learning, trust in the curriculum, and removing hurdles and providing resources systemically helped document how teachers were implementing *Core-Plus* in high poverty settings among teachers with different professional development opportunities. They also provided insights about teachers' implementation that are typically obscured in solely quantitative studies. For example, quantitative results imply that lower OTLAlg increased student achievement, however qualitative findings help explain this result by the supplements used and addition of a mathematics course at schools with low OTLAlg indices. Additionally, quantitative findings suggest that teachers enrolled in the summer workshops more faithfully implemented *content* from the textbook, but these data did not provide evidence on their *instructional practices*. The three themes highlighted that while professional development experience increased teachers' use of textbook content, instructional coaches were an important component to facilitating change in teachers'

instructional practices and in helping them more faithfully implement the fundamental principles of the curriculum. One finding worth further consideration was the recognition that teachers with strong content knowledge and more experience teaching *Core-Plus* were more likely to use supplements in their instruction, as they frequently utilized reform-based instructional practices. This result suggests that while beginning teachers may need to focus on learning to implement and trust the curriculum, once mastered, teachers could be developing more nuanced judgments of how and when to supplement effectively. The course of development of teachers, as individuals and as members of professional communities, must be part of the models to understand curricular effects. This study has shown convincingly that the complexities involved in decisions and actions teachers make when implementing a curriculum are more likely to be overlooked if one only analyzes quantitative data and thus support the case for multiple methods in the study of curricular effectiveness (NRC, 2004). The qualitative findings from this study greatly enhance the statistical results by providing both supporting and contradictory evidence detailing how teachers used the curriculum.

Implications

To help strengthen mathematics education across the state and nation, findings from this research can help guide future efforts in mathematics education. Specifically there are implications for professional developers, policy experts, researchers, administrators, and teachers. Recommendations for each of these five areas will be described below.

Policy

Policy experts should consider integrated mathematics as a viable curricular option, but should also consider a more nuanced question, "what works, for whom, and under what conditions?" (Bryk, et al., 2011; Confrey & Maloney, 2011; Means & Penuel, 2005). First, policy makers should look not only at student achievement on standardized tests, but also on tests aligned with the particular curricula under consideration. Data from this study showed that integrated mathematics students performed as well as, if not better than, subject-specific students. This is compelling evidence in favor of the use of integrated mathematics, but one must also determine if there would be differences in student achievement on reasoning assessments or ones that do not allow the use of a calculator or ones that are administered via a computer or on open-ended assessments. Prior research has shown differences in student achievement between *Core-Plus* and subject-specific students based on the type of assessment given (Huntley, et al., 2000; Schoen & Hirsch, 2003b; Schoen, et al., 1998; Tarr, et al., 2010).

Similarly, the demographics of the sample under study should be considered to determine if findings from a study are generalizable to the population of students represented in the area that seeks to implement a curriculum. It is important to consider the population of students and teachers that would use a particular curriculum and the support that would be needed to successfully implement it. The focus group was the only sample that did not show differences in student achievement based on minority status, an important finding suggesting that characteristics of the focus group schools implementing *Core-Plus* should be studied as

successful models. It appears that North Carolina New Schools Project involvement and the structure of these schools was critical at achieving similar achievement among all students, regardless of gender, grade, or minority classification.

Not only is it important to consider, "what works, for whom, and under what conditions?" (Bryk, et al., 2011; Confrey & Maloney, 2011; Means & Penuel, 2005), but also a structure needs to be put in place to provide researchers with data to conduct evaluation studies designed to answer these questions. This structure would not only benefit those researching curricular effectiveness, but also those choosing curricula to implement in districts and schools. This study has shown the importance of having a measure for prior achievement, which was lacking on the Algebra I analysis due to the confines of data access. Further, it suggests the importance of having longitudinal data to analyze both how students progress overtime and how teachers' average student achievement changes overtime based on their experiences and implementation. Districts and states should consider platforms that can collect and store student, teacher, and school level data so that researchers can design and conduct high quality curricular evaluations.

Professional Development

Teacher content knowledge was consistently related to student achievement and improving mathematics content knowledge of teachers should be a central goal of any professional development. Transitioning from a teacher-centered classroom to a student-centered classroom, where students make conjectures, argue with their peers, and explain their thinking, puts more demands on teachers' content knowledge. Teaching is no longer

about providing the right answer and more about facilitating rich student discourse along with negotiating an unprescribed path. Wilson and Lloyd (2000) suggest teachers with weak content knowledge may have greater difficulty changing their instructional practices, whereas teachers with strong and flexible content knowledge were able to concentrate on implementing specific instructional practices as well as the new mathematics content. It is imperative that professional development for integrated mathematics teachers be designed to embed content support into lessons on specific investigations in the textbook. Particularly weak content areas that need development should be identified and addressed.

Findings from this study confirm that professional development needs to extend beyond content dimensions into attention to the implementation of the curriculum and related practices at the classroom level. Teachers implementing reform mathematics for the first time need first-hand experiences working through the investigations as students ((Darling-Hammond & McLaughlin, 1995; Lloyd, 1999; Schoen & Hirsch, 2003b; Ziebarth, 2001). The two Group F teachers, who did not have workshop experience before implementing the curriculum, neither trusted the structure of the textbook nor believed in the investigative approach for learning mathematics. As a result they did not understand the contributions of collaborative learning and were uncomfortable with the manner in which content was presented in the textbook. Professional development should be designed to provide teachers with time to experience the curriculum as students working through the investigations in the textbook. Again, these experiences should be tied to enriching teachers' content knowledge and discussions about student misconceptions about the content.

Further, teachers implementing *Core-Plus* for the first time, both with and without workshop experience, struggled with trusting the spiraling nature of the curriculum. Professional development should be designed with the dual purpose of providing a broad overview for the entire curricular sequence and for follow-up discussions after teachers have worked through an investigation. Evaluations of the NCIM project urged project directors to consider adding an overview of the first three *Core-Plus* courses at the start of the week to present teachers with a glimpse at how mathematical content is refined throughout the sequence (Confrey, et al., 2008; Confrey, Maloney, & Krupa, 2009; Confrey, Maloney, Krupa, Thomas, & Corely, 2010). The necessity of such an overview was confirmed in this study, however providing the overview is not sufficient for teachers trusting the curriculum. After working through an investigation, there should be a discussion about the mathematical content that was learned and content that will be further developed in subsequent investigations or later in the course sequence. This will give teachers a better understanding for content development and might decrease the amount of "overteaching" that occurs in the classroom.

It is also important for professional development to introduce resources that are embedded within a curriculum. For example, *CPMP-Tools* is a computer software tool developed to support student inquiry during certain *Core-Plus* investigations. Teachers that did not have workshop experience and/or instructional coach support often skipped these investigations altogether because they were uncomfortable with the technology or could not get access to it in their classrooms. Summer workshops provide an opportunity to get

teachers excited about using innovative tools in their instruction and could lead to them being more proactive about asking for resource support and using these tools with their students.

Finally, prior research has documented the importance of having follow-up coaching support (Ball, 1996; Darling-Hammond, 1997; Joyce & Showers, 2002). Findings from the focus group have positive implications for professional development that offer on-going instructional coach support throughout the school year following a summer workshop. It is also important to note that instructional coach support absent workshop attendance may not influence teachers' instruction or help change their beliefs about how students learn mathematics.

Researcher

This research confirms recommendations by the NRC (2004) that using multiple methods when conducting curricular evaluations and including measures of prior achievement when analyzing student achievement is necessary. The NCR recommends including comparative studies, content analyses, and case studies into curricular evaluations. From the use of multiple methods, curricular evaluations have evolved into better ways to understand the complexity of curricular implementation and how teacher variables relate to student achievement. HLM models of student Algebra II achievement included a predictor for prior achievement, namely student Algebra I scores. These models accounted for much more student-level variance, around 30% more, than the Algebra I models that did not include a prior achievement measure.

This study has implications for researchers that want to assume teachers properly implement curricula after being exposed to professional development and that measures of implementation are not important to curricular effectiveness studies. Several researchers have documented the wide variance in teacher implementation of textbook material (Boaler & Staples, 2008; Huntley, et al., 2000; McNaught, et al., 2010; Swafford, 2003), and results from the TOC-log analysis attested the same thing for the focus group. Further, it was clear that the NCIM professional development did not ensure teachers were implementing *Core-Plus* with high fidelity. Therefore, researchers should not assume proper textbook implementation based on teachers professional development experience and efforts should be made to document implementation when conducting a curricular evaluation.

Further, continually in this research teacher content knowledge was a significant predictor of student achievement and teacher beliefs helped explain variance in teachers' implementation of *Core-Plus.* It is important for researchers to collect data on teacher-level characteristics to document what teacher variables have an effect on student achievement. Similarly, researchers should also focus on teacher-level characteristics that influence how teachers implement curriculum. Such research can help designers of professional development, teacher licensure requirements, and pre-service education.

Administrators

The main implication for principals and school officials is that the implementation of *Core-Plus* requires support. These supports include professional development for teachers and a supportive environment created by school leaders. The implementation of a new type

of curriculum such as integrated mathematics requires careful planning and attention to the goal: improved student learning and understanding. All participants at the school, teachers, parents, and staff need to be part of the decision and the plan of implementation. Data from studies such as this one should be summarized and understood as part of making the case to the community. The case for teacher professional development has been discussed previously, but it is important for principals to understand the major importance professional development plays in teachers' implementation of new curricula. Principals should also consider the types and forms of professional development for the long-term success of effecting change in teachers' instructional practice and these professional development offerings should be sustainable and scalable.

Principals should consider attending the professional development alongside their teachers for one day or more. This would not only show teachers they are supported by their school leaders, but also serve to educate principals on the main principles of the curriculum. When principals attend workshops designed to assist teachers with implementing *Core-Plus*, they gained insight on how to offer ongoing support to teachers in reform mathematics classrooms (Confrey, et al., 2008; Ziebarth, 2001). This issue promises to come to the forefront as plans to link teacher evaluation to student outcomes because increasingly likely. Principals must increasingly be sure they understand how to provide valid and specific instructional guidance to their staffs. From these experiences principals could see firsthand the importance of securing technology tools and materials for teachers to use when implementing the investigations. Several teachers noted they were unable to implement

textbook content because they lacked necessary resources. Principals should consider what tools the schools needs to invest in to facilitate an environment for proper implementation.

This research has highlighted the benefits of unique scheduling offered at some of the focus group schools. These schools served racially diverse, high need students well and unique features of their daily schedules provided teachers with more time to implement the curriculum and provide extra assistance to their students. Three of the teachers worked at schools that added a discrete mathematics course in between Courses 2 and 3. This sent the message that content in these courses was important and the school was supportive of ensuring the content was adequately taught. Other schools had extra tutoring built into their school day or utilized special schedules to ease students' transition to reform-based mathematics. Principals should consider alternative models to traditional schedules to ensure students and teachers are supported as the implement new curricula. Finally, they should consider advocating for curricula with similar instructional practices to be utilized at the middle school level to ease the transition time for students.

Teachers

Findings from this research have implications for teachers in the areas of content knowledge, use of collaborative groups, and providing reading support to students. First, prior research on NCIM teachers showed content weaknesses during their instruction (Thomas, 2011) and this current study found mathematical content weaknesses to be significantly correlated with student achievement. Teachers need to acknowledge where they have deficiencies in their mathematics content knowledge and try to improve in these areas.

They could seek assistance from other colleagues or attend professional development workshops. The main goal should be to improve their mathematical knowledge, but this learning should be embedded in the context of curricular units. This would not only benefit teacher content knowledge, but also increase their comfort with the curriculum and could be extended to incorporate a deeper understanding for the suggested instructional practices.

Second, utilizing collaborative groups is a major tenant of the *Core-Plus* curriculum, yet was implemented with varying levels of success with teachers in the focus group. It is important for teachers to consider the arrangement of desks in their classrooms. Teachers that had their desks in rows or in pairs facing the front of the room more frequently utilized direct instruction or pre-taught before launching an investigation. Teachers that embraced collaborative groups as a way for students to learn mathematics were more likely to provide their students with time to struggle with the mathematics, which allowed students to discover new concepts and engage in rich discussions with their peers.

Finally, focus group teachers discussed the importance of having students read the *Core-Plus* textbook, yet the difficulty they had transitioning students to being comfortable reading in mathematics. Teachers should be aware of strategies for helping students feel more confident reading in mathematics and for comprehending what they read. Teachers should maintain group roles throughout a class period to make sure that students are not switching readers. Even more importantly teachers should be aware that paraphrasing textbook content to make problems more comprehensible to students often encourages them to give up on reading and can decontextualize textbook problems. Additionally, teachers can

scaffold student reading, having them read a sentence and then interpret the meaning of what they read. This strategy is effective for individuals, groups, and for whole-class settings.

Limitations

Though an effort was made to ensure limitations of this study were minimal, they are inherent and unavoidable in any study (Punch, 2000). Limitations of this study included the data collection instruments, choice of outcome measure, and sample of participants for the focus group. First, several of the data collection instruments used in the quantitative analyses with the workshop and focus group samples involved self-reported teacher data. The teacher beliefs survey and TOC-logs were completed by teachers, and self-reported data should be interpreted with caution (Porter, 2002). To account for this with the focus group teachers, these data were triangulated with classroom observations, instructional coach reports, and interviews with the teachers.

Second, the outcome measures used in this study limits the conclusions that can be drawn from these analyses. The choice to use the North Carolina Algebra I and Algebra II End-of-Course assessments only allows the researcher to make claims about students' achievement on the multiple-choice standardized assessment and was not designed to test curricular validity, or rather how the assessment aligns to the curricular goals (NRC, 2004). This choice of outcome measure was important in the case of the NCIM project schools, which had been identified as low-performing based on N.C. accountability measures, because the funded professional development sought to improve teaching and learning at these schools and student performance was measured by their achievement on the state exams.

The final limitations of this study included the sample of participants for the focus group being a convenience sample with a small number of teachers. Teachers were recruited based on their professional development experience and course assignments. While HLM controlled for the non-randomization of the sample in the quantitative analysis, the sample of teachers chosen for the qualitative analysis may warrant limitations to generalizing the findings of this study to a broader set of teachers. Further, due to the final small sample size of teachers in the focus group HLM analyses (n=21), school-level effects could not be reliably modeled. The researcher acknowledges that focus group schools vary with respect to %FRL, ADM, and Title One status and these differences could explain variation that was attributed to the teacher-level effects.

Recommendations for Future Research

This study has indicated that integrated mathematics students have higher Algebra I achievement than subject-specific students and there were no significant differences in Algebra II achievement between students of either pathway. Further, data documented the wide variance in teachers' implementation of *Core-Plus* and factors that related to teachers implementation of these materials. This research has detailed the complexities of evaluating curricula and factors relating to its implementation and future research should be conducted to strengthen this study.

Recommendations for future research include analyzing longitudinal student data, tracking students' progress through high school for each pathway, utilizing a similar design with the new CCSSM assessments, replicating the focus sample with a larger number of

teachers within each group, and utilizing additional outcome measures. First, future research should longitudinally track individual students over the first three *Core-Plus* courses with a comparable group of subject-specific students. This research would provide more evidence about the knowledge students acquire throughout the sequence.

Second, an analysis of graduation rates, dropout rates, and post-secondary plans (college or career) should be conducted to determine how each pathway is preparing students for future success. Schools in this study were generally newer schools that had not had a graduating class of seniors. As these schools begin to grow, future research should analyze how curricular enrollment influences high school graduation rates and students next steps after high school.

The release of the CCSSM and accompanying assessments should enhance the quality of future research. These nationwide efforts will help future research study curricular effectiveness across states with a much larger population of students. Researchers are encouraged to leverage this national movement to develop high quality curricular effectiveness studies to analyze student achievement and teacher implementation in different geographic regions.

Third, the focus sample highlighted factors that related to teacher implementation of *Core-Plus,* but due to the small sample size there was not adequate statistical power to incorporate school-level characteristics into the HLM models. Further, variance was noted in teachers' implementation of the materials within groups. Another recommendation for future research would be to replicate the focus group study with a larger number of teacher

participants. This research would continue to help researchers understand the influence of professional development on teachers' implementation and for helping teachers implement new mathematics curricula.

A final recommendation for future research is to include additional outcome measures, most importantly a fair test (NRC, 2004), aligned with content from both curricula under study. It is well documented that, on average, students in classes using reform curricula perform comparably to students in subject-specific classes but with a higher performance on tests of conceptual understanding (Huntley, et al., 2000; Schoen & Hirsch, 2003a, 2003b; Schoen, et al., 1998; Tarr, et al., 2010). Findings from this research support the first part of that claim, but do not offer insight into students' performance on tests of conceptual understanding. Future research should be conducted to determine if integrated mathematics students have developed additional skills that were not tested on the EOCs. For example, from her synthesis of high school mathematics curricula Swafford (2003) concluded, "It is not that students in [reform] curriculum learn traditional content better, but that they develop other skills and understandings while not falling behind on traditional content" (p. 468). Findings from this study support that statement, yet future research should determine additional skills integrated mathematics students developed.

Closing Remarks

This research found that students exposed to integrated mathematics outperformed subject-specific students on the Algebra I EOC, which was highly aligned with content in *Core-Plus* textbook, and performed no differently on the Algebra II EOC, which was not

aligned with the *Core-Plus* materials. Due to the documented variance in teachers' implementation of *Core-Plus*, it is difficult to make claims about curricular effectiveness when implementation was less than ideal and students and teachers were still transitioning to reform-based instruction. However, quantitative results point to the benefits of using integrated mathematics in schools with high need and a high percentage of minority students. Given the effects of teacher content knowledge on student achievement and the factors that related to decisions teachers made about implementing the curricula, future professional development can be designed to strengthen the implementation of integrated mathematics, which should increase the effectiveness of the integrated mathematics curriculum on student achievement. Put simply, teachers need time and support to develop instructional practices aligned with successful implementation of reform practices.

REFERENCES

Altheide, D. L. (1987). Ethnographic content analysis. *Qualitative Sociology, 10*(1), 65.

Arbaugh, F., Lannin, J., Jones, D. L., & Park-Rogers, M. (2006). Examining instructional practices in Core-Plus lessons: Implications for professional pevelopment. *Journal of Mathematics Teacher Education, 9*(6), 517-550.

Balfanz, R., Mac Iver, D., & Byrnes, V. (2006). The implementation and impact of evidence-based mathematics reforms in high-poverty middle schools: A multi-site, multi-year study. *Journal for Research in Mathematics Education, 37*(1), 33-64.

Ball, D. L. (1988). Unlearning to teach mathematics. *For the Learning of Mathematics, 8*(1), 40-48.

Ball, D. L. (1996). Teacher learning and the mathematics reforms: What we think we know and what we need to learn. *Phi Delta Kappan, 77*(7).

Ball, D. L., & Cohen, D. K. (1996). Reform by the book: What is - or might be - the role of curriculum materials in teacher learning and instructional reform? *Educational Researcher, 25*(9), 6-14.

Ball, D. L., & Cohen, D. K. (1999). Developing practice, developing practitioners: Toward a practice-based theory of professional education. *Teaching as the learning profession: Handbook of policy and practice, 1*, 3ñ22.

Becker, J. R. (2001). Classroom coaching: An emerging method of professional development.

Boaler, J. O., & Staples, M. (2008). Creating mathematical futures through an equitable teaching approach: The case of Railside School. *Teachers College Record, 110*(3), 608-645.

Bowzer, A. (2008). *Professional identity and curricular construction: A study of teacher interaction with mathematics curricula of two types.* Unpublished doctoral dissertation, University of Missouri, Columbia.

Brown, A. L. (1992). Design experiments: Theoretical and methodological challenges in creating complex interventions in classroom settings. *Journal of the Learning Sciences, 2*(2), 141-178.

Bryk, A. S., Gomez, L. M., & Grunow, A. (2011). Getting ideas into action: Building networked improvement communities in education. In M. Hallinan (Ed.), *Frontiers in Sociology of Education*. New York, NY: Springer.

Chávez, O. (2003). *From the textbook to the enacted curriculum: Textbook use in the middle school mathematics classroom.* Unpublished doctoral dissertation, University of Missouri, Columbia.

Chávez, O., Papick, I., Ross, D., & Grouws, D. (2010, April). *The essential role of curricular analyses in comparative studies of mathematics achievement: Developing "fair" tests.* Paper presented at the Annual Meeting of the American Education Research Association, Denver.

Chval, K. B., Chavez, O., Reys, B. J., & Tarr, J. (2008). Considerations and limitations related to conceptualizing and measuring textbook integrity. In J. Remillard & B.

Herbel-Eisenmann (Eds.), *Mathematics teachers at work: Connecting curriculum materials and classroom instruction* (pp. 70-84). London: Routledge.

Cobb, P., Box, P., & Jackson, K. (under review). Towards an empirically grounded theory of action for improving the quality of mathematics teaching at scale.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*.

Common Core State Standards Initiative. (2010). Common Core State Standards for Mathematics. Retrieved from http://corestandards.org

Confrey, J., & Maloney, A. (2011). Engineering [for] effectiveness in mathematics education. *Highly Successful STEM Schools or Programs for K-12 STEM Education: A Workshop*. Washington, D.C.: National Academies Board on Science Education and Board on Testing and Assessment.

Confrey, J., Maloney, A., & Krupa, E. E. (2008). *N.C. Integrated Mathematics NC-STEM MSP: Year-End Evaluation Report*. North Carolina State University. Raleigh, NC.

Confrey, J., Maloney, A., & Krupa, E. E. (2009). *N.C. Integrated Mathematics NC-STEM MSP: Year-End Evaluation Report*. North Carolina State University. Raleigh, NC.

Confrey, J., Maloney, A., Krupa, E. E., Thomas, S., & Corely, D. (2010). *N.C. Integrated Mathematics NC-STEM MSP: Year-End Evaluation Report*. North Carolina State University. Raleigh, NC.

Costa, A. L., & Garmston, R. J. (1994). Cognitive coaching: A foundation for renaissance schools. Norwood, MA: Christopher: Gordon Publishers, Inc.

Coxford, A. F., Fey, J. T., Hirsch, C. R., Schoen, H. L., Burrill, G., Hart, E. W., et al. (2001). *Contemporary Mathematics in Context: A unified approach*. Chicago, IL: Everyday Learning.

Coxford, A. F., & Hirsch, C. R. (1996). A common core of math for all. *Educational Leadership, May*, 22-25.

Creswell, J. (1998). *Qualitative inquiry and research design: Choosing among five traditions*: Sage Publications, Inc.

Creswell, J., & Clark, V. (2007). *Designing and conducting mixed methods research*: Sage Publications, Inc.

Darling-Hammond, L. (1997). *The right to learn: a blueprint for creating schools that work*. San Franciso: Jossey Bass.

Darling-Hammond, L., & McLaughlin, M. W. (1995). Policies that support professional development in an era of reform. *Phi Delta Kappa, 76*, 597-604.

Dunn, O. J. (1964). Multiple comparisons using rank sums. *Technometrics, 6*(3), 241-252.

Elliott, A., & Hynan, L. (2010). A SAS macro implementation of a multiple comparison post hoc test for a Kruskal-Wallis analysis. *Computer Methods and Programs in Biomedicine*.

Emerson, R., Fretz, R., Shaw, L., & Thompson, I. (1995). *Writing ethnographic fieldnotes*: University of Chicago Press Chicago.

Field, A. (2009). *Discovering statistics using SPSS* (Vol. 2). Thousand Oaks, CA: Sage Publications

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

Glasser, B., & Strauss, A. (1967). *The discovery of grounded theory*. New York: Aldine.

Harwell, M., Post, T., Cutler, A., Maeda, Y., Anderson, E., Norman, K., et al. (2009). The preparation of students from National Science Foundation-funded and commercially developed high school mathematics curricula for their first university mathematics course. *American Educational Research Journal, 46*(1), 203.

Harwell, M., Post, T., Maeda, Y., Davis, J., Cutler, A., Andersen, E., et al. (2007). Standards-based mathematics curricula and secondary students' performance on standardized achievement tests. *Journal for Research in Mathematics Education, 38*(1), 71.

Herbel-Eisenmann, B. (2009). Introduction to the project, the people, and the reflective activities. In B. Herbel-Eisenmann & M. Cirillo (Eds.), *Promoting Purposeful Discourse: Teacher Research in Mathematics Classrooms* (pp. 3-29). Reston, VA: National Council of Teachers of Mathematics.

Hirsch, C. R., Fey, J. T., Hart, E. W., Schoen, H. L., & Watkins, A. (2008). *Core-Plus Mathematics: Contemporary Mathematics in Context*. Columbus, OH: Glencoe/McGraw-Hill.

Hollander, M., & Wolfe, D. (1999). *Nonparametric statistical methods* (2nd ed.). New York: John Wiley & Sons, INC.

Horizon Research. (2000). *Inside the classroom observation and analytic protocol*. Chapel Hill, NC: Horizon Research.

Hufferd-Ackles, K., Fuson, K. C., & Sherin, M. G. (2004). Describing levels and components of a math-talk learning community. *Journal for Research in Mathematics Education, 35*(2), 81-116.

Huntley, M. A., Rasmussen, C. L., Villarubi, R. S., Sangtong, J., & Fey, J. T. (2000). Effects of standards-based mathematics education: A study of the Core-Plus Mathematics Project algebra and functions strand. *Journal for Research in Mathematics Education, 31*(3), 328-361.

Joyce, B., & Showers, B. (2002). *Student achievement through staff development*: Association for Supervision & Curriculum Development.

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31-36.

Krupa, E. E. (2011). *Textbook implementation in rural secondary integrated mathematics classrooms.* Paper presented at the 33rd Annual Conference of North American Chapter of the International Group for the Psychology of Mathematics Education, Reno, Nevada.

Krupa, E. E., & Confrey, J. (2010). *Teacher change facilitated by instructional coaches: A customirzed approach to professional development.* Paper presented at the 32nd Annual Conference of North American Chapter of the International Group for the Psychology of Mathematics Education, Columbus, Ohio.

Krupa, E. E., & Confrey, J. (in press). Using instructional coaching to customize professional development in an integrated high school mathematics program. In J. M. Bay-Williams (Ed.), *Professional Collaborations in Mathematics Teaching and Learning: Seeking Success for All (yearbook; 2012)*. Reston, VA: National Council of Teachers of Mathematics.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics, 33*(1), 159.

Lieberman, A., & McLaughlin, M. W. (1992). Networks for educational change: powerful and problematic. *Phi Delta Kappan, 73*, 673-677.

Lloyd, G. M. (1999). Two teachers' conceptions of a reform-oriented curriculum: Implications for mathematics teacher development. *Journal for Research in Mathematics Education, 2*(3), 227-252.

Lloyd, G. M. (2002). Mathematics teachers' beliefs and experiences with innovative curriculum materials. In G. C. Leder, E. Pehkonen & G. Torner (Eds.), *Beliefs: A Hidden Variable in Mathematics Education?* (Vol. 31, pp. 149-159). Netherlands: Springer.

Loucks-Horsley, S., Love, N., Stiles, K. E., Mundry, S., & Hewson, P. W. (2003). *Designing professional development for teachers of science and mathematics*: Corwin Press.

Loucks-Horsley, S., Stiles, K. E., Mundry, S., Love, N., & Hewson, P. W. (2010). *Designing professional development for teachers of science and mathematics* (3rd ed.). Thousand Oaks, CA: Corwin Press.

McCaffrey, D. F., Hamilton, L. S., Stecher, B. M., Klein, S. P., Bugliari, D., & Robyn, A. (2001). Interactions among instructional practices, curriculum and student achievement: The case of standards-based high school mathematics. *Journal for Research in Mathematics Education, 32*(5), 493-517.

McNaught, M. D., Tarr, J. E., & Grouws, D. A. (2008). *Assessing curriculum implementation: insights from the comparing options in secondary mathematics: Investigating curriculum (cosmic) project.* Paper presented at the Annual Meeting of the American Education Research Association, New York.

McNaught, M. D., Tarr, J. E., & Sears, R. (2010, April). *Conceptualizing and measuring fidelity of implementation of secondary mathematics textbooks: Results of a three–year study*. Paper presented at the Annual Meeting of the American Education Research Association, Denver.

Means, B., & Penuel, W. R. (2005). Scaling up technology-based educational innovations. In C. Dede, J. P. Honan & L. C. Peters (Eds.), *Scaling up technology-based educational innovations*. San Francisco: Jossey-Bass.

Morse, J. (2003). Principles of mixed methods and multimethod research design. *Handbook of mixed methods in social and behavioral research*, 189-208.

Nathan, M. J., & Knuth, E. J. (2003). A Study of whole classroom mathematical discourse and teacher change. *Cognition and Instruction, 21*(2), 175-207.

National Council of Teachers of Mathematics. (1991). *Professional standards for teaching mathematics*. Reston, VA: NCTM.

National Research Council. (2004). *On evaluating curricular effectiveness: Judging the quality of K-12 mathematics evaluations*. Washington, DC: The National Academies Press.

NCTM. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: National Council of Teachers of Mathematics.

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

North Carolina New Schools Project. (2011) Retrieved April 26, 2011, from http://newschoolsproject.org/our-strategy/

O'Donoghue, T. A., & Punch, K. (2003). *Qualitative educational research in action: Doing and reflecting*: Routledge.

Olson, J., & Barrett, J. (2004). Coaching teachers to implement mathematics reform recommendations. *Mathematics Teacher Education and Development, 6,* 73-91.

Porter, A. C. (2002). Measuring the content of instruction: Uses in research and practice. *educational researcher, 31*(7), 3-14.

Post, T., Harwell, M., Davis, J., Maeda, Y., Cutler, A., Andersen, E., et al. (2008). Standards-based mathematics curricula and middle-grades students' performance on standardized achievement tests. *Journal for Research in Mathematics Education, 39*(2), 184.

Post, T., Medhanie, A., Harwell, M., Norman, K., Dupuis, D., Muchlinski, T., et al. (2010). The impact of prior mathematics achievement on the relationship between high

school mathematics curricula and postsecondary mathematics performance, course-taking, and persistence. *Journal for Research in Mathematics Education, 41*(3), 274-308.

Public Schools of North Carolina, State Board of Education, & Department of Public Instruction. (2011). The North Carolina State testing results: 2009-2010. Raleigh, NC

Punch, K. F. (2000). *Developing effective research proposals* (2nd ed.): Sage Publications Ltd.

Putnam, R. T., & Borko, H. (1997). Teacher learning: Implications of new views of cognition. *International handbook of teachers and teaching, 2*, 1223-1296.

Raudenbush, S., & Bryk, A. (2002). *Hierarchical linear models: Applications and data analysis methods* (Second ed.). Thousand Oaks, CA: Sage Publications.

Remillard, J. T. (2000). Can curriculum materials support teachers' learning? Two fourth-grade teachers' use of a new mathematics text. *The Elementary School Journal, 100*(4), 331-350.

Schoen, H. L., Bean, D. L., & Ziebarth, S. (1996). Embedding communication throughout the curriculum. In P. C. Elliot & M. J. Kenny (Eds.), *Communication in Mathematics: K-12 and Beyond* (pp. 170-179). Reston, VA: National Council of Teachers of Mathematics.

Schoen, H. L., Cebulla, K. J., Finn, K. F., & Fi, C. (2003). Teacher variables that relate to student achievement when using a standards-based curriculum. *Journal for Research in Mathematics Education, 34*(3), 228-228.

Schoen, H. L., Cebulla, K. J., & Winsor, M. S. (2001). *Preparation of students in a standards-oriented mathematics curriculum for college entrance tests, placement tests, and beginning mathematics courses.* Paper presented at the Annual Meeting of the American Educational Research Association, Seattle, WA.

Schoen, H. L., Fey, J. T., Hirsch, C. R., & Coxford, A. F. (1999). Issues and options in the math wars. *Phi Delta Kappan, 80*(6), 444-453.

Schoen, H. L., Finn, K. F., Griffin, S. F., & Fi, C. (2001). *Teacher variables that relate to student achievement in a standards-oriented curriculum.* Paper presented at the The Annual Meeting of the American Educational Research Association, Seattle, WA.

Schoen, H. L., & Hirsch, C. R. (2003a). Responding to calls for change in high school mathematics: Implications for collegiate mathematics. *The American Mathematical Monthly, 110*(2), 109-120.

Schoen, H. L., & Hirsch, C. R. (2003b). The Core-Plus Mathematics Project: Perspectives and student achievement. In S. L. Senk & D. R. Thompson (Eds.), *Standards-Based School Mathematics Curricula: What Are They? What Do Students Learn?* (pp. 311-344). Hillsdale, NJ: Lawrence Earlbaum Associates, Inc.

Schoen, H. L., Hirsch, C. R., & Ziebarth, S. W. (1998). *An emerging profile of the mathematics achievement of students in the Core-Plus Mathematics Project.* Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Schoen, H. L., & Pritchett, J. (1998). *Students' perceptions and attitudes in a standards-based high school mathematics curriculum.* Paper presented at the Annual Meeting of the American Educational Research Association, San Diego, CA.

Schon, D. A. (1987). *Educating the reflective practitioner*: Jossey-Bass San Francisco.

Scott, D. (1994). *"Fidelity" mentality is no reality: A message for curriculum developers* Paper presented at the Australian Association for Research in Education, Newcastle, Australia.

Senk, S., & Thompson, D. (2003). *Standards-based school mathematics curricula: What are they? What do students learn?* Mahway, NJ: Lawrence Erlbaum.

Sherin, M. G. (2002). A balancing act: Developing a discourse community in a mathematics classroom. *Journal of Mathematics Teacher Education, 5*(3), 205-233.

Showers, B., & Joyce, B. (1996). The evolution of peer coaching. *Educational Leadership, 53*, 12-16.

Singer, J. D. (1998). Using SAS PROC MIXED to fit multilevel models, hierarchical models, and individual growth models. *Journal of Educational and Behavioral Statistics, 23*(4), 323.

Stein, C. C. (2007). Let's talk: Promoting mathematical discourse in the classroom. *Mathematics Teacher, 101*(4), 285-289.

Stein, M. K., Remillard, J., & Smith, M. S. (2007). How curriculum influences student learning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning* (Vol. 1, pp. 319-369).

Stein, M. K., & Smith, M. S. (1998). Mathematical tasks as a framework for reflection: From research to practice. *Mathematics Teaching in the Middle School, 3*(4), 268-275.

Stein, M. K., & Smith, M. S. (2010). The influence of curriculum on students' learning. In B. J. Reys, R. E. Reys & R. Rubenstein (Eds.), *Mathematics Curriculum: Issues, Tends, and Future Directions*. Reston, VA: National Council for Teachers of Mathematics

Strategic Ed Solutions. (2008) Retrieved October 15, 2008, from http://www.strategicedsolutions.org/programs_submitted/1222195514/Curriculum-CorePlusMathematics.pdf

Strauss, A., & Corbin, J. (1990). Open coding. *Basics of qualitative research: Grounded theory procedures and techniques*, 101-121.

Swafford, J. (2003). Reaction to high school curriculum projects research. In S. L. Senk & D. R. Thompson (Eds.), *Standards-Based School Mathematics Curricula: What Are They? What Do Students Learn?* (pp. 457). Mahwah, NJ: Lawrence Erlbaum Associates.

Tarr, J. E., Ross, D., McNaught, M. D., Chávez, O., Grouws, D. A., Reys, R., et al. (2010, April). *Identification of student– and teacher–level variables in modeling variation of mathematics achievement data.* Paper presented at the Annual Meeting of the American Education Research Association, Denver.

Taylan, R. D., & Tarr, J. E. (2011). *On curricular effectiveness: Professional development activities, emphasis, and impact*. Paper presented at the Poster presented at the

Research Presesssion of the Annual Meeting of the National Council of Teachers of Mathematics, Indianapolis, IN.

Thomas, S. M. (2011). *Implementing reform-based mathematics instruction after participating in professional development: Two high school teachers' experiences*. Unpublished doctoral dissertation, North Carolina State University, Raleigh, N.C. Retrieved from http://www.lib.ncsu.edu/resolver/1840.16/6678

Thompson, D. R., & Senk, S. (2010). Myths about curriculum implementation. In B. J. Reys, R. E. Reys & R. Rubenstein (Eds.), *Mathematics Curriculum: Issues, Tends, and Future Directions*. Reston, VA: National Council for Teachers of Mathematics

Tyack, D., & Cuban, L. (1995). *Tinkering toward utopia: A century of public school reform*. Cambridge, M.A.: Harvard University Press.

United States Department of Education. (1999) Retrieved December 2, 2008, from http://www.ed.gov/PressReleases/10-1999/mathpanel.html

White, D. (2003). Promoting productive mathematical classroom discourse with diverse students. *The Journal of Mathematical Behavior, 22*(1), 37-53.

Wilson, M. S., & Lloyd, G. M. (2000). Sharing mathematical authority with students: The challenge for high school teachers. *Journal of Curriculum and Supervision, 15*(2), 146-169.

Wilson, S. M., & Berne, J. (1999). Teacher learning and the acquisition of professional knowledge: An examination of research on contemporary professlonal development. *Review of Research in Education, 24*(1), 173.

Ziebarth, S. W. (2001). PRIME-TEAM: The University of Iowa Local Systemic Change Project. *Mathematics Teacher, 94*(4), 336.

Ziebarth, S. W. (2003). *A Report on advances in secondary mathematics curriculum development in the United States and imminent new directions: Core-Plus Mathematics as a case study.* Paper presented at the Hawaii International Conference on Education, Honolulu, HI.

APPENDICES

Appendix A
Data Sources and Analysis Used to Answer the Research Questions

Research Question	**Data Used**	**Data Analysis**
Q1a (Traditional vs. CPMP) Is there evidence of differential impact on student performance on state tests between teachers using integrated curricular materials and those who use subject-specific curricula across the state of North Carolina?	Spring 2010 Algebra I and II EOC scores	Three-level HLM model with students nested within teachers, accounting for student background
Q1b (Workshop: Traditional vs. CPMP) For teachers who participated in the project workshops, is there evidence of differential impact on student performance on state tests between teachers using *Core-Plus* and subject-specific curricula?	Spring 2010 Algebra I and II EOC scores Teacher Mathematics Content Knowledge Teacher Background	Three-level HLM model with students nested within teachers, accounting for student background and teacher characteristics

Q2a (CPMP implementers) Among teachers using the *Core-Plus* curricular materials, what evidence of systematic differences is there on the effects of student learning based on varying levels of teacher participation in the NCIM project?	Spring 2010 Algebra I and II EOC scores Teacher Mathematics Content Knowledge Teacher Background Teacher Beliefs Table-of-Contents Logs	Three-level HLM model with students nested within teachers, accounting for student background, teacher background, teacher beliefs, and textbook implementation
Q2b (Variance in CPMP implementation) Among teachers using the *Core-Plus* curricular materials, to what extent and how can the variance in student outcomes be explained by teacher implementation?	Spring 2010 Algebra I and II EOC scores Teacher Mathematics Content Knowledge Teacher Background and Beliefs Table-of-Contents Logs Teacher Interviews Classroom Observations Instructional coach Reports	A multiple comparisons test based on pair-wise ranking, followed by grounded approach to analyzing transcripts and observation reports to identify broad categories

Appendix B

Example of a TOC Record

INSTRUCTIONS

This instrument is designed to document what mathematics content you teach this year from the *Core-Plus, Course 2* textbook. Its design is similar to the format of the Table of Contents of your textbook. That is, it is organized by investigations within lessons within units, with each page displaying the contents of one textbook unit.

We ask that you complete the survey for the first class of the day in which you teach from the *Core-Plus, Course 2* textbook, unless that class is an atypical class (e.g., accelerated, remedial, etc.). For each investigation, select the option that best describes the extent to which the content was taught (if at all) from your textbook: (a) Taught primarily from *Core-Plus* textbook, (b) Taught from *Core-Plus* textbook with some supplementation, (c) Taught primarily from alternative(s) to *Core-Plus*, and (d) Did not teach content.

We ask that you update the survey at the end of each quarter and at that time send us a photocopy of the survey using one of the four postage paid envelopes provided.

Please fill in your name and the school at which you teach:

Name: _____

School: _____

Choose the option that best describes the extent to which the mathematics content of each lesson was taught (if at all) from your *Core-Plus, Course 2* textbook.

Unit 1 **Matrix Models**		Taught primarily from *Core-Plus* textbook	Taught from *Core-Plus* textbook with some supplementation	Taught primarily from alternative(s) to *Core-Plus*	Did not teach content
Lesson 1 *Building and Using Matrix Models*					
Inv 1	*There's No Business Like Shoe Business*	☐	☐	☐	☐
Inv 2	*Analyzing Matrices*	☐	☐	☐	☐
Inv 3	*Combining Matrices*	☐	☐	☐	☐
Lesson 2 *Multiplying Matrices*					
Inv 1	*Brand Switching*	☐	☐	☐	☐
Inv 2	*More Matrix Multiplication*	☐	☐	☐	☐
Inv 3	*The Power of a Matrix*	☐	☐	☐	☐
Inv 4	*Properties of Matrices*	☐	☐	☐	☐
Lesson 3 *Matrices and Systems of Linear Equations*					
Inv 1	*Smart Promotions, Smart Solutions*	☐	☐	☐	☐
Inv 2	*Comparing Solution Methods*	☐	☐	☐	☐
Lesson 4 *Looking Back*		☐	☐	☐	☐

Choose the option that best describes the extent to which the mathematics content of each lesson was taught (if at all) from your *Core-Plus, Course 2* textbook.

Unit 2 **Patterns of Location, Shape, and Size**		Taught primarily from *Core-Plus* textbook	Taught from *Core-Plus* textbook with some supplementation	Taught primarily from alternative(s) to *Core-Plus*	Did not teach content
Lesson 1 *A Coordinate Model of a Plane*					
Inv 1	*Plotting Polygons and Computing Distances*	☐	☐	☐	☐
Inv 2	*Things Are Not Always What They Seem to Be*	☐	☐	☐	☐
Inv 3	*Families of Lines*	☐	☐	☐	☐
Lesson 2 *Coordinate Models of Transformations*					
Inv 1	*Modeling Rigid Transformations*	☐	☐	☐	☐
Inv 2	*Modeling Size Transformations*	☐	☐	☐	☐
Inv 3	*Modeling Combinations of Transformations*	☐	☐	☐	☐
Lesson 3 *Transformations, Matrices, and Animation*					
Inv 1	*Spinning Flags*	☐	☐	☐	☐
Inv 2	*Stretching and Shrinking*	☐	☐	☐	☐
Lesson 4 *Looking Back*		☐	☐	☐	☐

Choose the option that best describes the extent to which the mathematics content of each lesson was taught (if at all) from your *Core-Plus, Course 2* textbook.

Unit 3 **Patterns of Association**		Taught primarily from *Core-Plus* textbook	Taught from *Core-Plus* textbook with some supplementation	Taught primarily from alternative(s) to *Core-Plus*	Did not teach content
Lesson 1 *Seeing and Measuring Association*					
Inv 1	*Rank Correlation*	☐	☐	☐	☐
Lesson 2 *Correlation*					
Inv 1	*Pearson's Correlation Coefficient*	☐	☐	☐	☐
Inv 2	*Association and Causation*	☐	☐	☐	☐
Lesson 3 *Least Squares Regression*					
Inv 1	*How Good is the Fit?*	☐	☐	☐	☐
Inv 2	*The "Best-Fitting" Line*	☐	☐	☐	☐
Lesson 4 *Looking Back*		☐	☐	☐	☐

Choose the option that best describes the extent to which the mathematics content of each lesson was taught (if at all) from your *Core-Plus, Course 2* textbook.

Unit 4 **Power Models**		Taught primarily from *Core-Plus* textbook	Taught from *Core-Plus* textbook with some supplementation	Taught primarily from alternative(s) to *Core-Plus*	Did not teach content
Lesson 1 *Same Shape, Different Size*					
Inv 1	*Starting from Cube One*	☐	☐	☐	☐
Inv 2	*The Shape of $y = ax^2$ and $y = ax^3$*	☐	☐	☐	☐
Lesson 2 *Inverse Variation*					
Inv 1	*Travel Times*	☐	☐	☐	☐
Inv 2	*Sound and Light*	☐	☐	☐	☐
Inv 3	*The Shape of Inverse Variation Models*	☐	☐	☐	☐
Lesson 3 *Quadratic Models*					
Inv 1	*Going Up… Going Down*	☐	☐	☐	☐
Inv 2	*Profit Prospects*	☐	☐	☐	☐
Inv 3	*The Shape of Quadratic Models*	☐	☐	☐	☐
Inv 4	*Solving Quadratic Equations*	☐	☐	☐	☐
Inv 5	*How Many Solutions Are Possible?*	☐	☐	☐	☐

Choose the option that best describes the extent to which the mathematics content of each lesson was taught (if at all) from your *Core-Plus, Course 2* textbook.

Unit 4 **Radicals and Fractional Power Models**		Taught primarily from *Core-Plus* textbook	Taught from *Core-Plus* textbook with some supplementation	Taught primarily from alternative(s) to *Core-Plus*	Did not teach content
Lesson 4 *Quadratic Models*					
Inv 1	*The Power of a Brace*	☐	☐	☐	☐
Inv 2	*Powerful Radicals*	☐	☐	☐	☐
Inv 3	*Cube Roots*	☐	☐	☐	☐
Inv 4	*Operating with Powers*	☐	☐	☐	☐
Lesson 5 *Looking Back*		☐	☐	☐	☐

Choose the option that best describes the extent to which the mathematics content of each lesson was taught (if at all) from your *Core-Plus, Course 2* textbook.

Unit 5 **Network Optimization**		Taught primarily from *Core-Plus* textbook	Taught from *Core-Plus* textbook with some supplementation	Taught primarily from alternative(s) to *Core-Plus*	Did not teach content
Lesson 1 *Finding the Best Networks*					
Inv 1	*Optimizing a Computer Network*	☐	☐	☐	☐
Inv 2	*Optimizing a Road Network*	☐	☐	☐	☐
Lesson 2 *Shortest Paths and Circuits*					
Inv 1	*Shortest Routes*	☐	☐	☐	☐
Inv 2	*Graph Games*	☐	☐	☐	☐
Inv 3	*The Traveling Salesperson Problem*	☐	☐	☐	☐
Lesson 3 *Looking Back*		☐	☐	☐	☐

Choose the option that best describes the extent to which the mathematics content of each lesson was taught (if at all) from your *Core-Plus, Course 2* textbook.

Unit 6 **Geometric Form and Its Function**		Taught primarily from *Core-Plus* textbook	Taught from *Core-Plus* textbook with some supplementation	Taught primarily from alternative(s) to *Core-Plus*	Did not teach content
Lesson 1 *Flexible Quadrilaterals*					
Inv 1	*Using Quadrilaterals in Linkages*	☐	☐	☐	☐
Inv 2	*Linkages and Similarity*	☐	☐	☐	☐
Lesson 2 *Triangles and Trigonometric Ratios*					
Inv 1	*Triangles with a Variable-Length Side*	☐	☐	☐	☐
Inv 2	*What's the Angle?*	☐	☐	☐	☐
Inv 3	*Measuring Without Measuring*	☐	☐	☐	☐
Lesson 3 *The Power of the Circle*					
Inv 1	*Follow That Driver!*	☐	☐	☐	☐
Inv 2	*Radian Measure*	☐	☐	☐	☐
Inv 3	*Modeling Circular Motion*	☐	☐	☐	☐
Inv 4	*Patterns of Periodic Change*	☐	☐	☐	☐
Lesson 4 *Looking Back*		☐	☐	☐	☐

Choose the option that best describes the extent to which the mathematics content of each lesson was taught (if at all) from your *Core-Plus, Course 2* textbook.

Unit 7 **Patterns in Chance**		Taught primarily from *Core-Plus* textbook	Taught from *Core-Plus* textbook with some supplementation	Taught primarily from alternative(s) to *Core-Plus*	Did not teach content
Lesson 1 *Waiting Times*					
Inv 1	*Waiting for Doubles*	☐	☐	☐	☐
Inv 2	*Independent Trials*	☐	☐	☐	☐
Inv 3	*The Distribution of Waiting Times*	☐	☐	☐	☐
Lesson 2 *The Multiplication Rule*					
Inv 1	*Multiplying Probabilities*	☐	☐	☐	☐
Lesson 3 *Probability Distributions*					
Inv 1	*Theoretical Waiting-Time Distributions*	☐	☐	☐	☐
Inv 2	*Probability Distributions and Rare Events*	☐	☐	☐	☐
Lesson 4 *Expected Value of Probability Distribution*					
Inv 1	*What's a Fair Price?*	☐	☐	☐	☐
Inv 2	*Fair Price and Expected Value*	☐	☐	☐	☐
Inv 3	*Expected Value of a Waiting-Time Distribution*	☐	☐	☐	☐
Lesson 5 *Looking Back*		☐	☐	☐	☐

Choose the option that best describes the extent to which the mathematics content of each lesson was taught (if at all) from your *Core-Plus, Course 2* textbook.

Capstone **Looking Back at Course 2**		Taught primarily from *Core-Plus* textbook	Taught from *Core-Plus* textbook with some supplementation	Taught primarily from alternative(s) to *Core-Plus*	Did not teach content
Lesson *Forests, the Environment, and Mathematics*					
Inv 1	*Forestry and Mathematics*	☐	☐	☐	☐
Inv 2	*Land Use Change in Rapid-Growth Areas*	☐	☐	☐	☐
Inv 3	*Valuing Urban Trees*	☐	☐	☐	☐
Inv 4	*Forests, the Greenhouse Effect, and Global Warming*	☐	☐	☐	☐
Inv 5	*Measuring Trees*	☐	☐	☐	☐
Inv 6	*Producing Wood Products*	☐	☐	☐	☐
Inv 7	*Geographic Information Systems (GIS)*	☐	☐	☐	☐
Inv 8	*Further Analysis*	☐	☐	☐	☐
Reports: Putting It All Together		☐	☐	☐	☐

Appendix C
Example of Content Knowledge Assessment

Last Four Digits of Your Social Security Number: _____

NCIM Content Knowledge Assessment

*Directions: Solve the following problems. Show your work. You may use a calculator. Please turn in the additional graph paper that you use. If you have never **taught** the topic of a problem, indicate so by writing "NT" in the left margin next to the problem.*

1. The City Water Company commission has established water rates as follows: The Water Company charges their customers $3.75 for the first 1/3 gallon of water, and $.15 for each additional 1/3 gallon of water.

    a. If Darrius used 50 gallons of water in January, how much would his water bill cost?

    b. If he spent $39.60 in December, approximately how many gallons of water did Darrius use?

2. At a baseball game, Rory bought 5 bags of peanuts and 3 soft pretzels for $23.70. At the same game, Leslie bought 2 bags of peanuts and 4 soft pretzels for $21.10. How much does the combination of 1 bag of peanuts and 1 soft pretzel cost?

3. Simplify the following expression using positive exponents.

$$(-3x^{9}y^{2})^{-2}(5x^{-2}y^{4})$$

4. Terrance graphed a quadratic function with zeros at 0 and -4.

    a. Write a possible equation for Terrance's function.

    b. Given your equation in part a., identify its domain and range.

c. What transformations of this equation also go through these two zeros of the function?

5. There is a stream that is 7 feet wide. Jay is standing on one side of the stream. The other side is 5 feet higher than where Jay is standing. Can Jay cross the stream using a 12-foot ladder? Please explain your reasoning.

6. The Aquatics Club is considering selling weekly swimming passes. A market research firm has concluded that the number of weekly passes *n* they will sell is related to the price per pass *p* in dollars by the formula $n = -200p + 850$.

   a. Explain why the income *I* from sales of weekly swimming passes can be determined by the formula $I = -200p^2 + 850p$.

   b. Will the Aquatics Club always make a profit when selling swimming passes? Explain your reasoning

   c. For what price(s) will the income from weekly swimming passes be $500? Show your work or explain your reasoning.

7. Three functions f(x) and their parent functions g(x) are given below. For each pair of g(x) and f(x), describe the transformations of the graph of g(x) that result in the graph of f(x):

   a. $f(x) = x^2 + 5$

   b. $f(x) = \sqrt{x - 4}$

   c. $f(x) = -|2x|$

8. Kenny collected data about a person's average daily caloric intake and their weight in pounds. The least squares regression line and correlation coefficient are given below. Use them to answer the following questions.

x, average daily caloric intake	2000	1567	2342	3450	1247	2876	1430	900
y, weight in pounds.	180	156	182	275	116	231	132	75

$y = .07x + 25.02$

$R^2 = .97$

   a. If Mark consumes an average of 1600 calories in a day, about how much would you expect him to weigh?

   b. Explain what the .07 stand for in the least squares regression line equation?

   c. Is a linear regression line a good model for this relationship? Explain why or why not.

9. A computer manufacturing company can make a 250-gigabyte (Basic) or 500-gigabyte (Deluxe) hard drive.

   Basic hard drive cost $60 each; each requires 60 minutes to assemble. The basic model can be sold for an $15 profit.

   Deluxe models cost $85 to make; each requires only 45 minutes to assemble. The deluxe player can be sold for a $12 profit.

   If I have $800 to invest, and 450 hours of time, how many of each should I make to maximize my profit?

10. At an amusement park there is a thrilling ride that has a long rotating arm with riders in capsules at each end. The arm is 130 feet long and you get buckled in the capsule that starts at ground level. Assuming simple rotating motion and treating each capsule as a single point at the end of the arm, find your height above the ground when the arm has made the following counterclockwise rotations from its starting vertical position.

    a. 270°

    b. 60°

    c. 115°

Appendix D
Background Form

Background Questionnaire for NC Integrated Mathematics Project

1. Name _____
   First Middle Last Name called by
2. School _____ District _____
3. Email Address most often used _____
4. **Last Four Digits** of Social Security Number _____
5. What is your gender? Female Male
6. How do you identify yourself? American Indian/Alaska Native Asian
   Black/African American Native Hawaiian/Other Pacific Islander White
   Hispanic/Latino Multiracial Other:_____
7. Have you taught *Core-Plus* before? Yes No If yes, how many years?_____
   If yes, what course(s)? *Core-Plus 1* *Core-Plus 2* *Core-Plus 3* *Core-Plus 4*
8. Have you attended a *Core-Plus* training before? Yes No
   If yes, when and where? _____How many days?_____
9. List your majors for all the degrees you have completed.
   Bachelors degree Major_____
   Masters degree Major_____
10. Do you currently hold a North Carolina Teaching License? Yes No
    If yes, what grade level and **subject** is your licensure area? _____
    _____
11. How many years have you been teaching? _____high school _____middle school ____other
12. What courses did you teach during the 2008-2009 school year? _____
    _____
13. What type of daily schedule does your school use? (block, acc. block, 50-min. period) _____
14. Number of: Mathematics Classes taught this year _____ Students taught this year _____
15. How often do you use calculators for instruction? Daily Weekly Monthly Yearly Never
    Do you have a class set? Yes No What type of calculators do you use most often? _____
16. List the 3 most valuable computer-based software or web-based resources you use in your math classes, if you do not use any write *none*? _____
    _____
17. What non-instructional activities do you participate in at your school? _____
    _____

Appendix E
Teacher Beliefs Survey

Please indicate how well prepared you currently feel to do each of the following in your mathematics instruction: 4–very well prepared, 3 – fairly well prepared, 2 – somewhat prepared, 1 – not adequately prepared

		4	3	2	1
1.	Take students' prior understanding into account when planning curriculum and instructions.				
2.	Have students work in cooperative learning groups.				
3.	Use the textbooks as a resource rather than the primary instructional tool.				
4.	Teach groups that are heterogeneous in ability.				
5.	Teach students who have limited English proficiency.				
6.	Encourage participation of females in mathematics.				
7.	Encourage participation of minorities in mathematics.				

About how often do you typically do each of the following in your mathematics instruction in this class? 5 – almost daily, 4 – once or twice a week, 3 – once or twice a month, 2 – a few times a year, 1 – never

		5	4	3	2	1
8.	Arrange seating to facilitate small group work					
9.	Encourage students to use manipulatives in solving a problem					
10.	Ask students to look for alternative methods for solving a problem					
11.	Require students to explain their reasoning when giving an answer					
12.	Allow students to work at their own pace					
13.	Review student notebooks/journals					
14.	Give tests requiring open-ended response (e.g., descriptions, explanations)					
15.	Have students complete performance tasks individually					
16.	Have students complete performance tasks in groups					
17.	Explain new mathematics content to the whole class					
18.	Give short answer tests (e.g., multiple choice, true/false, fill in the blank)					

About how often do students in this mathematics class take part in the following types of activities? 5 – almost daily, 4 – once or twice a week, 3 – once or twice a month, 2 – a few times a year, 1 – never

		5	4	3	2	1
19.	Work on solving real-world problems					
20.	Record, represent, and/or analyze data					
21.	Collaborate with other students in solving a problem					
22.	Write a description of a plan, procedure, or problem-solving process					
23.	Work on extended mathematics investigations or projects (a week or more in duration)					
24.	Make formal presentations about their projects to the rest of the class					
25.	Students write reflections about something they learned					
26.	Students collect their best mathematics work in a portfolio					
27.	Answer textbook/worksheet questions					
28.	Read from a mathematics textbook in class					
29.	Practice mathematics facts, rules, or formulas					

Adapted from: Horizon Research, Inc. (2000). *National survey of science and mathematics education: teacher questionnaire.* Chapel Hill, NC. And McCaffrey, D. F., Hamilson, L. S., Stecher, B. M., Klein, S. P., Bugliari, & Robyn, A. (2001). Interactions among instructional practices, curriculum, and student achievement: The case of standards-based high school mathematics. *Journal for Research in Mathematics Education*, *32*(5), 493-517.

Appendix F

North Carolina Integrated Mathematics Project
Instructional coach Observation Report

Please fill in the template below and save the file as your last name underscore teacher you observed underscore observation number. For example, smith_white_3.doc would be instructional coach Mrs. Smith observing Mr. White for the third time. Email final report to eekrupa@ncsu.edu .

Instructional coach Name:
Teacher Observed:
School:
Date:
Class period and length of period:
Lesson Taught: **Investigation:** **Topics**:
How much time, in minutes, did you spend observing the teacher:
How much time, in minutes, did you spend in follow-up conversation:
Instructional coach activities used during this visit:

Use of Time and Pacing

1. **Was time used effectively for what the class needed?** Rate from 1 (ineffective or inappropriate) to 4 (effective and appropriate)

2. **Is this teacher on pace for the year?**

Instructional Behaviors

3. **How was the class started; did they review or collect homework**?

4. **How did the teacher pose the question/investigation for the day**?

5. W**as there any closure? Please explain**.

6. **Describe the content delivery quality**. (Accuracy, Clear, Understanding, Complete) **List instructional areas of difficulty.**

Collaborative Learning

7. **Are they using small groups?** _____Yes ____No

8. **How have the groups been selected?** _____Student Choice _____Teacher Assignment

9. **What criterion was used to create groups**?

10. **Describe the collaborative learning**?

11. **What was the teacher's role**?

12. **How did groups report back to the class**?

Use of Technology

13. What technology was used in class?

None	Calculator	Computers	Other (please specify)

Comment:

14. In your opinion, how effective was the technology use? (1, not effective, to 4, highly effective)

Formative Assessment

15. What formative assessment methods were used?

16. **Did the teacher recognize when students were lost, un-engaged, or showed evidence of a misunderstanding?**

Frequently	Occasionally	Seldom	Never

Comment:

Management Strategies

17. **Over the whole period percent of the class was typically on-task?**

90-100%	70-90%	50-70%	<50%

Comment:

18. **What strategies did teachers use to respond to the student's off task behaviors?**

Other

19. **Based on the Action Plan the teacher created at the Summer Workshop (if they attended), what was their teaching behavior focus for the lesson?**

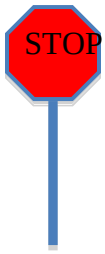
20. **What evidence of progress did the teacher make on their teaching behavior focus during the lesson?**

21. **Any other general observations**?

22. **What assistance do teachers need**? (materials, content support, management)

23. **What challenges is the teacher having to successfully implement the CORE-Plus curriculum?**

24. **What you will report to the teacher** (please provide this below)



Before you save, save the file as your last name underscore teacher you observed underscore observation number. For example, smith_white_3.doc would be instructional coach Mrs. Smith observing Mr. White for the third time. Email final report to eekrupa@ncsu.edu .

Appendix G

Teacher Interview Protocol

Career History:

1. How long have you been teaching at (name of school)?
2. Were you teaching before? Where? How long? What subjects have you taught?

Experiencing mathematics:

1. Please describe your enjoyment of mathematics as a student.
2. What are your feelings towards mathematics now?
3. Can you identify students in your class(es) who you think possess similar feelings? How do you interact with them?
4. How do you think students best learn mathematics?

Instructional Practices:

1. What do you feel is the teacher's role (your role) in the classroom?
2. How do you assign students to work in groups?

Reform Mathematics:

1. What comes to mind when you think about reform mathematics and the Core-Plus curriculum?
2. Please describe your experiences teaching Core-Plus?
3. What are your feelings toward Core-Plus?
4. How comfortable are you with teaching the Core-Plus curriculum?
5. Have you taught a traditional mathematics curriculum before? Describe your experience teaching a traditional mathematics curriculum.
6. Is there anything about Core-Plus that does not fit with your visions of teaching and learning mathematics?
7. Do you feel there are any difference between how students learn mathematics in Core-Plus classrooms and student in traditional mathematics classrooms? If so, what are they?

Implementation

1. Describe how you plan for teaching with Core-Plus?
    a. Do you find yourself planning more (or less) than you have in the past when you taught using a different curriculum?
    b. What types of things do you think about when you're planning?
    c. How do you decide which investigations/lessons to teach?
    d. Are there lessons that you've found more difficult than others to implement? Would you provide examples?
    e. Do you collaborate with other teachers when planning? If so, can you describe how that came about, and what your collaboration involves?

2. How are students responding to Core-Plus?
    a. Do you think there's an adjustment period? If so, please describe why you think it occurs, and how long it typically lasts?
    b. What do students have to get used to?

3. How are parents responding to Core-Plus?

4. How comfortable are you teaching Core-Plus?
    a. What do you think is necessary to increase your comfort with the curriculum?
    b. Have you encountered any hindrances? If so, please describe them.
    c. Have you ever struggled teaching with the content in any Core-Plus lesson or investigation? If so, how did you work through it?
    d. Do you use other materials besides the Core-Plus curriculum materials?
        i. If so, which ones?
        ii. Why do you feel like these materials are necessary? What's your motivation for their use?

5. How often do students talk about mathematics in your class(es)?
    a. Do you try to get students to talk more about mathematics? How? Can you describe that process?
    b. What types of norms did you establish at the beginning of the school year?
    c. How did the students respond?

Appendix H
HLM Model with Predictors at Level One, Two, and Three

Level-1 Model:
$$A\lg 1score_{ijk} = \pi_{0jk} + \pi_{1jk}(Grade_{ijk} - \overline{X}_1) + \pi_{2jk}(Daysabs_{ijk} - \overline{X}_2) + \pi_{3jk}(Gender_{ijk} - \overline{X}_3) + \pi_{4jk}(AmIn_{ijk} - \overline{X}_4) + \pi_{5jk}(Asian_{ijk} - \overline{X}_5) + \pi_{6jk}(Hispanic_{ijk} - \overline{X}_6) + \pi_{7jk}(Black_{ijk} - \overline{X}_7) + \pi_{8jk}(Multi_{ijk} - \overline{X}_8) + e_{ijk}$$

Level-2 Model:
$$\pi_{0jk} = \beta_{00k} + \beta_{01k}(Title - \overline{X}_9) + r_{0jk}$$
$$\pi_{1jk} = \beta_{10k}$$
$$\pi_{2jk} = \beta_{20k}$$
$$\pi_{3jk} = \beta_{30k}$$
$$\pi_{4jk} = \beta_{40k}$$
$$\pi_{5jk} = \beta_{50k}$$
$$\pi_{6jk} = \beta_{60k}$$
$$\pi_{7jk} = \beta_{70k}$$
$$\pi_{8jk} = \beta_{80k}$$

Level-3 Model:
$$\beta_{00k} = \gamma_{000} + \gamma_{001}(ADM - \overline{X}_{10}) + \gamma_{002}(NeedyPercent - \overline{X}_{11}) + \gamma_{003}(TitleOne - \overline{X}_{12}) + u_{00k}$$
$$\beta_{01k} = \gamma_{010}$$
$$\beta_{10k} = \gamma_{100}$$
$$\beta_{20k} = \gamma_{200}$$
$$\beta_{30k} = \gamma_{300}$$
$$\beta_{40k} = \gamma_{400}$$
$$\beta_{50k} = \gamma_{500}$$
$$\beta_{60k} = \gamma_{600}$$
$$\beta_{70k} = \gamma_{700}$$
$$\beta_{80k} = \gamma_{800}$$

Appendix I
Coding of Variables

Integrated Mathematics 1, Subject-Specific 0

Male 1, Female 0

NCNSP 1, Non-NCNSP 0

Title One 1, Non-Title One 0

Licensure in 6-12 Mathematics 1, No licensure 0

Mathematics or Mathematics Education Major 1, Non-mathematics major 0

Male Teacher 1, Female Teacher 0

Appendix J
Permission to Use NRC Figure

THE NATIONAL ACADEMIES PRESS

Marketing Department
Rights & Permissions

April 27, 2011 **Reference #: 04271100**

Erin Krupa
The Friday Institute for Educational Innovation
Campus Box 7249
1890 Main Campus Drive
Raleigh, NC 27606

Dear Ms. Krupa:

You have requested permission to reprint the following material copyrighted by the National Academy of Sciences in a dissertation:

Figure 3-1, ***On Evaluating Curricular Effectiveness: Judging the Quality of K-12 Mathematics Evaluations, 2004***

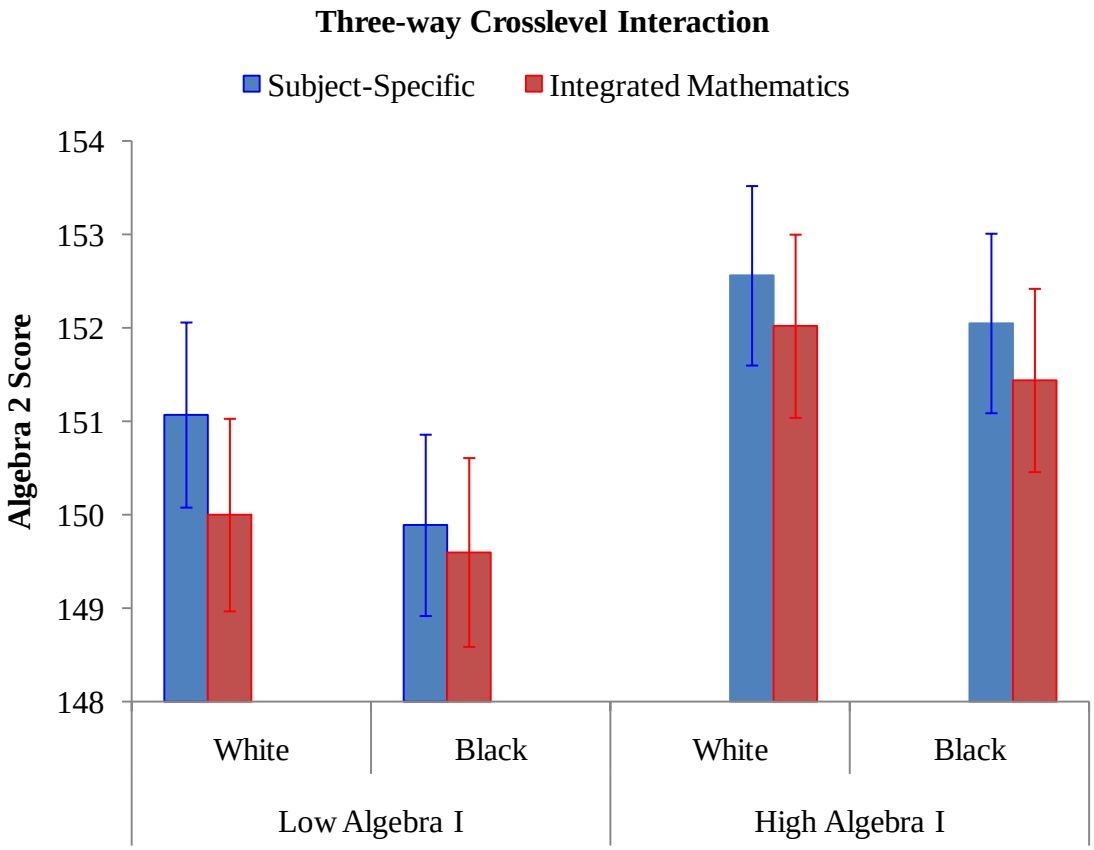
Your request is granted for the material cited above provided that credit is given to the copyright holder.

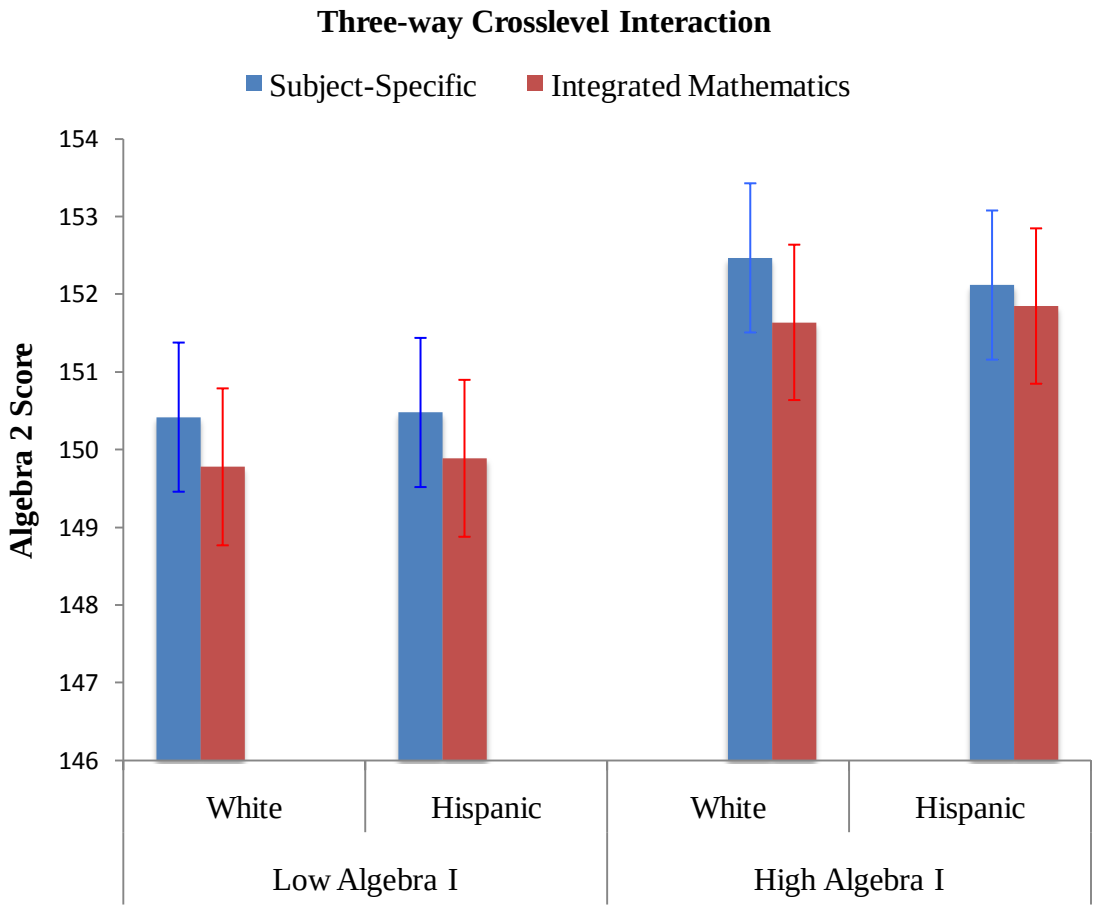
Suggested Credit (example):
Reprinted with permission from (title), (year) by the National Academy of Sciences, Courtesy of the National Academies Press, Washington, D.C. (This credit may be edited pursuant to the publisher's house style and format so long as the essential elements are included).

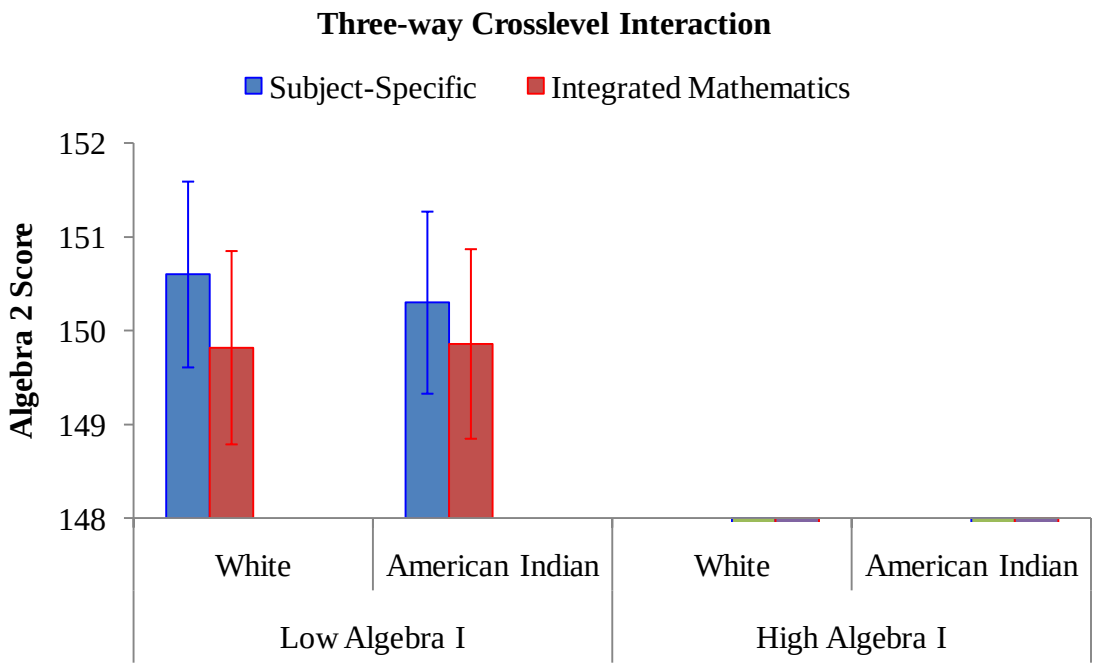
Thank you,

Barbara Murphy
Permissions Coordinator
National Academies Press

Appendix K
Interaction Between Race, Prior Achievement, Curriculum, and Algebra II Score




Three-way Crosslevel Interaction
Subject-Specific
Integrated Mathematics
Algebra 2 Score
154
153
152
151
150
149
148
147
146
White
Hispanic
White
Hispanic
Low Algebra I
High Algebra I


Three-way Crosslevel Interaction
Subject-Specific
Integrated Mathematics
Algebra 2 Score
152
151
150
149
148
White
American Indian
White
American Indian
Low Algebra I
High Algebra I