

1-1-2015

# Relationships between the Algebraic Performance of Students in Subject-Specific and Integrated Course Pathways

Derrick Saddler

*University of South Florida*, dsaddler@usf.edu

Follow this and additional works at: <http://scholarcommons.usf.edu/etd>

 Part of the [Curriculum and Instruction Commons](#), and the [Science and Mathematics Education Commons](#)

---

## Scholar Commons Citation

Saddler, Derrick, "Relationships between the Algebraic Performance of Students in Subject-Specific and Integrated Course Pathways" (2015). *Graduate Theses and Dissertations*.  
<http://scholarcommons.usf.edu/etd/5571>

This Dissertation is brought to you for free and open access by the Graduate School at Scholar Commons. It has been accepted for inclusion in Graduate Theses and Dissertations by an authorized administrator of Scholar Commons. For more information, please contact [scholarcommons@usf.edu](mailto:scholarcommons@usf.edu).

Relationships between the Algebraic Performance of Students in Subject-Specific and Integrated  
Course Pathways

by

Derrick Saddler

A dissertation submitted in partial fulfillment  
of the requirement for the degree of  
Doctor of Philosophy  
in Curriculum and Instruction with an emphasis in  
Mathematics Education  
Department of Teaching and Learning  
College of Education  
University of South Florida

Co-Major Professor: Gladis Kersaint, Ph.D.  
Co-Major Professor: Jeffrey Kromrey, Ph.D.  
Denisse R. Thompson, Ph.D.  
Sarah Van Ingen, Ph.D.

Date of Approval:  
April 9, 2015

Keywords: Content Organization, Algebra, Subject-Specific, Integrated, Course Pathways

Copyright © 2015, Derrick Saddler

## **Acknowledgements**

It is with sincere regards that I acknowledge those who played a major role in the completion of my doctoral journey. First and foremost, I give honor to God for giving me the strength to complete this journey. I am a testimony that all things can be done through God who strengthens us. Therefore, it is God who gets the Glory.

I also want to thank the member of my dissertation committee, Dr. Gladis Kersaint, Dr. Jeffrey Kromrey, Dr. Denisse Thompson, and Dr. Sarah Van Ingen. Although each of you has your individual personalities and research focus, I believe the committee, as a whole is a representation of the individual I have become. Thank you all for requiring nothing but excellence from me. In particular to Dr. Kersaint, one of my co-major professors, you have been there every step of the way. However, from my first course with you until the final defense, your expectations and the role you played were very clear and consistent, if I was going to finish this journey, it will not be because you held my hand and walked me through it. For that, I will be forever grateful. Not to be outdone, Dr. Kromrey once you became one of my co-major professors, I found that your expectations and the role you played were consistent with Dr. Kersaint's. You both are the best! Thank you! In addition to my committee, I would like to thank all my instructors and colleagues whom I encountered and had the opportunity to work with throughout my doctoral studies.

Last and most definitely not least, I must acknowledge my wife, Vivica. Vivica you have been wonderfully supportive and understanding throughout this journey. For this reason, there was never a doubt in my mind that I would not finish. Thank you for being a great person, and I am blessed and highly favored to have you as my wife. I love you very much. I pray that in God's will, he continue to bless our family.

## Table of Contents

List of Tables .....	iii
List of Figures .....	iv
Abstract .....	v
Chapter 1: Introduction .....	1
Statement of the Problem .....	2
Purpose of the Study .....	3
Research Question and Hypotheses .....	4
Brief Review of Literature .....	5
Significance of the Study .....	7
Definition of Terms .....	9
Chapter 2: Studies Related to Content Organization and Students' Learning .....	11
Comparing Options in School Mathematics: Investigating Curricula Project .....	12
Algebra I and Integrated I .....	15
Geometry and Integrated II .....	16
Algebra II and Integrated III .....	17
Summary of Findings from COSMIC Project .....	18
Critique of the COSMIC Project and Implications for Research .....	19
Integrated and Subject-Specific Curricula as a Holistic Body of Work .....	20
Analysis of Specific Content Topics or Test Items .....	21
Methodological Shortcomings .....	22
A Longitudinal Study to Examine Cognitive Growth .....	24
Summary .....	28
Chapter 3: Data and Method .....	29
Data Source: High School Longitudinal Study .....	30
Sample Population .....	31
Mathematics Assessment in Algebraic Reasoning .....	32
Theoretical Framework: Counterfactual Model of Causal Inference .....	34
Statistical Assumptions and Complex Survey Samples .....	39
Chapter 4: Results and Interpretations .....	42
Chapter 5: Discussion and Suggestions for Future Research .....	50
Suggestions for Future Research .....	51
Related to the COSMIC Project .....	51
Related to Algebra and College Readiness .....	52

Investigation of Other Course Pathways.....	53
Investigation of Various Student Groups.....	53
Related to Data Accuracy and Validity .....	54
Related to Curriculum Materials.....	54
References.....	56
Appendix A: List of variables used from HSLs dataset.....	62
Appendix B: SAS Syntax Code .....	66

## List of Tables

Table 1. Students' Demographics Before and After Matching (Percentage of Sample in Parentheses) .....	36
Table 2. Students' Mean Socioeconomic Status, Weight, and Propensity Score Before and After Matching (Standard Deviation in Parentheses) .....	37
Table 3. Students' Means 9 <sup>th</sup> Grade Performance Before and After Matching (Standard Deviation in Parentheses) .....	37
Table 4. Statistics of Students' Mean Performance on Measure with Pretest and Posttest Scores .....	42

## List of Figures

Figure 1. Students' growth on the overall algebra assessment .....	43
Figure 2. Students' proficiency growth on content related to algebraic expressions .....	44
Figure 3. Students' proficiency growth on content related to multiplicative and proportional thinking .....	45
Figure 4. Students' proficiency growth on content related to linear equivalents .....	46
Figure 5. Students' proficiency growth on content related to systems of equations .....	47
Figure 6. Students' proficiency growth on content related to linear functions .....	48



## **Abstract**

The purpose of this study was to compare the algebraic performance gains of high school students who enroll in an integrated mathematics course pathway (i.e., Integrated Mathematics I-II-III) to the algebraic performance gains of high school students who enroll in a subject-specific course pathway (i.e., Algebra I-Geometry-Algebra II). Several studies have been performed in which researchers examined relationships between mathematics outcomes and the course-taking patterns of high school students enrolled in subject-specific course pathways. However, there is little extant research in which researchers have investigated effects of content organization on students' learning and achievement. Therefore, this study addresses calls for more studies that examine the high school mathematics performance of students who learn from subject-specific and integrated course pathways. Data from a large scale observational study known as the High School Longitudinal Study of 2009 was used to compare relationships between the course pathways and students' performance on an assessment of algebraic skills. A pretest-posttest study design was used to statistically compare gain scores of high school students who learn from subject-specific course pathways to the gain scores of a comparable group of high school students who learn from integrated course pathways. Propensity score matching was used to reduce the threat of selection bias due to nonrandom assignment. The results revealed no statistical differences exist in the algebraic performance gains between high school students who learn mathematics from integrated course pathways and high school students who learn from subject-specific course pathways. Suggestions for future research are discussed.

## **Chapter 1: Introduction**

Many believe the future economic competitiveness of the United States depends upon strengthening students' skills in science, technology, engineering, and mathematics (STEM). If students' aspirations are to pursue a bachelor's degree in a STEM discipline, then students must successfully pass required college courses, such as College Algebra, Pre-Calculus, and Calculus. The need to pass College Algebra in order to successfully pursue a bachelor's degree in a STEM discipline demonstrates the importance of algebra and the role it plays as a gatekeeper to higher level mathematics, and in turn careers in a STEM discipline. Because of its fundamental role for exploring most areas of mathematics, science, and engineering, algebra readiness is characterized as one of the most important "gatekeepers" in school mathematics (Cai et al., 2011).

Research suggests more than 33% of students entering college are not ready to take College Algebra (ACT, 2012; Johnson, 2006). Even more, there is a significant number of students who take courses such as Calculus while in high school, but do not demonstrate readiness for College Algebra (ACT, 2012). As a result, many unprepared students are required to enroll in developmental or remedial courses upon entering college to meet the College Algebra benchmark. These developmental courses include Intermediate Algebra and Introductory Algebra. Intermediate Algebra is the prerequisite course to College Algebra, and Introductory Algebra is the prerequisite course to Intermediate Algebra. The majority of the content taught in Introductory Algebra and Intermediate Algebra courses is taught in a high school Algebra I course or earlier (ACT, 2012). In addition, many of the topics rated as most

important by instructors of College Algebra and subsequent advanced mathematics courses are typically covered in high school Algebra II or earlier (ACT, 2012). In other words, by the time high school students have completed the traditional Algebra I-Geometry-Algebra II course sequence, it is expected that they have had an opportunity to learn the content that is taught in College Algebra, Intermediate Algebra, and Introductory Algebra courses. If high school students who enroll in the traditional Algebra I-Geometry-Algebra II course sequence have had the opportunity to learn content taught up to a College Algebra course, but are still required to take remedial courses when they enroll in college, then it should not be a surprise there are growing concerns about United States students' inadequate preparation in algebra. As a result, algebra curricula content is a focal point of mathematics education research, and in turn a focal point of this study.

### **Statement of the Problem**

The primary goal of a high school education is to ensure students are prepared and ready for college coursework and/or the workforce. In an effort to reach this goal and better prepare students for college-level credit-bearing mathematics courses, the recently adopted Common Core State Standards (CCSS) delineates the specific content students should learn while they are in high school (National Governors Association Center for Best Practices [NGA Center] and the Council of Chief State School Officers [CCSSO], 2010a). As educational leaders consider how to implement the Common Core State Standards for Mathematics (CCSSM), an important consideration is how to organize the high school mathematics program into courses that provide a strong foundation for success at the post-secondary level. To assist in addressing this need, a group of mathematics experts, teachers, mathematics faculty from postsecondary institutions, mathematics teacher educators, and workforce representatives were convened to develop model

course pathways in mathematics based on the CCSSM (NGA & CCSSO, 2010b). One of the model course pathways is the U.S. traditional high school mathematics that includes Algebra I-Geometry-Algebra II. The primary content in a subject-specific course pathway is related directly to the name of each course. Researchers refer to this pathway as a *subject-specific* approach (Grouws, Tarr, Chávez, Sears, Soria & Taylan, 2013; Tarr, Grouws, Chávez, & Soria, 2013). Another model course pathway suggested in the CCSSM is Mathematics I-II-III. This course pathway is typically offered in other countries, and in curricula materials developed in response to standards-based reform documents (Senk & Thompson, 2003). The content in each of the integrated mathematics courses includes number, algebra, geometry, probability and statistics. Because multiple mathematical strands are integrated in each course, researchers refer to this pathway as an *integrated mathematics* approach (Grouws et al., 2013; Tarr et al., 2013). The high school portion of the CCSSM can be implemented using either a subject specific or integrated course pathway, and the decision regarding which approach to use is made by state or local education agencies. However, no information is provided as part of the CCSSM document about the advantages or disadvantages of either approach. In addition, there is little research in which researchers compare the mathematics performance of high school students who learn from subject-specific and integrated course pathways (NMAP, 2008). As a result, there are calls for more studies that examine the high school mathematics performance of students who learn from subject-specific and integrated course pathways (Grouws et al., 2013; Tarr et al., 2013). This study addresses these calls.

### **Purpose of the Study**

A significant difference between subject-specific and integrated course pathways is the manner in which the mathematics content is organized. However, regardless of where the

specific mathematics topics are taught in the high school curriculum, an expectation is students should be able to proceed successfully at least through the content of Algebra II (NMAP, 2008). This study assesses this expectation with respect to algebraic curricular content. As such, the purpose of this study is to compare the algebraic performance of high school students who enroll in integrated mathematics courses to the algebraic performance of high school students who enroll in subject-specific courses.

### **Research Question and Hypotheses**

The question this study investigated is “How do the algebraic performance gains of students enrolled in integrated course pathways relate to the algebraic performance gains of students enrolled in subject-specific course pathways?” To address this question a pretest-posttest design is used to statistically compare gain scores of high school students who learn from Algebra I-Geometry-Algebra II (i.e., subject-specific) course pathways to the gain scores of a comparable group of high school students who learn from Integrated I-II-III (i.e., integrated) course pathways. More specifically, the following 6 hypotheses are tested.

1. No statistical difference in mean gain scores on the algebra assessment will exist between students who learn from subject-specific and integrated course pathways.
2. No statistical difference in mean gain scores on tasks related to algebraic expressions will exist between students who learn from subject-specific and integrated course pathways.
3. No statistical difference in mean gain scores on tasks related to multiplicative and proportional thinking will exist between students who learn from subject-specific and integrated course pathways.

4. No statistical difference in mean gain scores on tasks related to linear equivalents will exist between students who learn from subject-specific and integrated course pathways.
5. No statistical difference in mean gain scores on tasks related to systems of linear equations will exist between students who learn from subject-specific and integrated course pathways.
6. No statistical difference in mean gain scores on tasks related to linear functions will exist between students who learn from subject-specific and integrated course pathways.

The first hypothesis will determine whether statistical differences exist on a holistic outcome assessment that measured algebraic reasoning and problem solving skills. However, an issue when interpreting gain scores is groups of students can have similar gains on an overall assessment, but the gains may be in different content areas. For example, two students who take a mathematics assessment that measures skills and ability in algebra, geometry, statistics, and arithmetic may have similar gains on a pretest and posttest. It is conceivable that one student may have significant gains on the algebra and geometry items, whereas the other student may have significant gains on arithmetic and statistics items. Because of possibilities of this nature, analysis of specific test items should also be conducted when comparing gain scores. Therefore, the last 5 hypotheses will determine whether statistical differences exist on specific algebra related concepts.

### **Brief Review of Literature**

Findings from recent studies (Chavez et al., 2013; Grouws et al., 2013; Tarr et al., 2013) suggest students who enroll in an integrated course pathway performed better on algebra related

content than students who enrolled in a subject-specific course pathway. The findings suggest content organization may have an effect, in favor of integrated mathematics, on high school students' algebraic performance. In addition, the findings support claims that learning mathematics from an integrated approach is a viable option for high school students, when compared to learning from a subject-specific approach (Reys & Reys, 2009). However, further research in this area is warranted as demonstrated by calls for more studies in which the organization of high school content is investigated (Tarr et al., 2013). In particular, there are gaps, issues and methodological shortcomings that have been identified and need to be addressed. For example, several studies examine the effects of content organization over a short period of time (i.e., a year). In these cases, the researchers compared the performance of students enrolled in Algebra I and Integrated I (Grouws et al., 2013), Geometry and Integrated II (Tarr et al., 2013) and Algebra II and Integrated III (Chavez et al., 2013). A longitudinal study of three or more years that employs a pretest-posttest design can provide new evidence related to students' cognitive growth from a complete curricular sequence, rather than a single course (Cai et al., 2013). In addition, results of these studies are based on a holistic outcome measure. The findings may suggest whether one curriculum pathway is more effective than another, but the findings do not indicate how or in what ways a curriculum pathway is more effective. An analysis of specific test items may identify more important differences in student performance, compared to results based on an entire set of test items (Huntley et al., 2001). Therefore, studies on the effectiveness of curricular approaches should include analysis of specific test items. Finally, the students who participated in the studies were not randomly assigned to course pathways. Instead, they were observed in their natural occurrence. Therefore, it is possible that the students in the groups had systematic differences prior to the study. These systematic

differences can result in selection bias in the estimation of treatment effects. As a result there is a need to reduce the threat of selection bias. Selection bias can be reduced by using the ‘Counterfactual Model of Causal Inference’.

In sum, results from previous studies that compare subject-specific and integrated approaches have examined effects over a short period of time, focused on specific curriculum materials, based results on a holistic measure of achievement, and have methodological shortcomings. This study addressed these areas of concern by 1) examining the achievement of students who learn from subject-specific and integrated course pathways as a comprehensive multiyear curriculum, 2) examining students’ performance on specific assessment items, rather than just on a holistic measure, and 3) using a framework to reduce the threat of selection bias. The results from this study provide evidence about the progression of students’ performance as they proceed through the first three years of high school.

### **Significance of the Study**

Researchers have used National Center for Educational Statistics (NCES) observational datasets, such as the National Education Longitudinal Study of 1988 (NELS) and Education Longitudinal Study of 2002 (ELS), to compare outcome measures based on students’ course-taking patterns (Bozick et al., 2007; Gamoran & Hannigan, 2000; Attewell et al., 2006). However, the course taking patterns are based solely on mathematics courses related to subject-specific courses. That is, these studies focus on courses related to the Algebra I-Geometry-Algebra II-Trigonometry-Calculus sequence. The studies do not include courses taken as part of the integrated course sequence (i.e., Mathematics I-II-III-IV). A primary reason the studies did not include integrated mathematics courses may be because the NCES observational data that were examined did not include information to reveal if the students enrolled in integrated



mathematics courses. The observational data only elicited information on courses in the subject-specific pathway.

This study uses data from an observational study from the NCES known as the High School Longitudinal Study of 2009 (HSLs) to compare the algebraic performance of students who enroll in subject-specific and integrated course pathways. In contrast to prior NCES observational studies (i.e., ELS, NELS), the HSLs elicits information to determine whether students were enrolled in integrated or subject-specific mathematics courses. In addition, prior NCES observational studies assessed students' skills and abilities on their overall mathematics performance (ELS, 2002; NELS, 1988). However, the HSLs is the first NCES observational study to assess students' skills and abilities on their overall algebra performance. As a result, the algebraic performance gains of high school students enrolled in subject-specific and integrated course pathways can be compared.

Many researchers suggest a randomized control trial is the best technique to establish a causal link between treatments and outcomes (Gall, Gall, & Borg, 2007). However, Graham (2010) suggests causal inferences can be made using non-experimental or observational data. Because students who participated in the HSLs09 were not randomly assigned to course pathways, it is likely the groups differ systematically, resulting in the threat of selection bias. To reduce the threat of selection bias, a theoretical framework known as the Counterfactual Model of Causal Inference' is used. This theoretical framework is explained in depth in Chapter 3. An important component of the 'Counterfactual Model of Causal Inference' is that it employs a statistical technique known as propensity score matching (Rosenbaum & Rubin, 1983). Propensity scores were introduced to the educational research community in the mid-1980s, but have been rarely used by educational researchers until recently (Graham, 2010). Still, few

published studies that focus on mathematics outcomes have employed propensity score methods to address selection bias. This study contributes to the extant research literature that employs the use of propensity scores, and can serve as a model for future mathematics education researchers who seek to use observational data to suggest causal inference. The propensity score matching procedure strengthens results of studies of curricular effectiveness because it approximates a random controlled experiment. However, causal inferences should be made with caution.

In sum, this study is significant to mathematics education policy because it provides information about the relationship between content organization and students' algebraic achievement. In turn, the results can be used to inform discussions regarding which course pathway to implement in high schools. In addition, the study is a model of how an observational data set and propensity score matching can be used by the mathematics education community.

### **Definition of Terms**

***Subject-specific Course Pathway.*** A high school mathematics curricular approach that is typically used in the United States. The pathway consists of two algebra courses and a geometry course, with some data, probability and statistics included in each course. The curricular sequence of the courses in this pathway is typically Algebra I-Geometry-Algebra II.

***Integrated Course Pathway.*** A high school mathematics curricular approach that is typically used in countries other than the United States and in curricula materials developed in response to standards-based reform documents (Senk & Thompson, 2003). The pathway consists of a sequence of three courses, in which each course includes topics in number, algebra, geometry, probability and statistics. The curricular sequence of the courses in this pathway is referred to as Integrated Mathematics I-II-III.

***Selection Bias.*** A statistical bias in which there is an error in choosing the individuals or groups to take part in a scientific study. It is possible that group differences on outcome variables will be a result of preexisting group differences rather than to a treatment effect. If the selection bias is not taken into account in non-experimental studies, then it is possible that erroneous conclusions may be drawn.

***Effect size.*** In statistics, an effect size is a measure of the strength of a phenomenon. An effect size that is calculated from data is a descriptive statistic that conveys the estimated magnitude of a relationship without making any statement about whether the apparent relationship in the data reflects a true relationship in the population. Effect sizes complement inferential statistics such as  $p$ -values, as they quantify the magnitude of a phenomenon rather than simply provide evidence of significant differences like the  $p$ -values does.

## **Chapter 2: Studies Related to Content Organization and Students' Learning**

This chapter reviews and critiques of studies in which researchers compared the mathematics performance of students who enroll in subject-specific or integrated course pathways. The objective of this literature review is to identify gaps and issues in the research as a means to inform future studies of curricular effectiveness, and in particular, studies that investigate effects of content organization on students' learning. The first section is a review of the most recent studies of content organization in which a collaboration of researchers compared the performance of students who enroll in subject-specific and integrated courses. In particular, the researchers compared Algebra I and Integrated I (Grouws et al., 2013), Geometry and Integrated II (Tarr et al. 2013), and Algebra II and Integrated III (Chavez et al., 2013). Collectively, the findings suggest that students who enroll in integrated mathematics courses perform better on outcome measures than students who enroll in subject-specific courses. However, there are concerns related to the study that need to be addressed. In particular, the studies examine specific curriculum materials, base findings on a holistic measure, examine students' learning over a short period of time (i.e., one academic year), and are threatened by an issue known as selection bias. In the second section I elaborate on these concerns, and in turn provide implications for research. The third and final section is a summary of the review.

In conducting this review of the literature, I omitted studies conducted to examine relationships between high school curricula and college preparation (Harwell et al., 2009, 2012, 2013; LeBeau et al., 2012; Post et al., 2010). These studies suggest students who learn from

subject-specific course pathways initially enroll in more difficult mathematics courses upon entering college when compared to students who learn from integrated course pathways.

However, as students enrolled in subsequent college mathematics courses, the researchers found no statistical relationship between high school curricula and college mathematics achievement, course-taking patterns, and attainment of a bachelor's degree in mathematics or engineering.

Collectively, the results suggest more attention is needed at the high school level to improve students' learning of mathematics. Hence, relationships between high school curricula and students' learning while in high school are the focus of this review.

### **Comparing Options in School Mathematics: Investigating Curricula Project**

The Comparing Options in School Mathematics: Investigating Curricula (COSMIC) study is the most recent project in which researchers compare the mathematics performance of high school students who learn from subject-specific and integrated course pathways (Chavez et al., 2013; Grouws et al., 2013; Tarr et al., 2013). The researchers conducted three studies in which they investigated the effects of content organization and curriculum implementation on high school students' mathematics learning. More specifically, the researchers statistically compared the mathematics performance of students enrolled in Algebra I and Integrated I (Grouws et al., 2013), Geometry and Integrated II (Tarr et al., 2013), and Algebra II and Integrated III (Chavez et al., 2013).

To help ensure a balance between curriculum types with regard to the number of days of instruction, schools were selected to participate in the study only if they offered both integrated and subject-specific course pathways. In addition, this requirement allowed the researchers to control for contextual school factors, such as policies related to homework and technology, length and organization of class periods, and professional development provided to teachers.

Another characteristic of the schools selected to participate in the study is the integrated curriculum implemented in the schools was the Core-Plus Mathematics Program (CPMP) (Coxford et al., 2003). CPMP is the most widely used integrated mathematics curriculum in the United States. Curriculum materials implemented in the subject specific courses include textbooks published by Glencoe, Prentice-Hall, and McDougal-Littell.

Data regarding curriculum implementation were collected using teacher surveys, a table of contents record, and classroom visit protocols. The researchers administered two teacher surveys to collect information about teacher characteristics. The teacher surveys were administered at the beginning and middle of the school year. In addition, to capture the nature and extent to which the textbook was used, teachers were asked to complete a table of content record. On a daily basis, the teachers were asked to record content taught primarily from the textbook, content taught from the textbook with some supplementation, content taught from an alternative source, or content that was not taught. Finally, the researchers also visited classrooms to document the use of textbook materials and classroom activities. The collected data were used to create four indices related to curriculum implementation and three indices related to teacher characteristics. The four indices related to curriculum implementation were 1) a classroom learning environment factor, 2) an implementation fidelity factor, 3) a teaching and collaborative learning factor, and a 4) an opportunity to learn factor. The three indices related to teacher characteristics were 1) a NCTM standards – familiarity, agreement and implementation factor, 2) a teacher experience factor, and 3) a professional development factor.

The researchers used multiple relevant grade level measures of student learning that included a state-mandated 8<sup>th</sup> grade test, nationally standardized assessments known as the Iowa Test of Educational Development (ITED), and two project developed tests referred to as the

Problem Solving and Reasoning Test (PSRT) and the Test of Common Objectives (TOCO). The state mandated 8<sup>th</sup> grade test was used as a measure of prior achievement. The ITED assessments were 40-question multiple-choice tests that assess students' computational and problem-solving skills in a number of mathematical contexts. The test is also referred to as the Mathematics: Concepts and Problem Solving test (Feldt, Forsyth, Ansley, & Alnot, 2003). The PSRT assessments consist of topics deemed appropriate based on content analyses and feedback from external reviewers. The tests were designed to assess nontrivial mathematical reasoning and problem-solving skills that focused on aspects of algebra, geometry, and statistics. The TOCO assessments consist mainly of constructed-response items that assess concepts and skills common to the respective curriculum types studied.

To examine students' performance, in each study the researchers fit each outcome measure to a three-level hierarchical linear model (HLM). A hierarchical linear model is used to take into account the structure of the data in which students are nested in classrooms and classrooms are nested in schools. Although the high school course in which students enrolled was the primary independent variable of interest, other variables were included in the analyses as control variables. At the student level, the control variables included sex, ethnicity, an 8<sup>th</sup> grade basic skills mathematics score to serve as a measure of prior achievement, and a variable indicating whether a student had an individual education plan (IEP). At the classroom level, control variables included curriculum type (i.e., integrated or subject-specific), the percentage of students receiving free or reduced-price lunch, classroom learning environment factor, curriculum implementation factor, an opportunity to learn factor, and a teaching experience variable. In addition to these main effects, the model includes the two-way interactions of curriculum with classroom learning environment factor, curriculum implementation factor, an

opportunity to learn factor, and percentage of students receiving free and reduced lunch. At the school level, a control variable that represented the average of the Level 2 free and reduced lunch variables was included. The researchers also computed the effect sizes ( $g$ ) by dividing the differences between group means by the pooled post-test standard deviation. The following sections report findings from the three studies.

**Algebra I and Integrated I.** In the first study, Grouws et al. (2013) examined the performance of students from 10 schools who were enrolled in an Integrated Mathematics 1 ( $n=1256$ ) or Algebra I ( $n=1365$ ) course. The researchers performed an analysis of variance (ANOVA) on the state-mandated eighth grade tests to test for mean group differences in students' prior achievement. They found no statistical mean difference in prior achievement between the students who enrolled in the integrated course and students who enrolled in the subject-specific course.

Grouws et al. (2013) administered the three grade level end-of-year assessments to study participants. The researchers administered an ITED (Form B, Level 15) standardized assessment appropriate for 9<sup>th</sup> grade. The inter-rater reliability of the Problem Solving and Reasoning Test and the Test of Common Objectives scoring for the study was 96% and 96.5%, respectively.

After fitting the outcome measures and collected data to the three-level model, Grouws et al. found students who enrolled in Integrated I performed statistically better on the ITED, PSRT, and TOCO than students who enrolled in Algebra I with small effect sizes of 0.166, 0.453, and 0.308, respectively. Other factors found to be related to students' performance were opportunity to learn, teacher experience, prior achievement, and the cross-level interaction between curriculum and prior achievement. In particular, prior achievement was found to be statistically related to the ITED, PRST, and TOCO with medium effect sizes of 0.586, 0.531, and 0.559,



respectively. Teacher experience was found to be statistically related to the ITED and TOCO with effect sizes of 0.16 and 0.381, respectively. Opportunity to learn was found to be statistically related to the ITED and TOCO with effect sizes of 0.187 and 0.203, respectively. The interaction between curriculum and opportunity to learn was found to be statistically related to ITED with an effect size of -0.178. The interaction between curriculum and prior achievement was found to be statistically related to the TOCO with an effect size of 0.078.

**Geometry and Integrated II.** In the second study, Tarr et al. (2013) examined the performance of students from 11 schools who were enrolled in an Integrated Mathematics 2 (n=1171) or Geometry (n=2087) course. Similar to the study by Grouws et al. (2013), the researchers performed an analysis of variance (ANOVA) on the state-mandated eighth grade tests to test for mean group differences in students' prior achievement. Tarr et al. found the mean prior achievement score for students who studied from the subject-specific curricula was statistically higher than the mean prior achievement score for students who studied from the integrated curricula. However, ancillary analysis revealed that differences in prior achievement occurred in only two schools.

Tarr et al. (2013) administered the three grade level end-of-year assessments to study participants. The researchers administered an ITED (Form B, Level 16) standardized assessment appropriate for 10<sup>th</sup> grade students. The inter-rater reliability of the Problem Solving and Reasoning Test and the Test of Common Objectives scoring for the study was 94% and 97.3%, respectively.

After fitting the outcome measures and collected data to the three-level model, Tarr et al. found students who enrolled in Integrated II performed statistically better on the ITED than students who enrolled in Geometry with a small effect size of 0.294. However, no statistical

mean differences in performance on the PRST and TOCO were found between the two groups. Other factors found to be related to students' performance were opportunity to learn, prior achievement, the cross-level interaction between curriculum and opportunity to learn, and the cross-level interaction between curriculum and prior achievement. In particular, opportunity to learn was found to be statistically related to the ITED, PSRT, and TOCO with effect sizes of 0.254, 0.163, and 0.264, respectively. Prior achievement was found to be statistically related to the ITED, PRST, and TOCO with medium effect sizes of 0.586, 0.531, and 0.559, respectively. The interaction between curriculum and opportunity to learn was found to be statistically related to ITED and TOCO with effect size of -0.235 and -0.23, respectively. The interaction between curriculum and prior achievement was found to be statistically related to the TOCO with an effect size of 0.10.

**Algebra II and Integrated III.** In the third study, Chavez et al. (2013) examined the performance of students from 10 schools who were enrolled in an Integrated III (n=892) or Algebra II (n=1350) course. Similar to related studies (Grouws et al., 2013; Tarr et al., 2013), the researchers performed an analysis of variance (ANOVA) on the state-mandated eighth grade tests to test for mean group differences in students' prior achievement. Chavez et al. found no statistical mean difference in prior achievement between the students who enrolled in the integrated course and students who enrolled in the subject-specific course.

Chavez et al. (2013) administered multiple grade level end-of-year assessments to study participants. However, the researchers only administered an ITED and TOCO. The ITED (Form B, Level 17) standardized assessment was appropriate for 11<sup>th</sup> grade students. The inter-rater reliability of the Test of Common Objectives scoring for the study was 96%.

After fitting the outcome measures and collected data to the three-level model, Chavez et al. found students who enrolled in Integrated III performed statistically better on the TOCO than students who enrolled in Algebra II with a small effect size of 0.33. However, no statistical mean differences in performance on the ITED were found between the two groups. Other factors found to be related to students' performance were prior achievement and teacher orientation. In particular, teacher orientation, which is a measure of teacher beliefs about reform-oriented practices, was also found to be statistically related to ITED and TOCO with effect sizes of 0.14 and 0.15, respectively. Prior achievement was also found to be statistically related to the ITED and TOCO.

**Summary of Findings from COSMIC Project.** Collectively, findings from the COSMIC project suggest students who learned from integrated courses performed as well or better on the outcome measures than students who learned from subject-specific courses (Chavez et al., 2013; Grouws et al., 2013; Tarr et al., 2013). In particular, on the ITED assessments, the researchers found statistical mean differences in students' performance in the first and second year studies in favor of students enrolled in the integrated courses (Grouws et al., 2013; Tarr et al., 2013). However, in the third year study no statistical mean differences in performance on the ITED were found between students enrolled in the two curricular approaches (Chavez et al., 2013). The findings based on the ITED assessments suggest differences exist in students' learning during the first two years of high school, but the differences disappear by the time the students complete the eleventh grade.

On the project developed tests, and more specifically the TOCO assessments, the researchers found statistical differences in the first and third year studies in favor of students enrolled in the integrated courses (Grouws et al., 2013; Chavez et al., 2013). However, in the

second year study no statistical differences in performance on the TOCO were found between students enrolled in the two curricular approaches (Tarr et al., 2013). The findings suggest statistical differences exist in students' learning during the first and third years of high school, but no statistical differences exist in students' learning during the second year of high school. Stated differently, the findings suggest no statistical differences exist in students' performance on common geometry concepts, but differences exist in students' performance of common algebraic concepts in favor of students who enroll in integrated courses.

In sum, findings from the COSMIC project reveal that content organization, along with curriculum implementation and prior achievement were key factors in improving students' learning in the first three years of high school. However, the COSMIC researchers called for more studies in which researchers investigate relationships between content organization and students' mathematics learning. The next section identifies gaps in the COSMIC research that lead to implications for future research.

### **Critique of the COSMIC Project and Implications for Research**

Although the COSMIC researchers established a solid foundation for researchers who seek to investigate relationships between content organization and students' learning, there are gaps and issues with the research that can be addressed in future studies. In particular, the studies mainly focused on the examination of specific curricula, examined effects over a short period of time (i.e., a year), based results on holistic measures, and had some methodological shortcomings that need to be addressed. This section elaborates on these concerns and suggests implications for future research.

**Integrated and Subject-Specific Curricula as a Holistic Body of Work.** A possible reason differences were found on the Test of Common Objectives that favor students who

learned from the integrated curricula may be attributed to the way in which the measure was developed (Chavez et al., 2013; Grouws et al., 2013). Specifically, the TOCO is based on a content analysis of one integrated curriculum (i.e., Core-Plus) and one subject-specific curriculum (Glencoe) (Chávez, Papick Ross, & Grouws, 2010). Although all students who learned from the integrated curricula only used Core-Plus curriculum materials, not all students who enrolled in subject-specific pathway learned from Glencoe curriculum materials (Chavez et al., 2013; Grouws et al., 2013; Tarr et al., 2013). In particular, some students who enrolled in subject-specific courses learned from curriculum materials published by other commercial developers, such as Prentice Hall and McDougal-Littell. Although the subject-specific curriculum materials may cover similar core mathematics topics and the content may be organized in a similar manner, there continue to be differences related to how the content is sequenced in particular textbooks. For example, Huntley and Terrell (2014) examined five high school textbook series to compare the treatment of solving one-step equations, multi-step equations, functions and graphs. The researchers examined two integrated textbook series and three subject-specific textbook series. With regards to content organization, they found differences between and within subject-specific textbook series and integrated textbook series. Because the TOCO is based on content from the Glencoe and Core-Plus textbooks, it is possible the measure may have been positively biased to students who enrolled in courses in which these curriculum materials were implemented, and consequently negatively biased to students who enrolled in courses in which the Prentice-Hall or McDougal Littell curriculum materials were used. Therefore, it is possible the differences in performance on the TOCO may be attributed to differences between the subject-specific curricula used by students in the study.

Rather than investigate the content organization of specific curriculum material, another option is to compare the performance of students who learn from subject-specific and integrated course pathways as a holistic body of work. Many of the curriculum materials implemented in subject-specific courses are developed by commercial publishers, whereas many of the curriculum materials developed in integrated mathematics courses are developed in response to standards-based reform documents (Senk & Thompson, 2003). Therefore, a study that involves a focus on course pathways (i.e., integrated and subject-specific) and a reliable measure can add evidence about the impact of subject-specific and integrated course pathways as a body of work, as opposed to evidence related to specific curricula materials. Because many countries use integrated mathematics courses in their high school mathematics programs, a study of course pathways as a body of work can provide insight to whether many of the difficulties that U.S. students have with mathematics are consequences of the content organization of subject-specific courses (Reys & Reys, 2009).

**Analysis of Specific Content Topics or Test Items.** The results from the COSMIC studies are based on holistic measures. Although these findings can suggest whether one curriculum is better than another in improving students' learning, they do not reveal more nuanced differences such as whether students benefited from specific content. The researchers could have identified more important differences in students' performance if they performed an analysis of specific test items or concepts. For example, Huntley, Rasmussen, Villarubi, Sangtong, and Fey (2000) conducted a comparative study on the effects Core-Plus curriculum materials and subject-specific curricula materials have on student understanding, skill, and problem-solving ability with specific algebraic and functions concepts. The researchers examined students' performance on evaluating expressions, testing equivalence, solving

equations, and solving inequalities. They found the students who studied from the subject-specific approach performed statistically better than students who studied from the integrated approach on each of the concepts when problems did not include application contexts and students were not allowed to use calculators. In contrast, when calculators were allowed and problems included realistic contexts, students who studied from the integrated approach performed statistically better than students who studied from the subject-specific approach when solving equations and inequalities. In sum, an analysis of test items or concepts can identify student learning advantages or disadvantages presented by different curriculum materials.

**Methodological Shortcomings.** Although statistical differences were found between students who learned from different types of curricular approaches (Cai et al., 2011; Chavez et al., 2013; Grouws et al., 2013; Harwell et al., 2009, 2012; Tarr et al., 2013), causal inferences based on the results should be made with caution. The best way to establish a causal link between an instructional intervention and an outcome measure is to randomly assign students to treatment groups (Gall, Gall, & Borg, 2007; Graham, 2010). The technique of random assignment ensures that groups are equivalent on observed and unobserved characteristics. In addition, the technique allows researchers to conclude that differences in outcome measures between two groups are attributable to the treatment alone, and not some other preexisting characteristics. However, when students are not randomly assigned to groups, there is a possibility that group differences on the outcome variables will be a result of preexisting group differences rather than to a treatment effect. This problem is referred to as selection bias. In particular, selection bias is a problem for mathematics education researchers who seek to make inferences about the effects of different instructional methods in mathematics on student outcomes (Graham, 2010).

The students who participated in the COSMIC project were not randomly assigned to course pathways. Instead, they were observed in their natural occurrence. Therefore, it is possible that students with certain attributes were assigned to one pathway over another. To investigate this possibility, Grouws et al. (2013) and Tarr et al. (2013) used logistic regression to generate a *propensity score* for each student to examine the threat of selection bias. The propensity scores represent the probability that a given student is assigned to a particular course pathway based on prior achievement, gender, ethnicity, IEP status, and LEP status. Grouws et al. (2013) found Hispanic students were more likely to be assigned to the subject-specific course pathway. Tarr et al. (2013) found African American students, Hispanic students, and students who had individual educational plans (IEP) were more likely to be in the subject-specific course pathway. The findings based on propensity scores suggest the studies are plagued with the threat of selection bias and differences in outcomes can be attributed to these preexisting differences. In particular, Grouws et al. (2013) and Tarr et al. (2013) found that Hispanic and African American students performed statistically lower than White students on all measures. Because minority students were more likely to be assigned to the subject-specific course pathway, the results of studies from the COSMIC project (Chavez et al., 2013; Grouws et al., 2013; Tarr et al., 2013) could have been negatively biased, causing integrated students to appear to perform better on some of the outcome measures.

The use of propensity scores revealed the studies are threatened by selection bias. However, the researchers could have investigated further and matched students based on their propensity scores. This procedure is called propensity score matching. Researchers suggest that propensity score matching can address the issue of selection bias (Rosenbaum & Rubin, 1983) because the goal of the procedure is to match, as closely as possible, each student from a



treatment group with a student from a control group. As a result, any differences between groups that are found on the outcome measures can be attributable to the treatment.

A limitation to the propensity score procedure is it only matches students based on observed variables, not unobserved variables (Rubin, 1997). Therefore, researchers cannot be completely certain that individuals in a treatment group do not differ from individuals in a control group on unobservable variables. Consequently, the use of propensity scores may not completely solve the problem of selection bias. However, using the method when random assignment is not possible can strengthen the results of a study, because the procedure can approximate a randomized controlled trial. Because of this, future studies of curricular effectiveness are encouraged to use propensity score matching to address the issue of selection bias.

**A Longitudinal Study to Examine Cognitive Growth.** The researchers of the COSMIC project focused each of their studies on a single academic year. This time period allowed the researchers to take into account factors related to curriculum implementation. However, the time frame allowed the researchers to examine the influence of a single textbook within the curriculum sequence. An examination of students' learning over a longer period of time can add new knowledge related to the entire curriculum sequence. Specifically, a longitudinal study that incorporates a pretest-posttest design can be used to investigate students' learning growth over the course of the first three years of high school. Cai, Wang, Moyer, Wang, and Nie (2011) conducted a study of this nature in which they compared the algebra learning of middle school students who learned from different curricula approaches.

The Longitudinal Investigation of the Effect of Curriculum on Algebra Learning (LieCal) was a project in which researchers' collected longitudinal data to compare the algebra learning

growth of middle school students who studied mathematics from two different types of mathematics curricula (Cai et al., 2011). The mathematics curricula they included in the study was the Connected Mathematics Project (CMP) and other curriculum materials that will be referred to as non-CMP curricula.

Cai et al. (2011) conducted the study in 14 middle schools. Seven of the schools used the CMP curriculum. From these schools, 695 students in 25 classes participated in the study. The other seven schools used a non-CMP curriculum. Of these schools, 589 students in 22 classes participated in the study. Cai et al. (2011) followed the 1284 students as they progressed from 6th grade to 8th grade. Study participants were administered a 6-question open-ended test and a 32-question multiple-choice test at the beginning of their sixth grade as a baseline measure. The participants were also administered similarly equated tests at the end of their sixth, seventh, and eighth grade school years. The assessments were used to assess the students' skills on a broad spectrum of algebra related concepts. Cai et al. reported results related to open-ended tasks, translations tasks, computation tasks, and equation solving tasks. The open-ended tasks assessed students' conceptual understanding and problem-solving skills. The translation tasks assessed students' ability to represent problem situations. The computation and equation-solving tasks assessed students' procedural knowledge and symbol manipulation skills. Collectively, the tasks represented problems that assessed students' conceptual understanding and procedural knowledge.

Cai et al. (2011) used repeated measures ANOVA to investigate whether the CMP curriculum can increase middle school students' learning growth in algebra when compared to the non-CMP curricula. Findings from the repeated measures ANOVA revealed both groups of students experienced statistical growth on all tasks over the three year time period. However, a

curricular effect was found only on the equation solving task, which suggests non-CMP students performed statistically better than the CMP students on these tasks. Cai et al. (2011) also examined the interaction between time and curriculum. The researchers found an interaction between time and curriculum on the open-ended and translation tasks, which suggests the annual growth rate of the CMP students was statistically greater than non-CMP students. However, on the computation and equation solving tasks, the findings did not reveal an interaction between time and curriculum, which suggest CMP and non-CMP students had similar growth rates over the three year period.

In addition to the repeated measures ANOVA, Cai et al. (2011) used a two-level growth curve model with students nested in classrooms to examine the effect of curricula while controlling for instructional and student variables on the four dependent measures. In both levels of the model, the researchers used curriculum, gender, and ethnicity as control variables. In addition, the researchers used measures of instructional practices (i.e., procedural and conceptual emphases) as control variables in the second level of the model. On the open-ended tasks, the growth curve model revealed no statistical difference in students' initial performance. However, the model suggested the growth rate of CMP students on the open-ended tasks increased statistically greater than the growth rate of non-CMP students over the three middle school years. On the translation tasks, the growth curve model revealed CMP students initially performed statistically lower than non-CMP students. However, the model suggested the growth rate of CMP students on the translation tasks increased statistically greater than the growth rate of non-CMP students' scores over the three middle school years. In addition, the model revealed the level of conceptual emphasis in classroom practices had a positive impact on the growth rate of students' performance. However, the differences in growth rate between CMP and non-CMP

students on the translation tasks diminished after the researchers controlled for the instructional variables. In other words, when teachers' instructional emphasis was similar, no difference existed in CMP and non-CMP students' growth rate on the translation tasks. On the computational tasks, the growth curve model revealed no statistical difference in students' initial performance or the growth rate of students' scores over the three year period. However, after controlling for ethnicity, the complete model revealed CMP students had a statistically lower growth rate than non-CMP students. On the equation solving tasks, the growth curve model revealed CMP students initially performed statistically lower than non-CMP students. However, the model revealed no statistical difference existed in students' growth rate over the three year period.

Collectively, the findings based on the repeated measures ANOVA and the growth curve models suggested the intended curricular approach to mathematics plays a key role in middle school students' learning of some algebraic related tasks. In particular, the longitudinal study provided evidence related to the influence of an entire curricular sequence on students' cognitive growth, as opposed to a single textbook in the curricular sequence. I believe the study by Cai et al. (2011) can be replicated with a population of high school students and relevant outcome measures, as no longitudinal study of at least three years involving high school students was found. In addition, because the researchers examined specific algebraic concepts, a study of this nature can provide new knowledge related to the advantages and disadvantages of the complete curricular sequence.

## **Summary**

In sum, this chapter reviewed and critiqued studies in which researchers investigated the effects of content organization on students' mathematics learning. Collectively, the results from

these studies suggest that content organization, along with curriculum implementation and prior mathematics achievement, represent key factors in efforts to increase students' learning in the first three years of middle school and high school mathematics. However, more research on the effects of content organization on students' learning is needed, especially at the high school level. Specifically, and similar to the study by Cai et al. (2011), longitudinal investigations of course pathways at the high school level that involve the analysis of specific concepts or test items will add new knowledge to the extant evidence base. In addition, instead of additional studies that focus on specific curriculum materials, research is needed that examines integrated and subject specific course pathways as a body of work. Finally, future studies should address the threat of selection bias that has plagued studies of curricular effectiveness by using a statistical technique known as propensity score matching.

### **Chapter 3: Data and Method**

The purpose of this study was to compare the algebra performance of high school students who learned from subject-specific course pathways (i.e., Algebra I-Geometry-Algebra II) to a comparable group of students who learned from an integrated course pathway (i.e., Integrated I-II-III). The rationale for the study is a call for more research in which integrated and subject-specific course pathways are investigated (Tarr, Grouws, Chávez, & Soria, 2013). In particular, Tarr et al. call for longitudinal studies of at least three years that examine the achievement of students who learn from the different course pathways, and examine student learning of specific assessment items as opposed to results based only on holistic measures. This study addresses the call for research by statistically comparing gain scores of high school students who learn from Algebra I-Geometry-Algebra II course pathways to a comparable group of high school students who learn from Integrated I-II-III course pathways.

A major difference between subject-specific and integrated course pathways is the manner in which the mathematics content is organized. The subject specific course pathway is a high school mathematics curricular approach that is typically used in the United States. The pathway consists of two algebra courses and a geometry course, with some data, probability and statistics included in each course. The curricular sequence of the courses in this pathway is typically Algebra I-Geometry-Algebra II. The integrated mathematics course pathway is a high school mathematics curricular approach that is often used in countries other than the United States. This pathway typically consists of a sequence of three courses, in which each course

includes topics in number, algebra, geometry, probability and statistics. The curricular sequence of the courses in this pathway is referred to as Integrated Mathematics I-II-III. The course pathways and an assessment of algebraic skills are used to address the following research question, “How do the algebraic performance gains of students enrolled in integrated course pathways relate to the algebraic performance gains of students enrolled in subject-specific course pathways?”

The rest of this chapter reviews the data and methods used to conduct this study. The first section is a description of the observational data used to conduct the study. The section includes a description of the sample of cases used from the data and the outcome measure administered to the students who participated in the observational study. The second section describes the methods used to conduct this study. Specifically, the section describes how a theoretical framework known as the Counterfactual Model of Causal Inference is applied to this study.

### **Data Source: High School Longitudinal Study**

The data used in this study came from a large scale observational study conducted by the National Center for Education Statistics (NCES) known as the High School Longitudinal Study of 2009 (HSLs). The HSLs collected data on the high school and postsecondary experiences of a nationally representative sample of high school students beginning with their ninth grade year. The target population for the HSLs included all ninth-grade students who attended public and private schools in the United States. The HSLs dataset is a complex sample survey that includes 21,444 students who were selected from 944 schools. The study participants were administered a survey instrument and assessment at the beginning of their ninth grade year and at the end of their eleventh grade year. The content of the student questionnaire included

demographic information, such as race, gender, socioeconomic status, and students' high school mathematics courses. Data from the HSLs student questionnaire were the primary data used in this study. Therefore, the unit of analysis for this study is the student.

**Sample Population.** The central variable for this study is high school mathematics course pathways. The selected cases from HSLs represent two well-defined groups of high school students enrolled in the same course pathway and same school for the first three years of high school. One group represents high school students who learned mathematics from an Algebra I-Geometry-Algebra II (i.e., subject-specific) course pathway and the other group represents high school students who learned mathematics from an Integrated Mathematics I-II-III (i.e., integrated) course pathway. Because the HSLs data set only inquires about courses students are enrolled in during their ninth grade year and their eleventh grade year, there were no data indicating what course the students were enrolled in during their tenth grade year. Therefore, the course pathways were inferred based on students' ninth and eleventh grade mathematics course enrollment.

To identify students in the subject-specific group, cases from the HSLs dataset that indicated students enrolled in Algebra I in the ninth grade (S1ALG1M09) and Algebra II in the eleventh grade (S2ALG2M12), and not in Integrated Mathematics I in the ninth grade (S1INTGM109) and Integrated Mathematics III in the eleventh grade (S2INTGM312) were selected. This yielded a sample with 4956 students who learned from a subject-specific course pathway. Similarly, to identify students in the integrated group, cases that indicated students enrolled in Integrated Mathematics I in the ninth grade (S1INTGM109) and Integrated Mathematics III in the eleventh grade (S2INTGM312) and not in Algebra I in the ninth grade



(S1ALG1M09) and Algebra II in the eleventh grade (S2ALG2M12) were selected. This yielded a sample with 73 students who learned from an integrated course pathway.

**Mathematics Assessment in Algebraic Reasoning.** The HSLs administered a mathematics assessment to the participants at the beginning of their ninth grade year and again at the end of their eleventh grade year. A Mathematics Advisory Panel reviewed, refined, and validated the framework and reviewed and approved each proposed item (Ingels et al., 2011). The objective of the test was to provide a measure of student achievement in algebraic reasoning at two points in time. The 40 item assessment measured students' performance on algebraic skills, reasoning and problem solving. The assessment employed a two-stage design in which the first stage contains 15 items that were common for all students. Based on their performance in the first stage, students were subsequently routed to a second stage that consisted of 25 items. The second stage included questions with varying levels of ability ranked as low, moderate, or high. The assessment included a mixture of ninth and eleventh grade items in both stages of the test (Ingels et al., 2011).

The algebraic reasoning framework of the assessment was designed to assess a cross-section of understandings representative of six domains of algebraic content and four algebraic processes. The six domains of algebraic content were 1) the language of algebra, 2) proportional relationships and change, 3) linear equations, inequalities, and functions, 4) nonlinear equations, inequalities, and functions, 5) systems of equations, and 6) sequences and recursive relationships. The four algebraic processes were 1) demonstrating algebraic skills, 2) using representations of algebraic ideas, 3) performing algebraic reasoning, and 4) solving algebraic problems. The HSLs outcome measures include an item response theory (IRT) based estimate of the score for each participant on the full set of items. In addition, the HSLs includes seven sets of

clustered-items that represent a broad spectrum of algebraic concepts. Each set of clustered-items represents four questions from the assessment and relates to specific content. Each set of clustered-items included on the pretest and posttest was included in the analysis. The sets of clustered-items measured student proficiency with algebraic concepts and represent the outcome variables for the study (Ingels et al., 2010). The algebraic concepts included the ability to 1) evaluate simple algebraic expressions and translate between verbal and symbolic representations of expressions, 2) solve proportional situation word problems, find the percent of a number, and identify equivalent algebraic expressions for multiplicative situations, 3) link equivalent tabular and symbolic representations of linear equations, identify equivalent lines and find the sum of variable expressions, 4) solve systems of equations algebraically and graphically and characterize lines represented by a system of linear equations, and 5) find and use slopes and intercepts of lines, and use functional notation. The five levels are hierarchical in the sense that mastery of a higher level typically implies proficiency at the lower levels (Ingels et al., 2011). A more in depth description of the outcome variables can be found in the Appendix. The probability of proficiency for a given student at a given level is calculated as the probability of getting correct at least three of the four items in a given cluster marking a proficiency level (the probability of a student getting at least three items correct out of four is expressed as the sum of (1) the probability of getting all four items correct and (2) the probability of getting any three items correct (Ingels et al., 2011). The IRT-estimated reliability of the HSLs test is 0.92 after sample weights are applied. This 0.92 reliability applies to all scale scores derived from the IRT estimation including the probability of proficiency scores.

## **Theoretical Framework: Counterfactual Model of Causal Inference**

This quantitative study employs a nonequivalent comparison group design. The design is similar to a true experiment because subjects in each group took a pretest and a posttest. However, unlike a true experiment, subjects in the nonequivalent comparison group design were not randomly assigned to treatment and control groups. Consequently, the main threat to the internal validity of a nonequivalent comparison group design is the possibility that group differences on the outcome variables will be a result of preexisting group differences rather than to a treatment effect (Gall, Gall, & Borg, 2007). This threat is referred to as selection bias. The main problem causing selection bias in non-randomized control trials is nonequivalence of treatment and control groups. If differences between students who enroll in subject-specific and integrated course pathways can be eliminated, then presumably the threat of selection bias will be eliminated.

To reduce the threat of selection bias due to non-random assignment of students, a theoretical framework known as the “Counterfactual Model of Causal Inference” was employed (Rosenbaum & Rubin, 1983; Attewell et al, 2006). The first step in the “Counterfactual Model of Causal Inference” is to create a logistic regression model to calculate the probability students learn from their respective course pathway. The calculated probability from the logistic regression model is called a propensity score. The propensity score takes a value between 0 and 1. Included in the logistic regression model are covariates related to students’ prior achievement (X1TXMSCR, X1TXMPROF1, X1TXMPROF2, X1TXMPROF3, X1TXMPROF4, X1TXMPROF5), gender (X1SEX), race (X1BLACK, X1HISPANIC, X1WHITE) socioeconomic status (X1SES), and a student longitudinal weight (W2W1STU). The inclusion of these variables was based on the inclusion of these variable in previous research studies (Cai

et al., 2011; Cai et al., 2013; Chavez et al., 2013; Grouws et al., 2013; Tarr et al., 2013; DuGoff, Schuler, & Stuart, 2014). In addition, preliminary analyses with various other covariates revealed these variables created a model with the best balance.

The next step is to use a caliper to create a 1:1 match of students in the groups. The pairs are selected at random from subject-specific and integrated course pathway students whose difference in propensity scores is less than 0.1 of each other. The caliper matching procedure matches each student with a given propensity score who learned from an integrated course pathway with a student who has a nearly identical propensity score, but enrolled in a subject-specific course pathway. The students in the integrated course pathway function like a treatment group. The students in the subject-specific course pathway function like a control group. Stated differently, the subject-specific student in each pair provides a "counterfactual" estimate of what the outcome for the integrated student would have been if that student had learned from a subject-specific pathway.

The goal of the first two steps in the "Counterfactual Model of Causal Inference" is to match, as closely as possible, each student who learned from an integrated course pathway with a student who learned from a subject-specific course pathway. A statistical software known as G\*Power was used to perform an *a priori* power analysis (effect size = .50, alpha level = 0.05, power = 0.99) for this study. Preliminary frequency analyses suggested the sample size would be 88 students or 44 matched pairs. The effect size value of 0.50 was determined to be the best estimate to use in the power analysis, along with an alpha value of 0.05, to achieve a power of 0.99 with this sample size. Upon completion of numerous iterations of the propensity score matching procedure, a sufficient sample size of 146 students or 73 matched pairs was yielded.

Numerous logistic regression models were specified with a goal of identifying the best model with the best balance on all observed variables. Shadish and Steiner (2010) suggest balance is achieved when the index is very close to zero for each of the pretest covariates and also for the propensity score itself. However, no rule exists for how close to zero will achieve adequate balance. Therefore, as suggested by Shadish and Steiner (2010), I used Cohen’s *d* for continuous measures and an odds ratio for categorical variables to determine acceptable balance. Specifically, acceptable balance was achieved when Cohen’s *d* for all continuous measures is  $d < 0.20$ , and the odds ratio for all categorical variables is between 0.80 and 1.25. The various logistic regression models included variables related to student and school characteristics. However, models that included school related variables created a greater imbalance on many of the variables after the propensity score matching procedure. The model with the best balance included variables related only to students’ gender, race, socioeconomic status, weights, and prior achievement. On the following pages the descriptive statistics of all the covariates included in the logistic regression model before and after matching are presented. Table 1 presents the demographic statistics, Table 2 presents statistics related to socioeconomic status, propensity score and weighting, and Table 3 presents statistics related to students’ mean performance.

Table 1  
*Students’ Demographics Before and After Matching (Percentage of Sample in Parentheses)*

Characteristics	Before Matching			After Matching		
	Subject-specific (n=4956)	Integrated (n=73)	Effect size ( <i>OR</i> )	Subject-specific (n=73)	Integrated (n=73)	Effect size ( <i>OR</i> )
Male	2450 (0.49)	38 (0.52)	0.94	36 (0.49)	38 (0.52)	0.94
Female	2506 (0.51)	35 (0.48)	1.05	37 (0.51)	35 (0.48)	1.05
Black	526 (0.11)	10 (0.14)	0.78	9 (0.12)	10 (0.14)	0.90
White	2893 (0.58)	43 (0.59)	0.99	43 (0.59)	43 (0.59)	1.00
Hispanic	748 (0.15)	9 (0.12)	1.26	8 (0.11)	9 (0.12)	0.91
Other Races	789 (0.16)	11 (0.15)	1.06	13 (0.18)	11 (0.15)	1.18

Note. Other races represent students classified as Asian, Native Hawaiian, American Indian, and Multi-racial.

Table 2

*Students' Mean Socioeconomic Status, Weight, and Propensity Score Before and After Matching (Standard Deviation in Parentheses)*

	Before Matching			After Matching		
	Subject-specific	Integrated	Effect size( <i>d</i> )	Subject-specific	Integrated	Effect size ( <i>d</i> )
SES	0.11 (0.72)	-.067 (0.69)	0.24	-0.097 (0.71)	-.067 (0.70)	0.04
Weight	209.64 (295.7)	170.48 (156.4)	0.16	189.85 (180.8)	170.48 (156.4)	0.11
Propensity	-	-	-	0.021 (.01)	0.021 (0.01)	0.00

Table 3

*Students' Means 9<sup>th</sup> Grade Performance Before and After Matching (Standard Deviation in Parentheses)*

Score	Before Matching			After Matching		
	Subject-specific	Integrated	Effect size ( <i>d</i> )	Subject-specific	Integrated	Effect size ( <i>d</i> )
Overall	39.42 (9.03)	41.93 (10.76)	0.25	42.01 (10.17)	41.93 (10.76)	0.01
Proficiency 1	0.92 (0.19)	0.91 (0.23)	-0.04	0.92 (0.23)	0.91 (0.23)	0.04
Proficiency 2	0.64 (0.30)	0.70 (0.32)	0.19	0.71 (0.30)	0.70 (0.32)	0.03
Proficiency 3	0.41 (0.31)	0.51 (0.35)	0.31	0.51 (0.34)	0.51 (0.35)	0.01
Proficiency 4	0.14 (0.15)	0.21 (0.23)	0.33	0.20 (0.20)	0.21 (0.23)	0.03
Proficiency 5	0.07 (0.04)	0.09 (0.07)	0.34	0.08 (0.06)	0.09 (0.07)	0.10

Collectively, the effect sizes of the differences in students' prior achievement, socioeconomic status, propensity score, and weighting after matching were trivial, which suggests the groups were balanced on all covariates included in the propensity score model. Theoretically, the differences in proficiency between students who learned from subject-specific and integrated course pathways represent an unbiased estimate (Graham, 2010; Rosenbaum & Rubin, 1983). Therefore, in the final step the mean outcome measures of students in the groups were statistically compared.

To examine whether differences exist in students' outcomes, the mean performance gain of students who enrolled in a subject-specific course pathway was statistically compared to the mean performance gain of students who enrolled in integrated course pathways. Mean gain

scores for each student were determined by calculating the difference between their posttest and pretest scores. After gain scores were calculated for each student, the difference in the gains of students in each matched pair was calculated. Finally, a single group *t*-test was used to determine if the mean difference in gain scores was statistically different from zero. The SAS statistical software package (SAS 9.4) was used to perform the single group means *t*-test. The PROC SURVEYMEANS procedure, along with the appropriate weight (W2W1STU), was used to test the following 6 hypotheses.

1. No statistical difference in mean gain scores on the algebra assessment will exist between students who learn from subject-specific and integrated course pathways.
2. No statistical difference in mean gain scores on tasks related to algebraic expressions will exist between students who learn from subject-specific and integrated course pathways.
3. No statistical difference in mean gain scores on tasks related to multiplicative and proportional thinking will exist between students who learn from subject-specific and integrated course pathways.
4. No statistical difference in mean gain scores on tasks related to linear equivalents will exist between students who learn from subject-specific and integrated course pathways.
5. No statistical difference in mean gain scores on tasks related to systems of linear equations will exist between students who learn from subject-specific and integrated course pathways.

6. No statistical difference in mean gain scores on tasks related to linear functions will exist between students who learn from subject-specific and integrated course pathways.

The hypotheses compared gain scores of students who enrolled in a subject-specific course pathway to the gain scores of students who enrolled in integrated course pathways. More specifically, the first hypothesis determined whether statistical differences existed on the holistic outcome measure, and the next 5 hypotheses determined whether statistical differences existed on specific algebra related concepts. In particular,  $p$ -values were used to determine whether statistically significant differences exist in students' performance on the algebra assessment.  $p$ -values inform whether a statistically significant difference exists between measures, but the results do not reveal the magnitude of the differences. Moreover, although no statistically significant differences on outcome measures may exist between treatment groups, it is still possible the results tentatively suggest some practical significance (Hess, 2003). Therefore, effect sizes ( $d$ ) were calculated to investigate whether the results suggest practical significances. The effect sizes ( $d$ ) were computed by dividing the differences between group mean gain scores by the pooled pre-test standard deviation (Morris, 2008).

### **Statistical Assumptions and Complex Survey Samples**

The SAS software package was used to statistically examine students' performance on the algebra assessment. However, SAS performs statistical calculations under the assumption the participants are a simple random sample, which is not the case for students who participated in the HSLs. The HSLs data were collected through a stratified, two-stage random sample design with schools selected at the first stage and students randomly selected from within the schools at the second stage. In addition, certain groups in the HSLs target population were



deliberately oversampled in the sample survey design to allow reasonable precision in the estimation of parameters and to ensure sufficient samples exist. As a result, these characteristics of the HSLs do not meet the requirements of a simple random sample in which there is only one stage of sampling and participants have equal probability of selection. Therefore, the HSLs dataset is referred to as a complex sample survey. Hypothesis testing based on data from a complex sample survey is different from hypothesis testing based on data from a simple random sample (Bell-Ellison & Kromrey, 2007). In particular, two areas of concern when using large scale observational datasets are bias and variance estimates.

The two-stage sampling process will increase the power of the statistical test, and ensures the generalization of effects from a random sample to a well-defined population (Schneider, Carnoy, Kirkpatrick, Schmidt, & Shavelson, 2007). However, the stratified sampling procedure can yield nested data in which the variance among students within each school is different from the variance among students in general (Bell-Ellison & Kromrey, 2007). Therefore, to appropriately estimate variances, the Balanced Repeated Replications (BRR) method, which is recommended by the HSLs (Ingels et al., 2011) was employed. The BRR method divides each sampling stratum into two primary sampling units and creates subsamples by randomly selecting one of the units from each stratum to represent the entire stratum (Bell-Ellison & Kromrey, 2007).

Another characteristic of the HSLs complex survey design is some subpopulations of students are deliberately oversampled. Therefore, the probability of selection to participate in the study is not equal for each student. This is a contrast to a simple random sample in which observations have equal probability of selection, and indicates an inherent selection bias within the dataset. Therefore, to produce subpopulation estimates for students enrolled in subject-

specific and integrated course pathways, SAS statistical software in conjunction with the first follow-up student longitudinal weight (W2W1STU) from the HSLs data was included in the analysis. Because some subpopulations were oversampled, it should be understood the observations more likely to be selected received a smaller weight than observations less likely to be selected.

In addition to the statistical assumption related to the complex survey design of the HSLs, three assumptions related to a  $t$ -test are 1) the scores are normally distributed, 2) homogeneity of the variances, and 3) observations are independent. Because there are an equal number of students in each group, the results of the  $t$ -test are robust to the first two assumptions. The third assumption is addressed by the BRR method.

## Chapter 4: Results and Interpretations

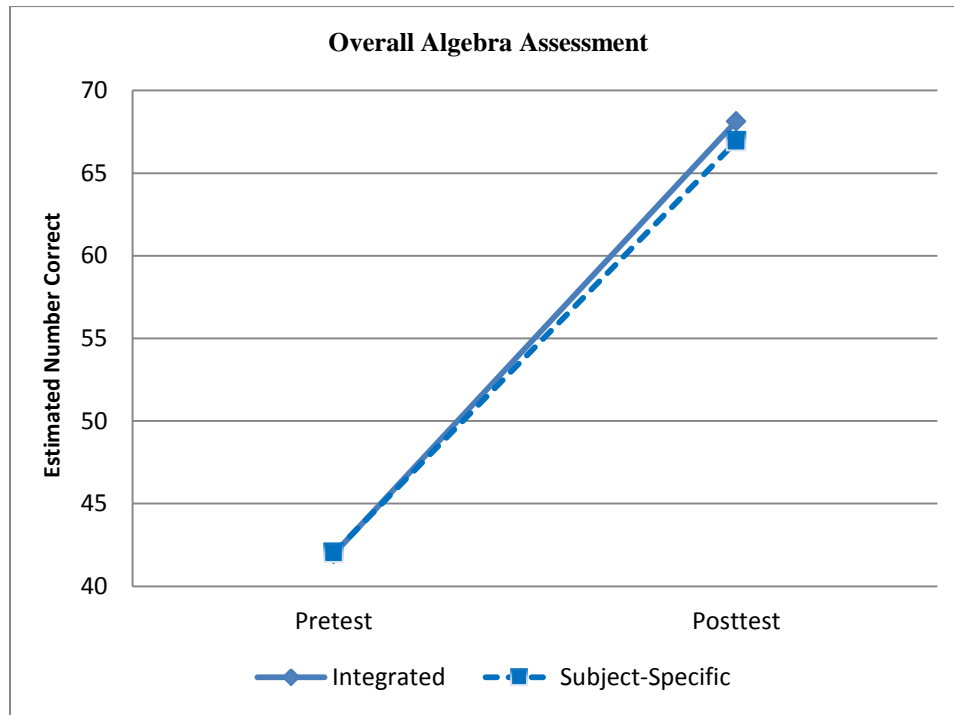
In this chapter, I present the findings from the statistical analyses. First, Table 4 presents statistics (i.e., mean and standard deviation) of students' performance on the pretest and posttest measures, and the mean gain scores of the groups. Following Table 4, the results from the six hypotheses are discussed and interpreted. Specifically, the first hypothesis tested whether the difference in mean gain scores on the overall algebra assessment was statistically different from zero. Because analysis of specific test items can provide more insight on the possible influence of content organization, hypotheses two through six tested whether the difference in mean gain scores on specific algebra concepts assessed on the pretest and posttest was statistically different from zero. The results of the hypothesis tests will include graphical representations,  $p$ -values, and effect sizes. The graphical representations are included to provide a visual of the results. The  $p$ -values are used to report the statistical significance of the results, and the effect sizes are used to report the practical significance of the results (Hess, 2003).

Table 4

*Statistics of Students' Mean Performance on Measures with Pretest and Posttest Scores*

	Integrated			Subject-Specific		
	<u>Pretest</u> Mean (SD)	<u>Posttest</u> Mean (SD)	<u>Gain</u> Mean (SD)	<u>Pretest</u> Mean (SD)	<u>Posttest</u> Mean (SD)	<u>Gain</u> Mean (SD)
Overall	41.9 (10.7)	68.1 (17.70)	26.2 (11.9)	42.0 (10.2)	66.9 (15.4)	24.9 (11.1)
Proficiency 1	0.91 (0.22)	0.94 (0.18)	0.03 (0.19)	0.92 (0.23)	0.95 (0.17)	0.04 (0.24)
Proficiency 2	0.70 (0.32)	0.81 (0.31)	0.11 (0.25)	0.71 (0.3)	0.84 (0.26)	0.12 (0.25)
Proficiency 3	0.51 (0.35)	0.73 (0.35)	0.22 (0.30)	0.51 (0.34)	0.73 (0.31)	0.22 (0.29)
Proficiency 4	0.21 (0.23)	0.35 (0.34)	0.14 (0.25)	0.20 (0.20)	0.29 (0.3)	0.08 (0.23)
Proficiency 5	0.09 (0.08)	0.21 (0.30)	0.12 (0.26)	0.08 (0.06)	0.16 (0.26)	0.07 (0.24)

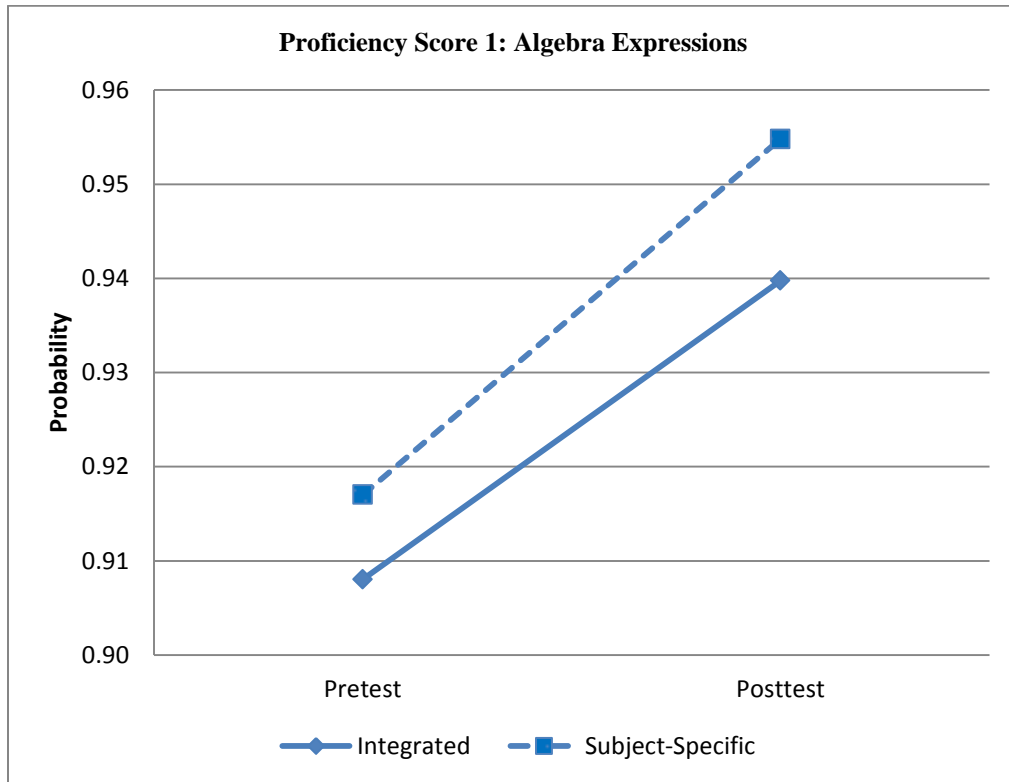
Hypothesis 1 states no statistical difference in mean gain scores on the algebra assessment will exist between students who learn from subject-specific and integrated course pathways.



*Figure 1. Students' growth on the overall algebra assessment.*

Figure 1 suggests no major difference exists in the mean gain scores of students in the two groups on the overall assessment. The overall assessment included 72 questions. Students who enrolled in integrated courses had slightly greater gains ( $M = 26.19, SD = 11.88$ ) than students who enrolled in subject-specific courses ( $M = 24.9, SD = 11.13$ ). However, the results of the  $t$ -test ( $M = 1.29, SD = 16.10, t(200) = 0.13, p = 0.89$ ) reveal no statistically significant difference in gain scores. Therefore, the null hypothesis that no statistical difference will exist in students' performance gains on the overall assessment is not rejected. Furthermore, Cohen's effect size value ( $d = .12$ ) suggests a low practical significance.

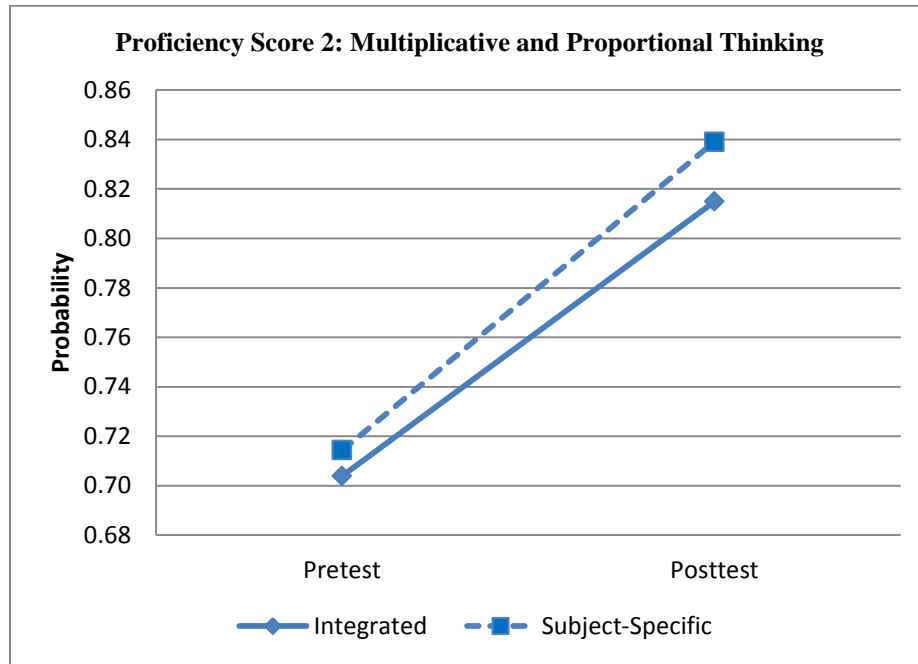
Hypothesis 2 states no statistical difference in mean gain scores on tasks related to algebraic expressions will exist between students who learn from subject-specific and integrated course pathways.



**Figure 2.** Students' proficiency growth on content related to algebraic expressions.

Figure 2 suggests no major difference exists in the mean gain scores of students in the two sample groups on items related to algebraic expressions on the algebra assessment. Students who enrolled in subject-specific courses had slightly greater gains ( $M=0.04$ ,  $SD =0.24$ ) than students who enrolled in integrated courses ( $M =0.03$ ,  $SD =0.19$ ). However, the results of the  $t$ -test ( $M = -0.01$ ,  $SD = 0.29$ ,  $t(200) = -0.36$ ,  $p = 0.72$ ) reveal no statistically significant difference in gain scores. Therefore, the null hypothesis that no statistical difference will exist in students' performance on items related to algebraic expressions is not rejected. Furthermore, Cohen's effect size value ( $d = .03$ ) suggests a low practical significance.

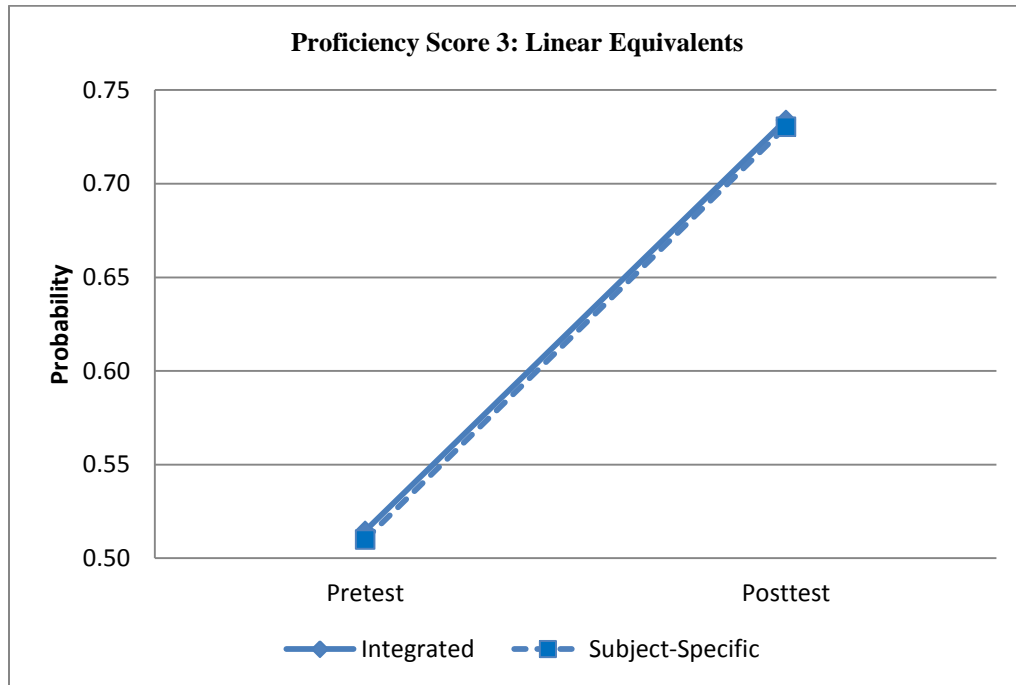
Hypothesis 3 states no statistical difference in mean gain scores on tasks related to multiplicative and proportional thinking will exist between students who learn from subject-specific and integrated course pathways.



**Figure 3.** Students' proficiency growth on content related to multiplicative and proportional thinking.

Figure 3 suggests no major difference exist in the mean gain scores of students in the two groups on items related to multiplicative and proportional thinking on the algebra assessment. Students who enrolled in subject-specific courses had slightly greater gains ( $M=0.12$ ,  $SD =0.25$ ) than students who enrolled in integrated courses ( $M =0.11$ ,  $SD =0.25$ ). However, the results of the  $t$ -test ( $M = -0.01$ ,  $SD = 0.31$ ,  $t(200) = .03$ ,  $p = 0.97$ ) reveal no statistically significant difference in gain scores. Therefore, the null hypothesis that no statistical difference will exist in students' performance on items related to multiplicative and proportional thinking is not rejected. Furthermore, Cohen's effect size value ( $d = .04$ ) suggests a low practical significance.

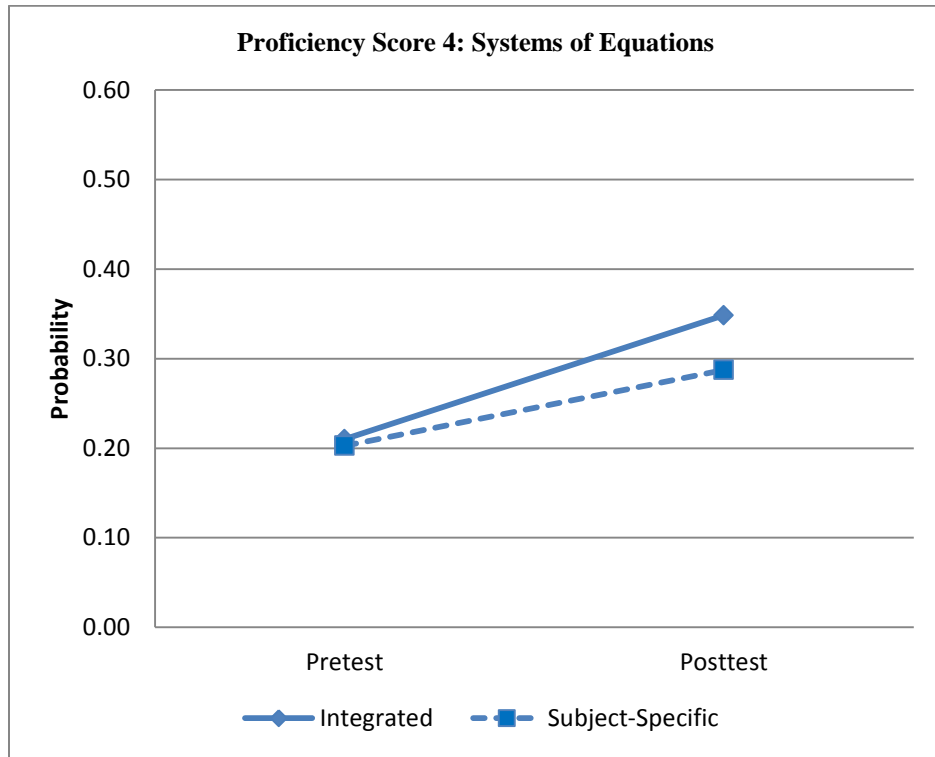
Hypothesis 4 states no statistical difference in mean gain scores on tasks related to linear equivalents will exist between students who learn from subject-specific and integrated course pathways.



**Figure 4.** Students' proficiency growth on content related to linear equivalents.

Figure 4 suggests no major difference exists in the mean gain scores of students in the two sample groups on items related to linear equivalents on the algebra assessment. The gains of students who enrolled in subject-specific courses is similar to the gains ( $M=0.22$ ,  $SD = 0.29$ ) of students who enrolled in integrated courses ( $M =0.22$ ,  $SD =0.30$ ). Furthermore, the results of the  $t$ -test ( $M = 0.00$ ,  $SD = 0.39$ ,  $t(200) = 0.43$ ,  $p = 0.67$ ) reveal no statistically significant difference in gain scores. Therefore, the null hypothesis that no statistical difference will exist in students' performance on items related to linear equivalents is not rejected. Furthermore, Cohen's effect size value ( $d = .001$ ) suggests a low practical significance.

Hypothesis 5 states no statistical difference in mean gain scores on tasks related to systems of linear equations will exist between students who learn from subject-specific and integrated course pathways.

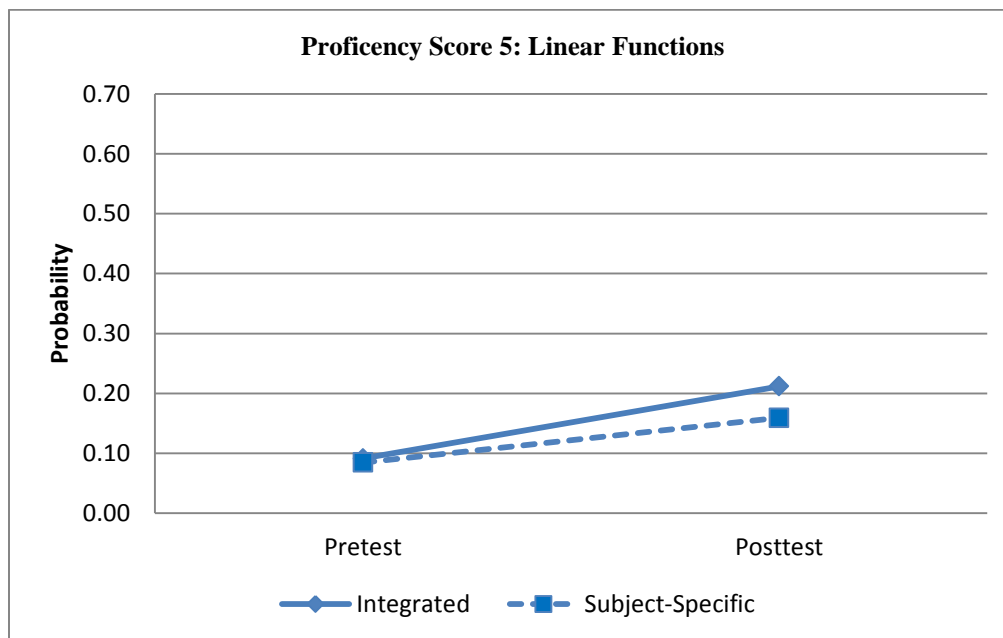


**Figure 5.** Students' proficiency growth on content related to systems of equations.

Figure 5 suggests no major difference exists in the mean gain scores of students in the two sample groups on items related to systems of equations on the algebra assessment. Students who enrolled in integrated courses had greater gains ( $M = 0.14$ ,  $SD = 0.25$ ) than students who enrolled in subject-specific courses ( $M = 0.08$ ,  $SD = 0.23$ ). However, the results of the  $t$ -test ( $M = 0.05$ ,  $SD = 0.35$ ,  $t(200) = -0.05$ ,  $p = 0.96$ ) reveal no statistically significant difference in gain scores. Therefore, the null hypothesis that no statistical difference will exist in students' performance on items related to systems of equations is not rejected. Furthermore, Cohen's effect size value ( $d = .25$ ) suggests a low practical significance.



Hypothesis 6 states no statistical difference in mean gain scores on tasks related to linear functions will exist between students who learn from subject-specific and integrated course pathways.



**Figure 6.** Students' proficiency growth on content related to linear functions.

Figure 6 suggests no major difference exists in the mean gain scores of students in the two sample groups on items related to linear functions on the algebra assessment. Students who enrolled in integrated courses had greater gains ( $M = 0.12$ ,  $SD = 0.26$ ) than students who enrolled in subject-specific courses ( $M = 0.07$ ,  $SD = 0.24$ ). However, the results of the  $t$ -test ( $M = 0.05$ ,  $SD = 0.33$ ,  $t(200) = -0.43$ ,  $p = 0.66$ ) reveal no statistically significant difference in gain scores. Therefore, the null hypothesis that no statistical difference will exist in students' performance on items related to linear functions is not rejected. Furthermore, Cohen's effect size value ( $d = .69$ ) suggests a moderate to high practical significance in this sample.

Collectively, the results of the  $t$ -tests reveal no statistically significant differences on the algebra assessment exist between students in the two groups. However, the results suggest

content organization may have a practical significance on students' proficiency with certain algebraic concepts. In particular, on the overall assessment, the estimated effect size is less than 0.20, which suggests content organization may have low practical significance with regards to students' overall algebraic performance. This effect size supports the  $p$ -value related to the hypothesis that no statistical differences exist between students' performance on the overall assessment. However, the effect sizes vary in magnitude when items related to specific concepts are investigated. Specifically, on concepts related to proficiencies 1, 2, 3 and 4, the calculated effect sizes were less than or equal to 0.25, which suggests content organization may have low practical significance with regards to students' performance on the concepts related to the proficiency levels. In addition, these effect sizes support the  $p$ -values related to the hypothesis that no statistical differences exist between students' performance on the items related to these respective proficiency levels. In contrast to concepts related to the lower proficiency areas (i.e., proficiency 1, 2, 3, 4), the calculated effect size for concepts related to proficiency 5 was 0.69. This medium effect size does not support the  $p$ -value related to the hypothesis that no statistical differences exist between students' performance on items related to proficiency 5. In particular, the effect size suggests content organization may have moderate practical significance with regards to students' performance on the concepts related to proficiency 5. These results related to proficiency 5 warrant further investigations into the effect of content organization on students' performance on concepts related to linear functions. The next chapter provides suggestions for future research.

## **Chapter 5: Discussion and Suggestions for Future Research**

The purpose of this study was to investigate the following question, “How do the algebraic performance gains of students enrolled in integrated course pathways relate to the algebraic performance gains of students enrolled in subject-specific course pathways?” Collectively, on the overall algebra assessment and specific content items, the results of the six hypothesis tests reveal no statistically significant differences exist in the algebraic performance gains between the high school students who learned mathematics from integrated course pathways and the high school students who learned from subject-specific course pathways. As a result, the six null hypotheses were not rejected. Consistent with expectations of the National Mathematics Advisory Panel (NMAP, 2008), the findings suggest students can perform comparably through algebraic content regardless of whether the students enroll in a subject-specific or integrated course pathway. Furthermore, the results are consistent with claims that learning mathematics from an integrated approach is a viable option for high school students, when compared to learning from a subject-specific approach (Reys & Reys, 2009).

The results from this study contribute to the knowledge base related to the effect of content organization on students’ algebraic performance. However, there is still a need for more studies in which relationships between content organization and students’ mathematics learning are investigated. In the rest of this chapter, I provide suggestions for future research.

## **Suggestions for Future Research**

**Related to the COSMIC Project.** Interestingly, the results of this study revealed no statistically significant differences exist between the groups' performance on items related to linear functions (i.e., proficiency 5). However, the medium effect size related to proficiency 5 suggests it is possible that content organization may have some practical significance on students' performance on items related to linear functions (i.e., proficiency 5). As such, further studies on the effect of content organization on students' performance on concepts related to linear functions are warranted. In particular, the non-statistically significant findings from this study contradict findings from prior studies of content organization that suggest high school students who study from an integrated curriculum are advantaged over students who study from subject-specific curricula (Chavez et al., 2013; Grouws et al., 2013; Tarr et al., 2013) during the first three years of high school. The results of studies from the COSMIC project are based on holistic measures. The researchers did not investigate the performance of students on items related to specific concepts. Based on the results of the current study, it is possible the statistical differences found in the COSMIC study are related to students' performance on specific concept items, such as linear functions. Given the medium effect size found between the sample groups on items related to linear functions in the current study, it would be interesting to observe whether the COSMIC researchers would find consistent results from their study, if specific items on the assessment were examined, and particularly on items related to linear functions.

Despite strengths of this study, such as the longitudinal design, the theoretical framework, and the algebra assessment scaled to different proficiency levels, like many other NCES data sets, the HSLs is observational. As such, students were not randomly assigned to course pathways. Because students were not randomly assigned to course pathways, caution

should be taken when making causal inferences based on the results of this study. Similar to this study, students who participated in the COSMIC project (Chavez et al., 2013; Grouws et al., 2013; Tarr et al., 2013) were not randomly assigned to course pathways. As a result, the studies were threatened by the issue of selection bias. In particular, the researchers used propensity scores and found that Hispanics and African-American students were more likely to be assigned to subject-specific mathematics courses. Findings from the studies reveal that Hispanic and African-American students performed statistically lower than White students. Therefore, it is possible that the greater percentage of minority students enrolled in subject-specific courses negatively biased the results to suggest that students in integrated courses statistically outperformed students in subject-specific courses. The COSMIC researchers could have reduced the threat of selection bias by employing the technique of propensity score matching, as was done in this study. It would be interesting to observe whether the findings from the COMSIC project would be the same if the data were re-analyzed and propensity score matching is employed.

**Related to Algebra and College Readiness.** The findings of this study are based only on the first three years of high school mathematics, and up to Algebra II or Integrated III. However, research (ACT, 2012; Aughinbaugh, 2012) suggests students who take courses beyond the level of Algebra II increase their chances of being prepared for college level courses. Therefore, further research is needed that explores students' algebraic performance as they complete at least four years of high school mathematics. In particular, as more data for the HSLS observational study are collected, researchers should investigate relationships between specific algebra related content topics and high school students' probability of enrolling in college-level mathematics courses when they graduate, rather than having to enroll in remedial

mathematics courses. A study of this nature can be used to determine the extent to which course pathways and performance on specific algebraic concepts predict students' readiness for college level courses.

**Investigation of Other Course Pathways.** In addition to longitudinal studies of more than three years of high school mathematics, other areas of research related to content organization are needed. For example, the results of this study are only generalizable to subject-specific and integrated course pathways. Yet, there are other course pathways in which the content is organized in a different manner that should be explored. For example, the HSLs indicate some students may have enrolled in an *Algebra-Specific* course pathway (Algebra I-Algebra II-Algebra III) in which the primary emphasis of each course is algebra related content. An exploration of the HSLs data reveals over 300 students were enrolled in an Algebra I course in their ninth grade and an Algebra III course during their eleventh grade year. Therefore, future studies should investigate the effects of an *Algebra-Specific* course pathway on students' algebraic performance compared to the effects of non-Algebra-specific course pathways (i.e., Subject-Specific or Integrated) on students' algebraic performance. If a goal is to prepare students to enroll in college level mathematics, specifically College Algebra and subsequently a career in a STEM discipline, then researchers should explore students' mathematics learning and performance when the primary learning focus of the course pathway is algebra related content because algebra is considered the gatekeeper to higher mathematics.

**Investigation of Various Student Groups.** The small sample size used in this study did not allow for the comparison of different student sub-groups. An examination of different student sub-groups would provide additional insights about the effects of specific course pathways on different groups of students based on race, gender, or socioeconomic status.

Research of this nature would also add insight related to whether content organization may have an influence on various student subgroups. Therefore, investigations on the effect of content organization on various student groups are needed.

**Related to Data Accuracy and Validity.** Because the study used a pre-existing data set, it was limited by the data that were collected. As such, another area for future research relates to the validity and accuracy of the observational data. For instance, the intent of this study was to examine the performance of high school students during their first three years of high school. However, the HSL09 only included information about courses in which students enrolled during their 9<sup>th</sup> and 11<sup>th</sup> grade years in school. No information was available about the courses in which students were enrolled during their 10<sup>th</sup> grade year. As a result, inferences were made about the pathway in which the student studied based on course assignments in the 9<sup>th</sup> and 11<sup>th</sup> grades. In regards to course enrollment, researchers suggest transcript data would be a more reliable source of information, compared to student self-reported data (Attewell et al., 2006). Therefore, future analysis of course pathways should include data based on the students' high school transcripts, if possible.

**Related to Curriculum Materials.** Another limitation of using the HSL09 data set is no information was provided about the actual curriculum materials implemented in schools. This type of information is important because algebraic content in curriculum materials implemented between and within the specific course pathways are organized differently (Huntley & Terrell, 2014; Chavez et al., 2009). Knowledge of the implemented curriculum materials can potentially reveal how the different ways mathematics content is sequenced can influence student learning. Unfortunately, because this information was not available, the specific curriculum materials used in the classrooms was not a predictor variable in this study. In addition, this study does not take

into account other external factors that can moderate the effect of course pathways, such as the fidelity with which the content was implemented or the amount of the content taught. Although the findings from this study can be used to inform decisions related to the course pathway implemented in schools, the results of this study should not be used to inform decisions on the curriculum materials that should be implemented in schools or classrooms. More studies that compare the effect of specific curriculum materials are needed. Moreover, because differences exist between and within textbooks developed for the specific course pathways (Chavez et al., 2009; Huntley & Terrell, 2014) future studies on the effect of textbooks should focus on specific textbook series, compared to grouping multiple textbooks into a single category.



## References

- Attewell, P., Lavin, D., Domina, T., & Levey, T. (2006). New evidence on college remediation. *Journal of Higher Education*, 77(5), 886–924.
- Aughinbaugh, A. (2012). The effects of high school math curriculum on college attendance: Evidence from the NLSY97. *Economics of Education Review*, 31(6), 861-870.
- Bell-Ellison, B. A., & Kromrey, J. D. (2007). *Alternatives for analysis of complex sample surveys: A comparison of SAS®, SUDAAN®, and AM® Software* (Unpublished manuscript, University of South Florida).
- Bozick, R., and Ingels, S.J. (2007). *Mathematics course-taking and achievement at the end of high school: Evidence from the Education Longitudinal Study of 2002 (ELS:2002)* (NCES 2008-319). National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education. Washington, DC.
- Cai, J., Moyer, J., Wang, N., Hwang, S., Nie, B., & Garber, T. (2013). Mathematical problem posing as a measure of curricular effect on students' learning. *Educational Studies in Mathematics*, 83(1), 57–69.
- Chávez, O., Papick, I., Ross, D. J., & Grouws, D. A. (2010). Developing fair tests for mathematics curriculum comparison studies: The role of content analyses. *Mathematics Education Research Journal*, 23(4), 397–416.

- Cai, J., Wang, N., Moyer, J. C., Wang, C., & Nie, B. (2011). Longitudinal investigation of the curricular effect: An analysis of student learning outcomes from the the LieCal Project in the United States. *International Journal of Educational Research*, 50, 117-136.
- Chavez, O., Tarr, J. E., Grouws, D. A., & Soria, V. M. (2013). Third-year high school mathematics curriculum: Effects of content organization and curriculum implementation. *International Journal of Science and Mathematics Education*. (13)1, 97-120.
- Chávez, Ó., Grouws, D. A., Tarr, J. E., Ross, D. J., & McNaught, M. D. (2009). *Mathematics curriculum Implementation and linear functions in secondary mathematics: Results from the comparing options in secondary mathematics project*. Paper presented at the American Education Research Association, San Diego, CA.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Coxford, A. F., Fey, J. T., Hirsch, C. R., Schoen, H.L., Burrill, G., Hart, E. W. et al. (1998). *Contemporary mathematics in context: A unified approach (Courses 1--4)*. New York: Glencoe/McGraw-Hill.
- Dewey, J. (1916). *Democracy and education*. New York: The Macmillan Company.
- DuGoff, E. H., Schuler, M., & Stuart, E. A. (2014). Generalizing observational study results: Applying propensity score methods to complex surveys. *Health Services Research*, 49(1), 284-303.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational Research: An Introduction*. Pearson.
- Gamoran, A., & Hannigan, E. C. (2000). Algebra for everyone? Benefits of college-preparatory mathematics for students of diverse abilities in early secondary. *Educational Evaluation and Policy Analysis*, 22, 241-254.

- Graham, S. E. (2010). Using propensity scores to reduce selection bias in mathematics education research. *Journal for Research in Mathematics Education*, 41(2), 147-168.
- Grouws, D. A., Tarr, J. E., Chavez, O., Sears, R., Soria, V., & Taylan, R. D. (2013). Curriculum and implementation effects on high school students' mathematics learning from curricula representing subject-specific and integrated content organizations. *Journal for Research in Mathematics Education*, 44(2), 416-463.
- Harwell, M. R., Post, T. P., Cutler, A., Maeda, Y, Anderson, E., Norman, K. W., & Medhanie, A. (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, 203–231.
- Harwell, M. R., Medhanie, A., Post, T. P., Norman, K. W., & Dupuis, D.N. (2012). Preparation of students completing a Core-Plus or commercially developed high school mathematics curriculum for intense college mathematics coursework. *The Journal of Experimental Education*, 80:1, 96-112.
- Harwell, M. R., Post, T. R., Medhanie, A., Dupuis, D., & LeBeau, B. (2013). A multi-institutional study of the impact of high school mathematics curricula on college mathematics achievement and course taking. *Journal for Research in Mathematics Education*, 44(5), 742-774.
- Hess M.R. (2003). *Effect sizes, significance tests, and confidence intervals: assessing the influence and impact of research reporting protocol and practice* [dissertation]. Tampa (FL): University of South Florida.
- Holliday, B., Cuevas, G. J., Moore-Harris, B., Carter, J. A., Marks, D., Casey, R. M., . . . Hayek, L. M. (2005). *Algebra 1*. New York: Glencoe.

- 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.
- Huntley, M. A., & Terrell, M. (2014). One-step and multi-step linear equations: a content analysis of five textbook series. *ZDM - The International Journal on Mathematics Education*, 46(5), 751-766.
- Ingels, S. J., Pratt, D. J., Herget, D. R., Burns, L. J., Dever, J. A., Ottem, R., . . . LoGerfo, L. (2011). *High School Longitudinal Study of 2009 (HSLs:09): Base-year data file documentation*. U.S. Department of Education; Institute of Education Sciences; National Center for Education Statistics.
- Johnson, P. (2007). What are we developing? A case study of a college mathematics program. *School Science & Mathematics*, 107(7), 279-292.
- Lanehart, R. E., Rodriguez, P., Kim, E. S., Ballara, A. P., Kromrey, J. D., & Lee, R. S. (2012). *Propensity score analysis and assessment of propensity score approaches using SAS procedures*. SAS Global Forum.
- LeBeau, B., Harwell, M., Monson, D., Dupuis, D., Medhanie, A., & Post, T. R. (2012). Student and high school characteristics related to completing a science, technology, engineering or mathematics (STEM) major in college. *Research in Science & Technological Education*, 30 (1), 17-28.
- Lubienski, S. T. (2008). On “gap gazing” in mathematics education: The need for gaps analyses. *Journal for Research in Mathematics Education*, 39, 350–356.

- Morris, S. B. (2008). Estimating effect sizes from pretest–posttest–control group designs. *Organizational Research Methods, 11*, 364–386
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010a). *Common core state standards for mathematics*. Retrieved from <http://www.corestandards.org>
- National Governors Association Center for Best Practices & Council of Chief State School Officers. (2010b). *Common core state standards for mathematics, Appendix A: Designing high school mathematics courses based on the common core state standards*. Retrieved from [http://www.corestandards.org/assets/CCSSI\\_Mathematics\\_Appendix\\_A.pdf](http://www.corestandards.org/assets/CCSSI_Mathematics_Appendix_A.pdf)
- Post, T. R., Medhanie, A., Harwell, M. R., Norman, K., Dupuis, D., Muchlinski, T., . . . Monson, D. (2010). Prior achievement, course-taking patterns, and persistence. *Journal for Research in Mathematics Education, 41*(3), 274–308.
- Reys, R. E., & Reys, R. R. (2009). Two high school mathematics curricula paths—which one to take? *Mathematics Teacher, 102*(8), 568–570.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika, 70*, 41-55.
- Rubin, D. B. (1997). Estimating causal effects from large data sets using propensity scores. *Annals of Internal Medicine, 127*, 757-763.
- Schneider, B., Carnoy, M., Kirkpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). *Estimating causal effects using experimental and observational designs (report from the Governing Board of the American Educational Research Association Grants Program)*. Washington, D.C.: American Educational Research Association.

Thompson, D. R., & Senk, S. L. (Eds.) (2003). *Standards-based school mathematics curricula: What are they? What do students learn?* Mahwah, NJ: Lawrence Erlbaum.

Tarr, J. E., Grouws, D. A., Chávez, Ó., & Soria, V. M. (2013). The effects of content organization and curriculum implementation on students' mathematics learning in second-year high school courses. *Journal for Research in Mathematics Education*, 44(4), 683-729.

## **Appendix A: List of Variables used from HSLs Dataset**

**List of variables used from HSLS dataset.**

S1ALG1M09	Student 9 <sup>th</sup> Grade Course is Algebra I	Part of the base year survey. This variable indicates whether the student was enrolled in Algebra I in the ninth grade.
S1INTGM109	Student 9 <sup>th</sup> Grade Course is Integrated Mathematics I	Part of the base year survey. This variable indicates whether the student was enrolled in Integrated Mathematics I in the ninth grade.
S2ALG2M12	Student 11 <sup>th</sup> Grade Course is Algebra II	Part of the first follow-up survey. This variable indicates whether the student was enrolled in Algebra II in the eleventh grade.
S2INTGM312	Student 11 <sup>th</sup> Grade Course is Integrated Mathematics III	Part of the first follow-up survey. This variable indicates whether the student was enrolled in Integrated Mathematics III in the eleventh grade.
X1SEX	Gender	Sex of the sample member, taken from the base year student questionnaire, parent questionnaire, and/or school-provided sampling roster. If the sex indicated by any of these three sources was inconsistent, X1SEX was coded based on manual review of the sample member's first name.
X1RACE	Race	X1RACE characterizes the sample member's race/ethnicity by summarizing the following six dichotomous race/ethnicity composites: X1HISPANIC, X1WHITE, X1BLACK, X1ASIAN, X1PACISLE, and X1AMINDIAN. The dichotomous race/ethnicity composites are based on data from the student questionnaire, if available; if not available from the student questionnaire, they are based on, in order of preference, data from the school-provided sampling roster or data from the parent questionnaire.
X1SES	Socioeconomic status	Part of the base year survey. This composite variable is used to measure a construct for socioeconomic status. X1SES is calculated using parent/guardians' education (X1PAR1EDU and X1PAR2EDU), occupation (X1PAR1OCC2 and X1PAR2OCC2), and family income (X1FAMINCOME). For cases with non-responding parent/guardians, 5 imputed values are generated (X1SES1-X1SES5), X1SES is computed as the average of the 5 imputed values, and the imputation flag is set as X1SES_IM=1 (values for parent/guardian education, occupation, and income are set to -8).



X1TXMSCR	Mathematics IRT-estimated number right score	The math IRT-estimated scale score is a criterion-referenced measure of achievement. The criterion is the set of skills defined by the HSLS:09 framework and represented by the 72 items in the HSLS:09 math item pool. The estimated scale score for math is an estimate of the number of items students would have answered correctly had they responded to all 72 items in the item pool. The ability estimates and item parameters derived from the IRT calibration can be used to calculate each student's probability of a correct answer for each of the items in the pool. These probabilities are summed to produce the IRT-estimated number-correct scale score.
X1TXMPROF1	Proficiency with Algebraic Expressions	The variable is a measure of students' proficiency with algebraic expressions. It is part of the base year survey. Students able to answer questions related to algebraic expression have an understanding of algebraic basics, including evaluating simple algebraic expressions and translating between verbal and symbolic representations of expressions.
X1TXMPROF2	Proficiency with Multiplicative and Proportional Thinking	Part of the base year survey. The variable is a measure of students' proficiency with Multiplicative and Proportional Thinking. Students able to answer questions related to multiplicative and proportional thinking can solve proportional situation word problems, find the percent of a number, and identify equivalent algebraic expressions for multiplicative situations.
X1TXMPROF3	Proficiency with Linear Equivalents	Part of the base year survey. The variable is a measure of students' proficiency with linear equivalents. Students able to answer questions related to algebraic equivalents can link equivalent tabular and symbolic representations of linear equations, identify equivalent lines and find the sum of variable expressions.
X1TXMPROF4	Proficiency with Systems of Linear Equations	Part of the base year survey. The variable is a measure of students' proficiency with systems of linear equations. Students able to answer the questions related to systems of equations can solve such systems algebraically and graphically and characterize lines (parallel, intersecting, collinear) represented by a system of linear equations.
X1TXMPROF5	Proficiency with Linear Functions	Part of the base year survey. The variable is a measure of students' proficiency with linear functions. Students able to answer the questions related to linear functions can find and use slopes and intercepts of lines, and use functional notation.
W2W1STU	First follow-up student longitudinal weight	Student weight used in analysis of data from base year and first follow-up responding students.

X2TXMSCR	Mathematics IRT-estimated number right score	The math IRT-estimated number correct is a criterion-referenced measure at the time of the First Follow-up assessment. The criterion is the set of skills defined by both the HSLs:09 Base Year and First Follow-up framework and represented by the 118 items used to score HSLs:09 First Follow-up math assessment. The estimated number correct for math is an estimate of the number of items students would have answered correctly had they responded to all 118 items in the item pool.
X2TXMPROF1	Proficiency with Algebraic Expressions.	Part of the first follow-up survey. The variable is a measure of students' proficiency with algebraic expressions. Students able to answer questions related to algebraic expression have an understanding of algebraic basics, including evaluating simple algebraic expressions and translating between verbal and symbolic representations of expressions.
X2TXMPROF2	Proficiency with Multiplicative and Proportional Thinking	Part of the first follow-up survey. The variable is a measure of students' proficiency with Multiplicative and Proportional Thinking. Students able to answer questions related to multiplicative and proportional thinking can solve proportional situation word problems, find the percent of a number, and identify equivalent algebraic expressions for multiplicative situations.
X2TXMPROF3	Proficiency with Linear Equivalents	Part of the first follow-up survey. The variable is a measure of students' proficiency with linear equivalents. Students able to answer questions related to algebraic equivalents can link equivalent tabular and symbolic representations of linear equations, identify equivalent lines and find the sum of variable expressions.
X2TXMPROF4	Proficiency with Systems of Linear Equations	Part of the first follow-up survey. The variable is a measure of students' proficiency with systems of linear equations. Students able to answer the questions related to systems of equations can solve such systems algebraically and graphically and characterize lines (parallel, intersecting, collinear) represented by a system of linear equations.
X2TXMPROF5	Proficiency with Linear Functions.	Part of the first follow-up survey. The variable is a measure of students' proficiency with linear functions. Students able to answer the questions related to linear functions can find and use slopes and intercepts of lines, and use functional notation.

## **Appendix B: SAS Syntax Code**

```

data student;
set lib.Hsls_student;
stud_id = stu_id + 0;
Course_Pathways = s2intgm312 + 0;
run;

```

**data** courses; \* This step is used to create a dataset with students who began the specific course pathway in 9th grade ;

```

set student;
if
(s1alg1m09=1 and s2alg2m12=1 and s1intgm109=0 and s2intgm312=0 )or
(s1alg1m09=0 and s2alg2m12=0 and s1intgm109=1 and s2intgm312=1 );
run;

```

```

proc logistic descending data = courses;
title 'Propensity Score Estimation';
model Course_Pathways = x1sex x1black x1hispanic x1white
x1txmprof1 x1txmprof2 x1txmprof3 x1txmprof4 x1txmprof5 x1ses w2w1stu
/lackfit outroc = ps_r; output out= ps_p XBETA=ps_xb STDXBETA= ps_sdx PREDICTED =
ps_pred;
run;

```

```

data one;
set ps_p;
ranvar = ranuni(0);
proc sort data = one ;
by Course_Pathways ranvar;
proc transpose data = one out = data1;
by Course_Pathways;
data id_t (rename=(COL1-COL73 = tid1-tid73));
* Note: N of columns is number of obs in treatment group;
set data1; if Course_Pathways = 1 and _NAME_ = 'stud_id';
data ps_t (rename=(COL1-COL73 = tps1-tps73));
set data1; if Course_Pathways = 1 and _NAME_ = 'ps_pred';
data id_c (rename=(COL1-COL4956 = cid1-cid4956));
* Note: N of columns is number of obs in control group;
set data1; if Course_Pathways = 0 and _NAME_ = 'stud_id';
data ps_c (rename=(COL1-COL4956 = cps1-cps4956));
set data1; if Course_Pathways = 0 and _NAME_ = 'ps_pred';
data all;
merge id_t ps_t id_c ps_c;
caliper = .1; * Note: caliper for matching is specified here;
array treat_id {*} tid1-tid73;
array ctl_id {*} cid1-cid4956;

```

```

array treat_p {*} tps1-tps73;
array ctl_p {*} cps1-cps4956;
array used_i {*} used1 - used4956;
array matched_t {*} m_tid1-m_tid73;
array matched_c {*} m_cid1-m_cid4956;
match_N = 0;
do i = 1 to 73;
  min_diff = 1;
  best_match = 0;

  do j = 1 to 4956;
    if used_i[j] = . then do;
      if ABS(treat_p[i] - ctl_p[j]) < caliper then do;
        if ABS(treat_p[i] - ctl_p[j]) < min_diff then do;
          min_diff = ABS(treat_p[i] - ctl_p[j]);
          best_match = j;
        end;
      end;
    end;
  end;
  if best_match > 0 then do;

    match_N = match_N + 1;
    used_i[best_match] = 1;
    matched_t[match_N] = treat_id[i];
    matched_c[match_N] = ctl_id[best_match];
  end;
end;
data matches;
set all;
array matched_t {*} m_tid1-m_tid73;
array matched_c {*} m_cid1-m_cid4956;
do match = 1 to match_N;
  Integrated_IDN = matched_t[match];
  SubjectSpecific_IDN = matched_c[match];
output;
end;
keep match Integrated_IDN SubjectSpecific_IDN ;
proc print;
var match Integrated_IDN SubjectSpecific_IDN ;
title 'Matched Observations in Integrated and Subject Specific Groups';
run;

proc sort data = ps_p;
by Course_Pathways;
proc boxplot data = ps_p;

```

```

symbol width = 2;
plot ps_pred*Course_Pathways /
cboxes=black
cframe = white
idsymbol = circle
idcolor = black
font='times new roman' height=3.5
boxwidth=6
boxstyle=schematic
waxis = 2;
run;

```

```

data Integrated;
set ps_p;
Integrated = course_pathways ;
Integrated_idn = stud_id ;
Integratedpreoverallscore = x1txmscr ;
Integratedpreprof1 = x1txmprof1 ;
Integratedpreprof2 = x1txmprof2 ;
Integratedpreprof3 = x1txmprof3 ;
Integratedpreprof4 = x1txmprof4 ;
Integratedpreprof5 = x1txmprof5 ;
Integratedpostoverallscore = x2txmscr ;
Integratedpostprof1 = x2txmprof1 ;
Integratedpostprof2 = x2txmprof2 ;
Integratedpostprof3 = x2txmprof3 ;
Integratedpostprof4 = x2txmprof4 ;
Integratedpostprof5 = x2txmprof5 ;
Integratedsex = x1sex ;
Integratedses = x1ses ;
Integratedrace = x1race ;
Integratedblack = x1black ;
Integratedwhite = x1white ;
Integratedhispanic = x1hispanic ;
Integratedwt2 = w2w1stu ;
Integratedps_pred = ps_pred;

```

```

data subject_specific;
set ps_p ;
subject_specific = course_pathways ;
SubjectSpecific_idn = stud_id ;
subject_specificpreoverallscore = x1txmscr ;
subject_specificpreprof1 = x1txmprof1 ;
subject_specificpreprof2 = x1txmprof2 ;
subject_specificpreprof3 = x1txmprof3 ;
subject_specificpreprof4 = x1txmprof4 ;

```

```
subject_specificpreprof5 = x1txmprof5 ;
subject_specificpostoverallscore = x2txmscr ;
subject_specificpostprof1 = x2txmprof1 ;
subject_specificpostprof2 = x2txmprof2 ;
subject_specificpostprof3 = x2txmprof3 ;
subject_specificpostprof4 = x2txmprof4 ;
subject_specificpostprof5 = x2txmprof5 ;
subject_specificsex = x1sex ;
subject_specificses = x1ses ;
subject_specificrace = x1race ;
subject_specificblack = x1black ;
subject_specificwhite = x1white ;
subject_specifichispanic = x1hispanic ;
subject_specificwt2 = w2w1stu;
subject_specificps_pred = ps_pred;
```

```
proc sort data = Integrated ;
by integrated_idn ;
```

```
proc sort data = matches ;
by integrated_idn ;
```

```
data combine1 ;
merge matches Integrated;
by integrated_idn ;
if match = . then delete;
run;
```

```
proc sort data = subject_specific ;
by subjectspecific_idn ;
```

```
proc sort data = combine1 ;
by subjectspecific_idn ;
```

```
data combine2 ;
merge subject_specific combine1 ;
by subjectspecific_idn ;
if match = . then delete;
run;
```

```
data final;
set combine2;
Integrateddiffoverall = Integratedpostoverallscore-Integratedpreoverallscore;
Integrateddiffprof1 = Integratedpostprof1-Integratedpreprof1;
Integrateddiffprof2 = Integratedpostprof2-Integratedpreprof2;
```

```
Integrateddiffprof3 = Integratedpostprof3-Integratedpreprof3;  
Integrateddiffprof4 = Integratedpostprof4-Integratedpreprof4;  
Integrateddiffprof5 = Integratedpostprof5-Integratedpreprof5;
```

```
subject_specificdiffoverall = subject_specificpostoverallscore-subject_specificpreoverallscore;  
subject_specificdiffprof1 = subject_specificpostprof1-subject_specificpreprof1;  
subject_specificdiffprof2 = subject_specificpostprof2-subject_specificpreprof2;  
subject_specificdiffprof3 = subject_specificpostprof3-subject_specificpreprof3;  
subject_specificdiffprof4 = subject_specificpostprof4-subject_specificpreprof4;  
subject_specificdiffprof5 = subject_specificpostprof5-subject_specificpreprof5;
```

```
diffingains_overall = subject_specificdiffoverall - Integrateddiffoverall ;  
diffingains_prof1 = subject_specificdiffprof1 - Integrateddiffprof1;  
diffingains_prof2 = subject_specificdiffprof2 - Integrateddiffprof2 ;  
diffingains_prof3 = subject_specificdiffprof3 - Integrateddiffprof3 ;  
diffingains_prof4 = subject_specificdiffprof4 - Integrateddiffprof4 ;  
diffingains_prof5 = subject_specificdiffprof5 - Integrateddiffprof5 ;  
run;
```

\* The following procedures are used to normalize the weights ;

```
proc means data = final n sum mean;  
var Integratedwt2 subject_specificwt2;  
output out = q mean = mn_wt;  
run;  
data ps_normalizedweight;  
set final;  
if _n_ = 1 then set q;  
retain mn_wt;  
set final;  
wt2 = Integratedwt2/mn_wt; * Normalized weight;  
run;  
proc means data = ps_normalizedweight n sum mean ;  
var wt2;  
run ;
```

```
proc surveymeans data = ps_normalizedweight nobs mean T varmethod = BRR ;  
repweights w2w1stu001 - w2w1stu200;  
var  
diffingains_overall diffingains_prof1 diffingains_prof2 diffingains_prof3  
diffingains_prof4 diffingains_prof5 ;  
weight wt2 ;  
run ;
```



```
proc means data= ps_normalizedweight ;  
var  
diffingains_overall diffingains_prof1 diffingains_prof2  
diffingains_prof3 diffingains_prof4 diffingains_prof5  
;  
run;
```

```
proc means data= ps_normalizedweight ;  
var  
Integratedpreoverallscore  
Integratedpreprof1  
Integratedpreprof2  
Integratedpreprof3  
Integratedpreprof4  
Integratedpreprof5  
;  
run;
```

```
proc means data= ps_normalizedweight ;  
var  
Integratedpostoverallscore  
Integratedpostprof1  
Integratedpostprof2  
Integratedpostprof3  
Integratedpostprof4  
Integratedpostprof5  
;  
run;
```

```
proc means data= ps_normalizedweight ;  
var  
subject_specificpreoverallscore  
subject_specificpreprof1  
subject_specificpreprof2  
subject_specificpreprof3  
subject_specificpreprof4  
subject_specificpreprof5  
;  
run;
```

```
proc means data= ps_normalizedweight ;
```

```
var
subject_specificpostoverallscore
subject_specificpostprof1
subject_specificpostprof2
subject_specificpostprof3
subject_specificpostprof4
subject_specificpostprof5
;
run;
```

```
proc means data= ps_normalizedweight ;
var
subject_specificses integratedses
;
run;
```

```
proc means data= ps_normalizedweight ;
var
subject_specificwt2 Integratedwt2
;
run ;
```

```
proc means data= ps_normalizedweight ;
var
subject_specificps_pred Integratedps_pred
;
run ;
```

```
proc freq data=ps_normalizedweight;
title 'all_welldefined_students';
tables Course_Pathways*integratedrace Course_Pathways*subject_specificrace
Course_Pathways*integratedsex Course_Pathways*subject_specificsex
Course_Pathways*integratedblack Course_Pathways*subject_specificblack
Course_Pathways*integratedwhite Course_Pathways*subject_specificwhite
Course_Pathways*integratedhispanic Course_Pathways*subject_specifichispanic
;
run;
```

```
proc freq data=courses;
title 'All Students before matching';
```

```
tables Course_Pathways*x1race Course_Pathways*x1sex ;  
run;
```

```
data courses_integrated ;  
set courses ;  
if course_pathways =1 ;  
run;
```

```
proc freq data=courses_integrated ;  
title 'Students in Integrated courses';  
tables course_pathways ;  
run;
```

```
proc means data = courses_integrated ;  
var  
course_pathways x1txmscr x1txmprof1 x1txmprof2 x1txmprof3 x1txmprof4 x1txmprof5  
;  
run ;
```

```
proc means data = courses_integrated ;  
var  
course_pathways x2txmscr x2txmprof1 x2txmprof2 x2txmprof3 x2txmprof4 x2txmprof5 ;  
run ;
```

```
proc means data = courses_integrated ;  
var  
course_pathways x1ses w2w1stu ;  
run ;
```

```
data courses_subject_specific ;  
set courses ;  
if course_pathways = 0 ;  
run;
```

```
proc freq data=courses_subject_specific ;  
title 'Students in Subject-Specific courses';  
tables course_pathways ;  
run;
```

```
proc means data = courses_subject_specific ;  
var
```

```
course_pathways x1txmscr x1txmprof1 x1txmprof2 x1txmprof3 x1txmprof4 x1txmprof5 ;  
run ;
```

```
proc means data = courses_subject_specific ;  
var  
course_pathways x2txmscr x2txmprof1 x2txmprof2 x2txmprof3 x2txmprof4 x2txmprof5 ;  
run ;
```

```
proc means data = courses_subject_specific ;  
var  
course_pathways x1ses w2w1stu ;  
run ;
```