WHAT DOES THE NATION’S REPORT CARD SAY ABOUT EARLY PHYSICS MODELS?
USING THE NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS HIGH SCHOOL TRANSCRIPT STUDY 2009 (NAEP HSTS 2009) TO EXAMINE INVERTED CURRICULUM

by

DAVID H. KHALIQI

M.A., University of Colorado Colorado Springs, 2004

A dissertation submitted to the Graduate Faculty of the
University of Colorado Colorado Springs
in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
Department of Education Leadership, Research, and Foundations

2014
This dissertation for the Doctor of Philosophy degree by

David H. Khaliqi

has been approved for the

Department of Education Leadership, Research, and Foundations

by

________________________________________
Dick Carpenter, Chair

________________________________________
Al Ramirez

________________________________________
Sylvia Martinez

________________________________________
Tom Christensen

________________________________________
Joe Taylor
Khaliqi, David Habib (Ph.D., Education Leadership, Research, and Foundations)

What does the nation’s report card say about early physics models? Using the National Assessment of Educational Progress High School Transcript Study 2009 (NAEP HSTS 2009) to examine inverted curriculum

Dissertation directed by Professor Dick Carpenter

This study, using data from the National Assessment on Educational Progress (NAEP) High School Transcript Study 2009, examined the effect of inverted curriculum (IC) on 12th grade NAEP math assessment performance, algebra performance, and advanced math course taking among US high school students. Inverted curricular models introduce physics early in the high school sequence at the 9th grade rather than later at the 11th or 12th. Results of two-level hierarchical linear modeling (HLM) analysis revealed that students taught with a full-year, algebra-focused IC model significantly outperformed non-IC, conceptual-based IC, and half-year IC students on 12th grade NAEP math assessment performance, and algebra GPA. Further analyses revealed that 9th grade students taught with a conceptual-based IC model significantly underperformed non-IC students on NAEP math assessment performance. Implications of the findings are discussed related to educational practice and policy.
# TABLE OF CONTENTS

**CHAPTER**

I. INTRODUCTION ........................................................................................................................................ 1  
- Between groups research questions.......................................................... 5  
- Within groups research questions............................................................... 6  
- Operational Definitions.................................................................................. 6  
- Need for the study.......................................................................................... 7  

II. REVIEW OF THE LITERATURE ............................................................................................................... 10  
- Student difficulties in algebra......................................................................... 12  
- History of the IC movement............................................................................ 18  
- Experimental hypotheses.................................................................................. 22  
- Theoretical framework.................................................................................... 41  

III. METHODS ............................................................................................................................................. 48  

IV. RESULTS .................................................................................................................................................. 82  
- Between groups analysis of results................................................................. 84  
- Within groups analysis of results................................................................... 92  

V. DISCUSSION .......................................................................................................................................... 100  
- Between groups discussion of results............................................................. 103  
- Within groups discussion of results............................................................... 108  
- Limitations....................................................................................................... 110  
- Policy recommendations.................................................................................. 111  
- Conclusions...................................................................................................... 112  

REFERENCES.............................................................................................................................................. 116
TABLES

Table

1. School Population Demographics ............................................................ 55
2. Student Sample Size and Percent Distribution ........................................ 56
3. Student Demographic and Performance Data Disaggregated by 9th grade Course Taken ................................................................. 61
4. Number and Percentage of Missing Values Per Variable ....................... 80
5. 12th Grade NAEP Math Achievement Between IC and Non-IC Students .......... 85
6. Cumulative High School Algebra GPA Between IC and Non-IC Students .......... 88
7. Cumulative Advanced Math Carnegie Units Earned Between IC and Non-IC Students ........................................................... 91
8. 12th Grade NAEP Math Achievement Between Half-Year and Full-Year 9th Grade Physics Students and Traditional IC Physics and Conceptual IC Physics ................................................. 93
9. High School Algebra GPA Between Half-Year and Full-Year Physics Students and Between Traditional IC Physics and Conceptual IC Physics .................................................... 95
10. Number of Advanced Math Carnegie Units Taken Between Half-Year and Full Year Physics Students ..................................................... 97
FIGURES

Figure

1. Conceptual problem model framework.............................................42
CHAPTER 1

INTRODUCTION

The most recent version of *Rising Above the Gathering Storm* (Augustine et al., 2007) identified persistently low math achievement among US students as the largest factor reducing the ability of the country to compete in an increasingly global economy. All technological innovation requires strong mathematics skills, yet among US high school seniors, only 26% performed at or above the proficient level on the 2013 National Assessment of Educational Progress (NAEP). A proficient score means that students have demonstrated competency over mathematical concepts. Alarmingly, 64% performed at or above the basic level, which denotes only partial mastery of the knowledge and skill required for proficient work at a given grade level. Fewer than 10% achieved advanced, with the remainder below basic. The 2013 scores are identical to those from 2009 indicating that US student math performance is stagnant at the 12th grade. Although 2009 and 2013 show improvement over the 2005 12th grade results in which 23% of US 12th grade students scored at or above proficient, they still point to the great need for improved math performance by US students (Grigg, Donahue, & Dion, 2010).

In addition to producing low performing students, the US educational system is failing to engage students to persist in mathematics study beyond algebra 2 with 35% taking pre-calculus and 16% taking calculus in 2009 (US Department of Education, 2013). Beyond course-taking patterns, US students do not retain what mathematical skills they acquire in high school. This is demonstrated by the fact that 58% of first-year college students require math remediation despite being fully eligible to attend college (Shulock, 2010). Among students not planning on attending a four-year college, algebra has become the greatest barrier to high school graduation. A study from *Education Week*
and the Editorial Projects in Education Research Center reported that 1.3 million high school students dropped out in 2010—roughly 7,200 per school day—with algebra difficulties being cited as the primary reason (Cavanagh, 2011). The Massachusetts Department of Elementary and Secondary Education identified several significantly negative societal impacts due to dropping out, including (a) negative impacts on students’ self-esteem and psychological well-being; (b) loss of potential income by up to 35%; and (c) a 250% increase in the likelihood of incarceration (Darling-Hammond, 2007). Beyond the 4-year college-going student population, many technical careers now require solid algebra skills for entry into the technical workforce. Now that algebra forms the core of the high school mathematics curriculum, improving student algebra performance is critical if the US is to continue to lead the world in technological innovation.

Given that algebra is considered a “gate-keeper” class for post-secondary success in 2-year and 4-year college (Helfand, 2006), the area of purported greatest need in high school math educational reform is in how algebra is taught to US high school students. Schmidt, Wang, and McKnight (2005) observed that curricular reforms can have significant effects on student mathematics performance and postulated that often poor student attitudes toward math can be ameliorated by reforming how math is taught. Trumper (2006) showed that curricular reforms can be especially effective in improving algebra performance.

This dissertation examines one lesser-known but possibly transformative curricular reform idea known as inverted curriculum (IC). This reform idea is designed to improve student algebra performance, enhance concept retention, and increase the number of advanced math courses taken by introducing physics early in the high school sequence at the 9th or 10th grades rather than later at the 11th or 12th, hence the term
inverted. IC models vary on the methodological continuum from discovery learning (i.e. learners construct their own understanding) to direct instruction (i.e. basic principles are passed from teacher to student), but all espouse the reform notion that physics instruction is better placed early in high school rather than later. In general, IC designs address the issues of poor US student math performance and persistence in advanced math by moving physics from late in the high school curriculum (non-IC model) to the first science class in the high school curriculum (IC model). According to IC proponents this change in science sequence will improve algebra understanding and concept retention (Kieran, 1992); algebra performance (Pasero, 2008); and applied problem solving skills by providing a context (i.e. force and motion) for algebraic concepts (i.e. equations and variables) to be used in real-world situations (Bell, Blair, Crawford, & Lederman, 2003).

Currently, IC models are adopted on an *ad-hoc* basis school-by-school or even student-by-student. Students within these schools may take physics at the same time as 9th grade algebra, thereby conforming to an IC model. This informal arrangement creates a natural experiment in which student course-taking patterns imitate an IC model and can therefore be used as an experimental group in the analysis of IC model efficacy on improving student math performance and increasing the number of advanced math courses taken by students compared to non-IC students. Since it was the student and not the school in some cases who decided to use an IC model, special care will be needed to minimize selection bias. In this study, unobserved variable bias was present but minimized by examining student course taking patterns while controlling for as many student, family, and school-based covariates as possible that can sometimes affect student performance. This study used the 2009 NAEP and High School Transcript Study (HSTS) to examine student course taking patterns throughout high school. The 2009 NAEP HSTS
is an excellent data source with which to examine the efficacy of IC models on improving student math performance and increasing the number of advanced math courses taken by students because it provides non-self-reported student course-taking data and links all students to their NAEP 12th grade math assessment score.

This study was done in two phases. The first phase involved a between-groups analysis by comparing performance data and course-taking patterns between IC and non-IC students. The second phase performed a within-group validation study by comparing the disaggregated effects of physics course type and amount of instruction on student math performance among IC students only. The goal of this study was to understand the efficacy of IC models in improving algebra performance and student interest in taking advanced math courses using national student data that conform to IC course-taking patterns.

This study was needed because prior research on the efficacy of IC models to improve algebra performance and student interest in mathematics focused on single school implementation, thereby lacking the robust external validity needed to understand the strengths and weaknesses of such programs on a national scale. When studies evaluated a single site, important questions remained unanswered regarding the broader efficacy of IC models in improving algebra performance and math interest. These unanswered questions include whether the observations at the single implementation site hold true nationwide, thereby making IC an effective and efficient reform method to improve algebra performance and math interest. If IC is to be a robust reform methodology, then it must be subjected to the scrutiny of research outside of the small, self-selecting populations normally studied.
This study differs from previous studies (see Chapter 2) because it does not evaluate a specific program at a specific site. Instead, it looks at high school course-taking patterns nationwide that conform to general IC models to determine whether these specific course-taking patterns result in improved math performance. The hypothesis of this study was that if IC is an effective curriculum reform model, then improvements in student math performance should be detectable in student data that conform to IC models whether the student’s school adopted a formal IC model or not. Put in formal terms, this study examined two experimental hypotheses. The first experimental hypothesis stated there is a statistically significant difference in student performance and course-taking patterns between students taking IC and those in non-IC curriculum, with IC students demonstrating greater performance and taking more advanced math courses than non-IC students. The second experimental hypothesis stated there is a statistically significant difference in student performance and course-taking patterns among IC students based on the type of early physics (i.e. algebra-based or concept-based physics) and the amount of early physics taken (i.e. half-year or full-year), with students taking more physics earlier in high school showing greater performance and taking more advanced math courses than those taking less physics. These hypotheses were tested using six research questions divided into two sections. The first three questions covered the between-groups analysis, while the second three questions covered the within-group analysis. The six research questions included:

**Between-groups Research Questions: Comparing IC to Non-IC Students**

1) Is there a significant difference in performance on the 12th grade NAEP mathematics assessment between students who receive early physics instruction (IC) in 9th grade and those who do not (non-IC)?
2) Is there a significant difference in cumulative high school algebra GPA between IC and non-IC students as documented on the student’s 12th grade transcript by the HSTS?

3) Is there a significant difference in the number of advanced math Carnegie credits earned by IC and non-IC students as documented on the student’s 12th grade transcript by the HSTS?

**Within-groups Research Questions: Comparing IC Students to Each Other**

4) Among IC students, is there a significant relationship between the length and type of physics class taken on 12th grade NAEP math assessment performance?

5) Among IC students, is there a significant relationship between the length and type of physics class taken on the overall algebra GPA earned?

6) Among IC students, is there a significant relationship between the length and type of physics taken on the number of high school math Carnegie credits earned?

**Operational Definitions Used in this Study:**

- Inverted curriculum (IC) models: those curricular models that move physics from the junior or senior year to freshman year (Li, Oranje, & Jiang, 2009);

- Non-IC curriculum: the standard curricular model that places physics in the junior or senior year of a student’s high school career (Haber-Schaim, 1984);

- Algebra: the mathematical concepts and skills that allow students to evaluate and solve functions; solve equations using abstract symbols; use equations to model mathematical relationships; and analyze change over time (National Council of Teachers of Mathematics, 2008).
• Discovery learning: a type of instructional methodology that places students in problem solving situations where the learner is given sparse background information and thereby draws on his own experience and prior knowledge to construct his own knowledge and arrive at a solution with little to no guidance (Alfieri, Brooks, Aldrich, & Tenenbaum, 2011).

• Direct instruction: a type of instructional methodology in which the teacher delivers the requisite knowledge base to the student who is then expected to apply that knowledge base in problem solving situations (Paas, Renkl, & Sweller, 2004); and

• Schema: an organized pattern of thought or behavior (John Sweller, 1988).

The aforementioned research questions were addressed using data from the 2009 NAEP HSTS dataset. These data were collected to allow researchers to examine the course-taking patterns of students who graduated from public and private high schools in the United States in 2009. Additionally, it was designed so that data on students’ course-taking patterns can be linked to the 2009 NAEP assessment results. Fortuitously, the NAEP 2009 assessments focused on mathematics, which makes the use of these data particularly appealing to address the questions specific to this study. More information regarding the structure of the 2009 NAEP HSTS data is covered in Chapter Three.

The Need for Improved Algebra Outcomes

Algebra has become the core of secondary mathematics education in the United States (Chambers, 1994) given its important role in future college and career success (Vogel, 2008). The National Mathematics Advisory Panel found that completion of algebra 2 doubles the probability of college graduation. Unfortunately, 39% pass rates and the sharp decline in mathematics achievement when students begin studying algebra
highlight the need to improve algebra instruction and student performance outcomes (National Mathematics Advisory Panel, 2008).

Given that all states have made passing algebra a high school graduation requirement, poor algebra performance has become a significant barrier to high school graduation among US high school students. A January 2006 *Los Angeles Times* article profiled one low-income high school student who, over the course of six semesters, failed algebra six times. Midway through her senior year, she returned all her textbooks to the campus book room and left school permanently. This anecdote captures the widespread effect that algebra difficulty can have on every aspect of a student’s studies. Helfand (2006) observed that repeated failures have significant impacts on student confidence across all areas of study so that even when a student performs well in other subjects, poor algebra performance alone can decrease the likelihood of high school graduation.

The aforementioned anecdotes highlight how difficult the transition can be from arithmetic to algebra for young high school students. Given the high priority that algebra is given in US curricula, it is important to explore why these difficulties exist. The National Council of Teachers of Mathematics (NCTM) most recently defined algebra as a “way of thinking and a set of concepts and skills that enable students to generalize, model, and analyze mathematical situations” (National Council of Teachers of Mathematics, 2008, p. 1). The power of algebra is that it provides a systematic way to investigate relationships, helping to describe, organize, and understand the world using non-arithmetic symbols. Difficulties arise for students because algebra is the first time they are introduced to the use of abstract symbols, expressions, and equations. To ameliorate these difficulties, the NCTM organized algebra by four overarching concepts and skills: functions, algebraic symbols, mathematical modeling, and analyzing change.
Abstract variable meaning in algebra informed all four of the NCTM high school algebra standards. The first standard focused on developing students’ understanding of variables as functionally related quantities. The second standard promoted an emphasis on understanding the symbols used to represent algebraic variables. The third standard emphasized the learning of mathematical modeling to represent quantitative (both functional and nonfunctional) relationships. The fourth standard supported an emphasis on teaching rates of change in algebra. This cursory glance at the current standards appears to support the notion that IC could be an effective means to help students understand algebra better by providing observable and applied contexts for the abstractions in the form of physics-based events. According to IC proponents, physics education, more than any other high school science course, addresses the difficulties that students have with the abstraction of algebra by providing applied, observable, and measurable mathematical problem solving opportunities (Haber-Schaim, 1984).

In summary, the need to improve student algebra performance is one of the most pressing issues facing educators in the US. The link between student math performance and the nation’s future economic health is well-documented (Executive Office of the President, 2010; Kilpatrick, Mathis, & Gunn, 2011; Phillips, 2009). This study, therefore, examines the efficacy of IC as a viable reform model to improve student algebra performance and increase the number of advanced math courses taken in high school. The goal of this study is to contribute to the current understanding of IC models and their possible efficacy to improve algebra performance in order to take full advantage of its strengths and understand its limitations on a national scale.
CHAPTER 2
LITERATURE REVIEW

The purpose of this literature review is to examine the existing educational research related to the efficacy of inverted curricula (IC) models on improving algebra performance. IC is a science curriculum reform idea that inverts the traditional science curricular sequence by making physics the first science course taken in high school rather than the last course taken. IC proponents state that the inversion will improve student algebra performance and engagement by providing an applied physics context to algebra problems.

The educational research literature on IC models is primarily composed of evaluations of IC implementation at specific sites. Few studies have applied quantitative methods to the examination of IC model efficacy on improving student math performance and increased advanced math course taking. This literature review examines the foundational hypotheses of IC model theorists in order to understand how IC models might improve student algebra performance and persistence.

This literature review contains four sections. First, a review of the current literature on student difficulties and misconceptions in algebra leads into a review of the current literature on effective algebra education. Second, an historical overview of the IC movement illuminates how IC models might address student difficulties in algebra. Third, the hypotheses made by IC proponents are examined in light of prior research. These hypotheses are that (a) the defining outcome of quality math education is the creation of robust problem solvers; (b) presupposing the truth of the aforementioned hypothesis, applied, problem solving efficiently and effectively develops the requisite algebra skills needed for student success in future math study; and (c) inherent in physics
Instruction reside effective algebra instructional methods that improve student algebra performance and increase the number of advanced math courses taken (Lederman, 2005). Finally, the fourth section will use Nathan, Kintsch, and Young’s (1992) conceptual problem model as a theoretical framework for understanding how IC methods may support improved algebra performance and mathematical problem solving. The conceptual problem model is used as a theoretical framework for this dissertation because it presents a theory of mathematical comprehension based on research describing the mental representations produced during reading comprehension, a field of research called discourse processing (Van Dijk & Kintsch, 1983). Discourse processing articulates the specific steps that learners use when confronted with any type of verbal or written problem. Nathan et al. (1993) applied discourse processing to mathematical problem solving to form the conceptual problem model. As mentioned in Chapter One, the transition from arithmetic to the abstraction of algebra is difficult for beginning algebra students. They hypothesized that by applying the processes of discourse processing to mathematical problems, students can become meta-cognitively aware of the problem-solving processes they use. Additionally, students become aware of the weak areas in their own problem-solving abilities and identify those areas in need of remediation. Related to this study, the high value that IC models place on understanding applied math problems makes conceptual problem modeling an ideal framework to understand the cognitive processes that students use when confronted with a physics-based algebra problem. Additionally, that IC models may use discovery learning pedagogies in the form of ill-defined physics problems, the conceptual problem model effectively builds a pathway between the application of algebra concepts to physics-based word problems as espoused by IC proponents. Given the strong link between word-problem comprehension
and IC models, the explanatory power of the conceptual problem model is significant. Using the conceptual problem model as the reference frame, IC methods will be examined to evaluate whether these methods incorporate the key characteristics of research based problem solving methodologies identified in the model. Note that this conceptual model will only be used in the literature review to understand the theoretical relationship between IC and research based problem solving methodologies. The conceptual model will not be used to guide the empirical analysis in this study.

**Student Difficulties in Algebra**

Algebra is often the first math course in which students are asked to engage in abstract reasoning and problem solving (Vogel, 2008). Researchers have demonstrated that the abstract nature of algebra increases its difficulty over arithmetic (Carraher & Schliemann, 2007; Kieran, 1992). Students who have experienced mathematics at a concrete or procedural level typical in many classrooms must negotiate the difficult gap from concrete to abstract reasoning with little preparation (Freudenthal, 1983; Pimm, 2002). This inexperience with abstraction directly affects the ability of students to manage multiple representations of algebraic objects (Kieran, 1992; Vogel, 2008).

Related to the difficulty of students engaging with the abstract nature of algebra is the requirement of students that they learn a language of mathematical symbols that is also completely foreign to their previous experiences (Kilpatrick, Swafford, & Findell, 2001). The ambiguity with which the language of algebra is described and used during instruction often prevent students from connecting algebraic symbols to their intended meaning (Blanco & Garrote, 2007; Tolar et al., 2012). In some cases, students are completely unaware that any meaning was intended for the symbols (Knuth & Cai, 2011). In other cases, they may know that meaning exists, but limited understanding
prevents them from ascribing meaning to the symbols, or they may assign erroneous meaning to the symbols (McNeil et al., 2010). For example, McDermott, Rosenquist and Van Zee (1987) found that students were generally able to plot points and equations; however, in spite of this procedural fluency, students still lacked the ability to extract meaning from graphical representations. The researchers concluded that the difficulty lay in the connection of a graph to the larger idea being represented by the graph. Specifically, students were readily capable of demonstrating procedural fluency, but memory and procedural understanding were unable to guide students through problems involving interpretation (Skemp, 2006).

Building upon Skemp’s research, Schliemann, Carraher, and Caddle (2013) recognized that learning the structural characteristics of algebra created persistent challenges for students. For example, students often failed to recognize the differences between expressions and equations. They also had difficulty conceptualizing an equation as a single object rather than a collection of objects. The meaning of equality was often confused within algebra contexts as well. Taken together, these three example structural challenges prevented students from recognizing the utility of algebra for generalizing numerical relationships.

The literature reveals that the leap from concrete arithmetic to algebraic abstraction is a large one fraught with difficulties that require particular instructional strategies to support students through this transition. The next several sections examine whether IC models conform to the best practices in bridging these significant cognitive gaps, thereby improving student algebra performance.

**Effective algebra instruction.** Before evaluating the efficacy of IC models, we must come to an understanding of what constitutes effective algebra instruction in the
first place. The purpose of this section is to examine findings from the current literature that support three defining characteristics of effective algebra instruction. First, effective algebra instruction improves student conceptual knowledge of algebraic relationships. Second, effective instruction requires the student to apply this conceptual knowledge to become fluent in interpreting and solving abstract variables. Finally, effective algebra instruction requires the student to apply this fluency of algebraic skills to analogous concepts and situations that are not clearly defined in the initial problem.

Improvement in students’ conceptual knowledge of algebraic relationships has been shown to be a significant characteristic of effective algebra education. For example, a meta-analysis of 122 algebra instruction articles by Rakes, Valentine, McGatha, and Ronau (2010) revealed that studies focusing on conceptual understanding produced an observed weighted average effect size more than twice that produced by interventions focusing on procedural understanding. This finding indicates that strategies focused on deepening conceptual understanding improve student achievement without relying on traditional drill and practice routines. This finding was earlier observed by Hiebert, Morris, Berk, and Jansen (2007) and Baumert et al. (2010), all of whom suggested that focus on the development of conceptual understanding improved student achievement far better than a strategy with a focus on procedural understanding. Conceptual understanding, according to the proponents of IC, is better taught through physics problem solving, which is an excellent way to increase deeper conceptual understanding of algebraic concepts and move beyond skill development alone (Lederman, 2005). If students are to retain the content of their algebra courses, then they must first build a foundation of strong conceptual knowledge upon which to add algebraic skills in the future.
In addition to conceptual understanding, many researchers have pointed out the crucial need for students to learn, with absolute fluency, how to interpret, manipulate, and construct algebraic equations and expressions. For example, Knuth, Alibali, McNeil, Weinberg, and Stephens (2011) found that students interpreted variables through six progressive (i.e., hierarchical) levels: (1) as a single value, through trial and error evaluation; (2) as irrelevant (i.e., students ignore the variable in a contextual situation); (3) as an object or label; (4) as a specific unknown; (5) as a generalized number; and (6) as a functional relationship. While the first three levels represent concrete variable interpretations, the next three levels comprise the formal, abstract ways of interpreting variables. Understanding the ways that students interpret variables is an important component to the successful transition into physics-based problem solving.

The importance of operations within algebraic equations and expressions was highlighted by Lins and Kaput (2004) when they argued that variables denoting extensive quantities such as money or lengths and those denoting intensive quantities such as velocity or acceleration prompted students to use different interpretations of the operations that were permitted on them even when those operations should have been similar. Smith and Unger (1997) and later Kloos, Fisher, and Van Orden (2010) discussed intensive versus extensive quantities indicating that the development of proportional reasoning involving intensive quantities was a gradual and protracted process extending well into adolescence and adulthood. One reason Smith and Unger stated that intensive and extensive quantity pairs in physics were so difficult to distinguish was that they involved continuous quantities that could not be directly visually apprehended (such as density and mass, temperature and heat). The more continuous and/or intensive the quantity, the more difficult it was for students. Not surprisingly, they also found that
asking students to make connections between two domains not equivalently understood typically caused students to utilize concepts from the better understood domain at the expense of the less well-understood domain.

Related to physics problems, Dooren, Verschaffel, and Onghena (2013) observed that physics based problem solving activities were effective in improving student understanding of continuous quantities related to algebraic problem solving, particularly for the conceptual understanding development of younger secondary students. This finding corroborated the IC hypothesis that physics problem solving is effective at helping students understand the abstract nature of operations within an algebraic equation.

An effective method of improving student understanding of the use of continuous and extensive quantities in the domains of algebraic manipulations was postulated by Sweller, Ayres, and Kalyuga (2011) in the form of worked examples. These researchers observed that problem-solving success was not due to innate abilities but to a highly automatized and interconnected knowledge base or schemata that could be developed over time. These schemata are cognitive structures that help the problem solver categorize a situation according to relevant features and then indicate the appropriate solutions for that problem type.

Ultimately, as students learn to apply their newly formed schemata to the construction of algebraic equations and expressions, they need to be able to transfer these mathematical skills to analogous conceptual situations. Novick (1988) and later Gentner and Colhoun (2010) studied the use of analogous stories to determine what factors affected transfer between different types of problems. Both studies found that the surface features of stories interfered with students’ ability to recognize analogous solutions and
led to false solutions between situations that appeared analogous but were not. Reed, Stebick, Comey, and Carroll (2012) further studied the use of analogous problems for solving mathematical word problems. They demonstrated that students who effectively formed their own schema about the solution of a problem (i.e. developed conceptual understanding) achieved the best results. Importantly, their results showed that students used solutions of applied (i.e. physics-based) algebra word problems to solve equivalent problems when the solutions were accompanied by an explanation of why a particular equation was used. Students used the schema they developed to understand the reasoning behind the solution of a problem. Memorizing a procedure did not teach them transferable problem solving strategies. The finding that conceptual understanding resulted in improved transfer of algebraic skills supported the IC idea that effective instruction in applied problem-solving (i.e. physics problems) is critical to developing flexible and robust algebraic skills.

In summary, this first of four sections had two goals: (1) to identify the areas of difficulty for students in learning algebra and (2) to examine the components of effective algebra education in improving algebra performance and increasing the number of advanced math classes taken. Related to the first goal, researchers identified the use of non-arithmetic abstract symbols and the manipulation or solving of expressions and equations to be the primary areas of difficulty for beginning algebra students. Related to the second goal, three components of effective algebra instruction were described from the literature. The three components most clearly identified by researchers were (1) developing a strong conceptual understanding of algebra rather than procedural only, (2) applying this conceptual understanding to the manipulation and solving of algebraic expressions and equations, and (3) applying conceptual knowledge and skills to ill-
defined and analogous applied problems. With these two goals accomplished, we now turn to the history of IC models to understand how the model was derived and to what degree the model is grounded in prior research.

**Historical Overview of the IC Movement**

The history of the IC model dates back more than 100 years to 1892 when the National Educational Association (NEA) organized a committee to determine the high school science sequence in public schools. The goal of the committee was to systematize the course sequence so that students would experience a more uniform preparation for college (NEA, 1893). At this time the high school subject offerings were a chaotic mess with different schools and colleges having widely varying requirements and curricula (Sheppard & Robbins, 2007). Early debate on the sequence focused on the relative placement of physics and chemistry. Biology was not a single subject but was taught as separate courses in botany, zoology, and physiology. These courses were offered in varying years in high school.

At the same time, in 1893, the Committee of Ten (CoT) published its report on secondary school studies, in which they recommended an IC model by placing physics before chemistry so that physics “may precede meteorology and physiography.” In 1895, an effort to implement the CoT findings was undertaken by the formulation of the Committee on College Entrance Requirements (CCER). The CCER followed the CoT recommendations and proposed that physics precede chemistry; however, biology was comprised of several individual classes and was therefore left to fill in the gaps around physics and chemistry (Bailey, 2011). More importantly the CCER made a critical recommendation that would influence science subject-taking patterns long into the future. In 1899, a report was issued that led to the establishment of the “credit system” for
subjects. Importantly, the report recommended that only one science credit be required for college admission. This decision would effectively make both physics and chemistry elective subjects, because as biology began to emerge as a single subject it was almost always placed before chemistry and physics in the sequence (Sheppard & Robbins, 2003).

By 1920, this haphazard drift was codified into policy when the Committee on Reorganization of Science in Secondary Schools formally included general biology in its recommendations for the high school science sequence. The committee placed biology in the second year of high school, with physics and chemistry offered in the third and fourth years but without a specified order. The list of course-sequencing recommendations made by the various committees from 1890 to 1920 showed clear preference for placing physics before chemistry. By 1930, the Dexter survey showed the distinct trend that physics was becoming the last science in the sequence as chemistry was more often being offered before physics (Dexter, 1926). As this trend continued, biology came to be taught predominantly in the 10th grade, with chemistry and physics in 11th and 12th grades, respectively. By the end of World War II, the biology, chemistry, physics sequence had become the status quo in US high schools and would remain so. Unfortunately, the transition from the original IC model to the current model was not a strategic policy decision, but yet another example of a haphazard drift into orthodoxy so common in educational policy.

The IC conversation remained dormant until 1970, at which time the Physics Teacher ran an article recounting a handful of schools, often led by their physics teachers, inverting their science sequences in the early 1970s (Palombi, 1979). The movement received a boost in 1984 from esteemed physicist and educator Uri Haber-Schaim, who
wrote a *Physics Teacher* article titled, “High School Physics Should be Taught Before Chemistry and Biology” (Haber-Schaim, 1984). In it, he laid out three tables of topics covered by high school science textbooks of the time. Two lengthy tables were titled, “Chemistry Prerequisites in Biology Texts” and “Physics Prerequisites in Chemistry Texts.” The third, much shorter, table was titled, “Chemistry Prerequisites in Physics Texts.” Interestingly, he did not include a table listing biology prerequisites for either physics or chemistry. Haber-Schaim used this prerequisite argument to say that physics should be taught first. Scientific understanding had expanded and connected the physical and life sciences in ways the CoT could not have foreseen. In particular, the increasingly apparent role of physics and chemistry as the context of biology was identified in the life-sciences reform literature as well (Labov, Reid, & Yamamoto, 2010).

The most recent champion of the movement has been Leon Lederman. Lederman’s work in this area began when he convened Project ARISE (American Renaissance in Science Education) in 1995. Since that time, he has written and spoken in a great number of forums advocating for the inversion. Along with Lederman, Haber-Schaim in his landmark 1984 article laid out what is probably still the most commonly given explanation for inversion to an early physics curriculum based on the changes that took place in the sciences themselves over the course of the twentieth century (Lederman, 2005). According to Haber-Schaim, the non-IC sequence of biology–chemistry–physics was introduced in the early years of the 20th century at a time when biology was largely descriptive botany and zoology. Chemistry was also descriptive and largely qualitative, with the exception of the laws of constant and multiple proportions. Physics, which was considered more demanding mathematically, was placed at the end of the sequence (Haber-Schaim, 1984). In those days biology required no chemistry, and chemistry
required no physics. Today, tenth-grade biology has substantial prerequisites in chemistry, and chemistry has substantial prerequisites in physics.

Myers (1987) echoed Haber-Schaim’s argument in further detail describing the advances in technology that allowed for a chemistry-based understanding of biological processes and a physics-based understanding of chemical processes. Relevant to this study, Myers was the first to introduce the idea of mathematical reinforcement, which he described as a use of the physics-first course to provide students with applications for their algebra skills. Under the non-IC sequence, Myers argued, students who take algebra in 9th grade will see it as little more than an abstraction until their 12th grade physics class. Moving physics to the front of the science sequence (he suggested 10th grade) allows students rapid reinforcement of their algebra through regular use. Today, most schools using IC offer it concurrently with algebra, an application of Myers’s idea.

One of the goals of Lederman’s Project ARISE was an effort to develop and disseminate a framework for a three-year core curriculum for high school science. A white paper on the framework was released in 1998, and one of its key elements was an inversion of the traditional order in which biology, chemistry, and physics are taught in high school (Lederman, 1998). This report was the first effort to aggregate qualitatively the experiences of schools that had performed the inversion. The rationale for the inversion was that the three sciences, especially biology, had undergone significant changes over the past hundred years. Biology and chemistry were no longer the purely descriptive sciences they once were. Comprehending chemistry in terms of the structure and behavior of atoms relied on an understanding of physical principles, and modern biology requires understanding the chemical functions of molecules such as DNA and proteins.
Proponents of IC asserted the primary advantage of the model was improvement of algebra skills because of the widespread and concrete use of algebra concepts in physics problems (Goodman, 2006; Myers, 1987). Embedded in this claim were three important hypotheses mentioned at the start of this review. To review, these were (a) the defining outcome of quality math education is the creation of robust problem solvers; (b) given the truth of hypothesis one, applied problem solving efficiently and effectively develops the requisite algebra skills needed for student success in future math study; and (c) inherent in physics instruction reside effective methods for algebra instruction to improve student performance and increase the number of advanced math courses taken (Lederman, 2005). The next section of this literature review examines the support for those hypotheses in the research literature.

**Hypothesis One: Problem Solving as the Defining Outcome of Math Education**

The first hypothesis put forth by IC proponents is that the chief end of mathematics education is to create problem solvers able to use algebraic principles in real-world situations. Proponents of IC models repeatedly state the primary advantage of early physics education is to create more robust problem solving skills in ninth or tenth grade students (Haber-Schaim, 1984; Myers, 1987). Lederman (2005), the primary proponent of the physics first style of IC, stated that the key advantage of having the first high school course be physics was its relationship to mathematics in that early physics provided mathematical definitions that were precise and concrete (i.e. position, velocity and acceleration). Ewald, Hickman, Hickman, and Myers (2005) noted that one of the four advantages of an inverted sequence was that students applied their mathematical skills to solve real-world problems. Trumper (2006) identified this assumption in an
international context in his discussion of Israeli schools adopting inverted curricular changes to improve student mathematics learning and performance.

Outside the IC advocacy community, other influential organizations in math education have espoused the view that the most important outcome of math education is the development of problem-solving abilities in students. The Common Core Mathematics Standards define a mathematically proficient student as one who can apply the mathematics s/he knows in order to solve problems arising in everyday life, society, and the workplace (Common Core State Initiative, 2010). The National Council of Teachers of Mathematics (2008) has identified the ability to model the mathematics inherent in physics as a key skill supporting all math learning for ninth grade students. Internationally, the TIMSS 2011 math frameworks stated that “learning mathematics for its own sake is probably not the most compelling reason for universal inclusion of mathematics in school curricula” (p. 19). The frameworks identify the prime reasons for having mathematics as a fundamental part of schooling to be the creation of citizens able to successfully use math in the workplace (Mullis, Martin, Ruddock, Sullivan, & Preuschoff, 2011). Notably, the NAEP frameworks do not specifically mention applied problem solving ability as a primary outcome goal for students. NAEP is the only major educational framework that does not specifically mention real-world problem solving as a primary outcome.

**The progression of thought on mathematical problem solving.** The question regarding whether problem solving is the ultimate mission of math education likely dates back to the dawn of mathematics itself. Early 4th century Babylonian manuscripts document problem solving as the primary goal of Babylonian math education. Interestingly, at these early stages of mathematical study, the primary focus was on
applied uses of mathematics on numbers and computations for use in civil engineering, astronomy, and economics (Berriman, 1956).

Later, the Greeks took up this question famously in the dialogues between Plato and his student Aristotle. Plato took the position that the objects of mathematics had an existence of their own, beyond the mind, in the external world and was indignant at “technicians” use of physical arguments to prove mathematical results in applied settings. In contrast, Aristotle drew clear lines between the ideal forms (i.e. non-empirical) envisioned by Plato and their empirical realizations in worldly objects. Aristotle’s view of mathematics was not based on a theory of an external body of knowledge independent of observable phenomenon. Rather, it was based on experienced reality, where knowledge is obtained from experimentation, observation, and abstraction (Lerman, 1983). Clearly IC models rest on an Aristotelian view of mathematics and not a Platonic.

The use of problem solving to characterize the nature of mathematics is not a settled issue among mathematicians and educators. Despite the recommendations of organizations such as the NCTM, math educators tend to carry strong Platonic views about the existence of mathematical concepts outside the human mind and sometimes resist incorporating applied problems (Brown & Mc Namara, 2011). It is important for IC proponents to realize that the Platonic math teacher may be reluctant to collaborate with a physics teacher to bring real world physics problems into the math classroom. This reluctance is a significant barrier to the implementation of IC models among faculty that hold to a Platonic view of mathematical knowledge.

While this debate has continued through the centuries, for the purposes of this paper, we take up the discussion as it emerged in the early 20th century. The 1920s witnessed a movement to design constructivist mathematics curricula around real-life
situations solved in a collaborative setting, a methodology called discovery learning. Students were to construct their understanding of mathematics in the context of applied problems through discussion and collaboration with peers. The ultimate 20th century theorist espousing this constructivist, problem-centric model of discovery learning was John Dewey. Dewey (1910) attempted to move beyond the “applied math vs. pure math” dichotomy by positing his notion of reflective inquiry, the predecessor to discovery learning. Dewey defined reflective inquiry as the process by which problems are identified, studied collaboratively with peers, and solved using shared conclusions. Reflective inquiry laid the groundwork for other constructivist-based problem solving methodologies for decades to come.

Another important 20th century thinker who espoused mathematical problem solving as the primary means for deeper student understanding was Polya (1957) through his seminal work, *How To Solve It*. In this work, Polya established a modern heuristic to describe problem solving. He primarily conceived of mathematical learning as a hands-on and highly active. The goal of mathematical engagement was to assimilate the mathematics into “one’s flesh and blood” to be ready for instant use. He saw independent problem solving of challenging problems to be far more efficacious to student math engagement than “more aphorisms to follow.”

In the decades following Polya’s work, many researchers began deconstructing the underlying frameworks of problem solving to understand the various components of this complex construct more deeply. Halmos (1980) theorized that the meaning of mathematics only became clear after the postulates and axioms were applied to meaningful problems for the students, similar to Polya’s primary thesis. Empirical support for mathematical problem solving as a teaching methodology have been put forth
by Confrey (1990) and Hiebert et al. (1996), who found that there were significant qualitative differences in the understanding that students developed in problem focused instructional situations. More recently, Araz and Sungur (2007) utilized structural equation modeling to model the relationships among reasoning ability, learning approach, prior knowledge, motivational variables, and achievement in math classes that utilized problem solving activities. Results showed that reasoning ability, learning approach, task value, and prior knowledge had direct effects on subject achievement. Moreover, reasoning ability and task value were found to have indirect effects, which were mediated by learning approach.

**Research challenging hypothesis one.** In considering whether problem-solving is the defining outcome of math education the first research examining the possible shortcomings of problem-solving based pedagogies was accomplished by de Groot (1966) examining how learners move from novice to expert understandings of complex problem structures such as chess moves. This research was the first to demonstrate that expert knowledge resides in long-term memory while short-term memory is used in the process of obtaining basic levels of knowledge at the novice stages of development. Importantly, he was the first to demonstrate that the development of each memory type was mutually exclusive as cognitive load increased. This finding meant that if a novice algebra student was faced with an ill-defined, discovery learning type of task before conceptual understanding was established, then the discovery learning task worked against the student’s ability to learn and apply the needed procedures to solve the problem, creating undue frustration for the student.

Sweller (1988) expanded the notion of cognitive load when he identified the role that cognitive load and schema formation play in problem solving and long-term
knowledge acquisition. He identified schema acquisition as the process by which a problem solver learned to recognize a problem state as belonging to a particular category of problem states that require particular moves. He demonstrated that the mechanisms required for problem solving and schema acquisition are substantially different and as a consequence, the cognitive effort required by conventional problem solving may not assist in schema acquisition. Since schema acquisition is possibly the most important component of problem solving expertise, the development of expertise may be retarded by a heavy emphasis on problem solving. Sweller concluded that problem solving could serve as an effective means of learning at later stages but may need to be used sparingly for novice algebra learners.

Kirschner and Clark (2006) built on Sweller’s work by studying the role that working and long-term memory play in problem solving tasks. They explained that discovery learning-based instruction requires the learner to search a problem space for problem-relevant information, making heavy demands on working memory. These demands can often work in opposition to the accumulation of knowledge in long-term memory. It is this accumulation of knowledge in long-term memory that the authors refer to as authentic learning. The reason working memory load did not contribute to the accumulation of knowledge in long-term memory was because while working memory was being used to search for problem solutions, it was not available and could not be used to learn. The authors pointed out that it was possible to search for problem solutions for extended periods of time with quite minimal alterations to long-term memory.

Wirth, Künsting, and Leutner (2009) replicated the findings of Kirschner and Clark related to problem solving expertise but found that differences in overall cognitive load were not sufficient for explaining differences in learning outcome. As a result,
problem solving strategies were found to be effective in fostering authentic learning that led to deeper levels of expertise.

Based on this review of the literature related to the role that problem solving has in math education, it appears that the first hypothesis of IC proponents is supported by the research literature. Even those researchers that questioned problem-solving as the primary outcome of math education simply question the degree and timing of its introduction into the math curriculum, not its overall value.

**Hypothesis Two: Algebra is Best Learned by Solving Applied Problems**

Given that research bears out that problem solving is an effective means of learning mathematics in general, let us now examine whether prior research supports the hypothesis that applied problem solving is an efficient and effective means of learning algebra. IC proponents claim that embedding physics concepts in the algebra curriculum and vice versa provides important context for students. The validity of the hypothesis that physics is an effective tool with which to accomplish this is examined later.

Many researchers and educators believe that the learning of algebraic ideas should always be anchored in real-world problem solving situations with which students are familiar (Jinfa Cai, Ng, & Moyer, 2011). The Common Core Math Standards reflect this view that algebra concepts are best retained when taught in the context of applied problems (Common Core State Initiative, 2010). The Common Core math curriculum engages students in mathematical problems embedded in authentic contexts in a manner similar to various international standards (Mullis et al., 2011).

Hiebert et al. (1996) argued that changes in curriculum should always allow students to problematize the subject. Rather than mastering skills and applying them, students should be encouraged to place their knowledge in the context of relevant
problems that support their own sense-making schemas. Based on observations of tenth grade students, the researchers observed that when students engaged in problem based algebra instruction those students retained the algebra concepts for a longer period of time and were able to discuss algebraic concepts on a more advanced level.

Beyond the notion of systematizing problem solving in curriculum, Nathan, Kintsch, and Young (1992) observed the skills and processes that support problem solving as a viable algebra instructional model. As will be noted later in this chapter regarding the theoretical framework, these researchers observed that effective algebraic problem solving skills included language comprehension processes; accurate mental representations; inferences and real-world knowledge; and the necessary formal calculations for deriving a solution. The authors argued that problem solving tasks are heavily reading-oriented making poor text comprehension the source of many errors rather than computational fluency. Interestingly, the authors hearken back to Sweller (1988) in observing that inability to access relevant long-term knowledge leads to serious errors. In particular, students with low algebra skills omitted steps and necessary mathematical relationships in the process of solving problems. These omissions were all based on the reading inferences needed to describe the problem situation, which point to the importance of schema formation and conceptual understanding by students.

Newell and Simon (1972) applied their research in algebraic problem-solving to the skills and processes students used in particular tasks such as chess, symbolic logic, and algebra-like puzzles. Their study dealt with student performance and de-emphasized the details of the processing of the problem. The researchers identified student problem solving as a highly complex process involving internal cognitive structures that significantly support visualizing an abstract problem in concrete terms. The development
of this visualization skill is one of the primary benefits of IC models, according to IC proponents. Building on Newell and Simon’s research, Boero (2002) identified a primary strength of applied algebraic problem solving to be the process of student transformation, which he labeled anticipation. By anticipation he was referring to the process by which students move from novice to expert by developing the ability to foresee some aspects of the final problem solving goal. As students become more mature in a subject they become adept at predicting and using the particular transformation needed to achieve the end goal in algebraic representation.

Quantitative studies on the role of problem solving in algebra education are relatively few, but several are noteworthy. Kapur and Dickson (2006) examined a methodology they termed productive failure in which students were confronted with an ill-structured problem without the provision of various support structures to solve the problem. Findings included more robust group discussions within the treatment group but a general inability to successfully solve the problems on their own. Interestingly, the productive failure class outperformed the traditional class by 10% on the post treatment content quiz containing material that was not explicitly covered during instruction. The observation that students struggled with an applied problem yet still improved their academic performance was a significant finding in support of the IC hypothesis that problem solving was an effective methodology for algebra learning. Based on these findings, within an IC model, students might struggle with a physics problem to the point of getting the solution wrong yet still learn and retain important algebra concepts.

Within the context of student groups, Wiedmann, Leach, Rummel, and Wiley (2012) examined the role of group learning on constructivist problem solving. The researchers found that mixed ability groups were needed to successfully use “learning by
invention” strategies in which students were given an algebraic concept and then expected to invent their own formula to solve the problem. Groups consisting of both high and low math ability members generated a broader range of solution attempts and demonstrated improved retention of the mathematical concepts from later lessons. These findings highlight the notion that strategic mixing of abilities within problem solving groups is an effective strategy that opens IC physics problem solving to students of diverse abilities, not just the advanced students.

The largest quantitative analysis of applied problem solving based curricula occurred in 2002 with the Longitudinal Investigation of the Effect of Curriculum on Algebra Learning (LieCal). This project was designed to compare longitudinally the effects of a problem-based curriculum to the effects of non-IC middle school curricula using a sample of 1,300 middle school students. Experimental group students showed significantly higher growth rates than the control group on situational representation, equation-solving, and generalizing tasks across the three middle school years. These findings suggested that the use of problem-based curricula was effective in improving key algebra skills. Importantly, students were able to retain these skills over three years (Lappan, Fey, Fitzgerald, Friel, & Phillips, 2006).

An important component of algebra problem solving is the ability to represent mathematical relationships and situations graphically. Heckler (2010) found that students who were prompted to represent a problem graphically were less likely to obtain a correct solution than those who were not prompted to solve the problem in any particular way. Analysis of the solution methods revealed that those students prompted to use a diagram tended to use the formally taught problem-solving method, and those students not prompted to draw a force diagram tended to use more intuitive methods. Students who
were prompted to draw diagrams were also more likely to depict incorrect forces. These results meant that novice students were effective at using intuitive, situational reasoning rather than using new formal methods. This finding supports the IC notion that students are natural problem solvers and can find adequate solutions when called on to do so.

**Research challenging hypothesis two.** Most of the research challenging the notion that applied problem solving is the most effective means of teaching algebra focus on the role that direct instruction must play in early algebra education. The primary concern leveled against discovery learning based problem solving is that the cognitive load encountered often interferes with schema development and basic knowledge acquisition. These challenges hearken back to Sweller’s (1988) work on cognitive load and schema development mentioned in the discussion of hypothesis one. In addition to Sweller, Baddeley (1992) examined the role the overloaded working-memory capacity had on schema formation. He went on to demonstrate that discovery learning conditions often overloaded working memory and conflicted with establishing well-organized knowledge structure.

Because cognitive overload was shown to be a significant deterrent to student problem solving capacity, several areas of research have emerged examining problem solving capacity at the extreme ends of cognitive load capacity. For example, Paas, Renkl, and Sweller (2004) demonstrated that performance degrades at the cognitive load extremes of either excessively low load (under-load) or excessively high load (overload). Under conditions of both under-load and overload, students ceased to learn. Learning situations with low processing demands will benefit by adding more discovery-type problem solving conditions that increase the load and challenge the learner. On the other hand, learning situations with an extremely high load will benefit from more direct-
instruction practice conditions that reduce the load to more manageable levels.

Additionally, Kalyuga, Ayres, Chandler, and Sweller (2003) showed that knowledge of the learner’s level of expertise is important in accurately determining the cognitive load level of the designed activities and objectives. A cognitive load appropriate for a novice may be an under-load for an expert.

As insightful as cognitive load research has been for the past two decades, it does not address how a student’s prior knowledge and misconceptions affect problem solving capabilities. Rowell, Dawson, and Lyndon (1990) proposed a strategy in which resolution between students’ prior ideas and new conceptions occurred after new concepts were introduced and schema formation took place. Their approach was novel in that it drew on perspectives from Sweller’s cognitive load theory and Piaget’s (1979) equilibration theory. Equilibration theory is a discovery learning based strategy built upon the notion that introducing conflict with a student’s prior conceptions is an effective strategy for deeper understanding beyond the novice level. Again, this finding significantly supports the hypothesis by IC proponents that ill-defined physics problems are an effective tool to use to support student algebra learning.

Given the important role that student prior knowledge plays in algebra learning, several studies demonstrated effective methods that build on this pre-existent student knowledge base. One such method is the worked example in which a student works alongside a tutor or teacher to solve a problem (Brown, Collins, & Duguid, 1989). Some researchers have hypothesized that when students study worked examples (a form of direct instruction) rather than solve problems unaided (a form of discovery learning) student prior knowledge is more efficiently utilized, thereby supporting schema formation and preventing cognitive overload. For example Conati, Gertner, VanLehn,
and Druzdzel (1997) researched worked examples in the context of coached problem solving. The tutor gave immediate feedback regarding the correctness of the student’s work. The process outlined by Conati, et al. included an initial phase in which the student acquired preliminary knowledge of the skill via direct instruction from a text or tutor. During this instructional period the student did not try to apply the skill. The intermediate phase consisted of learning how to solve problems by applying the knowledge encountered. This phase was supported with considerable help from the tutor and included numerous examples being worked by the tutor or peers. The final phase occurred when a student performed the skill autonomously. These skills were then applied to ill-defined problems. Researchers found that students using this method performed better on autonomous performance items and retained that knowledge for a longer period that those students who skipped the intermediate phase.

Related to the timing of direct instruction in the teaching sequence, Atkinson, Derry, Renkl, and Wortham (2000) postulated that effective direct instruction at the lesson level did three things: (a) employed multiple worked examples for each conceptual problem type, (b) varied the example formats within problem type, and (c) employed surface features to signal deep structure. Also, examples were presented in close proximity to matched practice problems. They also found that learners were encouraged through direct training or by the structure of the worked example to actively self-explain examples. These researchers associated worked examples with early stages of skill development.

As students become more proficient in problem-solving using worked examples, they move into a more autonomous role that is more dependent on their own understanding rather than the understanding of someone else. This transition was
researched by Rittle-Johnson (2006) who explored the role that self-explanation can have in direct instruction. The basis of Rittle-Johnson’s framework was a combination of discovery learning and the direct instructional strategy of worked examples on mathematical equivalence. Both self-explanation and direct instruction helped children learn and remember a correct procedure. However, self-explanation improved student performance on math questions regardless of instructional condition.

As research in this area has progressed, it has become clear that models of direct instruction early in concept development followed by discovery learning based problem solving approaches later in the sequence are very effective in improving learning and performance on procedural problem solving tasks. For example, Eiriksdottir and Catrambone (2011), in describing effective methods for developing procedural fluency, found that direct instruction in the form of specific instructions helped initial performance, whereas discovery learning in the form of vague instructions helped retention and transfer. Eiriksdottir and Catrambone showed it was possible to meet both goals of good initial performance and long-term retention by combining different types of instruction and using a strategy called fading. In fading, learners were initially supported to accomplish a task (i.e. worked examples). This support was then gradually removed (faded) as they learned to do the task (i.e. self-explanation). Finally, students were confronted with an ill-defined problem with little to no direction in which they used the skills they just developed to come to a solution to the problem (i.e. discovery learning).

In summary, the IC assumption that algebra is best learned by solving applied problems is supported in the current research literature. Many studies point out that problem solving in algebra instruction provides context, meaning, and structure to the difficult transition from arithmetic to algebra. Examining the notion of problem solving
further, the current research expressed support for a balance between problem-solving via direct instructional models (i.e. worked examples and self-explanation) and problem-solving via discovery learning. When applied problem solving was juxtaposed between discovery learning and direct instruction, it appeared to significantly improve math concept retention and performance. Importantly, no research was found that undermined the role that problem solving had in improving algebra performance. Rather, the debate between direct instructional strategies and discovery learning strategies centered on the timing at which to introduce each strategy. Fortunately, IC models utilize both strategies, so it becomes an issue of balance between the two approaches within an inverted physics curriculum.

**Hypothesis Three: IC Strategies are Effective in Improving Algebra Learning**

The final foundational hypothesis made by IC proponents is that IC models are effective methods of improving algebra performance. The literature related to IC models abounds with effusive praise bordering on advocacy for the effectiveness of these models. Although the support is enthusiastic, few quantitative studies have examined IC model efficacy, and of those, most have low levels of external validity due to sampling biases (Goodman, 2006). With this caveat in mind, this section will review the literature on effective algebra instructional strategies and identify where these intersect with topics covered in an IC physics course.

**The link between IC and improved math performance.** IC proponents claim that using algebra-based concepts in the physics curriculum provides a concrete context for algebra students to base concepts (Lederman, 2005). Initially, however, Lederman believed that 9th grade students lacked the prerequisite algebra skills and that physics should be taught in a math-free, conceptual manner. More than five years later, Lederman
changed his position and stated that mathematics must be an integral part of this new approach to science education. He pointed out that mathematics must be brought into the curriculum early because so many students are “math-phobic” in later grades (Lederman, 2005).

In support of Lederman’s position, several researchers have examined the effect of IC models on improving algebra performance. For example, Hickman (1990) reported that 9th grade inverted curriculum students outperformed 11th grade non-IC students on the New York Regents exam. He anecdotally attributed this difference to the fact that the students were able to understand better the math involved because it coincided with their algebra class.

Ewald, Hickman, Hickman, and Myers (2005) delved deeper into physics curricula to observe that students developed proficient algebra skills through physics experimentation, data collection, and data analysis. This was because they interacted with physics phenomena to which they could easily relate. Additionally, the researchers found that students applied their growing mathematical skills to solve real-world problems and demonstrated improved mathematical understanding and achievement. Corroborating this relationship was Sherin (2001) in which he argued that mathematics and physics are inextricably intertwined given that mathematical expressions are part of the language of physics.

One of the few quantitative studies examining the relationship between math performance and IC was done by Glasser (2012) in which he found statistically significant improvements in student PSAT math performance among three cohorts of students who matriculated through a 9th grade physics course compared to a matched group of earlier student cohorts who matriculated through the non-IC ninth grade biology
course. While internal reliability of the study was high, the external reliability of the study is questionable because the setting for the experiment was a suburban private day-school. Questions remain as to whether the same effects would be observed in a more non-IC public school setting.

Liang et al. (2012) corroborated Glasser’s study by observing statistically significant differences in student inquiry, modeling, communication, and reflection of results within IC classrooms. The researchers went on to observe that particular teaching practices contributed most to the students’ enhanced problem-solving abilities, including working in small groups to conduct investigations, writing explanations about what students observed, making presentations to the class on their investigations, and critically reviewing peers’ work. Additionally, Malone (2008) reported that the students experiencing these types of practices in the environment of a modeling physics instruction classroom developed more expert-like knowledge structure and problem-solving skills associated with greater metacognitive awareness than those students learning algebra in a more non-IC classroom environment.

**The link between general physics and improved math performance.**

Researchers reported a strong link between algebra skill improvement and physics problem solving when student mathematical misconceptions were accounted for. For example, Kieran's (1992) and later Rittle-Johnson's (2006) investigation of students’ interpretations of the equals sign in simple mechanics equations demonstrated that many students viewed the equal sign as meaning “something happens” in a physical system. This coincides with a common student misconception that the equals sign means “do something” in general math (Knuth, Stephens, Mcneil, Alibali, & Alibali, 2013). Although Kieran indicated that it is not clear whether this interpretation is harmful, the
current literature indicates this is a significant and widespread misconception that many students bring with them to math class that may be reinforced by the addition of physics. This finding highlights the importance of identifying and correcting student misconceptions early in the instructional sequence. Interestingly, physics instruction can be an effective method to correct this misconception by physically demonstrating that the equals sign denotes equivalence, not action by comparing the equivalent weights of two physical objects.

In addition to misconceptions regarding the equals sign, several studies investigated students’ misconceptions of the variables in equations. Clement, Lochhead, and Monk (1981) followed by Niss (2012) found that students had difficulties in translating from a verbal representation of a physical process to a mathematical representation in terms of algebraic symbols. Besson (2010) demonstrated that students found it hard to parse the relationship between variables in multivariable problems involving Newtonian forces. Steinberg, Wittmann, and Redish (2002) found that students encountered difficulties with equations involving functions of more than one variable in physics problems involving velocity in multiple directions. The work done by these researchers corroborates research mentioned in Chapter One regarding equations, variables, and expressions being the main difficulties of learning algebra. The various lines of research cited in this review point to the critical importance of early, consistent, and deliberate instruction on these three critical concepts throughout a student’s algebra study.

Exploring students misconceptions further, Domert, Airey, Linder, and Kung’s (2012) study explored the ways students interpret sign conventions used in introductory classical mechanics equations. They observed that many students found it difficult to
interpret the positive or negative sign allocated to the vector components and that these difficulties are similar to difficulties that students have using a number line.

Outside of the mechanics of algebra, specific models of physics-based problem solving were explored. Hearkening back to the model of schema formation elucidated by Sweller (1988), Priest and Lindsay (1992) presented the schema-guided forward inference model. This model was designed to simulate the process from novice to expert knowledge in a variety of physics topics such as Newton’s second law or the conservation of momentum. Within this model, the problem was first recognized as being of a certain type, followed by the presentation of a schema that guided the learner to a solution. These schemata contained equations and techniques for instantiating the equations in particular circumstances.

While not algebra specific, Marrongelle, Sztajn, and Smith (2013) qualitatively examined how calculus students in an integrated calculus and physics class used physics to help them solve calculus problems. In this case study the researchers monitored eight students and found that some students autonomously introduced contexts to solve the problems successfully. The ability of a student to place mathematical information into a context with which they were familiar was an important first step in becoming an expert problem solver. This study showed that physics was an effective context in which the mathematical concepts were placed. This process was one of the most common advantages of IC models cited by proponents.

In summary, the third hypothesis put forth by IC proponents is supported in the research literature. Several studies documented the linkage between algebra and physics learning; however, important instructional methods to address student misconceptions must be used early and often. There were many overlapping topics between algebra and
physics shown to be effective pathways to improved algebra conceptual understanding and subsequent performance.

**Theoretical Framework**

Although observations of problem solving behavior abound in the literature, few theoretical frameworks encapsulate the specific processes by which proficient students comprehend problems, digest problem information, and access the relevant real-world knowledge that can be applied to the problem’s representation and solution. The process of model comprehension and the correct application of procedures is the primary skill set that IC proponents identify as benefits of students learning physics and algebra simultaneously (Pasero, 2003).

Much of the research related to mathematical problem solving has concluded that a significant component of a comprehension model is how a student’s mental representations of the problem situation inform and help to constrain the formal expressions necessary for a solution (Cobb, Stephan, McClain, & Gravemeijer, 2011). Based on these findings, this research uses the Nathan et al. (1993) conceptual problem model as the theoretical framework to understand the complex nature of mathematical problem solving. It is important to understand how mental representations inherent in physics-based problems might improve algebra performance and concept retention. Nathan et al. showed that arithmetic-word-problem comprehension could be understood within the framework of the general theory of discourse processing in that efficient problem solving is no different from understanding a story. To understand a story problem, the reader must have sufficient knowledge to understand the situation adequately and apply appropriate strategies to generate necessary inferences and elaborations to make the story complete.
The conceptual problem model theorizes that mental representations are formed when a student, after reading a problem statement, forms a propositional representation, termed the textbase, to capture the meaning of the passage. Simultaneously, the student forms a representation for the actions in a text, termed the situation model. The next phase involves the development of the problem schema and is critical in the development of proficient algebraic problem solving. The problem schema is the combination of the textbase with the situation model producing a coherent representation of the problem in the student’s mind. Successful production of the problem schema allows the student to begin the process of selecting from several possible problem-solving strategies, which naturally capture the set relations and arithmetic operations needed for solving the problem. From these arithmetic operations a set of algebraic problem schemas (templates for organizing problem-relevant information) can be developed providing the explicit, graphical cues to guide the construction of formalized algebraic representations of these problem models. Finally, equations may be derived from the algebraic representations specified in the problem schema (see Figure 1).

**Figure 1.** Conceptual problem model framework

In this view, the process of understanding and solving word problems involves three mutually constraining levels of representation that must be constructed by the student: (1) a representation of the textual input—the textbase; (2) a model of the
situation conveyed by the text in everyday terms—the situation model; and (3) the integration of the textbase and the situation model—the problem schema. The problem schema gives rise to the appropriate arithmetic and algebraic operations needed for the successful completion of the problem. In the problems considered by Nathan et al. (1993), situation models were often based on imagery or concrete representations of the textbase, much like the physics problems proposed by IC proponents.

In theory, the simultaneous study of algebra-based physics and algebra is supported by the conceptual problem model in several ways. First, physics learning confronts students with problems they observe every day in the world around them in the form of common physical processes (Lederman, 2005). Second, students take this common event and create both a textbase and a situation model describing the event, such as dropping an object from a height. This common event is easily pictured in the student’s mind and is easily described by the student, thereby forming the problem schema. Third, the student identifies the tools available to her in order to solve the problem. At this point the appropriate background knowledge must be in place (i.e. the acceleration due to gravity) either through discovery learning or direct instruction. Particular methodologies may include table construction, drawing a graph, modeling the event in real life, or any combination. The correct solution depends on the student correctly identifying the various components of the system, correctly synthesizing those components into a coherent whole, then identifying the tools and strategies that could be used to derive a solution.

The notions encompassed by the conceptual problem model make logical sense, but has this sequence been documented in the research literature? Prior to Nathan et al. (1993), Cummins, Kintsch, Reusser, and Weimer (1988) showed that incorrect problem-
solving behavior could be simulated by introducing faults to the students. Two classes of defects were introduced: (a) incorrect arithmetic algorithms and (b) linguistic deficiencies (analogous to the textbase). Language processing errors modeled students’ faulty behavior best. These researchers also showed that solution performance was associated with the ability to recall the problem statements. Students who could correctly recall the story textbases were more likely to produce correct solutions. Interestingly, students who incorrectly recalled problems apparently encoded a different problem textbase than the one presented, which they often then solved correctly. This study suggested that the development of an appropriate representation of the formal aspects of a problem is highly dependent on language comprehension skills, which in turn form correct textbase and situation models. This in turn affects word-problem-solving performance. The study prescribed that instruction focus on language comprehension processes as well as on the mathematical aspects of word-problem solving task. This finding is important when examined in the context of algebra learning because a significant portion of the transition from arithmetic to algebra is making the leap from concrete numbers to abstract variables. This transition is fundamentally a language comprehension skill activity requiring a student to form a textbase and situation model from abstract notation in the form of variables, equations, and expressions. Approaching this process from the vantage of language comprehension better explains the difficulties a beginning algebra student faces. The mechanics of the problem are secondary to forming a textbase and situation model in which to apply the mechanics. The difficulty is forming the textbase and situation model because the notation is foreign to the student.

McCloskey (1983) and later Ambrose, Bridges, DiPietro, Lovett, and Norman (2010) reported on instances where students lacked a situational understanding (e.g.
physics problems that challenge students’ naive notions). In these instances the situational understanding had to be taught first before the problem schema could be developed. This finding was in agreement with Sweller’s (1988) findings related to cognitive load in problem solving. When there was situational understanding, it provided the student with a solution-enabling algebraic interpretation for the problem text. This process functioned by helping the student access and apply the real-world knowledge associated with a situation when setting up the formal relations or problem schema. Singley (1989) postulated that the situation model further helps the student by providing situational constraints, against which formal constraints may be checked.

Of course, none of the steps making up conceptual problem solving matter if it has not been shown to improve mathematical concept retention. Gresalfi, Martin, Hand, and Greeno (2009) observed that when students created and manipulated mathematical objects in ways that they could explain (i.e. created a textbase) and justify (i.e. created a situation model or problem schema) they demonstrated improved conceptual understanding and recall up to six months after learning occurred as compared to a control group engaged in more passive learning structures. They theorized that such an approach afforded ample opportunities for students to be actively engaged in the learning process and to engage both concepts and skills.

In addition to improved concept retention, Rasmussen, Zandieh, King, and Teppo (2005) observed that students were able to create meaningful mathematical ideas as they engaged in challenging tasks. In particular, when students were able to use symbols, algorithms, and definitions placed in particular contexts, they could better understand those concepts and retain the associated conceptual understanding for up to a year after the instruction took place. In theory, when students engaged in building concepts from
the bottom up through a process of suitably guided reinvention, concept retention was improved. Kwon, Rasmussen, and Allen (2005) also observed that representations, such as graphs of solution functions, were better understood when students recreated mathematical concepts as part of the problem-solving process. They later showed that students in inquiry-oriented math classes usually used some sort of graphical visualization as they explored new mathematical concepts in class. The researchers hypothesized that retention of pictorial images was longer term than retention of semantic knowledge, such as procedures and algorithms. The efficacy of pictorial image use in math problem solving corroborates the importance of students creating a situation model within the conceptual problem model framework.

In summary, the conceptual problem model framework was supported by the research literature described in this section. This model provided a robust description of the necessary processes to improve problem-solving, concept retention, and student performance related to algebra instruction. The model was supported by diverse research questions and, most importantly, accurately conveyed the necessary balance between discovery based learning and direct instruction that was elucidated by the research literature. Using this theoretical framework provides an accurate description of the problem-solving process used by students and enlightens the following discussions related to the efficacy of IC models on algebra performance and course-taking patterns.

This literature review examined the theory of change concerning effective algebra education and compared the findings to foundational hypotheses laid out by IC proponents. In all cases, the literature supported these hypotheses. Based on the literature reviewed, students who matriculate through an IC model should experience improved algebra performance, concept retention, and overall math engagement. Moreover, Koller,
Baumert, and Schnabel (2001) pointed out that algebra proficiency then directly supports future advanced math course taking and improves general math performance.

Additionally, Singh, Granville, and Dika (2010) noted that algebra proficiency is a primary indicator of future academic success. Based on the extant literature, then, the theory of change concerning IC models is as follows: (1) IC models create a context for algebra learning (Hake & First, 2002; Lederman, 1998, 2005; Tabachnick & Fidell, 2007; Trumper, 2006); (2) this context for algebra learning improves algebra performance (Cobb, 1994; Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013; Nathan et al., 1992; Rakes et al., 2010); hence (3) improved algebra performance leads to students taking more advanced math courses leading to improved math achievement (Koller et al., 2001; Schmidt, Houang, & Shahrani, 2009; Schmidt et al., 2005; Shen & Tam, 2008; Shen & Pedulla, 2010). The remainder of this dissertation examines whether this theory of change is apparent in the 2009 NAEP HSTS data.
CHAPTER 3

METHODS

This study examines the effectiveness of inverted curriculum (IC) models on 12th grade student NAEP math assessment performance, algebra performance, and advanced math high school course taking. Inverted curriculum proponents claim that physics should be taken in 9th grade to coincide with algebra instruction. This dissertation will now examine the efficacy of these claims by examining whether IC models can improve student math GPA and algebra performance and increase the number of advanced math course taken by a high school student.

To review, IC proponents asserted that IC models improve algebra skills because of the widespread and concrete use of algebra concepts in physics problems (Goodman, 2006; Myers, 1987). This claim was based on three hypotheses: (1) The defining outcome of quality math education is the creation of robust problem solvers; (2) given the truth of hypothesis one, applied problem solving efficiently and effectively develops the requisite algebra skills needed for student success in future math study; and (3) inherent in physics instruction reside effective methods for algebra instruction to improve student performance and increase the number of advanced math courses taken (Lederman, 2005). These aforementioned hypotheses were examined in the context of how algebra learners use specific math problem-solving tasks as outlined in the conceptual problem model framework (Nathan, et al., 1992) (see Chapter 2). The research questions were used to determine whether IC models were effective in improving math performance outcomes by using physics instruction to provide the text-base, situation model, and problem schema identified in the conceptual problem model. Finally, data from the 2009 NAEP HSTS were analyzed to examine the effect of IC models on general math performance,
algebra performance, and advanced math course-taking patterns. This analysis was designed to examine the veracity of the three hypotheses postulated by IC proponents within the context of the conceptual problem model framework.

Importantly, the research questions did not directly examine the efficacy of the specific problem-solving tasks outlined in the conceptual problem model. Instead the questions examined specific math performance outcomes that, if improved, would substantiate the effectiveness of IC models in improving the math problem solving skills identified in the conceptual problem model, namely the formation of a text base, situation model, and problem schema. In this research, the conceptual problem model provides an explanatory framework with which to discuss the quantitative findings related to algebra learning and positive math performance outcomes.

As a reminder this study used data from the 2009 NAEP HSTS dataset to address the following research questions:

**Between-Groups Research Questions: Comparing IC to non-IC Students**

1) Is there a significant difference in performance on the 12th grade NAEP mathematics assessment between students who receive early physics instruction (IC) in 9th grade and those who do not (non-IC)?

2) Is there a significant difference in cumulative high school algebra GPA between IC and non-IC students as documented on the student’s 12th grade transcript by the HSTS?

3) Is there a significant difference in the number of advanced math Carnegie units earned by IC and non-IC students as documented on the student’s 12th grade transcript by the HSTS?
Within-Group Research Questions: Comparing IC Students to Each Other

4) Among IC students, is there a significant relationship between the length and type of physics class taken and 12th grade NAEP math assessment performance?

5) Among IC students, is there a significant relationship between the length and type of physics class taken and the overall algebra GPA earned?

6) Among IC students, is there a significant relationship between the length and type of physics taken and the number of high school math Carnegie credits earned?

The complete NAEP HSTS 2009 dataset included 41,200 12th grade students from 950 schools. The analysis specific to this study used 6,554 students from 552 schools from within the larger NAEP HSTS 2009 dataset. Specific filtering criteria are described later in this chapter.

NAEP HSTS 2009 Overview

NAEP HSTS 2009 was designed to allow an analysis of the course-taking patterns of US high school completers who graduated from US public and private high schools in 2009. It was further designed so that data on graduates’ course-taking patterns can be linked to the NAEP 2009 assessment results. NAEP provides results about subject-matter achievement, instructional experiences, and school environment. These results are reported for populations of students (e.g., 12th-graders) and selected subgroups of those populations (e.g., female students) (Grigg et al., 2010).

NAEP 2009. The NAEP assessment is often referred to as “the nation’s report card” because it is the only nationally representative, continuing assessment of what US
students know and can do in various subject areas. NAEP provides a comprehensive measure of students’ learning at critical junctures in their school experiences. In existence since 1969, NAEP assessments have been conducted periodically in such subjects as reading, mathematics, science, writing, US history, civics, economics, geography, and the arts. Importantly, NAEP does not provide scores for individual students or schools; instead, it offers results regarding subject-matter achievement, instructional experiences, and school environment for populations of 4th, 8th, and 12th grade students and student groups of those populations (e.g., female students, Hispanic students) (Grigg et al., 2010).

The NAEP program includes two distinct components: “main NAEP” and “long-term trend NAEP.” Main NAEP, the dataset being used in this research, includes assessment instruments used at both the national and state levels. Long-term trend NAEP is not used in this study because it is administered less frequently than main NAEP (Nord et al., 2011).

The main NAEP assessment reports data for students from both public and nonpublic schools in specific census-defined geographic regions of the country (northeast, south, mid-west, and west), as well as for several major demographic student groups (Sikali, 2011). Main NAEP assessments include constructed-response questions (questions that ask students to write responses ranging from a single word or figure to a few paragraphs) and questions that require the use of calculators and other materials (Grigg et al., 2010).

The 12th grade NAEP mathematics assessment measures student knowledge and abilities across four mathematical content areas: number properties and operations; measurement and geometry; data analysis; statistics and probability; and algebra. The distribution of questions among each content area differs by grade to reflect
developmentally appropriate knowledge and skills. Mathematical complexity is modeled by tasks that require students to reason, perform procedures, understand concepts, and solve problems at various levels of difficulty (Grigg et al., 2010). Sample question booklets can be referenced at: http://nces.ed.gov/nationsreportcard/about/booklets.aspx.

Student responses to questions in each of the four content areas are measured on a 300 point scale. In addition to reporting an overall mathematics score, scale scores are reported using three cut-point scores. The lowest category is basic, which requires a minimum score of 141. A basic level student is defined as one who is able to solve mathematical problems that require the direct application of concepts and procedures in familiar situations. The middle category is proficient, which requires a minimum score of 176. Proficient students are those who are able to select strategies to solve problems and integrate concepts and procedures. Finally, the highest scoring category is advanced, which requires a minimum score of 216. An advanced student is defined as one who is able to demonstrate in-depth knowledge of the mathematical concepts and procedures represented in the NAEP frameworks. These achievement levels are cumulative; therefore, students performing at the proficient level also display the competencies associated with the basic level, and students at the advanced level also demonstrate the skills and knowledge associated with both the basic and the proficient levels (Sikali, 2013).

**HSTS 2009.** The HSTS is a periodic survey designed to provide educational professionals and researchers with information regarding high school graduates’ course-taking patterns and grade point averages (GPAs). The HSTS 2009 collected these high school course-taking data in the form of complete transcripts from a nationally representative sample of US public and private high schools (Nord et al., 2011).
Additionally, the HSTS 2009 linked high school course-taking patterns to NAEP achievement by collecting transcript data from 12th grade students who took the 2009 NAEP assessments (Nord et al., 2011).

In addition to the course-taking and academic performance variables, contextual background data for HSTS 2009 were obtained from the NAEP 2009 questionnaires via school-level surveys completed by a school coordinator. These questionnaires included (a) The School Background Questionnaire containing information about the school, its teachers (for the 4th and 8th grade assessments), and its student body; (b) The Students with Disabilities (SD) Background Questionnaire containing information about students classified by their schools as having a disability; (c) The English Language Learners (ELL) Background Questionnaire containing information about students classified by their schools as having limited proficiency in English; and (d) The Student Background Questionnaire completed by each student containing information about the student’s background, demographic characteristics, and educational experiences (Nord et al., 2011). The conjunction of these two national datasets: NAEP and HSTS, which link subject-matter achievement with course-taking patterns and school environment into one nationally representative sample provides an opportunity to explore how 9th grade science course taking patterns influence overall high school math performance.

Quasi-Experimental Design

Unit of analysis. This study investigates how the organization of the high school curriculum affects student performance and as such falls into a category called “school-effects research” (Raudenbush & Williams, 1995) making hierarchical linear modeling (HLM) an appropriate methodology (Raudenbush & Bryk, 2002). More specifically, the units of analysis in this study are students nested or clustered within schools. Such
clustering comes with well-known statistical properties that must be addressed to realize non-spurious estimates. Hierarchical linear modeling is a well-known strategy designed for complex stratified and clustered samples with balanced or unbalanced designs and varying degrees of missing data patterns. This modeling technique is useful for the analysis of educational data from complex designs involving nested data made up of large samples (Hayes, 2006). HLM has a long and robust history in educational research. Some examples include Pasero’s (2008) use to compare the growth in student attitude about the nature of scientific knowledge in non-IC and IC programs. Shelley and Lubienski (2005) utilized HLM to identify reform oriented mathematical instructional practices and how they correlate with achievement for disadvantaged students. Shelley and Su (2011) applied HLM to NAEP data to investigate the role of social structures in the relationship between mathematics instruction and student achievement.

**School sample selection.** Schools were selected systematically from a stratified school frames sample with probabilities proportional to the number of enrolled 12th grade students. To ensure a sufficient number of Black and Hispanic students in the sample, these students were given twice the probability of selection of a low minority school of comparable size. Sample school population demographics are presented in Table 1.
Table 1

School Population Demographics

<table>
<thead>
<tr>
<th>Total school sample: $n = 552$</th>
</tr>
</thead>
<tbody>
<tr>
<td>% of schools with majority (&gt; 50%) of students eligible for:</td>
</tr>
<tr>
<td>Free/reduced lunch</td>
</tr>
<tr>
<td>Title 1 services</td>
</tr>
<tr>
<td>ELL services</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>School demographics</th>
</tr>
</thead>
<tbody>
<tr>
<td>High minority student population</td>
</tr>
<tr>
<td>Type</td>
</tr>
<tr>
<td>Public</td>
</tr>
<tr>
<td>Private</td>
</tr>
<tr>
<td>Missing</td>
</tr>
<tr>
<td>Urbanicity</td>
</tr>
<tr>
<td>Urban</td>
</tr>
<tr>
<td>Suburban</td>
</tr>
<tr>
<td>Town/rural</td>
</tr>
</tbody>
</table>

The sampling frame of schools was created by combining the 2005–2006 Common Core of Data (CCD) file of public schools and the 2005–2006 Private School Universe Survey (PSS) file of private schools. Additionally, special procedures were used to represent students in new public and parochial schools which opened between 2007 and 2009. The school sample included Bureau of Indian Education (BIE) schools, Department of Defense Education Activity (DoDEA) schools, magnet schools, and charter schools. Schools that provide only special education, vocational training, were part of hospital or treatment center programs, were part of juvenile correctional institutions, were home school or online entities, or were for adult education were excluded from the sampling frame.

**Student selection procedures.** In each sampled school, a sample of students was selected from all eligible grade 12 students provided on a student roster. The student sample was collected from 950 eligible schools sampling 41,200 12th grade students, who graduated by October 2009. Graduates’ transcripts were collected by field workers.
following all Family Educational Rights and Privacy Act (FERPA) and National Center for Education Statistics (NCES) policies for ethical research.

The 12th grade student sample for NAEP 2009 was a two-stage probability-based national sample that included NAEP mathematics results for the following 11 states: Arkansas, Connecticut, Florida, Idaho, Illinois, Iowa, Massachusetts, New Hampshire, New Jersey, South Dakota, and West Virginia. The first stage of sampling involved the selection of schools. The second stage involved selecting students within schools and their assignment to an assessment subject. Students were selected with equal probability using a systematic sampling approach. The within-school student sample sizes varied by explicit school sampling strata and are shown in Table 2.

Table 2

<table>
<thead>
<tr>
<th>Strata</th>
<th>Student n per school</th>
<th>Mathematics Percent</th>
<th>Science Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>Public schools</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Idaho, New Hampshire, South Dakota, West Virginia</td>
<td>80</td>
<td>49</td>
<td>1</td>
</tr>
<tr>
<td>Arkansas, Connecticut, Iowa</td>
<td>85</td>
<td>49</td>
<td>2</td>
</tr>
<tr>
<td>Massachusetts</td>
<td>95</td>
<td>47</td>
<td>4</td>
</tr>
<tr>
<td>New Jersey</td>
<td>100</td>
<td>47</td>
<td>5</td>
</tr>
<tr>
<td>Illinois</td>
<td>120</td>
<td>46</td>
<td>7</td>
</tr>
<tr>
<td>Florida</td>
<td>125</td>
<td>45</td>
<td>8</td>
</tr>
<tr>
<td>Remaining states</td>
<td>150</td>
<td>43</td>
<td>19</td>
</tr>
<tr>
<td>Private schools</td>
<td>150</td>
<td>43</td>
<td>19</td>
</tr>
</tbody>
</table>

SOURCE: U.S. Department of Education, Institute of Education Sciences, National Center for Education Statistics, National Assessment of Educational Progress (NAEP), 2009 Grade 12 Assessment. * Does not equal 100% due to reading assessments also being distributed

Data collection procedures. The data collection portion of HSTS 2009 took place in two phases. During Phase 1, from September 2008 through March 2009, field workers contacted schools in their region by phone and in person. Phase 1 involved introducing the study to the sampled schools, obtaining school and course information necessary to understand the content of the transcripts that would be collected during
Phase 2, flagging sampled students’ records to facilitate Phase 2 data collection, and arranging a return visit to the school to collect transcripts from graduating students. Phase 2 took place from the end of the school year: May through October 2009. During this phase, selected field workers visited the schools to collect the sampled students’ current transcripts.

**Data coding.** Transcripts contained lists of courses taken and the grades and credits earned for each course, and also contained a variety of other information, such as graduation date, class rank, names of tests taken by the graduate, and test scores.

Students’ transcripts averaged slightly more than 46 course records per graduate. For each course, the catalog course ID number, course name, grade level (9th, 10th, 11th, or 12th), year in which the course was taken, term (e.g., fall semester, summer school), grade received, and number of credits earned were recorded.

**Weighting procedures.** The NAEP HSTS 2009 structure is a complex structure using NAEP-linked and un-linked weights, which allow users to analyze the relationship between student academic performance and student course-taking patterns, as measured by their NAEP mathematics assessment outcomes. Nord et al. (2011) recommended the use of linked weights when computing estimates of high school graduates in the HSTS sample linked to a particular NAEP assessment. The problem of whether and how to incorporate weights in fitting HLM 7.0 to survey data is an area of active research (Chambers, 2003; Little, 2003; Pfeffermann, Moura, & Silva, 2004; Pfeffermann, Skinner, Holmes, Goldstein, & Rabash, 1998). More recently Braun, Jenkins, Grigg, and Tirre (2006) observed that there is no unanimity in the field with respect to this question, even as to whether weights should be used at all. Alternative suggestions have been made, but there is no consensus on a preferred approach. Typically, sampling weights are
used in the estimation process to account for the fact that the probabilities of selection were not identical for all students. For this HLM analysis, weights were handled following a suggestion of Pfeffermann et al. (1998) and Braun, Jenkins, Grigg, and Tirre (2006) that employed variable school weights at level 2 but no student weights at level 1.

**Plausible values.** NAEP assessment scores are collected in spiraled balanced incomplete block (BIB) design. BIB is a type of "sparse item-sampling design," which has the advantage of obtaining performance data on a large number of items within the mathematics assessment across the tested population. The result of the BIB sampling design is that individuals’ scores on the same test are imputed rather than directly observed. Because the test item scores are imputed values, NAEP provides five estimates of cognitive performance for each individual, rather than a single score. In the 2009 mathematics assessment, NAEP refers to these imputed values as plausible values. These plausible values are conditioned on information from two sources—performance on the items on which each student was assessed and background information about the student and his/her school. Scaling, using information from the first source, involved the use of Item Response Theory (IRT), a well-known psychometric technique used in large-scale dichotomous-response (right/wrong) assessment. Detailed information on construction of all variables used in the study (including plausible values for the 2009 mathematics assessment), the NAEP items from which they were drawn, and other psychometric information can be referenced at [http://www.nces.ed.gov/nationsreportcard/tdw/](http://www.nces.ed.gov/nationsreportcard/tdw/).

**Present Study Design**

The specific sample frame used in this study \( n = 6,554 \) was selected from the aforementioned NAEP HSTS 2009 data set based on several filtering criteria. First, student HSTS records from their 9th grade year were selected from the larger data set to
ensuring a sample that reflected an IC model. Within the 9th grade data, all course titles were analyzed to determine what type of science class each student took in his or her 9th grade year. Due to the large number and specificity of course titles in the HSTS, courses were selected only if they had generic names that clearly fit the categories of the study (i.e. “physical science” versus “aerospace science”). Additionally, all classes with an honors, advanced, or pre-AP designation were dropped to minimize the bias toward higher performing student outcomes as much as possible.

After sorting the selected courses into the appropriate categories each student was identified as having received physics instruction as a 9th grade student (IC model) or biology (non-IC model). The IC sample was further broken down into three sub-categories. The first category was whether a student took a full year of physics using a traditional physics curriculum. The second category was IC students who took a half-year of physics as opposed to a full year. The third category was IC students who were taught using a concept-based physics curriculum. Physics classes taught in a conceptual manner emphasize general physics principles while minimizing the mathematical concepts and calculations describing the principles. The traditional model of physics instruction tends to emphasize the mathematical concepts and calculations in support of the physics concepts. As mentioned in Chapter 2, IC proponents disagree on which model is the most effective treatment of physics for 9th grade student understanding of physics concepts and future math achievement. The difference between conceptually-based models and the math-based models will be discussed in detail in Chapter 4. Within these data, it was possible to delineate between concept-based instruction and math-based instruction using the name of the course. On the transcript, conceptual-based physics courses were listed as “conceptual physics,” “active physics,” or “physics first” after which text they used. All
three of these course titles are well recognized IC models that emphasize physics concepts over the mathematics involved in those concepts. Where these course titles appeared in the HSTS data, the course was counted as a conceptual-based class (see Table 3).
Table 3

Student Demographic and Performance Data Disaggregated by 9th grade Course Taken

<table>
<thead>
<tr>
<th>Science course taken in 9th grade:</th>
<th>Biology</th>
<th>Half-year physics</th>
<th>Full-year physics</th>
<th>Conceptual-based physics (subset of full-year physics)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Sample</td>
<td>n = 6,554 students</td>
<td>263 schools</td>
<td>2,136 students</td>
<td>2,336 students</td>
</tr>
</tbody>
</table>

Student Performance & Demographics

<table>
<thead>
<tr>
<th></th>
<th>Total Sample</th>
<th>n = 2,136 students</th>
<th>n = 2,336 students</th>
<th>n = 2,082 students</th>
<th>n = 396 students</th>
</tr>
</thead>
<tbody>
<tr>
<td>NAEP math score mean (SD)</td>
<td>158.7 (31.2)</td>
<td>159.8 (30.4)</td>
<td>146.5 (28.5)</td>
<td>160.5 (31.8)</td>
<td>159.0 (24.7)</td>
</tr>
<tr>
<td>Advanced math credit mean (SD)</td>
<td>3.9 (0.8)</td>
<td>2.4 (1.1)</td>
<td>1.4 (0.94)</td>
<td>2.1 (0.9)</td>
<td>1.9 (0.8)</td>
</tr>
<tr>
<td>Algebra 1 GPA (SD)</td>
<td>2.7 (0.8)</td>
<td>2.7 (0.8)</td>
<td>2.5 (0.8)</td>
<td>2.8 (0.8)</td>
<td>2.7 (0.8)</td>
</tr>
<tr>
<td>% PSAT &gt; 60</td>
<td>34.3</td>
<td>36.3</td>
<td>27.8</td>
<td>38.8</td>
<td>28.8</td>
</tr>
<tr>
<td>% Female</td>
<td>50.2</td>
<td>50.7</td>
<td>49.9</td>
<td>50.3</td>
<td>62.1</td>
</tr>
<tr>
<td>% Black</td>
<td>18.8</td>
<td>16.7</td>
<td>26.1</td>
<td>11.7</td>
<td>4.5</td>
</tr>
<tr>
<td>% Hispanic</td>
<td>14.4</td>
<td>16.7</td>
<td>19.2</td>
<td>10.7</td>
<td>22.7</td>
</tr>
<tr>
<td>% American Indian</td>
<td>6.9</td>
<td>2.4</td>
<td>1.4</td>
<td>0.6</td>
<td>3.0</td>
</tr>
<tr>
<td>% Asian</td>
<td>4.8</td>
<td>4.1</td>
<td>2.9</td>
<td>6.6</td>
<td>4.5</td>
</tr>
<tr>
<td>% ELL</td>
<td>5.2</td>
<td>1.4</td>
<td>1.1</td>
<td>8.4</td>
<td>0.0</td>
</tr>
<tr>
<td>% Free &amp; reduced lunch</td>
<td>32.1</td>
<td>25.7</td>
<td>35.0</td>
<td>22.8</td>
<td>19.7</td>
</tr>
</tbody>
</table>
After filtering on the aforementioned dimensions, these data were merged with the student file, NAEP performance file, and school data file to create a master file linking 9th grade science course-taking with student background variables, 12th grade NAEP performance, and school background variables. The final dataset was produced after selecting those cases that met the NCES minimum requirement for accuracy and completeness of transcript data.

A concern related to the sample frame was whether there were significant academic differences between 9th grade students who took biology compared to those who took physics. Additionally, were there significant differences between 9th grade students who took half-year physics compared to those who took full-year physics? Related to the former comparison, to check whether a 9th grade biology student was academically equivalent to a 9th grade full-year physics student this study utilized a one-way between-groups analysis of covariance (ANCOVA) to compare the performance level of biology and full-year physics students on the pre-SAT assessment in order to determine how closely matched the two student groups were on a performance measure prior to the completion of the IC intervention. The independent variable was the type of science taken as a 9th grade student (biology or full-year physics). The dependent variable consisted of scores on the pre-SAT assessment before the IC intervention was completed. The fixed factor covariates consisted of student gender, race, and the highest math course completed. Preliminary checks were conducted to ensure that there was no violation of the assumptions of normality, linearity, homogeneity of variances, homogeneity of regression slopes, and reliable measurement of the covariate. After adjusting for pre-intervention scores, there was no significant difference between the biology and full-year physics groups on the pre-SAT assessment, $F(1, 4218) = 3.31, p = .07$, partial
eta squared = .01. Related to the latter question of whether there were significant differences between the academic levels of full-year physics students compared to half-year I found a significant difference between half-year physics students and full-year physics students on pre-SAT assessment performance, \( F(1, 4418) = 4.54, p = .03 \), partial eta squared = .01. This pre-analysis supports the assumption that in this sample, 9th grade biology students are academically similar to 9th grade full-year physics students and are therefore safe to use as comparison groups for subsequent analyses. Conversely, 9th grade half-year physics students scored significantly lower on the pre-SAT assessment than 9th grade full-year physics students. These findings conform to the general assumption that a 9th grade student who takes a half-year of physics is generally less academically prepared compared to a 9th grade student who takes biology or a full-year physics course.

Variables that could confound the relationship between the presence of an IC model and achievement were included in this analysis. Lubienski (2006) identified several NAEP student- and school-level variables that correlated with mathematics achievement that could unwittingly bias results. In selecting variables to include, the intent was to focus strictly on biology, half-year physics, and full-year physics courses that account for the differences in student math achievement and advanced math course-taking patterns, as opposed to those factors that could be influenced by the school, such as the school discipline climate, teacher qualities, and parent involvement.

**Dependent measures.** The dependent measures used in this study were 2009 NAEP 12th grade mathematics assessment plausible values, student algebra GPA, and number of advanced math Carnegie units taken. Given that Carnegie units are
count data, they were analyzed as a Poisson model with constant exposure (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2011).

**Independent measures on students (Level-1).** The primary level-1 independent measures of interest used in this study are as follows:

- **TRADPHYS** – the first type of physics an IC student may have taken in 9\textsuperscript{th} grade. This variable was dummy coded to reflect an IC student who took a full-year traditional physics course (coded “1”). All other students were coded “0.” Traditional IC physics curricula generally require that students perform the algebraic calculations underlying the physics concepts. IC proponents theorize that algebra performance may be improved by teaching algebraic principles in a physics class placed at the beginning of the high school science sequence (Lederman, 1995).

- **CONCEPTPHYS** – the second type of physics an IC student may have taken in 9\textsuperscript{th} grade. This variable was dummy coded to reflect an IC student who was taught physics using a concept-based curriculum such as *Conceptual Physics* (Hewitt, 2002) or *Active Physics* (Eisenkraft, 1998) (coded “1”). These concept-based curricula teach physics in a conceptual fashion with little emphasis given to the concept’s related mathematical calculations. All other IC students were coded “0.” This variable was used to examine the effects of the type of IC curriculum on student performance compared to traditional, algebra-based models.

- **HALFPHYS** – the third type of physics an IC student may have taken in 9\textsuperscript{th} grade. This variable was dummy coded to reflect an IC student who took a half-year physics class (coded “1”). All other groups were coded “0.” This variable was used to examine the effect that amount of time
spent studying physics had on student performance compared to full-year IC models (coded “0”). This group was excluded from the analysis for research questions 1-3 but served as the reference group for research questions 4-6 in the within-groups analysis.

- **TRADBIO** – the science taken by the non-IC student sample in 9th grade. This variable was dummy coded to reflect a non-IC student who took a traditional course sequence of biology in 9th grade. These non-IC students were coded “1.” All other groups were coded “0.” This group of non-IC students was used as the reference group for research questions 1-3 in the between-groups analysis. Additionally, this group was excluded from the analysis related to research questions 4-6.

Other level-1 independent measures describing students, used as statistical controls, are as follows:

- **STDELL** – student classified by school as having a disability and/or being an English Language Learner (ELL) (dummy coded with without disability/ELL “1”, with disability/ELL “0”). Abedi, Lord, and Hofstetter (1998) showed that NAEP math performance is sensitive to student language and learning background variables.

- **FRL** – used as a proxy for SES via student eligibility for National School Lunch Program based on school records (dummy-coded with eligible coded “1,” not eligible coded “0”). Goddard, Goddard, and Tschannen-Moran (2007) demonstrated that free and reduced lunch values could serve as a viable proxy for SES. Additionally, Cooper and Dunne (2000) reported that SES has a significant impact on student’s ability to interpret open-ended math classroom discourse and contextualized problems.
PSATMAT – student’s highest pre-SAT math score received (coded in the original data as a score less than 40 “0,” score 41-50 as “1,” 51-60 as “2,” and greater than 60 as “3”). Moran, Oranje, and Freund (2009) established a statistically significant relationship between NAEP mathematics and SAT mathematics by computing a correlation of $r = 0.91$ between the two instruments using a sample of 15,200 students. This finding supports the use of SAT scores as predictive measures for NAEP and vice versa. In this study, 11th grade PSAT score will be used as a proxy measure for pre-NAEP assessment performance given its close alignment to SAT performance. This proxy variable will control for endogenous mathematical ability not accounted for by science course-taking.

- RACE – school-reported student race/ethnicity (dummy coded with African Americans, Latinos, and Native Americans coded "1," others [Whites and Asians] coded "0"). Race-related effects along with SES and parent educational levels have been shown to account for significant amounts of variation in student math performance (Rothstein, 2004).

- SEX – student gender (dummy coded with female coded “1,” male coded “0”). Steffens and Jelenec (2011) observed that gender stereotypes related to mathematical study significantly undermined women’s performance and interest in math independent of their academic ability to perform mathematical calculations. This finding indicates that student gender plays a significant role in math performance and is therefore an important variable to specify in any statistical model.
- **EDU** – highest education level attained by the parent (coded as graduated college “3,” some education after high school “2,” graduated high school “1,” and did not finish high school “0”). Ferguson (1998) showed that parent education is highly predictive of student motivation and success.

- **HIGHMATH** – highest math course taken by the student (coded as algebra one “1,” geometry “2,” algebra two “3,” advanced math “4”). Adelman (1999) observed that finishing a course beyond the level of algebra two improves student high school outcomes and more than doubles the odds that a student will complete a bachelor’s degree. This variable was removed from the analysis of research questions 3 and 6 due to the high degree of inter-correlation between taking advanced math courses and earning more advanced math Carnegie units.

**Independent measures on schools (Level 2).** Variables describing schools were also of two types—those describing demographic and structural characteristics of schools and those measuring the schools' academic organization. These variables are as follows:

- **STYPE** – a categorical variable indicating type of school (coded as private “0” and public “1”). Lubienski and Lubienski (2006) observed that school type can significantly affect student mathematical performance.

- **ENROLL** – an ordinal variable classifying a school’s size. This variable was coded as “1” for schools less than 400 students, “2” for schools between 400 and 800 students, “3” for schools between 800 and 1000 students, “4” for between 1000 and 1500, and “5” for schools above 1500
students. Leithwood and Jantzi (2009) observed significant school size effects on student math performance when socio-economic variables were also accounted for.

- **URBAN** – a dummy coded variable reflecting schools located in an urban school district (coded “1”). All other schools were coded “0.” Suburban districts were used as the reference group. Urbanicity has been shown to be a significant negative predictor of student success in mathematics and availability of advanced math courses (Walker, 2006);

- **TOWN** – a dummy coded variable reflecting schools located in a town of less than 100,000 people or a rural community of less than 50,000 (coded “1”). All other schools were coded “0” and schools in suburban districts were used as the reference group. George (2000) observed that students attending suburban schools general had a more positive attitude toward math and science and generally outperformed their urban and rural peers.

- **SUBURB** – a dummy-coded variable reflecting schools located in a suburban school district (coded “1”). All other schools were coded “0.” This variable was used as the reference group for urbanicity. Suburban districts are defined as those communities that directly border a larger urban core community.

- **%MINORITY** – a dummy-coded variable indicating the percent of a school’s population that belong to an ethnic minority group (coded as low minority “0” and high minority “1”). Schools with high minority student populations have been shown to significantly underperform schools with a low percentage of minority students (McBay & Davidson, 1993).
**HLM as applied to the research questions.** The bulk of this analysis focuses on the relationship between science course type and math achievement. Because of the nested nature of these data (students within schools), HLM was used to create six, two level hierarchical linear models. These models examined achievement by IC and non-IC schools while controlling for potential student and school-level confounding variables. A school-level weight was used at level 2; no level-1 weight was used because students were randomly selected within schools. School effects were modeled on the level-1 intercept with no estimation of cross-level interaction effects. This strategy was examined by calculating the intra-class correlation (ICC) using the Mathieu, Aguinis, Culpepper, and Chen (2012) power calculator post hoc to understand whether the model reached sufficient levels of statistical power to detect existing effects as recommended by Aguinis, Gottfredson, and Culpepper (2013).

The plausible values feature of HLM 7.0 was used, which prompts the program to run models for each of the five plausible values internally and produce their average value and correct standard errors. Within these models, grand mean centering was used on variables that lacked a meaningful intercept value. Grand mean centering reduces the co-variation between the random intercepts and slopes while providing a computational advantage over the raw metric approach. This reduction of co-variation helped alleviate potential level-2 estimation problems due to multi-collinearity (Hofmann & Gavin, 1998). A detailed explanation of the data analysis methods used by the HLM 7.0 software is available from Raudenbush and Bryk (2002). Finally, random intercepts models with fixed slopes were employed.

**Between-groups model: Comparing IC to non-IC students.** As described earlier, this study occurred in two phases. The first phase was the between-groups
analysis comparing IC students to non-IC students. The phase one between-groups level-1 and 2 null models are provided by the following:

\[ Y_{ij} = \beta_{0j} + \varepsilon_{ij}, \]
\[ \beta_{0j} = \gamma_{00} + u_{0j}, \]

where \( Y_{ij} \) takes the form of three dependent variables corresponding to the first three research questions. These dependent variables include: (1) the five plausible values related to NAEP math performance, (2) student math GPA, and (3) number of math Carnegie units earned. Additionally, \( \beta_{0j} \) represents the school means; \( u_{0j} \) the between-school variance; and \( \varepsilon_{ij} \) the within-school variance.

The first dependent variable, NAEP math performance, made use of five plausible value scores, which required specific analysis using HLM methods. The HLM 7.0 software takes the PVs into account by running each of the PVs internally and producing their average value and the correct standard errors (Raudenbush et al., 2011), hence the level 1 random intercept and random slope model (RIRSM) is as follows for each of five plausible values:

\[
\bar{PV}_{ij} = \beta_{0j} + \beta_{1j}(TRADPHYS)_{ij} + \beta_{2j}(CONCEPTPHYS)_{ij} + \sum_{q=2}^{Q} \beta_{qj} X_{qij} + r_{ij}
\]

where

\( \bar{PV}_{ij} \) is one of the five plausible values calculated from the NAEP 12th grade math assessment. All five plausible values are averaged and pooled into a single value;

\( \beta_{0j} \) is the aggregate mean of school \( j \)'s 12th grade NAEP math plausible values;

\( \beta_{1j} \) is the first independent variable of interest for the between-groups analysis being whether student \( i \) took full-year traditional physics in the 9th grade year at school \( j \);
\( \beta_{2j} \) is the second independent variable of interest for the between-groups analysis being whether student \( i \) was taught using a conceptually based IC physics course in the 9\(^{th} \) grade year at school \( j \);

\( \beta_{qj} (q = 3, \ldots, Q) \) are the aforementioned student level (level-1) coefficients;

\( X_{qij} \) is a vector of the aforementioned student level predictor \( q \) for student \( i \) in school \( j \);

\( r_{ij} \) is the student level random effect being the mean difference between student \( i \)'s 12\(^{th} \) grade NAEP plausible values and school \( j \)'s mean NAEP plausible values; and

\( \sigma^2 \) is the variance of \( r_{ij} \), that is the student level variance.

The level-2 model is as follows:

\[
\beta_{qj} = \gamma_{q0} + \sum_{s=1}^{S_q} \gamma_{qs} W_{sj}
\]

where

\( \beta_{qj} \) are the Level-1 coefficients defined by the level-1 model;

\( \gamma_{qs} (q = 0,1, \ldots, S_q) \) are the aforementioned level-2 coefficients (\( \gamma_{00} \) being the grand mean);

\( W_{sj} \) are the level-2 predictors; and

The second dependent variable, student algebra GPA is specified by the level-1 model as follows:

\[
GPA_{ij} = \beta_{0j} + \beta_{1j}(\text{TRADPHYS})_{ij} + \beta_{2j}(\text{CONCEPTPHYS})_{ij} + \sum_{q=2}^{Q} \beta_{qj} X_{qij} + r_{ij}
\]

where

\( GPA_{ij} \) is overall high school algebra GPA for student \( i \) at school \( j \);

\( \beta_{0j} \) is the aggregate mean of school \( j \)'s algebra GPA;
\( \beta_{ij} \) is the first independent variable of interest for the between-groups analysis being whether student \( i \) took full-year traditional physics in the 9\(^{\text{th}}\) grade year at school \( j \);

\( \beta_{2j} \) is the second independent variable of interest for the between-groups analysis being whether student \( i \) was taught using a conceptually-based IC physics course in the 9\(^{\text{th}}\) grade year at school \( j \);

\( \beta_{qj} \) (\( q = 3, \ldots, Q \)) are the aforementioned student level (level-1) coefficients;

\( X_{qij} \) is a vector of the aforementioned student level predictor \( q \) for student \( i \) in school \( j \);

\( r_{ij} \) is the student level random effect being the mean difference between student \( i \)’s high school algebra GPA and school \( j \)’s mean algebra GPA; and

\( \sigma^2 \) is the variance of \( r_{ij} \) that is the student level variance.

The level-2 model is as follows:

\[
\beta_{qj} = \gamma_{q0} + \sum_{s=1}^{S_q} \gamma_{qs} W_{sj}
\]

where

\( \beta_{qj} \) are the Level-1 coefficients defined by the level-1 model;

\( \gamma_{qs} \) (\( q = 0, 1, \ldots, S_q \)) are the aforementioned level-2 coefficients (\( \gamma_{00} \) being the grand mean);

\( W_{sj} \) are the level-2 predictors; and
The third dependent variable, number of math Carnegie units is specified by the level 1 Poisson model as follows:

\[ E(UNITS_{ij}|\beta_j) = \lambda_{ij} \]

\[ \log[\lambda_{ij}] = \eta_{ij} \]

\[ \eta_{ij} = \beta_{0j} + \beta_{1j}(TRADPHYS_{ij}) + \beta_{2j}(CONCEPTPHYS_{ij}) + \sum_{q=2}^{Q} \beta_{qj} X_{qij} + r_{ij} \]

where

- \( UNITS_{ij} \) is the number of math Carnegie units earned by student \( i \) at school \( j \);
- \( \beta_{0j} \) is the mean number of math Carnegie units earned at school \( j \);
- \( \beta_{1j} \) is the first independent variable of interest for the between-groups analysis being whether student \( i \) took full-year traditional physics in the 9th grade year at school \( j \);
- \( \beta_{2j} \) is the second independent variable of interest for the between-groups analysis being whether student \( i \) was taught using a conceptually-based IC physics course in the 9th grade year at school \( j \);
- \( \beta_{qj} \) (\( q = 3, \ldots, Q \)) are the aforementioned student level (level-1) coefficients except for the variable describing the highest math course taken;
- \( X_{qij} \) is a vector of the aforementioned student level predictor \( q \) for student \( i \) in school \( j \);
- \( r_{ij} \) is the student level random effect being the mean difference between student \( i \)'s math Carnegie units and school \( j \)'s mean math Carnegie units; and
- \( \sigma^2 \) is the variance of \( r_{ij} \) that is the student level variance.
The level-2 model is as follows:

\[ \beta_{qj} = \gamma_{q0} + \sum_{s=1}^{S_q} \gamma_{qs} W_{sj} \]

where

- \( \beta_{qj} \) are the Level-1 coefficients defined by the level-1 model;
- \( \gamma_{qs} (q = 0, 1, \ldots, S_q) \) are the aforementioned level-2 coefficients (\( \gamma_{00} \) being the grand mean);
- \( W_{sj} \) are the level-2 predictors; and

**Within-group model:** Comparing IC students to each other. The second phase of this study is the within-groups analysis comparing IC students to one another. Instead of comparing traditional physics and biology 9th grade students, these analyses compared 9th grade students who took traditional, algebra-based full-year physics and full-year concept-based physics as the sample frame using half-year traditional physics as the reference group. The phase two within-groups analysis estimates are provided by the following general models:

\[ Y_{ij} = \alpha_j + \varepsilon_{ij}, \]
\[ \alpha_j = \gamma_{00} + \mu_{0j}, \]

where \( Y_{ij} \) takes the form of the five plausible values related to NAEP math performance corresponding to the first research question. Additionally, \( \alpha_j \) represents the school means; \( \mu_{0j} \) the between school variance; and \( \varepsilon_{ij} \) the within-school variance.
The specific level 1 model is as follows:

$$P_{ij} = \beta_{0j} + \beta_{1j}(\text{TRADPHYS})_{ij} + \beta_{2j}(\text{CONCEPTPHYS})_{ij} + \sum_{q=3}^{Q} \beta_{qj} X_{qij} + r_{ij}$$

where $P_{ij}$ is one of the five plausible values calculated from the NAEP 12th grade math assessment. All five plausible values will be averaged and pooled into a single value;

$\beta_{0j}$ is the aggregate mean of school $j$’s 12th grade NAEP math plausible values;

$\beta_{1j}$ is the first independent variable of interest for the between-groups analysis being whether student $i$ took full-year traditional physics in the 9th grade year at school $j$;

$\beta_{2j}$ is the second independent variable of interest for the within-groups analysis being whether student $i$ was taught using a full-year concept-based IC physics curriculum in the 9th grade year at school $j$;

$\beta_{qj}$ ($q = 3, 4, \ldots, Q$) are the aforementioned student level (level-1) coefficients;

$X_{qij}$ is a vector of the aforementioned student level predictor $q$ for student $i$ in school $j$;

$r_{ij}$ is the student level random effect being the mean difference between student $i$’s 12th grade NAEP plausible values and school $j$’s mean NAEP plausible values;

and

$\sigma^2$ is the variance of $r_{ij}$ that is the student level variance.
The level-2 model is as follows:

$$\beta_{qj} = \gamma_{q0} + \sum_{s=1}^{S_q} \gamma_{qs} W_{sj}$$

where

- $\beta_{qj}$ are the Level-1 coefficients defined by the level-1 model;
- $\gamma_{qs}$ ($q = 0, 1, \ldots, S_q$) are the aforementioned level-2 coefficients ($\gamma_{00}$ being the grand mean);
- $W_{sj}$ are the level-2 predictors; and

The second dependent variable, student algebra GPA is specified by the level-1 model as follows:

$$GPA_{ij} = \beta_{0j} + \beta_{1j}(TRADPHYS)_{ij} + \beta_{2j}(CONCEPTPHYS)_{ij} + \sum_{q=2}^{Q} \beta_{qj} X_{qij} + r_{ij}$$

where

- $GPA_{ij}$ is overall high school algebra GPA for student $i$ at school $j$;
- $\beta_{0j}$ is the aggregate mean of school $j$’s algebra GPA;
- $\beta_{1j}$ is the first independent variable of interest for the between-groups analysis being whether student $i$ took full-year traditional physics in the 9th grade year at school $j$;
- $\beta_{2j}$ is the second independent variable of interest for the within-groups analysis being whether student $i$ was taught using a full-year concept-based IC physics curriculum in the 9th grade year at school $j$;
- $\beta_{qj}$ ($q = 2, \ldots, Q$) are the aforementioned student level (level-1) coefficients;
- $X_{qij}$ is a vector of the aforementioned student level predictor $q$ for student $i$ in school $j$;
\( r_{ij} \) is the student level random effect being the mean difference between student 
\( i \)'s high school algebra GPA and school \( j \)'s mean algebra GPA; and
\( \sigma^2 \) is the variance of \( r_{ij} \) that is the student level variance.

The level-2 model is as follows:

\[
\beta_{qj} = \gamma_{q0} + \sum_{s=1}^{S_q} \gamma_{qs} W_{sj}
\]

where

\( \beta_{qj} \) are the Level-1 coefficients defined by the level-1 model;
\( \gamma_{qs} (q = 0, 1, \ldots, S_q) \) are the aforementioned level-2 coefficients (\( \gamma_{00} \) being the grand mean);
\( W_{sj} \) are the level-2 predictors; and

The third dependent variable, number of advanced math Carnegie units is specified by the level 1 Poisson model as follows:

\[
E(UNITS_{ij} | \beta_j) = \lambda_{ij} \\
\log[\lambda_{ij}] = \eta_{ij} \\
\eta_{ij} = \beta_{0j} + \beta_{1j}(TRADPHYS_{ij}) + \beta_{2j}(CONCEPTPHYS_{ij}) + \sum_{q=2}^{Q} \beta_{qj} X_{qij} + r_{ij}
\]

where

\( UNITS_{ij} \) is the number of math Carnegie units earned by student \( i \) at school \( j \);
\( \beta_{0j} \) is the mean number of math Carnegie units earned at school \( j \);
\( \beta_{1j} \) is the first independent variable of interest for the between-groups analysis being whether student \( i \) took full-year traditional physics in the 9th grade year at school \( j \);
\( \beta_{2j} \) is the second independent variable of interest for the within-groups analysis being whether student \( i \) was taught using a concept-based IC physics curriculum in the 9th grade year at school \( j \);
\[ \beta_{qj} (q = 2,3, \ldots, Q) \] are the aforementioned student level (level-1) coefficients except for the highest math course taken;

\( X_{qij} \) is a vector of the aforementioned student level predictor \( q \) for student \( i \) in school \( j \);

\( r_{ij} \) is the student level random effect being the mean difference between student \( i \)'s math Carnegie units and school \( j \)'s mean math Carnegie units; and \( \sigma^2 \) is the variance of \( r_{ij} \) that is the student level variance.

The level-2 model is as follows:

\[ \beta_{qj} = \gamma_{q0} + \sum_{s=1}^{S_q} \gamma_{qs} W_{sj} \]

where

\( \beta_{qj} \) are the Level-1 coefficients defined by the level-1 model;

\( \gamma_{qs} (q = 0,1, \ldots, S_q) \) are the aforementioned level-2 coefficients (\( \gamma_{00} \) being the grand mean);

\( W_{sj} \) are the level-2 predictors; and

All models in this study assume that for each unit \( j \), the vector \((u_{0j}, u_{1j}, \ldots, u_{Qj})'\) is distributed as multivariate normal, with each element of \( u_{qj} \) having a mean of zero and variance of

\[ \text{Var}(u_{qj}) = \tau_{qq}. \]

The goal of using HLM was to generate regression coefficients that were sensitive to the clustered nature of the data and the aforementioned research questions. The significant student-level coefficients from the independent variables of interest were also inserted into the following Hedge’s \( g \) equation to calculate pooled effect size:
where

\[ g = \frac{\gamma}{\sqrt{(n_1 - 1)S_1^2 + (n_2 - 1)S_2^2 \over (n_1 + n_2 - 2)}} \]

\( \gamma \) is the HLM coefficient for the intervention’s effect, which represents the group mean difference adjusted for both level-1 and level-2 covariates;

\( n_1 \) and \( n_2 \) are the student sample sizes;

\( S_1 \) and \( S_2 \) are the student-level unadjusted standard deviations for the intervention group and the comparison group respectively.

The aforementioned Hedge’s \( g \) is the unstandardized regression coefficient of the two compared groups standardized by the unbiased estimate of the population standard deviation. This effect size model is recommended to compute student-level effect sizes when it is reasonable to assume that cluster size was unrelated to cluster means (US Department of Education, 2008).

**Missing Data**

The complex sparse-item, spiraling sampling design used for NAEP can produce a large amount of missing data on several variables (Sikali, 2011). This is because examinees completed data on the blocks of items that were administered but have missing values on the blocks that were not administered (Peugh & Enders, 2004). Given that the missing item blocks are unrelated to student achievement and other measured variables, the missing data becomes a purposive by-product of the data collection procedure and can therefore be considered missing completely at random (MCAR) (Graham, Hofer, & MacKinnon, 1996; Graham, Taylor, & Cumsille, 2001; Kaplan, 1995).
Since it is justified in the extant literature to classify NAEP missing data as MCAR, this study utilized the Bayesian multiple imputation (MI) algorithm as postulated by Schafer (1997). Additionally, Schafer (1997) suggested that nominal and ordinal variables can be used in the MI process to produce unbiased parameter estimates when data are MCAR.

Specific to this this study, the sample frame \( n = 6,544 \) exhibited a large amount of missing data on the aforementioned independent variables (see Table 4). Given that on some variables up to 89% of data were missing, it was necessary to run the MI procedure on the full NAEP HSTS 2009 data set \( n = 41,200 \) to maximize sample size, thereby increasing statistical power. After the MI was run, the full dataset was merged back into the smaller sample frame dataset (keyed on student ID) from which the aforementioned HLM analyses are conducted.

Table 4

<table>
<thead>
<tr>
<th>Variable</th>
<th>Number missing</th>
<th>% missing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Advance math course GPA</td>
<td>962</td>
<td>14.7</td>
</tr>
<tr>
<td>Math Carnegie units earned</td>
<td>1616</td>
<td>24.7</td>
</tr>
<tr>
<td>Science Carnegie units earned</td>
<td>890</td>
<td>13.6</td>
</tr>
<tr>
<td>Highest SAT math score</td>
<td>5863</td>
<td>89.6</td>
</tr>
<tr>
<td>School enrollment</td>
<td>654</td>
<td>10.0</td>
</tr>
<tr>
<td>Parental education</td>
<td>255</td>
<td>3.9</td>
</tr>
<tr>
<td>School type</td>
<td>458</td>
<td>7.0</td>
</tr>
</tbody>
</table>

In summary, this study used NAEP HSTS 2009 data to examine the effect of 9th grade physics course-taking on overall high school math performance using a series of six, 2-level hierarchical models. Specifically, the dependent variables being examined were (1) 12th grade NAEP math performance, (2) high school algebra GPA, and (3) the number of advanced math Carnegie credits earned in high school. The independent variables used were whether a student takes physics as a 9th grade
student and if so, what type of physics the student took. The final sample frame \((n = 6,554)\) was derived from the total NAEP HSTS 2009 sample \((n = 41,200)\) using a series of three filters. First, data were filtered to ensure only 9th grade students were included in the sample. Second, the data were filtered to ensure that each science course taken by the 9th grade student fit into an easily identifiable course description. Third, these data were filtered to exclude honors and advanced placement courses \((n = 22,110)\). The filtered sample was randomly split to create approximately equal groups of 9th grade students in biology \((n = 2,136)\), physical science \((n = 2,366)\), and physics \((n = 2,082)\). Missing data were imputed using Bayesian multiple imputation.

One-way between-groups analysis of variance (ANOVA) with planned comparisons between the split samples demonstrated no significant difference between the samples on 12th grade NAEP math performance \((p = 0.21)\) and pre-SAT performance \((p = 0.14)\). These results indicated that the variance between the split groups was minimal and that each group was homogenous compared to the others.

Finally, one-way between-groups analysis of covariance (ANCOVA) was performed to compare the performance level of biology, and physics students on the pre-SAT assessment indicated that there was no significant difference between the biology and physics groups on the pre-SAT assessment. These results support the assumption that 9th grade biology students are similar in academic performance to 9th grade physics students and are therefore safe to use as comparison groups for subsequent analyses in this sample.

The purpose of these data cleaning strategies was to create a reliable and manageable data set that is nationally representative of US high school student science course taking patterns and the effect of those course-taking patterns on algebra performance.
CHAPTER 4

RESULTS

The methods described in the previous chapter were designed to examine the efficacy of inverted curriculum (IC) models on student algebra performance over the course of their high school careers. This chapter presents the results of using the aforementioned methods to answer whether IC models are indeed effective in improving student algebra performance. Specifically, 2009 NAEP HSTS data were examined in relation to the following six research questions:

**Between-groups Research Questions: Comparing IC to Non-IC Students**

1) Is there a significant difference in performance on the 12th grade NAEP mathematics assessment between students who receive early physics instruction (IC) in 9th grade and those who do not (non-IC)?

2) Is there a significant difference in cumulative high school algebra GPA between IC and non-IC students as documented on the student’s 12th grade transcript by the HSTS?

3) Is there a significant difference in the number of advanced math credits earned by IC and non-IC students as documented on the student’s 12th grade transcript by the HSTS?

**Within-groups Research Questions: Comparing IC Students to Each Other**

7) Among IC students, is there a significant relationship between the length and type of physics class taken and 12th grade NAEP math assessment performance?

8) Among IC students, is there a significant relationship between the length and type of physics class taken and the overall algebra GPA earned?
9) Among IC students, is there a significant relationship between the length and type of physics taken and the number of high school math Carnegie credits earned?

Given the hierarchical structure of the HSTS 2009 dataset (i.e. students nested within schools), hierarchical linear modelling was conducted using HLM 7.0. The random intercepts model with fixed slopes was employed (see Tables 6 - 10). Dichotomous variables were retained in their original metric. Ordinal and continuous variables were centered on the grand mean. Sampling weights for schools (level 2) were employed in HLM analyses to make the sample reflective of the population. For all analyses, the solutions were generated on the basis of full information maximum likelihood estimation.

The model building followed a step-up strategy as suggested by Raudenbush and Bryk (2000). At the first stage, a fully unconditional model (or null or intercept-only model), containing only an outcome variable and no independent variables, was built. The intercept-only model is equivalent to a one-way random effects ANOVA (Raudenbush et al., 2011). The intercept-only model was used to identify the source of variation within the following outcome measures: (a) 12th grade NAEP math performance, (b) high school algebra GPA, and (c) high school advanced math Carnegie units earned. This was accomplished by partitioning the total variance in the outcome measures into their within-school (level 1) and between-school (level 2) components. At the second-stage, student demographics and course-taking variables were added to the fully unconditional models to examine the statistical significance of student demographic and course-taking predictors. At the final stage, the statistical significance of school predictors was examined. The proportion of reduction in variance as accounted for by the models served as a basis for making a
judgment about the relative importance of student and school level variables (Raudenbush et al., 2011). As recommended by Hox (2002), a variable was considered to have a small effect if it explained 1% variance, a medium effect if it explained 10% variance, and a large effect if it explained 25% variance.

Additionally, a Hedges g student-level effect size was calculated to determine the magnitude of the effect of the use of various IC models and a non-IC model on the aforementioned dependent variables. These analyses used the threshold of a 0.25 effect size as a “substantively important” effect per the guidance of the What Works Clearinghouse Procedures and Standards Handbook (US Department of Education, 2008).

**Between Groups Analysis Results**

Table 5 presents the analysis addressing the first research question: Is there a significant difference in performance on the 12th grade NAEP mathematics assessment between students who receive physics instruction in ninth grade (IC) and those who do not (non-IC)?
## Table 5

**Question 1: 12th Grade NAEP Math Achievement Between IC and Non-IC Students**

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Empty Coeff. (se)</th>
<th>Model 1 Coeff.(se)</th>
<th>Model 2 Coeff. (se)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercepts</td>
<td>152.19 (1.62)***</td>
<td>153.33 (1.70)***</td>
<td>153.29 (1.29)***</td>
</tr>
<tr>
<td>Student-level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELL status</td>
<td>-16.68 (3.47)***</td>
<td>-16.42 (2.28)***</td>
<td></td>
</tr>
<tr>
<td>(STDELL)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School lunch prog.</td>
<td>-6.52 (2.90)**</td>
<td>-6.12 (1.73)**</td>
<td></td>
</tr>
<tr>
<td>(FRL)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-SAT score</td>
<td>0.53 (1.01)</td>
<td>0.52 (0.55)</td>
<td></td>
</tr>
<tr>
<td>(PSAT)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race (RACE)</td>
<td>-2.75 (2.98)</td>
<td>-2.41 (2.03)</td>
<td></td>
</tr>
<tr>
<td>Gender (SEX)</td>
<td>-6.21 (2.07)**</td>
<td>-6.30 (1.53)**</td>
<td></td>
</tr>
<tr>
<td>Parent ed. level</td>
<td>0.58 (1.57)</td>
<td>-0.49 (0.98)</td>
<td></td>
</tr>
<tr>
<td>(EDU)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest math taken</td>
<td>17.77 (1.20)***</td>
<td>17.71 (0.84)***</td>
<td></td>
</tr>
<tr>
<td>(HIGHMATH)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IC traditional physics</td>
<td>10.06 (5.01)**</td>
<td>10.17 (5.06)***</td>
<td></td>
</tr>
<tr>
<td>(TRADPHYS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>compared to non-IC biology (TRADBIO)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IC conceptual physics</td>
<td>-17.4 (5.19)**</td>
<td>-16.25 (5.28)**</td>
<td></td>
</tr>
<tr>
<td>(CONCEPTPHYS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School-level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School type</td>
<td>-2.90 (7.43)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(STYPE)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrollment size</td>
<td>1.47 (3.23)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ENROLL)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban setting</td>
<td>1.90 (2.67)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>compared to suburban (URBAN)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Town and rural setting compared to suburban (TOWN)</td>
<td></td>
<td>-1.91 (2.67)</td>
<td></td>
</tr>
<tr>
<td>% Minority of school</td>
<td>-8.70 (2.60)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(MINORITY)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Random Effects

- $u_{00}$: 409.57
- $\sigma^2$: 492.97
- $\chi^2$: 2542.6***
- ICC: 0.45
- PVAF Level 1: -
- PVAF Level 2: -

* $p < .05$; ** $p < .01$; *** $p < .001$
Referring to Table 5, the independent variables used in the model indicated that students who took IC physics using a traditional physics curriculum, when compared to non-IC biology students, outperformed their non-IC biology peers on the 12th grade NAEP math assessment by 10.06 points ($\beta = 10.06, p < .05$) and yielded $g = 0.31$. In contrast, when conceptual IC physics students were compared to non-IC biology students, the conceptual IC students scored 16.25 points below their non-IC biology peers ($\beta = -16.25, p < .01$) and yielded $g = -0.55$ on the 12th grade NAEP math assessment. As a reminder, concept-based curricula such as “Conceptual Physics” (Hewitt, 2002) and “Active Physics” (Eisenkraft, 1998) place greater emphasis on the conceptual understanding of physics and minimize the mathematical calculations underlying the concepts.

The positive impact of traditional IC physics curricula when compared to non-IC biology students was statistically significant and substantively important ($< 0.25$) in effect size. Based on these analyses, using a traditional IC physics model appeared to improve NAEP math performance. Importantly, these significant positive impacts held true when accounting for school-level differences in the full model indicating that they were not a result of school differences but were indeed related to the courses taken by students. This finding was corroborated by the fact that the student-level model alone explained a large degree of variance in 12th grade NAEP math performance by describing 34% of the variance in NAEP math score at level 1. School factors explained much less variance (7%) at level 2 indicating that more of the variance in 12th grade NAEP math performance was explained by student variables. This finding means that more than a third of the variance in 12th grade math NAEP performance can be explained by student characteristics (i.e. pre-
SAT performance, parent education levels, etc.) with a highly significant variable being if the student took a traditional IC physics course in 9th grade.

Significant covariates that affected a student’s 12th grade NAEP math scores were (a) qualifying for English language learning (ELL) services ($\beta = -16.42$, $p < .001$), (b) qualifying for free and reduced lunch (FRL) services ($\beta = -6.12$, $p < .01$), (c) gender, with females scoring on average 6.3 points below males ($\beta = -6.30$, $p < .01$), and (d) the highest level of math taken where for every math course taken above geometry, 12th grade NAEP math scores increases 17.71 points ($\beta = 17.71$, $p < .001$). School-based factors that significantly affected a student’s assessment scores was a school’s percent minority population ($\beta = -8.69$, $p < .01$).

Table 6 presents the results related to the second research question: Is there a significant difference in cumulative high school algebra GPA between IC and non-IC students as documented on the student’s 12th grade transcript by the HSTS?
Table 6

**Question 2: Cumulative High School Algebra GPA Between IC and Non-IC Students**

<table>
<thead>
<tr>
<th></th>
<th>Empty Coeff. (se)</th>
<th>Model 1 Coeff. (se)</th>
<th>Model 2 Coeff. (se)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Fixed Effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>2.65 (0.05)***</td>
<td>2.70 (0.05)***</td>
<td>2.70 (0.05)***</td>
</tr>
<tr>
<td>Student-level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELL status (STDELL)</td>
<td>-0.06 (0.14)</td>
<td>-0.04 (0.15)</td>
<td></td>
</tr>
<tr>
<td>School lunch prog. (FRL)</td>
<td>-0.15 (0.08)</td>
<td>-0.12 (0.08)</td>
<td></td>
</tr>
<tr>
<td>Pre-SAT score (PSAT)</td>
<td>0.10 (0.05)*</td>
<td>0.10 (0.05)</td>
<td></td>
</tr>
<tr>
<td>Race (RACE)</td>
<td>-0.034 (0.06)</td>
<td>0.00 (0.09)</td>
<td></td>
</tr>
<tr>
<td>Gender (SEX)</td>
<td>0.08 (0.07)</td>
<td>0.07 (0.07)</td>
<td></td>
</tr>
<tr>
<td>Parent ed. level (EDU)</td>
<td>-0.07 (0.05)*</td>
<td>0.08 (0.05)</td>
<td></td>
</tr>
<tr>
<td>Highest math taken (HIGHMATH)</td>
<td>0.04 (0.06)</td>
<td>0.03 (0.06)</td>
<td></td>
</tr>
<tr>
<td>IC traditional physics (TRADPHYS)</td>
<td>0.32 (0.14)*</td>
<td>0.35 (0.14)***</td>
<td></td>
</tr>
<tr>
<td>IC conceptual physics (CONCEPTPHYS)</td>
<td>-0.23 (.33)</td>
<td>-0.12 (.31)</td>
<td></td>
</tr>
<tr>
<td><strong>School-level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School type (STYPE)</td>
<td>-0.35 (0.19)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrollment size (ENROLL)</td>
<td>-0.01 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban setting</td>
<td>0.16 (0.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>compared to suburban (URBAN)</td>
<td>-0.16 (0.11)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Town setting</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>compared to suburban (TOWN)</td>
<td>-0.32 (0.09)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Minority of school (MINORITY)</td>
<td>-0.32 (0.09)**</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random Effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_{00}$</td>
<td>0.23</td>
<td>0.23</td>
<td>0.21</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>1.32</td>
<td>1.31</td>
<td>1.3</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>582.04***</td>
<td>647.28***</td>
<td>551.81***</td>
</tr>
<tr>
<td>ICC</td>
<td>0.15</td>
<td>0.15</td>
<td>0.14</td>
</tr>
<tr>
<td>PVAF Level 1</td>
<td>-</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>PVAF Level 2</td>
<td>-</td>
<td>-</td>
<td>0.09</td>
</tr>
</tbody>
</table>

* $p < .05$; ** $p < .01$; *** $p < .001$
Referring to Table 6, the independent variables used in the model indicated that students who took IC physics using a traditional physics curriculum earned algebra GPAs 0.32 points higher than their non-IC biology peers ($\beta = 0.32, p < .01$) and yielded $g = 0.28$. There was no significant difference observed between conceptual IC physics students and non-IC biology students related to algebra GPA. The models explained a very small degree of variance in GPA by describing 1% of the variance at level 1 in a student’s high school algebra GPA and 9% of the variance at level 2. This finding means that a traditional IC model is statistically and substantively effective in improving student high school algebra performance but that the moderate amount of variance explained by this model indicates that there are other important variables affecting a student’s overall high school algebra GPA outside of the variables used in this model. Additionally, that more variance was explained by the level 2 model indicates that variance in student GPA is more thoroughly explained by school characteristics rather than student characteristics.

The significant level-1 covariates were (a) qualifying for free and reduced lunch (FRL) services ($\beta = -0.15, p < .05$, level 1 only); (b) Pre-SAT score ($\beta = 0.10, p < .05$), parent education level ($\beta = 0.07, p < .001$), attending an urban school ($p < .05$); and (c) attending a school with a high percentage of ELL students ($p < .05$). The significant level-2 covariates were school type, with private school students earning algebra GPAs 0.35 points higher than their public school peers ($\beta = -0.35, p < .05$), and the school’s percent minority population ($\beta = -0.32, p < .001$).

Table 7 presents the analysis of the data used to address the third research question: Is there a significant difference in the number of advanced math Carnegie units earned by IC and non-IC students as documented on the student’s 12th grade transcript by the HSTS? Because Carnegie units are considered count data, the
Poisson regression option using continuous exposure was used in HLM 7.0 to minimize sampling error (McCullagh & Nelder, 1989).
Table 7

Question 3: Cumulative Advanced Math Carnegie Units Earned Between IC and Non-IC Students

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Empty Coeff. (se)</th>
<th>Model 1 Coeff.(se)</th>
<th>Model 2 Coeff. (se)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>23.77 (0.01)***</td>
<td>23.44 (0.011)***</td>
<td>23.32 (0.01)***</td>
</tr>
<tr>
<td>Student-level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELL status (STDELL)</td>
<td>1.35 (0.04)**</td>
<td>1.25 (0.04)</td>
<td></td>
</tr>
<tr>
<td>School lunch prog. (FRL)</td>
<td>1.01 (0.02)</td>
<td>0.94 (0.02)</td>
<td></td>
</tr>
<tr>
<td>Pre-SAT score (PSAT)</td>
<td>1.02 (0.01)</td>
<td>1.07 (0.01)</td>
<td></td>
</tr>
<tr>
<td>Race (RACE)</td>
<td>0.98 (0.02)</td>
<td>0.98 (0.02)</td>
<td></td>
</tr>
<tr>
<td>Gender (SEX)</td>
<td>1.06 (0.01)</td>
<td>1.12 (0.01)</td>
<td></td>
</tr>
<tr>
<td>Parent ed. level (EDU)</td>
<td>1.27 (0.01)**</td>
<td>1.07 (0.01)**</td>
<td></td>
</tr>
<tr>
<td>IC traditional physics (TRADPHYS) compared to non-IC biology (TRADBIO)</td>
<td>1.02 (0.08)</td>
<td>1.05 (0.084)</td>
<td></td>
</tr>
<tr>
<td>IC conceptual physics (CONCEPTPHYS) compared to TRADBIO</td>
<td>1.28 (0.08)</td>
<td>1.32 (0.09)</td>
<td></td>
</tr>
<tr>
<td>School-level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School type (STYPE)</td>
<td>1.13 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrollment size (ENROLL)</td>
<td>1.05 (0.09)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban setting (URBAN)</td>
<td>1.08 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Town/rural setting (TOWN)</td>
<td>0.93 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Minority of school (%MINORITY)</td>
<td>1.03 (0.063)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Random Effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_{000}$</td>
<td>0.003</td>
<td>0.0029</td>
<td>0.003</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>-</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>197.58</td>
<td>186.17</td>
<td>182.6</td>
</tr>
<tr>
<td>ICC</td>
<td>0.0009</td>
<td>0.0009</td>
<td>0.0009</td>
</tr>
</tbody>
</table>

* p < .05; ** p < .01; *** p < .001
Referring to Table 7, the independent variables used in the model indicated no statistically significant difference between both types of IC students (traditional or conceptual) and their non-IC peers in total advanced math Carnegie credits earned. This finding means that IC models have no significant effect in improving the number of advanced math Carnegie credits earned over the course of a student’s high school career.

The significant level-1 covariates were a student’s ELL status ($\beta = 1.35, p < .01$) and parent education level ($\beta = 1.07, p < .001$). There were no significant level-2 covariates.

**Within-Groups Analysis Results**

The remaining tables present analyses of those data that pertain to the within-groups analysis comparing 12th grade students who took a half-year of physics in 9th grade to those who took a full year of physics 9th grade. Table 8 presents the analysis of the data used to answer the fourth research question: Among IC students, are there significant differences in 12th grade NAEP math assessment performance by half-year physics students as compared to full-year traditional and conceptual IC physics students?
Table 8

**Question 4: 12th Grade NAEP Math Achievement Between Half-Year and Full-Year 9th Grade Physics Students and Traditional IC Physics and Conceptual IC Physics**

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Empty (Coeff. (se))</th>
<th>Model 1 (Coeff. (se))</th>
<th>Model 2 (Coeff. (se))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>144.89 (1.63)***</td>
<td>147.83 (1.41)***</td>
<td>147.79 (1.39)***</td>
</tr>
<tr>
<td><strong>Student-level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELL status</td>
<td>-6.22 (4.68)</td>
<td>-6.04 (4.67)</td>
<td></td>
</tr>
<tr>
<td>(STDELL)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School lunch prog.</td>
<td>-2.49 (1.74)</td>
<td>-2.45 (1.74)</td>
<td></td>
</tr>
<tr>
<td>(FRL)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-SAT score</td>
<td>-0.01 (1.74)</td>
<td>-0.01 (0.61)</td>
<td></td>
</tr>
<tr>
<td>(PSAT)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race (RACE)</td>
<td>-1.98 (0.83)*</td>
<td>-1.95 (0.83)*</td>
<td></td>
</tr>
<tr>
<td>Gender (SEX)</td>
<td>-3.69 (1.09)**</td>
<td>-3.69 (1.09)**</td>
<td></td>
</tr>
<tr>
<td>Parent ed. level</td>
<td>0.16 (0.83)</td>
<td>0.14 (0.83)</td>
<td></td>
</tr>
<tr>
<td>(EDU)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Highest math taken</td>
<td>15.84 (0.83)***</td>
<td>15.82 (0.83)***</td>
<td></td>
</tr>
<tr>
<td>(HIGHMATH)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full year traditional IC physics</td>
<td>24.02 (10.18)**</td>
<td>23.68 (9.89)**</td>
<td></td>
</tr>
<tr>
<td>(TRADPHYS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>compared to half-year IC physics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(HALFPHYS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full year conceptual IC physics</td>
<td>7.57 (11.14)</td>
<td>7.41 (10.81)</td>
<td></td>
</tr>
<tr>
<td>(CONCEPTPHYS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>compared to HALFPHYS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONCEPTPHYS</td>
<td>-27.47 (7.95)**</td>
<td>-26.41 (7.91)**</td>
<td></td>
</tr>
<tr>
<td>compared to TRADPHYS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School-level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School type</td>
<td>8.14 (4.13)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(STYPE)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrollment size</td>
<td>1.37 (1.08)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(ENROLL)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban setting (URBAN)</td>
<td>4.92 (2.51)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Town/rural setting</td>
<td>4.92 (2.56)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(TOWN)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Minority of school</td>
<td>-2.27 (3.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(%MINORITY)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random Effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(\mu_0)</td>
<td>539.43</td>
<td>319.54</td>
<td>300.63</td>
</tr>
<tr>
<td>(\sigma^2)</td>
<td>488.46</td>
<td>362.33</td>
<td>362.33</td>
</tr>
<tr>
<td>(\chi^2)</td>
<td>13527.97***</td>
<td>10314.13***</td>
<td>9157.98***</td>
</tr>
<tr>
<td>ICC</td>
<td>0.53</td>
<td>0.46</td>
<td>0.45</td>
</tr>
<tr>
<td>PVAF Level 1</td>
<td>-</td>
<td>0.26</td>
<td>0.26</td>
</tr>
<tr>
<td>PVAF Level 2</td>
<td>-</td>
<td>-</td>
<td>0.06</td>
</tr>
</tbody>
</table>

* \(p < .05\); ** \(p < .01\); *** \(p < .001\)
Referring to Table 8, the independent variables used in the model indicated that students who took full-year IC physics using a traditional physics curriculum outperformed their half-year physics peers by 23.68 points on the 12th grade NAEP math assessment ($\beta = 23.68, p < .001$) yielding $g = 0.82$. There was no significant difference between concept-based IC physics students and their half-year IC peers. Additionally, when traditional IC physics students were compared to conceptual IC physics students, traditional IC students outperformed the conceptual students by 26.41 points ($\beta = -26.41, p < .01$) and yielded $g = -0.81$.

The level-1 model explained a large amount of variance in NAEP scores by describing 26% of the variance in 12th grade NAEP math score. The level-2 model explained an additional 6% of the overall variance in NAEP performance.

The significant covariates in the level-1 model were (a) race, with minority students earning lower scores than white students by 1.95 points ($\beta = -1.95, p < .05$), (b) gender, with females earning lower scores than males by 3.7 points ($\beta = -3.69, p < .01$), and (c) highest math course taken, where for every math course a student takes beyond geometry, the NAEP math assessment scores increases 15.8 points ($\beta = 15.84, p < .001$). There were no significant level-2 covariates in the model.

Table 9 presents the results related to the fifth research question: Are there significant differences in cumulative algebra GPA earned by half-year physics students as compared to full-year traditional and conceptual IC physics students?
Table 9

Question 5: High School Algebra GPA Between Half-Year and Full-Year Physics Students and Between Traditional IC Physics and Conceptual IC Physics Students

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Empty</th>
<th>Model 1</th>
<th>Model 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>2.55 (0.4)***</td>
<td>2.57 (0.04)**</td>
<td>2.56 (0.04)**</td>
</tr>
<tr>
<td>Student-level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELL status</td>
<td>0.30 (0.17)</td>
<td>0.31 (0.17)</td>
<td></td>
</tr>
<tr>
<td>School lunch prog.</td>
<td>-.05 (0.05)</td>
<td>-0.05 (0.05)</td>
<td></td>
</tr>
<tr>
<td>Pre-SAT score</td>
<td>0.02 (0.03)</td>
<td>0.02 (0.03)</td>
<td></td>
</tr>
<tr>
<td>Race (RACE)</td>
<td>0.02 (0.02)</td>
<td>-0.02 (0.02)</td>
<td></td>
</tr>
<tr>
<td>Gender (SEX)</td>
<td>0.22 (0.04)***</td>
<td>0.22 (0.04)***</td>
<td></td>
</tr>
<tr>
<td>Highest math taken (HIGHMATH)</td>
<td>0.25 (0.03)***</td>
<td>0.25 (0.03)***</td>
<td></td>
</tr>
<tr>
<td>Full year traditional IC physics (TRADPHYS) compared to half-year IC physics (HALFPHYS)</td>
<td>0.52 (0.23)*</td>
<td>0.51 (.25)*</td>
<td></td>
</tr>
<tr>
<td>Full year conceptual IC physics (CONCEPTPHYS) compared to HALFPHYS</td>
<td>.37 (0.36)</td>
<td>0.35 (0.36)</td>
<td></td>
</tr>
<tr>
<td>CONCEPTPHYS compared to TRADPHYS</td>
<td>-0.50 (0.22)*</td>
<td>-0.48 (0.22)*</td>
<td></td>
</tr>
<tr>
<td>School-level</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School type (STYPE)</td>
<td>-0.23 (0.09)*</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrollment size (ENROLL)</td>
<td>0.01 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban setting (URBAN)</td>
<td>0.21 (0.17)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Town/rural setting (TOWN)</td>
<td></td>
<td>-0.24 (0.07)**</td>
<td></td>
</tr>
<tr>
<td>% Minority of school (%MINORITY)</td>
<td></td>
<td>-0.03 (0.07)</td>
<td></td>
</tr>
<tr>
<td>Random Effects</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_{00}$</td>
<td>0.32</td>
<td>0.32</td>
<td>0.31</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td>0.58</td>
<td>0.53</td>
<td>0.53</td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>5628.94**</td>
<td>5475.57***</td>
<td>5093.78**</td>
</tr>
<tr>
<td>ICC</td>
<td>0.36</td>
<td>0.38</td>
<td>0.37</td>
</tr>
<tr>
<td>PVAF Level 1</td>
<td>-</td>
<td>0.09</td>
<td>0.09</td>
</tr>
<tr>
<td>PVAF Level 2</td>
<td>-</td>
<td>-</td>
<td>0.03</td>
</tr>
</tbody>
</table>

* $p < .05$; ** $p < .01$; *** $p < .001$
Referring to Table 9, the independent variables used in the model indicated that students who took full-year IC physics using a traditional physics curriculum earned an algebra GPA 0.51 points higher than their half-year IC peers ($\beta = 0.51, p < .05$) with $g = 0.57$ and 0.48 points higher than their concept-based IC physics peers with $g = 0.85$ ($\beta = -0.48, p < .05, g = 0.85$). There was no significant difference between half-year IC students and concept-based IC students related to algebra GPA. The full model level-1 variables explained a small amount of variance in GPA by describing 9% of the level-1 variance and 3% of the level-2 variance. This finding means that full-year IC models are more effective than half-year IC models and full-year concept-based models in improving student high school algebra GPA but that there is still a large amount of unexplained variance resident in the current model.

Significant level-1 covariates included gender with females earning slightly high algebra GPAs than males ($\beta = 0.02, p < .001$) and highest math taken ($\beta = 0.51, p < .001$). The significant level-2 covariate was urbanicity, with town/rural students earning 0.24 point lower math GPAs than their suburban peers ($\beta = 0.24, p < .01$).

Table 10 presents the analysis of the data used to address the sixth research question: Are there significant differences in the number of advanced math Carnegie credits earned by half-year physics students as compared to full-year traditional and conceptual IC physics students? Once again, this dependent variable was measured in Carnegie units which were described by a Poisson distribution thereby requiring a Poisson regression using continuous exposure.
Table 10

*Question 6: Number of Advanced Math Carnegie Units Taken Between Half-Year and Full Year Physics Students*

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Empty Coeff. (se)</th>
<th>Model 1 Coeff. (se)</th>
<th>Model 2 Coeff. (se)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept</td>
<td>23.17 (0.01)***</td>
<td>23.04 (0.01)***</td>
<td>23.10 (0.01)***</td>
</tr>
<tr>
<td><strong>Student-level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ELL status</td>
<td>1.22 (0.04)</td>
<td>1.21 (0.04)*</td>
<td></td>
</tr>
<tr>
<td>(STDELL)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School lunch prog.</td>
<td>0.99 (0.01)</td>
<td>1.02 (0.04)</td>
<td></td>
</tr>
<tr>
<td>(FRL)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pre-SAT score</td>
<td>0.98 (0.01)</td>
<td>0.99 (0.05)</td>
<td></td>
</tr>
<tr>
<td>(PSAT)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Race (RACE)</td>
<td>1.03 (0.01)**</td>
<td>1.03 (0.01)**</td>
<td></td>
</tr>
<tr>
<td>Gender (SEX)</td>
<td>1.03 (0.01)</td>
<td>1.03 (0.01)</td>
<td></td>
</tr>
<tr>
<td>Parent ed. level</td>
<td>1.03 (0.01)*</td>
<td>1.03 (0.01)*</td>
<td></td>
</tr>
<tr>
<td>(EDU)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full year traditional IC physics</td>
<td>1.08 (0.04)*</td>
<td>1.09 (0.05)*</td>
<td></td>
</tr>
<tr>
<td>(TRADPHYS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>compared to half-year IC physics</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(HALFPHYS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Full year conceptual IC physics</td>
<td>0.93 (0.04)</td>
<td>0.93 (0.41)</td>
<td></td>
</tr>
<tr>
<td>(CONCEPTPHYS)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>compared to HALFPHYS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>CONCEPTPHYS compared to TRADPHYS</td>
<td>0.98 (0.04)*</td>
<td>0.98 (0.04)*</td>
<td></td>
</tr>
<tr>
<td><strong>School-level</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>School type</td>
<td>1.01 (0.03)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(STYPE)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enrollment size of school</td>
<td>1.02 (0.01)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urban setting (URBAN)</td>
<td>1.06 (0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Town/rural setting</td>
<td>0.95 (0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(TOWN)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>% Minority of school</td>
<td>1.05 (0.02)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(%MINORITY)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Random Effects</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$u_0$</td>
<td>0.02</td>
<td>0.02</td>
<td>0.02</td>
</tr>
<tr>
<td>$\sigma^2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\chi^2$</td>
<td>1340.53***</td>
<td>1302.27***</td>
<td>1264.54***</td>
</tr>
<tr>
<td>ICC</td>
<td>0.006</td>
<td>0.006</td>
<td>0.006</td>
</tr>
<tr>
<td>PVAF Level 1</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>PVAF Level 2</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
</tbody>
</table>

* $p < .05$; ** $p < .01$; *** $p < .001$
Referring to Table 10, the independent variables used in the model indicated that when compared to half-year traditional IC physics, full-year IC students earned 8% more advanced math Carnegie credits ($\beta = 1.08$, $p < .05$). There was no significant difference between full-year conceptual IC physics students and half-year IC students. However, full-year conceptual IC students earned 2% fewer advanced math Carnegie credits compared to full-year traditional IC physics students ($\beta = 0.98$, $p < .05$).

The significant covariates in the level-1 model were race ($\beta = 1.03$, $p < .01$) and parent education level ($\beta = 1.03$, $p < .05$). There were no significant level-2 covariates that affected the number of advanced math high school Carnegie credits earned by $12^{\text{th}}$ grade students who had taken half-year or full-year physics as a $9^{\text{th}}$ grade student.

These results mean that the amount of physics taken is a significant predictor of the number of advanced math Carnegie credits earned by a $12^{\text{th}}$ grade student who took a full or half-year physics course as a $9^{\text{th}}$ grade student. Additionally, the type of IC physics taken appeared to significantly affect the number of advanced math Carnegie credits earned.

In summary, IC models that use traditional physics instruction and curriculum for a full year appear to be effective in improving all but one of the dependent variables in this study. Related to $12^{\text{th}}$ grade NAEP math performance, students taught using a full-year, traditional physics curriculum in $9^{\text{th}}$ grade scored 10.4 points higher than non-IC students with $g = 0.31$, 23.7 points higher than half-year IC students with $g = 0.82$, and 26.4 points higher than conceptual IC students with $g = 0.81$. Importantly, the non-IC students outperformed the concept-based IC students by more than 16 points. Related to algebra GPA, students taught using a
full-year traditional IC physics curriculum earned algebra GPAs 0.32 points higher than non-IC students with \( g = 0.28 \). Additionally, full-year traditional IC physics students earned algebra GPAs 0.52 points higher than half-year IC students with \( g = 0.57 \). There was no significant difference between traditional IC physics student algebra GPA compared to conceptual IC physics students. Finally, related to advanced math Carnegie credits earned, students taught using a full-year traditional IC physics curriculum earned 2% more advanced math Carnegie credits than non-IC students and conceptual IC physics. Additionally these traditional IC students earned 8% more advanced math credits than half-year physics students.
CHAPTER 5
DISCUSSION

This dissertation examined a particular science and math curriculum reform idea known as inverted curriculum (IC). This study examined the relationship between high school math performance among IC and non-IC students using data from the 2009 NAEP High School Transcript Study (HSTS). This study was done in two phases. The first phase involved a between-groups analysis by comparing performance data and course-taking patterns between IC and non-IC students. The second phase involved a within-group validation study by comparing the disaggregated effects of physics course type on student math performance among IC students only. The goal of this study was to use nationally representative student data to understand the efficacy of IC models on general math performance, algebra performance, and the number of advanced math Carnegie units taken in high school.

IC models are most commonly adopted on an ad-hoc basis in a school-by-school fashion. Students within these schools may take physics at the same time as 9th grade algebra, thereby conforming to an IC model. This study utilized this informal arrangement to create a natural experiment in which student course-taking patterns were used to identify those students who were taught using an IC model. These students were used as the experimental group in the analysis of IC model efficacy on improving student math performance and increasing the number of advanced math courses taken by students compared to non-IC students. The control group was a sample of students who took biology in the 9th grade and exhibited no significant differences on PSAT score compared to the IC student sample.

This natural experiment tested two hypotheses. The first hypothesis was that a statistically significant difference would be observed in student performance and
course-taking patterns between students taking IC and those in non-IC curriculum with IC students exhibiting greater performance and taking more advanced math courses than non-IC students. The second hypothesis was that a statistically significant difference would be observed in student algebra performance and math course-taking patterns among IC students based on the amount of early physics (i.e. full year or half-year physics) taken, with students taking more physics earlier in high school showing greater performance and taking more advanced math courses than those taking less physics. These hypotheses were tested using the following six research questions:

1) Is there a significant difference in performance on the 12th grade NAEP mathematics assessment between students who receive early physics instruction (IC) in 9th grade and those who do not (non-IC)?

2) Is there a significant difference in cumulative high school algebra GPA between IC and non-IC students as documented on the student’s 12th grade transcript by the HSTS?

3) Is there a significant difference in the number of advanced math Carnegie units earned by IC and non-IC students as documented on the student’s 12th grade transcript by the HSTS?

4) Among IC students, is there a significant relationship between the length and type of physics class taken and 12th grade NAEP math assessment performance?

5) Among IC students, is there a significant relationship between the length and type of physics class taken and the overall algebra GPA earned?
6) Among IC students, is there a significant relationship between the length and type of physics taken and the number of high school math Carnegie units earned?

Math performance in general and algebra specifically are widely considered “gate-keeper” classes for post-secondary degree completion in 2-year and 4-year college (Helfand, 2006). Additionally, Trumper (2006) showed that if math curricular reforms are to have lasting impact, they must be especially effective in improving algebra performance. A rather radical reform idea in the form of inverted curricula (IC) has been proposed by physics educators with the goal of improving student general math performance, algebra performance, and student engagement in math course taking. This reform is described as inverted because it takes physics, generally taken at the end of a student’s high school career, and moves it to the front of the high school sequence as a 9th grade science course. The primary argument used in support of an IC model is that it gives context to a student’s algebra learning by co-locating physics and algebra in the high school curriculum. This co-location allows the physics teacher to provide a real-world context for the algebraic calculations that the student is simultaneously studying.

The IC model is a radical reform effort because it requires a complete re-structuring of the high school science sequence requiring an increased number of physics teachers be hired or other science teachers be trained in teaching physics to 9th grade students. Additionally, the requisite personnel changes require a large degree of professional development support for the physics teachers to be successful. In particular, veteran physics teachers need instructional support to effectively teach younger students who may lack the strong mathematical background that the teacher had become accustomed to in the traditional sequence. Finally, new physics teachers
in the IC system will need significant professional development support to become adept in physics content and best teaching practices using a conceptual physics model.

IC models may incur high costs to schools with uncertain benefits for students. Because of these significant uncertainties it is vitally important to understand whether nationally representative data supports the presupposed efficacy of IC models on improving student math ability and advanced math course taking. The remainder of this chapter will address these uncertainties using 2009 NAEP HSTS data. The goal of this chapter is to discuss whether IC reform model is an effective way of improving US student math ability and participation.

**Between-Groups Research Questions – IC Compared To Non-IC Students**

The research questions were separated into two distinct sets of three questions. The first set of research questions involved a between-groups analysis. These questions compared students who were taught using an IC model to students who were taught using a non-IC or traditional model. Among the non-IC students, 9th grade students taking biology were the most closely matched comparison group based on PSAT performance and were therefore the comparison sample to the IC students.

**Question one: 12th grade NAEP math performance.** Twelfth grade math NAEP performance was used as a proxy for measuring student general math ability. NAEP has been shown to be a reliable predictor of general student math knowledge and future success (Grigg et al., 2010; Moran et al., 2009). Using non-IC biology as the reference group, this analysis of the 2009 NAEP HSTS data showed that 12th graders who were taught traditional physics as 9th graders using an IC model scored nearly a third of a standard deviation ($g = 0.32, 10.06$ points) higher than non-IC
students. However, students who were taught physics as a 9th grader using a conceptual-based curriculum scored more than a half a standard deviation ($g = -0.55$, -16.25 points) lower than non-IC students. The comparison between traditional and conceptual IC students is discussed later in this chapter related to research question four. These findings indicate that full-year traditional IC physics students perform better on the 12th grade NAEP math assessment than all other groups of students analyzed and may be useful in improving general student math performance. Conversely, concept-based IC approaches appear to be less efficacious in improving general math performance as compared to IC models that teach physics in a traditional manner and non-IC models which teach biology in 9th grade.

To review, IC students taught using a concept-based curriculum such as Conceptual Physics (Hewitt, 2002) or Active Physics (Eisenkraft, 1998) are generally taught physics concepts without performing the traditional mathematical calculations underlying these concepts. In contrast, traditional approaches usually emphasize instruction related to the mathematical calculations underlying the physics concepts. This traditional approach emphasizes mathematical computation, while the conceptual approach tends to minimize the mathematical calculations and instead focuses on understanding physics at a conceptual level.

These observations help to shed light on the debate within IC circles regarding the degree to which mathematical concepts should be embedded in a 9th grade physics course. Concept-based IC proponents argue that most 9th grade students have not yet developed the algebra skills required for a mathematically rigorous treatment of physics so it is best to cover the material in a conceptual fashion. Traditional algebra-based IC proponents argue that a complete physics course must include math if the concepts are to be fully understood and appreciated.
Though the debate continues in IC circles, this research indicates that a traditional, algebra-based IC physics course is preferable to a concept-based IC approach. In fact, it appears that not adopting an IC model may be preferable to adopting a concept-based IC model in some cases.

These results appear to support hypotheses 1 and 2, discussed in Chapter 2, by indicating that IC instruction, when it emphasizes mathematical problem-solving (as most traditional physics curricula do) is an effective method to improve general math competency (hypothesis 1) and applied problem solving (hypothesis 2), both measured by the 2009 NAEP assessment. The finding supports both hypotheses that traditional physics IC students significantly outperformed both conceptual-based IC students and non-IC students on the 12th grade NAEP math assessment. The NAEP assessment has two main parts. The first tests general math knowledge and ability, while the second tests mathematical problem solving ability (Sikali, 2013). That traditional physics IC students outperformed both concept-based IC and non-IC students by more than a half and slightly less than a third of a standard deviation respectively indicates that traditional physics IC models are effective in improving both general math ability and problem solving skills.

From a policy perspective, based on these results using 2009 NAEP HSTS data converting a school’s science curriculum to a traditional physics (i.e. non-conceptual) IC model can produce a statistically significant and substantially positive improvement in general math performance as measured by the 12th grade NAEP math assessment. Concept-based IC models appear less preferable because of the finding that non-IC 9th grade students taking biology outperform the concept-based 9th grade IC students.
**Question two – high school algebra GPA.** The second research question explored the relationship between IC models and student algebra GPA. Does a student who is taught using IC models have a high school algebra GPA higher than a non-IC student? This analysis of the 2009 NAEP HSTS data revealed that 12\textsuperscript{th} grade students taught using a traditional, algebra-focused physics IC model as a 9\textsuperscript{th} grade student earned algebra GPAs that were more than a quarter of a standard deviation ($g = 0.27, 0.32$ points) higher than their non-IC 12\textsuperscript{th} grade peers who took biology as a 9\textsuperscript{th} grader (see Table 6). There was no significant difference between conceptual IC students and non-IC 9\textsuperscript{th} grade biology students. The comparison between traditional IC physics students and conceptual IC physics students will be discussed later in this chapter related to research question five.

These findings support hypothesis 3, discussed in Chapter 2, by indicating that there was a significant improvement in algebra GPA among traditional physics IC students compared to their non-IC peers. To review, IC proponents claimed that using algebra-based concepts in the physics curriculum provided a concrete context for algebra students to base the mathematical concepts, which should result in improved algebra performance (Lederman, 2005). This finding is important in that it demonstrates improved algebra performance as measured by algebra GPA and highlights the observation that this effect is only observed among those students that learned physics using a traditional curriculum. While this study was unable to document the exact teaching strategies used in each class, it is commonly understood that the traditional IC physics classes give more robust treatment of the mathematical computation inherent in physics study. In fact, this very debate is what generated the concept-based IC model. These concept-based models were developed specifically
to avoid forcing young students to deal with the algebra involved in basic physics (Hewitt, 2002; Tobias, 2013).

Based on these results, it appears that traditional physics instruction in the 9th grade under an IC model is effective in supporting algebra achievement by providing a context for algebra students as was hypothesized earlier. Additionally, the support of hypothesis 3 by these findings corroborates the theoretical framework discussed in Chapter 2. This model is supported in this analysis by providing a mechanism by which students create a text base and situation model, thereby forming the problem schema required for effective mathematical problem solving. It is important to note that this conclusion is inferred from the significant improvement in algebra performance after taking a traditional physics IC course in 9th grade. To fully test this inference will require future research using data that include teacher pedagogy and classroom practice. From these teacher variables, it will be possible to understand the strength of the relationship between traditional physics instruction, teacher practice within that instruction, and the outcome on mathematical problem solving.

**Question three – number of advanced math Carnegie units.** The final between-groups research question pertained to the effect of IC models on the number of advanced math Carnegie units IC students took compared to their matched non-IC peers. Does a student who is taught using an IC model take more advanced math courses over the course of their high school career than a non-IC student? The model showed that after controlling for student and school characteristics, traditional physics IC students earned 2% more advanced math Carnegie credits than their non-IC peers. Additionally, there was no significant difference between the concept-based IC students and the non-IC biology students, thereby indicating that traditional
physics IC students had a greater likelihood of taking more advanced math courses
than the conceptual IC students.

These results mean that traditional physics IC models are effective at
encouraging students to take advanced math courses as compared to non-IC physics.
These results support the IC notion that providing traditional physics instruction to
9th grade students can motivate these students to take more math later in their high
school careers.

Within-Groups Questions – Amount and Type of Physics Comparisons

The second group of research questions delved into within-groups
differences among students who were taught using an IC model. Specifically, the
type of physics taught delimited the sample so that half the students were taught
using a half-year physics model ($n= 2,336$) while the other half of students in the
sample were taught using a full year physics model ($n = 2,082$). Within the full-year
physics sample, a sub-sample was taught using a conceptually based physics
curriculum ($n = 396$) while others were taught using a traditionally math-based
curriculum ($n = 1,686$). The goal of this within-groups analysis was to conduct a
validity study among IC students to examine if there are any significant effects due
to the amount of time and type of curriculum used with IC students.

**Question four – 12th grade NAEP math performance.** As with research
question one, NAEP math performance scores were used as a proxy for overall 12th
grade student math performance among IC students. Using half-year IC students as
the reference group, these analyses showed that full-year traditional IC physics
students scored more than three quarters of a standard deviation ($g = 0.82$, 23.68
points) higher than half-year physics students on the 12th grade NAEP math
assessment. Additionally, IC conceptual physics students scored more than a quarter
of a standard deviation ($g = 0.81, 26.41$ points) lower than IC traditional students.

There were no statistically significant differences between concept-based IC students and half-year physics IC students. Similar to the findings mentioned in question 1, the full-year traditional physics IC model is preferable to both the concept-based IC model and the half-year IC model to support improved general math performance.

**Question five – high school algebra GPA.** Among IC students, the amount of physics taken was a significant predictor of high school algebra GPA with full-year traditional IC students earning GPAs more than a half a standard deviation ($g = 0.57, 0.51$ points) higher than half-year traditional IC students. Full-year concept-based IC students earned GPAs that were $85\%$ of a standard deviation ($g = -0.85, 0.50$ points) lower than full-year traditional IC physics students. This finding corroborates previous findings by implying that if an IC model is adopted, the adoption should include full-year, traditional (i.e. mathematically rigorous) physics for $9^{th}$ grade students as opposed to the half-year option or the conceptual option if gains in high school algebra GPA are to be realized.

**Question six – advanced math Carnegie units earned.** The final research question of the within groups analysis pertained to whether $9^{th}$ grade students who took full-year physics earn more advanced math Carnegie units than their half-year physics peers. Using half-year traditional IC physics as the reference group, these results showed that full-year IC students earned $8\%$ more advanced math Carnegie credits than their half-year IC peers. Additionally, conceptual IC physics students earned $2\%$ fewer advanced math credits than their traditional IC physics peers. This finding corroborates the previous findings by indicating that among IC models, traditional IC physics that emphasizes math concepts and problem solving are superior to half-year and conceptual IC models.
Limitations

Two primary limitations exist within this study design. First, while great effort was taken to ensure the inclusion of the correct students, it became very difficult to sort through the various course titles that could have described an IC model. This difficulty was due to the diversity of course titles of many of the courses listed. Additionally, it was impossible to determine to what degree the courses were taught in an applied or problem-focused fashion. Because of these limitations, many courses were excluded from the study that may have incorporated large amounts of physics content which would include them in an IC model. To alleviate this limitation, only the courses with “traditional” titles were included. This became only those classes titled physical science or physics for inclusion into the sample frame of 9th grade, early physics IC models. Of course, there was also no quantifiable manner in which to gauge how much physics was taught in the physical science classes and in what manner (i.e. applied, hands-on, or direct instruction). It was assumed that a physical science course was a half year of physics and a half year of chemistry, but no documentation was available by which to confirm this assumption. Therefore, the sample frame may have excluded some cases that fit an IC model while including others that may not have fit an IC model.

The second limitation was inherent in the structure of the 2009 NAEP HSTS data structure. Due to the need to avoid creating too complex of a sample, NCES researchers did not collect teacher level data pertaining to teacher practice and pedagogy. These data would have been very helpful in answering several of this study’s research questions, but these data were simply not available. Ideally, this study would have utilized a three level HLM analysis to control for teaching methods and pedagogies as the level-2 variables and school factors becoming the
level-3 variables. Future research studies could address this missing level by using a dataset that includes the teacher practice level (i.e. TIMSS) to address some of the pertinent research questions mentioned in this study.

**Policy Recommendations**

Persistently low math achievement has been identified as the largest factor reducing the ability of the US to compete in an increasingly global economy (Augustine et al., 2007). Calls for the large-scale overhaul of the manner in which math is taught to US students have been persistent, loud, and numerous (Dunkle, 2012). Many educational policy makers have begun challenging these reform efforts based on the prohibitive costs and resources required to implement the reforms (Burke, 2011). These policy makers have challenged the notion that these large and often expensive reforms produce a sufficiently significant return on the investment in the form of improved math performance by students. These critics have begun to question whether the costs paid by citizens, school districts, and states produce the desired and needed outcomes. Most US citizens agree that something must be done to improve student math performance in the country. The basic question of this research was whether IC models are efficacious in improving student math ability and advanced math course-taking. The findings of this research indicated that offering full-year traditional physics in the 9th grade produces small but statistically significant gains in general math ability, algebra performance, and to a limited degree improved advanced math course taking. These improvements can be realized by establishing school-wide expectations that every student is a “physics student” and that all sciences emanate from this most basic and foundational science. Additionally, these findings implied that the quantitative nature of traditional physics consistently appeared to be a significant positive predictor for student math
achievement. Other research has indicated that this quantitative approach to science education must continue throughout subsequent science courses if these gains in mathematics are to be sustained (Bybee & Gardner, 2006).

**Conclusion**

This research set out to understand the impact of IC models on general math performance, high school algebra performance, and advanced math course-taking patterns using a two-level hierarchical linear model (HLM). The findings of this research demonstrated that full-year IC models using traditional physics curricula can significantly improve 12th grade general math performance compared to concept-based IC models and non-IC models as measured by 12th grade NAEP assessment data. Additionally, this same traditional IC physics model can significantly improve high school algebra GPA, and the number of advanced math Carnegie credits earned. These findings demonstrated that the type of IC model matters in improving student math performance. Specifically, when students were taught using a conceptual-based curriculum, they underperformed their traditional physics IC peers by 26 points and their non-IC peers by 16 points on the 12th grade NAEP math assessment. The trend of concept-based IC students underperforming their traditional IC peers held true on five out of the six dependent variables tested. If an IC model is adopted, it should be a full-year class that focuses on the underlying algebra concepts and computations.

These results indicate that teachers using traditional physics IC models may support the theoretical framework discussed in Chapter 2 by (1) presenting students with a physics-based problem text base; (2) challenging those students to develop a situation model that accurately portrays the physical principles embedded in the physics problem text base; (3) supporting students in the formation of a problem schema derived from algebraic principles that best model the physical phenomenon
presented; (4) applying an algebraic strategy to solve the physics problem identified in the text base; and (5) testing the algebraic strategy against other problems like the initial problem thereby beginning the process of concept retention. It is important to note that the aforementioned process is inferred from the literature and the findings of this study. Confirmation of these processes lies outside the scope of this study but would be worthwhile questions for further research.

The theory of change discussed in Chapter 2 may be supported by the findings of this research by describing the process that may occur when a student successfully applies the aforementioned problem schema to the physics problem situation model. Upon the formation of a parsimonious problem schema, the student will apply one or more algebraic principles to solve the problem and other problems like the initial one. By developing this problem schema and using it to successfully apply the requisite algebraic principles, the student experiences improved algebra performance in a variety of problem solving contexts. This successful application of algebraic principles across a variety of contexts improves student algebra performance in his/her formal mathematics courses. The student’s ongoing mathematical success incentivizes the student to take more advanced math courses throughout high school leading to improved math achievement. It may be inferred that when a teacher removes algebra-focused instruction using a conceptual IC model, the student’s problem schema formation is negatively impacted. This problem schema while being conceptually accurate does not require the student to apply the mathematical tools needed to test their problem schema for accuracy and parsimoniousness related to the text base and situation model as outlined in the theoretical framework (see Chapter 2). It may be inferred from the findings of this study that while conceptual understanding is vitally important for the correct
formation of the text base and situation model, the process of concept retention is most effectively supported through the application of algebraic tools to the problem at hand. This inference was corroborated but not directly supported by the findings of this study because teacher practice data were not recorded and hence lies beyond the scope of this particular study. Future related studies might consider quantifying the role that a teacher’s classroom practice plays on the aforementioned theory of change and theoretical framework related to the use of IC models.

The primary limitation of this research is that the 2009 12th grade NAEP assessment data did not collect teacher-level data. Hence this analysis did not take into account teacher practice data in the analysis of the efficacy of IC models on high school math performance. Teacher practice plays a large role in determining the strength of any educational reform effort (Leikin & Levav-Waynberg, 2007). Teacher level data were unavailable for this analysis due to the structure of the 2009 NAEP HSTS data set and were therefore excluded from this particular analysis.

Over the past 20 years there has been a proliferation of reports, committees, and policy papers stating the need to improve student math performance. All of these documents and statements have identified US student math performance as one of the most pressing issues facing educators today (Executive Office of the President, 2010; Kilpatrick et al., 2011; Phillips, 2009). This study contributed to the research literature on math education reform by examining the efficacy of IC models as a viable reform model to improve student algebra performance and increase the number of advanced math courses taken in high school. These results indicated that the type of IC model used plays a critical role in the efficacy of improved math performance. The full-year traditional (i.e. mathematical-based) physics IC model outperformed the full-year conceptual-based physics IC model and the non-IC model
on all but one of the dependent variables. This research may help to inform the
debate within IC circles regarding whether to include algebra concepts and
mathematical computation in a 9th grade physics class. Based on this research, the
answer is yes, algebra and mathematical computation should be an integral part of
any IC model if it is to produce gains in student math ability and performance.

The call to maintain the US role as a global technology leader continues to
catalyze heated and sometimes divisive debate. Pushing aside the rancor that can
sometimes develop around effective math teaching methods, it is my hope that this
research presents an encouraging argument: That early physics that emphasizes
mathematical rigor can be an effective means of improving US student math
performance. Additionally, traditional physics IC models may support students in
taking more advanced math courses beyond geometry. If the US is to rise above the
“gathering storm” of poor student math performance, an IC model using algebra-
focused physics curricula appears to be an effective mechanism by which to
accomplish this objective. This research contributes to the existing knowledge base
by clearly indicating that IC models that focus on the mathematical rigor inherent in
physics can produce positive gains in math performance. If, as a country we are to
rise above the aforementioned gathering storm of poor math performance, it appears
that an algebra-focused IC model is a vessel well suited for this daunting yet
important voyage for students.
References


