Opportunity to Learn, Engagement, and Science Achievement:

Evidence form TIMSS 2003 Data

Yun Mo

Dissertation Submitted To the Faculty of the Virginia Polytechnic Institute and State University in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy
in
Educational Research and Evaluation

Dr. Kusum Singh, Chair
Dr. M. David Alexander
Dr. Mido Chang
Dr. Yasuo Miyazaki

October 31, 2008
Blacksburg, Virginia

Keywords: Science Achievement (Performance), Self-Efficacy, Engagement, Attitude, Opportunity to Learn

Abstract

This study examined the relationships between opportunity to learn (OTL), science engagement, and science achievement in students' middle school level. This study used the Trends in International Mathematics and Science Study (TIMSS) data from the 2003 wave. The data were analyzed using structuring equation modeling and hierarchical linear modeling. It was hypothesized that students' engagement in science is a mediator between opportunity to learn and science achievement. Moreover, class and school level variability was also examined since the organization of the data was nested. The study examined the effects of OTL on students' emotional, cognitive, and behavioral engagement in science and subsequently on science achievement controlling for family socioeconomics status. The results of structural equation modeling supported some theoretical formulations of the conceptual model, and showed significant effect of OTL factors on students' science engagement, especially the behavioral engagement. Furthermore, science emotional and cognitive engagement showed positive effects on science achievement, but the effect of behavioral engagement on science achievement was complex. Detailed exploration and discussions were included in this study. The findings from hierarchical linear models suggested that students' science achievement was not only related to students' engagement, but also varied by class and school level OTL factors. The study had both theoretical and practical significances, providing valuable insights for the pedagogy of science.

Acknowledgements

First of all, I would especially like to express my sincere gratitude to my adviser, Dr. Kusum Singh, for her generous time and commitment; I am very thankful to her for the guidance and the encouragement she provided for my academic study and personal issues. Without her help, I could not have completed my study.

I also would like to express my sincere appreciation to Dr. M. David Alexander. I am greatly indebted to him for the financial assistance that he provided to me for completing my graduate study.

My thankfulness also goes to Dr. Mido Chang and Dr. Yasuo Miyazaki, for serving as my committee members and giving me invaluable academic advice whenever needed.

And finally, sincere gratitude goes to my husband, my parents, and my sister. Their endless love for me always encourages me to make progress.

Table of Contents

List of Tables

List of Figures

CHAPTER ONE

INTRODUCTION

This dissertation presents a quantitative study of the effects of opportunity to learn and engagement on science achievement of the middle school students in the United States. Science achievement of the U.S. students is an important topic of research and has attracted the attention of both researchers and policy makers in the recent years. There is a growing body of research on factors that are linked to science achievement. Educational researchers have studied the role of family, socio-economic factors, student demographic characteristics such as race and gender as well as such affective factors as attitude, motivation and interest. More recent research has focused on school factors, teacher effects and student engagement. Despite an extensive literature on science learning and science achievement, there are few consistent and robust findings on science learning. There is need for more research on factors that link classroom teaching to students' engagement in learning and achievement. The focus of this study is on the construct of opportunity to learn in classroom through teaching-learning activities and how the opportunity to learn is related to student engagement and achievement. Furthermore, socioeconomic status based differences in opportunity to learn, student engagement and achievement are explored to understand the role of these variables in explaining the achievement differences in science.

Background of the Study

For more than a decade, educators have emphasized the importance of science literacy (American Association for the Advancement of Science Washington DC., 1989; National Academy of Sciences - National Research Council Washington DC., 1996). Considerable investments have been made to investigate student acquisition of knowledge and skills in science nationally and internationally. Large scale studies such as the Trends in International Mathematics and Science Study (TIMSS) and the National Assessment of Educational Progress (NAEP) have primarily focused on math and science achievement and have collected data on various social, educational and school variables to better understand the relationship of these factors to student learning. Educators have drawn attention to the various factors that affect science achievement of students (Manning, 1998).

Earlier research on science achievement had focused on cognitive factors such as ability, IQ and other measures of innate aptitude. But recent research has found that IQ only explained about 25% of the variance in achievement (Jensen, 1998). Other domains such as affective and motivational characteristics of individuals also are important factors in science achievement. There are many studies that have focused on background and family factors in models of science achievement. But demographic variables such as gender, race, parents' education and socio-economic status cannot be manipulated by teachers, schools and policy makers. Although the studies of these variables deepen our understanding of achievement differences, these do not suggest educational reform and changes in educational policies.

Recent research on science achievement has incorporated factors that are related to schooling (Duschl, Shouse, & Schweingruber, 2007). For example, the opportunity to learn through innovative teaching practices and curriculum reform is a construct that links student engagement and learning to classroom teaching (Finn & Voelkl, 1993). There are individual level variables that can be changed through teaching, like students' self-efficacy, motivation, engagement. There is also the possibility of interaction effects and mediating effects among the individual and school level variables. Thus, including both individual and school level variables will lead to better understanding of the nature of relationships among these constructs.

Although science learning is important from kindergarten to high school, middle school years are particularly critical in inculcating interest and values of science learning. In the US middle school/junior high schools, students begin to enroll in class schedules where they take classes from several different teachers in a given day. The classes are usually a set of four core academic classes (English or "language arts," science, mathematics, and history or "social studies") with two to four other classes, either electives or supplementary or remedial academic classes. Most middle/junior school science classes only broadly cover the natural science, from basic physics, chemistry, earth sciences, environment science, and biology, while in high school students are required three standard options of science courses Biology, chemistry, and physics. Other science studies include geology, the environment and forensics. It is through science courses that students begin to understand the natural world and build their knowledge, skills and attitudes for its care. They also are able to use science knowledge learned from classroom for problem-solving and developing further knowledge. 8th grade is a connecting period for middle school and high school. In middle school science, students will be given the foundation of knowledge and skills necessary for the application of scientific concepts throughout their everyday lives and academic experiences. In addition, a mastery of these science skills will be necessary for continued success in high school science study. Teachers are preparing students for the more rigorous science classes students will take in high school. Therefore, 8th grade is vitally important that students start building good study habits and positive attitudes toward science. The present study includes both individual level and school level factors and focuses on the role of opportunity to learn and student engagement on science achievement

Significance of This Study

Our world is shaped by science, so students may learn science not only from classroom, but also from their daily lives. Science achievement in the middle school level is a critical for students' further learning and for science-related career aspirations because middle school is the time when schools change from general ideas to specific concepts in terms of teaching science especially in 8th grade.

The United States has a long-term leadership in science and technology. However, in recent decades, U.S. students have fallen far behind their peers in much of Western Europe and in some Asian countries. According to the Trends in International Mathematics and Science Study (TIMSS 2003) — the largest international study of student mathematics and science achievement— when U.S. students reach their last year of middle school, they rank only 9th out of 45 countries in science achievement. US middle school and high school students are unprepared in science and correspondingly uninterested in these careers. The result was that in National Science Board, graduate and postgraduate positions at the nation's leading universities were often filled with foreigner in 2005. During the past two decades, the share of Science & Engineering (S&E) master's degrees earned by international students rose from 19% to 28%. More than a third (36%) of all S&E doctorates awarded in the United States were international students, and 55% of S&E postdocs in academic institutions in fall 2005 were temporary visa holders. In industry, twenty-five percent of all college-educated workers in S&E occupations in 2003 were foreign born (National Science Foundation Washington DC. National Science Board., 2008). Realizing the critical situation, educators and policy makers are interested in improving students' science performance in high schools and middle schools.

There are a large number of studies emphasizing the effects of various student, family and school factors on science achievement. Besides ability, the most frequently used factors to explain science achievement are science attitude and science engagement. Students' engagement provides the greatest link to increasing student achievement (Darling-Hammond, 1997). High levels of achievement implicitly demand engagement; consequently, engagement is a potentially useful construct for organizing strategies to support achievement in schools (Hudley, Daoud, Hershberg, Wright-Castro, & Polanco, 2002). However, there is no clear and consistent theory of science achievement and the socio-cognitive factors related to science achievement.

Definitions of Factors

Students' Engagement

A generally used definition is "psychological investment in and effort directed toward learning, understanding, or mastering the knowledge, skills, or crafts that academic work is intended to promote" (Newmann, Wehlage, & Lamborn, 1992, p.12). It can be indicated by both emotional and behavioral factors, (Newmann, et al., 1992). In science learning, behavioral engagement can be specified as completion of science assignments; participation in science class and experiments and doing extra science work. Emotional science engagement can be defined as interest or efficacy toward science. In later research, (Fredricks, Blumenfeld, & Paris, 2004) added another component in engagement – cognitive engagement. They indicated that engagement contains behavioral, emotional, and cognitive components. Each of these three is thought to range on a continuum of investment or commitment from the simple to the complex. Behavioral engagement encompasses participation, task involvement, and pro-social conduct. Emotional engagement includes affect, interest, identification with school, and belonging. Cognitive engagement is centered on self-regulation, strategic thinking, and psychological investment. In this study, emotional engagement is measured by students' interest/efficacy in science; cognitive engagement is measured by value of science; and behavioral engagement is measured by students' science class activities.

Opportunity to Learn

One factor that has been considered in understanding differences in science achievement is opportunity to learn. The concept of "opportunity to learn" (OTL) was first introduced several decades ago and was defined by a narrow set of instructional components. Since then, educators and policy makers have incorporated many additional criteria into the OTL concept, some specifically to ensure an equal education for disadvantaged and minority students. (Steven, 1993, p233-234) examined the question of opportunity to learn from four perspectives:

- Content Coverage - These variables measure whether or not students cover the core curriculum for a particular grade level or subject matter.
- Content Exposure - These are variables that take into consideration the time allowed for and devoted to instruction (time-on-task) and the depth of the teaching provided.
- Content Emphasis - These are variables that influence which topics within the curriculum are selected for emphasis and which students are selected to receive instruction emphasizing lower order skills (i.e., rote memorization) or higher order skills (i.e., critical problem solving).
- Quality of Instructional Delivery - These variables reveal how classroom teaching practices (i.e., presentation of lessons) affect students' academic achievement.

Educators and policy makers agree that students urgently need better education and equal opportunity in science. National Science Education Standards were developed in 1996 to provide guidelines to schools and teachers. The Standards offer a coherent vision of what it means to be scientifically literate, describing what all students should understand and be able to do in science. These standards reflect the principles that learning science is an inquiry-based process, that science in schools should reflect the intellectual traditions of contemporary science, and that all Americans have a role in science education reform (National Academy of Sciences - National Research Council Washington DC., 1996). In this study the OTL factors that were considered are content exposure in teachers' instruction, topic coverage, teachers' quality and schools' enrichment/remedial programs.

Problem Statement

The present research examines the relationship of opportunity to learn, science engagement, and science achievement. It further explores some demographic-based background differences such as socioeconomic status (SES) in the relationship of opportunity to learn (OTL) to science engagement and science achievement. It is hypothesized that students' engagement in science is a mediator between opportunity to learn (OTL) and science achievement. Better opportunity would enhance the students' engagement in science learning and would increase their science performance. Moreover, since the organization of the data is nested, meaning student is nested in classroom and a classroom is nested in a school, class and school level variability is also examined. The research questions that guided the study are:

- What are the effects of OTL and students' science engagement on science achievement? Does students' science engagement act as a mediator between OTL and science achievement?
- Is there any evidence that the effects of engagement and OTL vary by class and school level?

Conceptual Model

To answer the research questions, a conceptual model is hypothesized that expanded on Greenwood's (1996) performance-based model wherein the effects of instruction (opportunity to learn) on school outcome were mediated by. Emotional, cognitive, and cognitive engagement factors served as mediators that directly affect science achievement. Family background factor and opportunity to learn factors are correlated and serve as exogenous variables which influence students' science engagement and achievement engagement (See Figure 1 for details). The conceptual model guided the analyses undertaken to confirm if the data fit the model.

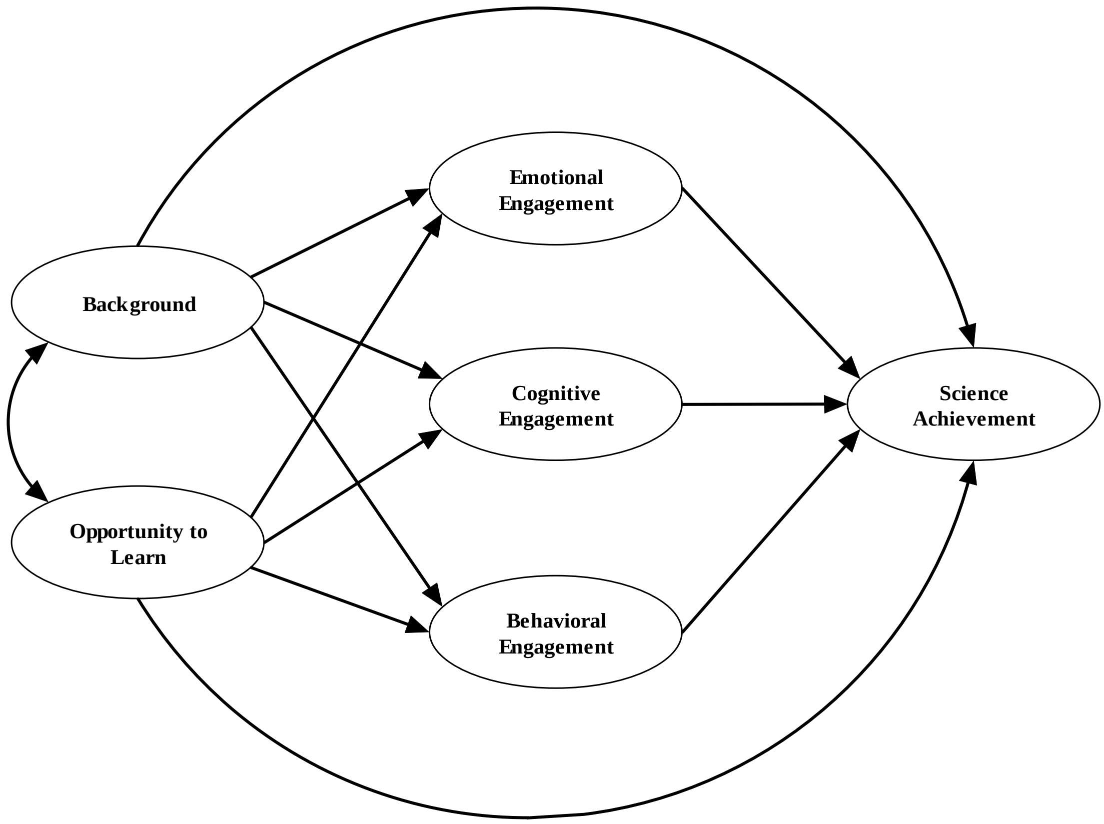

Figure 1. Hypothesized structural equation model

Overview of Methodology

Characteristics of the Dataset

This study uses TIMSS, the Trends in International Mathematics and Science Study, which was sponsored by the International Association for the Evaluation of Educational Achievement (IEA). TIMSS is designed to help countries all over the world improve student learning in mathematics and science. It collects educational achievement data at the fourth and eighth grades to provide information about trends in performance over time together with extensive background information to address concerns about the quantity, quality, and content of instruction. Since 1995, TIMSS has been assessing trends in students' mathematics and science achievement on a regular four-year cycle. Countries participated at the 4^{th} and 8^{th} grades. Results for TIMSS 2003 were reported in December 2004 and more than 60 countries provided data for TIMSS 2007. TIMSS 2003 comprises student achievement data in mathematics and science as well as student, teacher, school, and curricular background data for the 48 countries that participated in TIMSS 2003 at the eighth grade and 26 countries that participated in TIMSS 2003 at the fourth grade. The database includes data from over 360,000 students, about 25,000 teachers, about 12,000 school principals, and the National Research Coordinators of each country (Martin, 2005).

Structural Equation Modeling

This study mainly uses the structural equation modeling (SEM) to test the hypothesis that science achievement is affected by students' background constructs, opportunity to learn, science engagement constructs and the links between each construct.

SEM is a statistical method that takes a confirmatory approach to data analysis of a structural theory bearing on the relationships of some variables of interest (Byrne, 1998). In present study, a latent variable structural equation model is estimated in several steps using LISREL 8.8 computer program (Jöreskog & Sörbom, 2006). First, the measurement part of the model (constructs and their indicators) are specified and estimated; then structural relationships in the model are specified and estimated; finally, proper changes to the model specification are made according to estimated modification indices. Structural equation modeling is an especially appropriate method for analyzing non-experimental data. In addition to parameter estimates, the program provides fit indices to assess how well the model fits the data. Such fit indices make it possible to evaluate the adequacy of the theoretical model in explaining the data (Bollen, 1989; Schumacker & Lomax, 2004).

After selecting items that logically seemed related to science achievement and its effects, a measurement model is developed and estimated to determine if the selected items have significant loadings on hypothesized constructs. And then research hypotheses are tested.

Hierarchical Linear Modeling

Another analytical approach used in the study is Hierarchical Linear Modeling. In educational field, organizations and data structures are often hierarchical, or nested. That means individual units are grouped into larger groups; these groups of individuals are grouped into higher order organizations; the organizations may be grouped at still higher levels (Raudenbush & Bryk, 2002). Here is a common example: students are grouped in classrooms, or within teacher; classrooms are components of schools; schools belong to districts, which are part of the state educational system, etc. According to this nested structure, a three-level hierarchical linear model (Raudenbush & Bryk, 2002) is formulated to explain the class and school level variability, where the level-1 unit is students; the level-2 unit is the classes/teachers; and the level-3 unit is the school. Within the hierarchical linear model, each of the levels in the data structure is formally represented by its own sub-model. Each sub-model represents the structural relations occurring at that level and the residual variability at that level.

Organization of the Study

This study proceeds with a comprehensive review of the literature related to science achievement, students' engagement in science and opportunity to learn in chapter two. A detailed explanation of the literature search process is provided before review of the theoretical and empirical research. The final section of chapter two lists some limitations of current research on science achievement and the links between science achievement and the related factors. In chapter three, the methodology of the study is described in detail, including database selection, variable selection and procedures used to analyze data. Chapter four is a presentation of the preliminary analysis of raw data, including variable recoding and transformation if necessary; the five steps taken for model specification, model identification, model estimation, model testing and model re-specification are explained. In chapter five, summary of findings, discussion and implication of the results, limitations and directions for further research are presented.

CHAPTER TWO

LITERATURE REVIEW

This chapter surveys the theoretical literature and empirical research on opportunity to learn, engagement and science achievement. The first section describes the literature search process. The second section describes previous theories on academic achievement, including aptitude theory (Snow, 1992); self-efficacy theory (Bandura, 1997); and performance-based structural equation model (Greenwood, 1996). The definitions and measures of student engagement are presented in the third section. Three types of engagement perspectives (Fredricks et al., 2004) behavioral, emotional and cognitive engagement are emphasized in this section. The last section introduces the definitions and measures of opportunity to learn.

Explanation of the Search Process

In last decades, a substantial body of literature has emerged that examines the role of affective factors in learning. The relationships of affective factors to science learning have been an important area of educational research. Broad searches of this literature on science achievement and affective factors were carried out using several approaches. First an initial search on the following databases was conducted: the Education Resources Information Center (ERIC) database, Educational Abstract, Educational Full Text, Ingenta, PsycINFO, and WorldCat. The search process was aided by EndNote X.0. We also searched on Google for any published studies related to science achievement in secondary school. The utilized keys words included "science achievement (performance)", "self-efficacy", "engagement", "attitude”, and “opportunity to learn”.

After the initial search stage, the abstracts were downloaded from various databases. We merged the abstracts and deleted the duplicate entries. This process resulted in a set of non-redundant abstracts from various databases. From this set of abstracts, we created an EndNote library of those abstracts that described an empirical study about some aspect of affect in science achievement conducted between 1995 and current. Qualitative as well as quantitative studies were included. The next step involved an exhaustive reference search based on the studies selected previously (cross-referencing). This step included searching for specific studies and theories cited in these articles.

In addition to the review of the literature on science achievement, we also searched for methodology articles and textbooks on the subject of structural equation modeling and hierarchical linear model.

Review of Previous Theories

This study examined previous research that has looked at the non-ability factors in achievement outcomes and sought to integrate the work of psychologists and educational researchers on affective, conative, and motivational variables. Consistent with the focus of the present study on the relationship of self, engagement, and achievement, we drew on the work of several researchers to create the theoretical framework for the study. The work of Eccles, Wigfield, and Schiefele (1997) has shown that a number of non-ability factors, such as self-efficacy, interest, and value of the task, can influence engagement and achievement. The works of Snow and his colleagues (Snow, 1989; 1992; Snow, Corno, & Jackson, 1996) have shown that there are multiple pathways to achievement and commitment, and engagement to the task is an important precursor of achievement (Lau & Roeser, 2002). Snow et al., (1996) considered what they called conative and affective characteristics of the individual as important for learning. The conative domain referred to motivation, achievement orientations, and volitional factors, while the affective domain described emotions and attitudinal variables. Snow (1989, 1992, 1994) provided a general theory of cognitive and conative-affective processes that affect achievement. Chang, Singh, & Mo (2007) found support for affective factors in science learning.

Self-efficacy is a widely used indicator of psychological engagement in the literature (Hudley et al., 2002). Bandura (1997) defined self-efficacy as "beliefs in one's capabilities to organize and execute the courses of action required to produce given attainments" (p.3). Self-efficacy is derived from four sources of information including enactive mastery experiences (past performances), vicarious experiences, verbal persuasion, and physiological states (Bandura, 1997). "A strong sense of self-efficacy enhances human accomplishment and personal well-being in many ways. People with high assurance in their capabilities approach difficult tasks as challenges to be mastered rather than as threats to be avoided." (Bandura, 1994, p.71). Self-efficacy has significant influence on attainments and performance. Several investigators have reported a positive relationship between self-efficacy and performance(Feltz & Riessinger, 1990). These findings provide support to a position that perceived self-efficacy can be a mediating factor in enhancing performance (Bandura, 1997).

To investigate the links between instruction, engagement and student achievement, Greenwood (1996) tested the fit of several causal models as alternative explanations of student achievement. A performance-based model wherein the effects of instruction (a second-order factor composed of exposure and task quality) on school outcome were mediated by engagement produced the best fit to the data and provided evidence of a causal path between instruction, engagement, and academic achievement. In this study, this model is expanded by adding students' background information as an exogenous variable; and presenting engagement as multidimensional construct. In addition, the variability due classroom and school membership is taken into account.

Students' Engagement

Definition of Engagement

Engagement is a multidimensional construct and there are various definitions. It can be broken down into academic engagement and social engagement. There appear to be two dimensions to engagement: academic engagement, or engagement in learning, and social engagement, or engagement in the social dimensions of schooling (Rumberger & Thomas, 2000). Academic and social engagement each comprises participation and identification. Measures of students' academic engagement include participation in academic life (e.g., attend classes, skip classes, complete assignments, or number of hours spent on homework) (Newmann, Wehlage, & Lamborn, 1992). Measures of social engagement include social life such as sense of belonging (Anderman, 2002), being able to count on the support of friends or ease in making friends (Ladd, 1990). The purpose of this research is to find the correlation between engagement and science achievement, so academic engagement is emphasized in this study.

Newmann et al.,(1992, p.12) defined student engagement in academic work as "psychological investment in and effort directed toward learning, understanding or mastering the knowledge, skill, or crafts that academic work is intended to promote". They pointed out that academic engagement is a function of three factors. The first is students' needs for competence, or students' needs to achieve cognitive understanding and mastering learning tasks. The second factor is school membership, which is students' feeling and understanding that they are part of the school community. Factors that cultivate student' school membership include clarity of purpose, fairness, personal support, success, and caring. And the third, engagement is influenced by authentic work. Authentic tasks as "tasks that are meaningful, valuable, significant, and worthy of one's efforts" (Newmann et. al., 1992, p.23). The difference between being engaged and being on task was distinguished by Schlechty (2001, p. 64):

> Engagement is active. It requires that students be attentive as well as in attendance; it requires the student to be committed to the task and find some inherent value in what he or she is being asked to do. The engaged student not only does the task assigned but also does it with enthusiasm and diligence. Moreover, the student performs the task because he or she perceives the task to be associated with a near-term end that he or she values.

Engagement can be indicated by behavioral, emotional and behavioral factors and each of these three is thought to range on a continuum of investment or commitment from the simple to the complex (Fredricks et al., 2004). Behavioral engagement is usually defined as participating or involving in school or social activities (Finn, 1993; Finn & Voelkl, 1993; Furlong et al., 2003); doing school work and following the rules (National Center for School Engagement, 2006). For example, involving basic school attendance and completion of schoolwork (Finn, 1993, Finn & Rock, 1997); involving participation in class when requested by school officials (Newmann, 1992); and intrinsic involving active participation (Finn, 1993) are three levels of behavioral engagement.

Emotional engagement refers to affective reaction in the school or classroom, including interest, happiness, anxiety, identification with school, and belonging (Connell & Wellborn, 1991). Finn (1989) defined emotional engagement as identification with school which includes feelings of being important to the school and appreciation of success in school. Newmann et al. (1992) defined emotional engagement as emotional connectedness to school and teachers. The definition of “emotional engagement” here is somewhat duplicated on attitudes which examines feelings toward schools, teachers, subjects or the works (Fredricks et al., 2004). Conell (1994) has suggested that factors such as how much students like or dislike school, and how strongly they feel they belong in school are emotional engagement.

Cognitive Engagement is essentially defined as motivation, effort, and strategy use (Fredricks et al., 2004). This highlights a psychological investment in learning, a desire to go beyond the requirements, and a preference for challenge (Connell & Wellborn, 1991; Newmann et al., 1992). Cognitive engagement also can be viewed as motivated behavior which represents by the cognitive strategies students choose to learn and by their willingness to continue working on difficult tasks by regulating their own learning behavior (Chapman, 2003).

There are various methods to measure students' engagement in school. Jimerson, Campos and Greif (2003) classified the related items into five contexts; a) academic performance, b) classroom behavior, c) extracurricular involvement, d) interpersonal relationships, and e) school community. Furlong et al (2003) derived a relationship among engagement contexts and their relationship to student outcomes. They indicated that school engagement context should include: a) The Student Context; b) The Peer Context; c) The Classroom Context; and d) The School-wide Context. These contexts are similar to Jimerson et al.'s (2003) theory reviewed earlier. The main difference is considering "academic performance" as an outcome instead of other contexts.

In addition, Fredericks et al. (2004) have suggested that engagement should be studied for additive and interaction effects between the context/ environment and the individual. They describe school, classroom, and peer characteristics as antecedents of school engagement and acknowledge that future research is necessary to examine the complex interplay between the individual and the school environment and its contexts. Researchers and educational practitioners focus on these contexts because of an interest to better understand how to enhance positive student outcomes. So far, many theories have supported the positive linkage between student academic achievement and engagement (Finn, 1993; Finn & Rock, 1997; Gamoran & Nystrand, 1992). Engagement can be serve as both independent and dependent variables. The relationship between SES and engagement was found by Berends (1995). The study found that students whose families had a higher socioeconomic status showed higher levels of engagement, as evidenced by their reports of how much they liked school and how much time they spent on homework.

Overall, engagement is a multi-dimensional construct which involves some internal, unobservable indicators (e.g. emotional and cognitive engagement) and external, observable indicators (behavioral engagement). However, the definitions of engagement are varied and it has been measured by different types of items in terms of the instrument (Jimerson et al., 2003; Furlong et al., 2003). The present study followed Fredericks' et al.'s (2004) theory and composed three engagement constructs.

Measurement of Students' Engagement

Behavioral engagement

Finn (1989) described four levels of behavioral engagement. The first level is exhibited in students' complying with classroom and school rules. Regular classroom behaviors such as whether the student is attentive (Finn & Rock, 1997; Johnson, Crosnoe, & Elder, 2001), disruptive (Finn & Rock, 1997; Greenwood, Horton, & Utley, 2002), answering a question when called upon, and doing activities when teachers asked (Newmann et al., 1992) represent this level of behavioral engagement. Gamoran and Nystrand (1992) reported that behavioral engagement in assignment's completion was closely tied to achievement. The level-two participation includes student initiating questions and dialogue with the teacher, working hard, spending extra time in the classroom, and doing more coursework (Newcomb et al., 2002). This level also includes effort, persistence, and concentration, attention, asking questions, and contributing to class discussion (Finn, Pannozzo, & Voelkl, 1995). The third level is exhibited by students' participation in social, extracurricular, and athletic activities in school. Students who are involved in various extracurricular activities will be less inclined to commit antisocial acts and will have decreased opportunity to do so (Gottfredson & Hirschi, 1990). This level may have no direct effects on school performance, but it does increase the students' psychological well-being and decrease the possibility of risk behavior (Gottfredson & Hirschi, 1990). Finally, the level-four participation involves participating in school governance (Finn, 1993). In this science study, behavioral engagement is operationally defined as students' science class activities.

Emotional engagement

This type of engagement is unobserved and usually measured by students' self-report related to the school, the classes, and the people at school (Fredricks, et al., 2004). The measures of emotional engagement tend to relate to interests or identifications. For example, some measures ask students questions pertaining to their feelings about their teachers and their school, i.e. identification with school (Finn, 1989), student-teacher relations and values (Finn, 1989), and school belonging (Mo & Singh, 2008a). Furlong et al., (2003) include affect, interest, identification with school, and belonging as indicators of emotional engagement. In this study, emotional engagement is defined as interest/efficacy in science. Positive effects of self-efficacy and liking science on science achievement have been found by Mo and Singh (2008b) using TIMSS 2003 8th grade student sample.

Cognitive engagement

The measures of cognitive engagement conceptualized as psychological investment in learning are similar to those measures of intrinsic motivation. The concepts of intrinsic motivation emphasize the degree to which students are invested in and value learning (Fredricks et al., 2004). Another definition targets cognition and emphasizes strategic learning or self-regulation. Researchers that conceptualize cognitive engagement as self-regulation assess school engagement with measures of students' strategy use (Fredricks et al., 2004). The literatures described shows that cognitively engaged or self-regulated students use meta-cognitive strategies to evaluate or plan their effort on tasks (Pintrich & De Groot, 1990; Zimmerman, 1990). In these studies the students were asked to respond to questions about their knowledge and use of cognitive and metacognitive strategies (Pintrich & De Groot, 1990). For instance, they are asked about how they set learning goals, organize study efforts, how they plan, expectation of education, self-instruct, monitor and adjust their cognition (Fredricks et al., 2004; Pintrich & De Groot, 1990). In present study cognitive engagement is operationally defined as students' value of science for future education and occupational aspirations in science.

Overlaps of engagement constructs

The three types of engagement are related and somewhat overlapping. First, the term "effort" is included in definitions of both behavioral and cognitive engagement (Fredricks et al., 2004). They state, the engagement factors "are dynamically interrelated … they are not isolated processes" (p. 61). The division merely aids in understanding that "engagement" as a whole is a multi-dimensional construct. Second, "items that tap behavioral engagement and emotional engagement are often combined in a single scale" (Fredricks et al., 2004, p. 67). For example, Marks (2000) found student engagement was significant related to GPA, school support and classroom support for all elementary, middle, and high school students by measuring "how often do you try as hard as you can;

pay attention or feel bored in the class?" In fact, the researchers have discussed correlation in these types of engagement in past research and the potential for overlap is evident in the concepts as they are defined. Consequently, these components are correlated with each other, but relate differently to student outcomes such as achievement and attendance (NCES, 2006).

Effects of Engagement on Science Achievement

The focus of this research is on middle school students' science achievement and engagement. Though there are many studies related to this area, few of them consistently involved all the dimensions of engagement. Some studies mainly emphasized the behavioral engagement in science classroom. Butler (1999) narrowed science engagement to science behavior intention. He found that attitude toward science accounted for a significant amount of the variance in science behavioral (both laboratory and non-laboratory) intention. Attitude towards science can be also reflected emotional or cognitive engagement in science according to Fredericks et al.'s (2004) theory. It played an important role in students' science achievement in USA and other countries (Papanastasiou & Zembylas, 2004). The positive attitudes towards science that USA students have significantly and positively influence their science achievement. Butler's (2002) study confirmed that three types of engagement components are correlated with each other. Compared to emotional and cognitive engagement, the effect of behavioral engagement on science achievement might be small. Chang, et al. (2007) found there was a small effect of science behavioral engagement in class activities and time spent on science homework on science achievement using NELS 88 data. Similar, only a few instructional activities in science classroom such as students frequently did an experiment or practical investigation had significant effect on science achievement (House, 2004).

Despite many studies that support the role of engagement in achievement, it is not clear what kinds of activities lead to engagement and achievement. Not only that, it is unclear if all instructional activities promote students' engagement and motivation to learn. Some studies, using TIMSS data, have not found support for increase in students' achievement when students were involved in certain instructional activities. Li, Ruiz-Primo and Shavelson's (2006) research found that most of the coefficients of instructional experience in class were not significant on science achievement by using TIMSS 1999 data. Those activities included watch students watched teacher gave a demonstration of an experiment and students did an experiment or practical investigation by their own, etc. Moreover, even some negative relationships of class activities and science achievement have been found by Chang et al. (2007). The result indicated that classroom engagement measured by students design and conduct experiments or projects on their own and make their own choice of science topic or problem to study negatively affected science achievement. These results show inconsistent effects of classroom activities on science achievement and need to future investigation.

Opportunity to Learn (OTL)

Definition of OTL

As the name indicates the construct, opportunity to learn, was conceptualized to consider various school factors that provide conditions, pedagogy and curriculum for students. Opportunity to learn was originally defined as the overlap between the content students were taught and the content on which they were tested (Anderson, 1990). It first referred to equitable conditions or circumstances within the school or classroom that promote learning for all students. It includes the provision of curricula, learning materials, facilities, teachers, and instructional experiences that enable students to achieve high standards. This term also relates to the absence of barriers that prevent learning. Later, educators and policy makers have incorporated many additional criteria into the OTL concept, so the concept has expanded to include the quality of resources, school conditions, curriculum, and teaching that students experience, some specifically to ensure an equal education for disadvantaged and minority students (Guiton & Oakes, 1995).

Measurement of OTL

Research indicates that opportunity to learn is a critical issue that is often difficult to measure. Background information, like SES and ethnicity are generally used variables to indicate unequal opportunity. For example, students in low-income, high minority schools have less access to computers, equipment, and laboratories (Mere, Reiska, & Smith, 2006). Students' achievement is significantly different between high, low, low-moderate and moderate SES groups. Caucasian and Asian students' achievement is significantly better than their Latino and African American peers (Chang, Singh, & Mo, 2007).

Although SES and opportunity to learn are often correlated, the measures of OTL are school or class level factors. Most researchers agree that measures of opportunity to learn should include information about the resources, school conditions, curriculum, and teachers' quality (Winfield & Woodard, 1994). It is unfair to judge the low SES students as less achieving or less engaged because they have fewer resources. Studies of secondary schools report that low-income and high minority schools offer fewer advanced courses (Guiton & Oakes, 1995). The least qualified teachers teach minority and poor students (Darling-Hammond, 1997). Schools with the highest numbers of Latino/ African American students enrolled have the biggest shortages of textbooks and the lowest numbers of qualified teachers (Venezia & Maxwell-Jolly, 2007). Curriculum differentiation-the practice of making different knowledge access to different groups of students sometimes signifies responsiveness to students' needs, but raises questions about unequal opportunities as well (Oakes, Gamoran, & Page, 1992). Each of these areas: quality teachers, materials and equipment are related to student achievement, and are not fully in the control of the students and schools.

The relationship between SES, opportunity to learn and students' science engagement has been investigated in several researches. For example, Hayes and Deyhle (2001) indicated that students were likely to be more engaged when teachers' science instruction was open, supportive, and personally relevant to students. In Hayes and Deyhle's (2001) research, they used ethnographic methods to compare two different science instruction methods at two schools: one serving high-SES white students and the other serving low-SES minority students. The results showed science curriculum were very different between these two schools. At the first school, instruction was rigid and emphasized traditional methods to raise students' standardized test score. At the latter school, instruction emphasized enhancing students' engagement and enjoyment in science. However, former school had higher standardized test score. The researchers suggest that it was difficult to determine whether the instruction differentiation enhanced inequality in science education and which methods were good for promoting science achievement (Lee & Luykx, 2006).

For classroom instruction, four opportunity-to-learn variables were gleaned from previous studies that focus on teaching activities by Steven (1993). The variables are (1) content coverage; (2) content exposure; (3) content emphasis; and (4) quality of instructional delivery.

Content coverage refers to the topics covered at the class instruction. For example: how many of the items on the test match the content that was taught in class or the content is sufficient and timely prior to taking the test (Leinhardt & Seewald, 1981). It is the most frequently studied OTL measure in some studies, Winfield (1993) suggested three ways of measuring content coverage: teachers' self reports, direct observation of classroom instruction, and analyzing the content of curriculum materials. In the Trends of International Mathematics and Science Study (TIMSS 2003), test topics coverage was analyzed to provide researchers with information on content coverage (Martin, 2005; Wang, 1998).

Content exposure refers to the estimated total amount of time actually spent to cover the specific content (Leinhardt & Seewald, 1981). Terms like instruction time, amount of time devoted to a certain subject area, and amount of time in class periods, can be measures of this dimension of OTL (Stedman, 1994; Wang, 1998). Content exposure is highly related to students behavioral engagement such as time spent on appropriate academic tasks. Students' academic learning was strongly influence by the amount of time that spend on tasks (Cooper, Robinson, & Patall, 2006).

Content Emphasis determines which topics within the curriculum are selected for emphasis by teachers and which students are chosen to receive special instruction. In McDonnell's (1990) study, content emphasis was defined as teachers arrange different amount of time in emphasizing different objects and schools offer different curricula according to student ability level. Teachers choose what they want to emphasize and base their choices on personal experiences, personal proficiency, perception of certain topics as important, professional experiences, and influence of past professors, courses, textbooks and other authorities (Floden, 1981). The findings of content emphasis are complex. Content emphasis measured by science enrichment programs generally had been effective in increasing students' science knowledge and mastery (Freedman, 1997). Gamorand and Nystrand (1992) found that authentic teaching questions such as "do you ever have to take notes?" had negative effects on achievement. Furthermore, the effects were different from high- and low-ability classes. That authenticity had positive effects in high-ability classes and negative effect in low-ability classes. Students also can be grouped by culture and language. Teachers should know the complex dynamics between science instruction and students' everyday knowledge. "As teachers identify and incorporate students' cultural and linguistic experiences as intellectual resources for science learning, they provide opportunities for students to learn to use language, think and act as members of a science learning community" (Lee & Luykx, 2006, p.90).

Quality of Instructional Delivery: This dimension of OTL is usually obtained by classroom observation and measures how classroom teaching practices affect students' academic achievement. The quality of instructional delivery can be measured by whether teachers conducted class activities to meet the teaching objectives and teachers' effectiveness in presenting materials (Alkin, Doby, & Lindheim, 1990); or teachers monitor students' performance, and provides corrective feedback during the classes (Brophy & Good, 1986). Quality of instructional delivery was also measured by direct observation in studies conducted by Stevenson and Stigler (1992). The authors discussed the comparison of the learning gap and quality of teaching among teachers from China, Japan (high-performing countries), and the U.S. "Instructional coherence, how students' errors are used in the classroom, pace of instruction, and the quality of interaction between students and teachers are some of the aspects that were examined" (Wang, 1998, p,141). Stevenson and Stigler (1992) suggested teachers relate the different part of a lesson to one another, to explain the interrelatedness of the knowledge. Wyatt (1991) indicated that quality of instructional delivery was an important element in improving the academic achievement of poor and minority student in the US.

Effects of OTL on Science Achievement

While few studies concentrated on science especially, the results are quite significant. Positive relationship has been found between OTL and student understanding within two science domains using TIMSS 1995 data (Angell, Kjærnsli, & Lie, 2006). The findings of this international study indicated that OTL measured as content coverage of the taught topics was strongly significantly correlated to the countries' average percentage correct of selected science items. Between 25 and 50 percent of the variance between countries for individual items was explained by this OTL effects.

This study based on secondary dataset did not include various measures of the OTL discussed in the literature but the study selected items from the survey that capture various aspects of OTL available in TIMSS 2003 data. The teachers' self reports of classroom activities and instructional practices were used as OTL. In addition, content coverage, and whether a teacher has the teaching license in science were considered two classroom level OTL factors. At the school level, OTL measure considered in the study was whether the school offered enrichment or remedial science classes.

CHAPTER THREE

METHODOLOGY

This chapter outlines the methods of the study. It presents a description of the data set, survey instrument, and items in the data set. It also presents an overview of the analysis plan and the statistical procedures used. In this study, the relationships of various personal, classroom and school level factors to science achievement are tested. There are three main domains that are included in the model: the first one is students' engagement in science; the second factor is opportunity to learn; and the third factor is students' family background information.

The objective of this study is to test the hypothesis that science achievement is affected by students' engagement, opportunity to learn and students' socio-economic background and the links between these factors are significant. It is hypothesized that students' engagement in science is a mediator between opportunity to learn and science achievement. Two analytic methods, structural equation modeling and hierarchical linear models are applied in this study.

Data Source

This study uses the Trends in International Mathematics and Science Study (TIMSS) 2003 database. TIMSS is the largest international comparative study of math and science achievement to date. It was conducted on a four-year cycle, the first round of TIMSS was in 1995 and the second in 1999. TIMSS 2003 was the third assessment in the series of studies by the International Association for the Evaluation of Educational

Achievement (IEA) to measure trends in students' mathematics and science achievement. TIMSS 2003 was conducted by the International Study Center at Boston College and included 45 countries. In the United States, TIMSS is supported by the Education's National Center for Education Statistics (NCES) and the National Science Foundation (NSF). It measured the mathematics and science achievement of fourth and eighth-grade students and collected extensive information from students, teachers, and school principals about mathematics and science curricula, instruction, home contexts, and school characteristics and policies.

Sample

The international sample design for TIMSS is generally referred to as a two stage stratified cluster sample design. The first stage consists of a sample of schools; the second stage consists of a sample of one or more classrooms from the target grade in sampled schools. This data set is well suited for studying the relationship among constructs that indicate science engagement and science achievement. First, TIMSS data provided a wide array of student level items on science attitude and more domain specific self-efficacy in science, science classroom activities and science achievement (Martin, 2005). Second, TIMSS data provided teacher level and school level survey to measure opportunity to learn.

The full data include six sections: 1) Students' responses to each of the mathematics and science items administered in the study; 2) Student achievement scores in mathematics and science;3) Students' responses to the student questionnaires; 4)Teachers' responses to the teacher questionnaires; 5) Principals' responses to the school questionnaires;6) National Research Coordinators' responses to the curriculum questionnaires. Because this research is focused on US middle school science education, we only used the 8th grade students' questionnaire in the United States. There were a total of 8912 8th grade students who finished the students' survey from 212 public schools and 20 private schools. This study focused on the relationships of student engagement, opportunity to learn and science achievement in only public schools so only public sample was chosen in analyses. The sample size is 8544.

Measures

Reflective indicators of each latent variable were selected from the original TIMSS 2003 survey, based on theory and operational definitions used in empirical studies. The following sections list items representing different dimensions of science engagement and opportunity to learn.

Science psychological items: There are 12 science attitude items in students' questionnaire. Students were asked to indicate their level of agreement on a 4-point Likert scale, ranging from Scale: 1= agree a lot; 2= agree a little; 3= disagree a little; 4= disagree a lot, with the following statements

1) I usually do well in science.
2) I would like to take more science in school.
3) Science is more difficult for me than for many of my classmates.
4) I enjoy learning science.
5) Sometimes, when I do not initially understand a new topic in science, I know that I will never really understand it.

6) Science is not one of my strengths.

7) I learn things quickly in science.

8) I think learning science will help me in my daily life.

9) I need science to learn other school subjects.

10) I need to do well in science to get into the <university> of my choice.

11) I would like a job that involved using science.

12) I need to do well in science to get the job I want.

Classroom activities items: Science classroom activities items are chosen to measure behavioral engagement variables. Scales of these items are: 1=every or almost every lesson; 2= about half the lessons; 3= some lessons; and 4=never. Students were asked "how often do you do … activities?" The detailed item wording are as follow:

1) watch the teacher demonstrate an experiment or investigation in your science lessons?

2) formulate hypotheses or predictions to be tested in your science lessons?

3) design or plan an experiment or investigation in your science lessons?

4) conduct an experiment or investigation in your science lessons?

5) work in small groups on an experiment or investigation in your science lessons?

6) write explanations about what was observed and why it happened in your science lessons?

7) study the impact of technology on society in your science lessons?

8) relate what you are learning in science to your daily life in your science lessons?

9) present your work to the class in your science lessons?

10) review your homework in your science lessons?

11) listen to the teacher give a lecture-style presentation in your science lessons?

12) work problems on your own in your science lessons?

13) begin your homework in class in your science lessons?

14) have a quiz or test in your science lessons?

Several of the above questions are not only related to students' engagement, but also depend on teachers' instruction. For example, students cannot have a quiz if their teacher doesn't give them the quiz. So we used the teacher data to measure OTL.

Opportunity to learn is a set of items which measured content-related activities. In the other words, these items represented the frequency with which the teacher asked students to do various content-related activities in science classroom. In the same way, scales of these items are: 1=every or almost every lesson; 2= about half the lessons; 3= some lessons; and 4=never. The question is: In teaching science to the students in the TIMSS class, how often do you usually ask them to:

1) watch you demonstrate an experiment or investigation?
2) formulate hypotheses or predictions to be tested?
3) design or plan experiments or investigations?
4) conduct experiments or investigations?
5) work together in small groups on experiments or investigations?
6) write explanations about what was observed and why it happened?
7) put events or objects in order and give a reason for the organization?
8) study the impact of technology on society?

9) learn about the nature of science and inquiry?

10) relate what they are learning in science to their daily lives?

11) present their work to the class?

We also considered students' family socioeconomic status as background information in this study. It includes parents' education level, and number of books in student's home. Other common measures of socio-economic status were not available in the data.

Data Analysis and Model Estimation

Descriptive Statistics

Prior to conducting confirmatory factor analysis, preliminary descriptive statistics (mean, standard deviation, skewness and kurtosis) and reliability estimates were calculated. Preliminary item analysis was carried out and the reliability coefficients of each latent construct: engagement, opportunity to learn – as a whole and within each dimension of constructs – were estimated. Also, correlations among observed indicators were calculated and examined in order to do a preliminary assessment of the correlation pattern. AM software was used in analyzing the descriptive and correlation statistics. AM is a statistical software package for analyzing data from complex samples, especially large-scale assessments such as the National Assessment of Educational Progress (NAEP) and TIMSS. Exploratory factor analysis (EFA) was used to identify the underlying structure of the items and reduce the number of observed indicators for each scale in this study.

Structural Equation Modeling

Structural equation modeling (SEM) is a statistical method that grows out of multiple regression, but takes a confirmatory approach to analyze a structural theory of the relationships of some variables of interest (Byrne, 1998). Causal relationships among theoretical variables are represented by a series of structural equations (regression equations), and the relationships are represented in a model. Then, the goodness of fit between the theory-based model and the data (usually the variance-covariance matrix) is tested statistically. If the goodness of fit is adequate, then the postulated model is plausible and consistent with the data; if the goodness of fit is poor, then the model is not plausible and needs to be rejected or re-specified (Byrne, 1998).

Structural equation models can be developed in five steps (Schumacker & Lomax, 2004). These steps are: 1) model specification, 2) model identification, 3) model estimation, 4) model testing, and 5) model modification.

Model specification is the specification of the relationships among latent factors. The causal links are developed based on the hypothesized model the researcher formulates based on theoretical hunches.

Model identification is "to ask whether unique values can be found for the parameters to be estimated in the theoretical model" (Schumacker & Lomax, 1996). Structural models may be just-identified, over-identified, or under-identified based on degree of freedom (the difference between the number of parameters of variances/covariances and the number of parameters to be estimated). When the degree of freedom is zero, the model is just-identified; when the degree of freedom is negative, the model is under-identified; when the degree of freedom is positive, the model is over-identified that allow rejection of the model. Only over-identified models can be tested for fit.

Once the model is identified, the next step is to estimate model parameters. The main focus of the estimation process is to yield parameter values such that the discrepancy between the sample covariance matrix and the population covariance matrix implied by the model is minimal. The three most commonly used estimation models are: Maximum Likelihood, Generalized Least Squares, and Asymptotic Distribution Free (Baloglu, 2000). In this study, we used the maximum likelihood method.

The next step after estimating model parameters is to test the model. Model fit is tested by comparing the predicted model covariance with the sample covariance matrix. A model is deemed fit if the hypothesized population covariance matrix is similar to the sample covariance matrix. In evaluating the overall goodness-of-fit for the SEM model, Schumacker and Lomax's (2004) criteria are used: (a) the chi-square and p value, which if $p > .05$ indicates that there are no statistically significant discrepancies between the sample variance-covariance matrix and the reproduced implied covariance matrices. As this statistic is very sensitive to sample size and departures from multivariate normality, it may easily reject a well-fitting model (Hatcher, 1994); (b) Goodness-of-Fit Index (GFI), Adjusted Goodness-of-Fit Index (AGFI) and Comparative Fit Index (CFI) values close to .95 reflect a good fit and 1.0 indicate a perfect fit; (c) Normed Fit Index (NFI), which defines the null model as a model in which all the correlations or covariance are zero and value close to .95 reflects a good model fit; (d) Root-Mean-Square Error of Approximation (RMSEA) value less than .05 indicates a good model fit. These fit indices indicate how well the data support the model.

The last step in structural equation modeling is model modification. If the fit indices suggest a poor fit or the modification indices strongly suggest certain changes the model is respecified, a few theoretically sound changes should be made.

The standard approach to estimate a SEM includes two steps (Anderson & Gerbing, 1988). In the first phase of the two-step model building approach, measurement models for all latent variables in the model are estimated. In SEM, researchers are interested in the causal relationship between latent variables and observed variables, which is called the measurement model. Measurement model is tested by confirmatory factor analysis, which is specified based on theory or empirical studies that generate a statistical representation about the relationships among latent and observed variables. In the measurement model, both dependent and independent latent variables are specified (Baloglu, 2000). The second step is the structural part of the SEM. This structural part specifies the relationships between the exogenous and endogenous variables.

Evaluating the measurement model

Besides item reliability, evaluation of the measurement model by composite reliability and average variance extracted (Fornell & Larcker, 1981) is suggested. Composite reliability ($\rho_{(\eta)}$) is a measure of the overall reliability of a latent variable. It is calculated by:

$$\rho_{(\eta)} = \frac{\left(\sum_{i=1}^{p} \lambda_i\right)^2}{\left(\sum_{i=1}^{p} \lambda_i\right)^2 + \sum_{i=1}^{p} Var\left(\varepsilon_i\right)}$$

Where λ indicates standardized loading and ε represents indicator measurement error.

Because composite reliability does not measure the amount of variance that is captured by the construct in relation to the amount of variance due to measurement error, the average variance extracted $\rho_{ve(\eta)}$ is also needed to evaluate the measurement model.

$$\rho_{ve(\eta)} = \frac{\sum_{i=1}^{p} \lambda_i^2}{\sum_{i=1}^{p} \lambda_i^2 + \sum_{i=1}^{p} Var(\varepsilon_i)}$$

The average variance extracted (AVE) varies from 0 to 1 and it represents the ratio of the total variance due to the latent variable. According to Bagozzi (1991), a variance extracted of greater than 0.50 indicates that the validity of both the construct and the individual variables is high.

Hypotheses testing

The main purpose of this study is to test the links between opportunity to learn, students' engagement in science and science achievement. Because engagement is a multidimensional construct, the engagement construct is separated into following three components with separate hypotheses. The first hypothesis tested the effect of behavioral engagement on science achievement. The second hypothesis investigated the links among OTL, students' behavioral engagement, and science achievement and whether students' behavioral engagement was a mediator between OTL and achievement. The third hypothesis tested the effects of emotional and cognitive engagement on science engagement.

Hypothesis 1: Students' science classroom activities (behavioral engagement) are positively related to science achievement.

Hypothesis 2

a: teacher reports of instructional activities (OTL) are related to students' science achievement. or

b: teachers reported instructional activities (OTL) are related to students' reported activities (behavioral engagement), and these, in turn, affect student science achievement.

Hypothesis 3: Students' science psychological engagement factors (emotional and cognitive engagement) positively influence science achievement, controlling for family socioeconomic background variables.

An overall conceptual model is tested based on the above research hypotheses. Family socioeconomics background factor and opportunity to learn factors are correlated and serve as exogenous variables. Emotional, cognitive, and behavioral engagement and science achievement are endogenous variables where three types of engagement are mediators in this study.

Hierarchical Linear Modeling

Another analytic strategy for the study is the use of HLM. The purpose is to examine the class level and school level variability of the U.S. 8^{th} graders' science achievement scores in 2003. It was hypothesized that there are school or teacher effects which contribute towards explaining achievement differences.

The Hierarchical Linear Model provides an integrated strategy for handling problems such as aggregation bias in standard error estimates and erroneous probability values in hypothesis testing of school effects. For this study, HLM is chosen as the software program appropriate to study school and student effects relating to student outcomes.

Research on school effects has previously been conducted with a set of data analyzed at the individual student level with the assumption that classrooms and schools affect students equally. However, when the effects vary among individuals and their contexts, this type of statistical analysis can be misleading (Raudenbush & Bryk, 2002). Ordinary least squares analysis provides information about the total variance, but can only break this total variance into the between- and within-school effects. The between-school effect may be influenced by school level variables, such as the affluence of the school. This study endeavored to explain variations in student outcomes by first decomposing observed relationships into between- and within school components.

Previous studies have shown clearly that educational researchers need to account for the inherent hierarchical structure of data collected from schools and this literature includes the work of Raudenbush and Bryk (2002). In this study, three-level HLM analyses were conducted. The first stage was an unconditional model. The unconditional model is the simplest HLM model and contains no predictor variables from any level (Raudenbush & Bryk, 2002). The fully unconditional model is used to estimate how much variance is attributed to the school level, class level and the student level. According to Lee (2000), if the proportion of variance that exists at a higher level of aggregation is more than 10% of the total variance in the outcome, then it is necessary to consider a HLM analysis. The analyses in this study use these guidelines.

The unconditional model

The use of the hierarchical linear model involved the single cross-section of data with a three-level structure consisting of students (Level 1) nested within classes (Level 2) nested within schools (Level 3).

The simplest model was used first, that is, the fully unconditional model with no predictor variables specified. The outcome measure, science achievement, was free to vary across three different levels of analysis: student, class and school. This model is described below.

Student-level model

Science Achievement for each student was estimated as a function of the class average plus random error:

$$\text{Ach}_{ijk} = \pi_{0jk} + e_{ijk}$$

where

Ach_{ijk} represents the Science Achievement of each student *i* in class *j* and school *k*.

π_{0jk} represents the class mean Science achievement of class *j* in school *k*.

e_{ijk} represents the random error of student *i* in class *j* and school *k*

$i = 1, 2, 3, \ldots, n_{jk}$ students in class *j* and school *k*.

$j = 1,2, \ldots J_k$ classes within school *k*,

$k = 1, \ldots, K$ schools.

Class-level model

Science achievement classroom mean varies as a function of the school mean plus random error:

$$\pi_{0jk} = \beta_{00k} + r_{0jk}$$

where

β_{00k} represents the mean Science achievement in school *k*.

r_{0jk} represents the random error of class *j* within school *k*

School-level model.

Science school mean achievement varies randomly around a grand mean for all schools.

$$\beta_{00k} = \gamma_{000} + \mu_{00k}$$

where

γ_{000} represents the grand mean Science achievement for all schools.

μ_{00k} represents the random school effect, the deviation of school *k's* mean from the grand mean.

This three-level model partitions the total variability in the outcome measure, Science achievement, into its three components: students within teachers, teachers within schools and between schools. If these random effects are significant, they suggest that a hierarchical linear model of teachers and schools level is necessary.

Conditional model

The results of structural equation modeling did not indicate the significant direct effect of OTL measured by teachers' instructional activities on science achievement. Some domains of OTL, such as content coverage and content emphasis are taken into account in the HLM model.

Student-level

The level-1 model added engagement factors as predictors.

$$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}\,(ENGAGEMENT) + e_{ijk}$$

Class-level

The level-2 model included the teachers' characteristics, which represent OTL. For example: whether the science teacher had a full license or certificate, the number of students per class, and the percentage of students taught science topics.

$$\pi_{ijk} = \beta_{i0k} + \beta_{i1k}\,(LICENSE) + \beta_{i2k}\,(TOPICS) + r_{0jk}$$

where:

LICENSE: this is a dichotomous variable that asks science teachers: “Do you have a teaching license or certificate?” 1=Yes, 0=No.

TOPICS is a composite derived from the teachers’ survey. Teachers were asked to indicate whether and when students in the TIMSS class had been taught each topic. The options were: 1) Mostly taught before this year; 2) Mostly taught this year; and 3) Not yet taught or just introduced. The composite is computed as the percentage of students whose teachers CHECKED option 1 or 2 for each individual topic. For example, if there were a total of 100 science items on the TIMSS questionnaire and a teacher checked 80 items that were taught this year or last year, the percentage would be 80%. The scale of this composite is from 0 to 100%.

School-level

Level-3 model took school variables into account to investigate the OTL. Those school level OTL variables were: group ability; offer science enrichment/remedial; and percentage of free/reduced lunch.

$$\beta_{ijk} = \gamma_{ij0} + \gamma_{ij1}(\text{ENRICHMENT/REMEDIAL}) + \gamma_{ij2}\ (\text{REDUCED LUNCH}) + \mu_{ijk}$$

where

ENRICHMENT/REMEDIAL is a dichotomous composite created by “Does your school offer enrichment science for students in the eighth-grade?” and “Does your school offer remedial science for students in the eighth-grade?” 1=Yes, 0=No. The response will be “Yes” (1) if the answer for either of these two items is “Yes;” otherwise, the response is “No” (0).

REDUCED LUNCH is a US national variable which represents the percentage of students receiving a free/reduced lunch. The scale is 0 to 100.

Since not all predictors or effects were significant, in order to get parsimonious results, the conditional models were built by three steps following Raudenbush and Bryk (2002) and Singer and Willett's (2003) instruction of model specification.

Step 1: Start small at level 1, adding one variable at a time. Add each predictor after all others are removed, to avoid collinearity.

Step 2: Add level 2 predictors together and check p-values. Remove insignificant effects, first from the slopes and then the intercepts. The deviance difference test is used to evaluate the model changes.

Step 3: Add level 3 predictors based on research hypotheses and check p-values.

To sum, this chapter outlines the methods used in the study. It provides an overview of dataset, sample of the study, items and factors in the study. It lays out the analytic approaches taken to estimate the models.

CHAPTER FOUR

ANALYSIS AND RESULTS

This chapter presents the results of data analysis and hypothesis testing. The chapter is divided in three sections. The first section outlines descriptive statistics and correlations for all variables of this study. The second section of the chapter presents the structural equation modeling analyses (SEM). And the third section presents the results of the Hierarchical Linear Modeling (HLM) analyses. Both a brief statement about the analyses and the results are provided for each hypothesis test and the overall model. This part is presented in a brief narrative to allow the reader to follow the analytical decisions made sequentially at each step of the analysis.

Descriptive Statistics

TIMSS is one of the largest surveys of students' science learning, containing numerous indicators in social, psychological and educational domains. To understand the data in detail, descriptive statistics for each of the study variables, including mean, standard deviation and correlation matrices of science psychological items, students' classroom activities, teachers' instructional activities and family background, are shown below. All results presented in this chapter are based on descriptive data analyses, where the sample was weighted by SCIWGT, which is "computed by dividing the sampling weight for the student by the number of science teachers that the student has" (Martin, 2005). SCIWGT is an appropriate weight to analyze student and teacher data together.

Science Psychological Items

There were 12 items that pertain to psychological dimension of science learning. These are items that capture students' interest in science and how much they value science learning. The Cronbach's alpha reliability estimate of all 12 psychological items was .88. The item "Sometimes, when I do not initially understand a new topic in science, I know that I will never really understand it." was found to be factorially complex, loading on two factors, and the variance extracted was less than 0.5, thus the item was deleted. The Cronbach's alpha increased to .883 for the scale. So we kept other 11 items in the follow-up analysis. Some items were recoded in order that higher score indicated positive attitude toward science. Detailed descriptive statistics and correlation matrix are shown on Table 1.

Science Classroom Behaviors

The TIMSS survey contained many items about the classroom activities in science. Several student classroom activity items violated the linear assumption and had seriously skewed distributions. Some examples are, "How often do you watch the teacher demonstrate an experiment or investigation in your science lessons?" or "how often do you have a quiz in your science lessons?" We deleted those items and kept 9 indicators to represent student class behavior. These items had acceptable distributions (See Table 2 for details).

Opportunity to Learn Items (Teachers' Instructional Activities)

Corresponding to students' survey items, similar questions were asked of science teachers about the classroom instructional activities and the frequency of those activities. Since what the teachers do in the classroom provide the opportunity to learn, these items were considered as reflecting the teachers' instructional practices. These items capture what kinds of opportunity teachers provide students to learn science. Two items were deleted after preliminary reliability analysis due to very low reliabilities of these items (Detailed descriptive statistics and correlation matrix are shown on Table 3). After deleting items with low reliabilities we kept 9 items that reflected the teachers' instructional activities.

Table 1

Descriptive statistics and Correlations for Psychological Items

	Scale and item	*M*	*SD*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*	*10*	*11*
1	I usually do well in science.	3.22	.829	1										
2	I would like to take more science in school.	2.73	1.034	.54**	1									
3	Science is more difficult for me than for many of my classmates._R	2.99	.979	.44**	.26**	1								
4	I enjoy learning science.	2.95	.982	.57**	.71**	.28**	1							
5	Science is not one of my strengths._R	2.74	1.073	.51**	.40**	.55**	.43**	1						
6	I learn things quickly in science.	2.95	.931	.65**	.55**	.42**	.59**	.50**	1					
7	I think learning science will help me in my daily life.	2.98	.882	.38**	.50**	.15**	.50**	.25**	.39**	1				
8	I need science to learn other school subjects.	2.73	.922	.30**	.43**	.08**	.42**	.18**	.32**	.61**	1			
9	I need to do well in science to get into the <university> of my choice.	3.17	.912	.30**	.38**	.12**	.36**	.20**	.30**	.49**	.48**	1		
10	I would like a job that involved using science.	2.50	1.087	.38**	.55**	.20**	.52**	.33**	.42**	.52**	.46**	.50**	1	
11	I need to do well in science to get the job I want.	2.73	1.085	.30**	.42**	.14**	.40**	.23**	.32**	.51**	.47**	.59**	.73**	1
	4= agree a lot; 3= agree a little; 2= disagree a little; 1= disagree a lot													

N=8020

Note: **, p<.01; *, p<.05

Table 2

Descriptive statistics and Correlations for Students' Class Activities (Behavioral Engagement)

	Scale and item	*M*	*SD*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
1	conduct an experiment or investigation	2.71	.953	1								
2	work in small groups on an experiment or investigation	2.92	.962	.65**	1							
3	write explanations about what was observed and why it happened	2.91	.949	.61**	.62**	1						
4	study the impact of technology on society	2.40	.950	.39**	.33**	.41**	1					
5	relate what you are learning in science to your daily life	2.59	.994	.34**	.29**	.36**	.53**	1				
6	present your work to the class	2.30	.997	.33**	.30**	.32**	.40**	.41**	1			
7	listen to the teacher give a lecture-style presentation	3.07	.979	.24**	.23**	.26**	.27**	.27**	.23**	1		
8	work problems on you own	3.07	.882	.19**	.15**	.23**	.23**	.24**	.18**	.30**	1	
9	begin your homework in class	2.73	1.081	.15**	.14**	.17**	.24**	.24**	.18**	.22**	.37**	1
	4=every or almost every lesson; 3=about half the lessons; 2=some lessons; and 1=never											

(N=8082)

Note: **, *p*<.01; *, *p*<.05

Table 3

Descriptive Statistic and Correlations for Teachers' Instructional Activities (OTL)

	Scale and item	*M*	*SD*	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	*9*
1	ask them to formulate hypotheses or predictions to be tested	2.61	.786	1								
2	ask them to design or plan experiments or investigations	2.28	.665	.55**	1							
3	ask them to conduct experiments or investigations	2.66	.780	.44**	.52**	1						
4	ask them to work together in small groups on experiments or investigations	2.88	.780	.37**	.37**	.66**	1					
5	ask them to write explanations about what was observed and why it happened	2.80	.801	.56**	.49**	.67**	.66**	1				
6	ask them to put events or objects in order and give a reason for the organization	2.37	.728	.30**	.36**	.25**	.20**	.38**	1			
7	ask them to study the impact of technology on society	2.21	.692	.26**	.29**	.09**	.09**	.24**	.44**	1		
8	ask them to learn about the nature of science and inquiry	2.76	.808	.40**	.43**	.37**	.28**	.43**	.42**	.45**	1	
9	ask them to relate what they are learning in science to their daily lives	3.20	.777	.36**	.28**	.35**	.30**	.38**	.27**	.36**	.50**	1

4=every or almost every lesson; 3=about half the lessons; 2=some lessons; and 1=never

(N=6672)

Note: **, *p*<.01; *, *p*<.05

Family Background (SES)

There was no existing socio-economic status composite in TIMSS data. We used data on parents' education level and number of books in home to create a measure of family SES. Measures of parents' income and occupation were not included in TIMSS. (See Table 4 for details).

Table 4

Descriptive Statistics and Correlations for Family Background

	Scale and item	*M*	*SD*	*1*	*2*	*3*
1	Number of books in your home	3.30	1.317	1		
2	Highest level of education\Mother	5.53	1.985	0.36**	1	
3	Highest level of education\Father	5.63	2.002	0.40**	0.59**	1

(N=5557)
Note: **, p<.01; *, p<.05

Exploratory Factor Analysis

Exploratory factory analysis is a method that suggests the number of underlying factors based on empirical analysis. Principal components method with VARIMAX rotation was used to extract factors for each scale in this study. Each analysis was run separately. Screen plots and total explained variance (Table 5) were used to determine the number of factors. The results of the final exploratory factor analysis of each scale are presented in Table 6, Table 7, Table 8, and Table 9. The results of the analysis showed that science psychological engagement items can be reduced to two factors. Engagement in classroom activities can be explained by three factors. Two factors were extracted for teachers' instructional activities (OTL) and one single factor can explain family background. For each scale, more than 60% of total variance was explained by the extracted factors.

Table 5

Results of Exploratory Factor Analyses for Scales

Scale	Number of items	Factors extracted	Explained variance	Total explained variance
Psychological dimension	11	2	47.13%	
			15.24%	62.37%
Students' classroom activities	9	3	39.32%	
			14.44%	
			10.66%	64.42%
Opportunity to learn	9	2	45.69%	
			15.46%	61.15%
Family background	3	1	63.54%	63.54%

Table 6

EFA for Science Psychological Indicators

Observed indicators	Factors	
	Value of science (Cognitive engagement)	Interest/efficacy (Emotional engagement)
I usually do well in science.	.289	***.765***
I would like to take more science in school.	.564	***.544***
Science is more difficult for me than for many of my classmates._R	-.059	***.758***
I enjoy learning science.	.527	***.595***
Science is not one of my strengths._R	.101	***.789***
I learn things quickly in science.	.328	***.743***
I think learning science will help me in my daily life.	***.758***	.214
I need science to learn other school subjects.	***.752***	.096
I need to do well in science to get into the <university> of my choice.	***.747***	.087
I would like a job that involved using science.	***.759***	.283
I need to do well in science to get the job I want.	***.803***	.118

Table 7

EFA for Science Class Activities Indicators

Observed indicators	Factors		
	Scientific investigation	Innovation/impact of science	Homework
conduct an experiment or investigation	***.830***	.230	.110
work in small groups on an experiment or investigation	***.865***	.155	.077
write explanations about what was observed and why it happened	***.800***	.245	.160
study the impact of technology on society	.247	***.744***	.184
relate what you are learning in science to your daily life	.157	***.789***	.197
present your work to the class	.190	***.725***	.082
listen to the teacher give a lecture-style presentation	.202	.216	***.550***
work problems on your own	.101	.066	***.805***
begin your homework in class	.012	.139	***.751***

Table 8

EFA for OTL Indicators

Observed indicators	Factors	
	Scientific investigation Instruction	Impact of science
ask them to formulate hypotheses or predictions to be tested	***.591***	.400
ask them to design or plan experiments or investigations	***.585***	.424
ask them to conduct experiments or investigations	***.870***	.109
ask them to work together in small groups on experiments or investigations	***.841***	.023
ask them to write explanations about what was observed and why it happened	***.814***	.291
ask them to put events or objects in order	.170	***.693***

and give a reason for the organization		
ask them to study the impact of technology on society	-.047	***.829***
ask them to learn about the nature of science and inquiry	.318	***.719***
ask them to relate what they are learning in science to their daily lives	.321	***.574***

Table 9

EFA for Family Background Indicators

Observed indicators	SES
Number of books in your home	.702
Highest level of education\Mother	.831
Highest level of education\Father	.849

Structural Equation Modeling

Overview of the Methodology

Structural equation modeling (SEM) is an especially appropriate method for analyzing non-experimental data. In addition to parameter estimates, the programs such as *LISREL* provide fit indices to assess how well the model fits the data. Such fit indices make it possible to evaluate the adequacy of the theoretical model in explaining the data (Bollen, 1989; Schumacker & Lomax, 2004). In this study, we estimated a latent variable structural equation model in several steps using LISREL 8.8 computer program (Jöreskog & Sörbom, 2006). First, the measurement of each scale of the model (constructs and their indicators) was specified and estimated; second, the full measurement model was tested and modified; third, structural relationships in the model were specified and estimated; finally, some appropriate changes were made according to estimated modification indices.

Measurement Model

Following the results of the exploratory factor analysis, we conducted confirmatory factor analyses to test the measurement models and assess the construct validity of the various latent constructs. The model was constructed in three steps. First, the model for each construct was tested separately. Then, the full model was tested for all constructs simultaneously. Last, we evaluated the full measurement model following Fornell and Larcker's (1981) criteria. All models were estimated using the maximum likelihood method.

Measurement of the psychological dimensions in science learning

Table 10

Goodness-of-Fit Summary Table for Measurement Models of Psychological Items

	χ^2	*df*	$\Delta\chi^2$	Δdf	CFI	GFI	AGFI	RMSEA	SRMR
Initial Model	3126.61	43			.92	.85	.76	.15	.08
TD(5,3)	2421.10	42	705.51	1	.94	.88	.81	.13	.07
TD(11,10)	1854.54	41	566.56	1	.95	.90	.84	.12	.06
TD(6,1)	1491.65	40	362.89	1	.96	.93	.88	.10	.06
TD(8,7)	1288.53	39	203.12	1	.97	.94	.90	.09	.05
TD(11,9)	1061.74	38	226.79	1	.97	.95	.91	.08	.05

The first six indicators (I usually do well in science; like to take more science in school; Science is more difficult for me than for many of my classmates; enjoy learning science; Science is not one of my strengths; and I learn things quickly in science) were significantly loaded on the first latent variable while the rest five indicators (I think learning science will help me in my daily life; I need science to learn other school subjects; I need to do well in science to get into the <university> of my choice; I would like a job that involved using science; and I need to do well in science to get the job I want) were loaded on the second latent variable. The first factor reflected students' emotional engagement in learning science by measuring interest and efficacy in science. The second factor represented cognitive engagement by measuring the future value of science learning. To get the better fitting model (CFI=.97, GFI=.95, AGFI=.91, RMSEA=.05), five pairs of covariance of indicators were set free. Correlating the errors made sense because these items were conceptually related and their errors had some common variances. Chi-square difference test showed the model significantly improved at each step (Table 10). Bollen (1989) suggested 3 to 4 indicators are the best number of indicators for each latent construct, so the new composites of correlated indicators were created to simplify and optimize the model.

Measurement of classroom activities

Students' classroom learning activities were conceptualized as behavioral engagement in learning. The measurement model of student classroom activities by three related factors was well fitting (GFI=.98, CFI=.98, AGFI=.97, and RMSEA=.05, SRMR=.04). The first three indicators (Students conduct an experiment or investigation; work in small groups on an experiment or investigation; and write explanations about what was observed and why it happened) were measures of the first factor which represented core activities of scientific investigation. The 4^{th} to 6^{th} items (study the impact of technology on society; relate what learned in science to daily life; and present work to the class) were measures of the second factor which can be explained as reflecting the connection of science to society and daily life. And the last three indicators (listen to the teacher give a lecture-style presentation; work problems on the own; and begin homework in class) were measures of the factor three which mainly related to routine class work and homework.

Measurement of teachers' instructional activities (OTL)

Table 11

Goodness-of-Fit Summary Table for Measurement Models of OTL

	χ^2	*df*	$\Delta\chi^2$	Δdf	CFI	GFI	AGFI	RMSEA	SRMR
Initial Model	1769.54	26			.93	.90	.83	.13	.08
TD(7,6)	1573.21	25	196.33	1	.93	.91	.85	.13	.07
TD(4,3)	1360.93	24	212.28	1	.94	.93	.87	.12	.07
TD(5,2)	963.39	23	397.54	1	.96	.95	.90	.10	.06
TD(5,4)	807.29	22	156.10	1	.97	.96	.91	.09	.05
TD(5,3)	563.35	21	243.94	1	.98	.97	.94	.08	.04

Science teachers' instructional activities in the classroom represented one way of measuring Opportunity to Learn (OTL) in this study. What teachers do in the classroom provides students with opportunities to master the content of the course. The analysis showed two dimensions of the OTL this model. The first latent variable of OTL was measured by five observed indicators: "teachers ask students to formulate hypotheses or predictions to be tested"; "ask students to design or plan experiments or investigations"; "ask them to conduct experiments or investigations"; "ask them to work together in small groups on experiments or investigations"; and "ask them to write explanations about what was observed and why it happened". These reflected core science teaching activities. The second latent variable was measured by four items: "ask students to put events or objects in order and give a reason for the organization"; "ask them to study the impact of technology on society"; "ask them to learn about the nature of science and inquiry"; "ask them to relate what they are learning in science to their daily lives". The second factor reflected more diverse activities, mostly related to the nature and impact of science on society. The measurement of OTL was similar to the measurement of students' report of classroom activities since many indicators were highly correlated. Several nested models were tested using chi-square difference tests (Table 11) to make the model fit better (GFI=.97, CFI=.98, AGFI=.94, and RMSEA=.08). New composites of average of correlated items were created for future analyses.

Measurement of family background (SES)

There were only three measures of the socio-economic status of the family in the TIMSS data. Thus, the measurement model of family background was a just identified model which means degree of freedom was 0 and other fit indices were not available. Family background was measured by three indicators (father' education, mother's education and number of book in home) and all three loadings were significant.

Measurement of science achievement

Science achievement was measured by National Science Rasch Score. National Rasch score was computed by standardizing the science logits to create logit scores with a weighted mean of 150 and a weighted standard deviation of 10 within each country (Sofroniou & Kellaghan, 2004). It was computed to facilitate preliminary item analyses that were conducted prior to the IRT scaling. Because only US data were investigated in this study, Rasch score was used as a criterion variable in studies of item discrimination in a single country.

Full measurement model

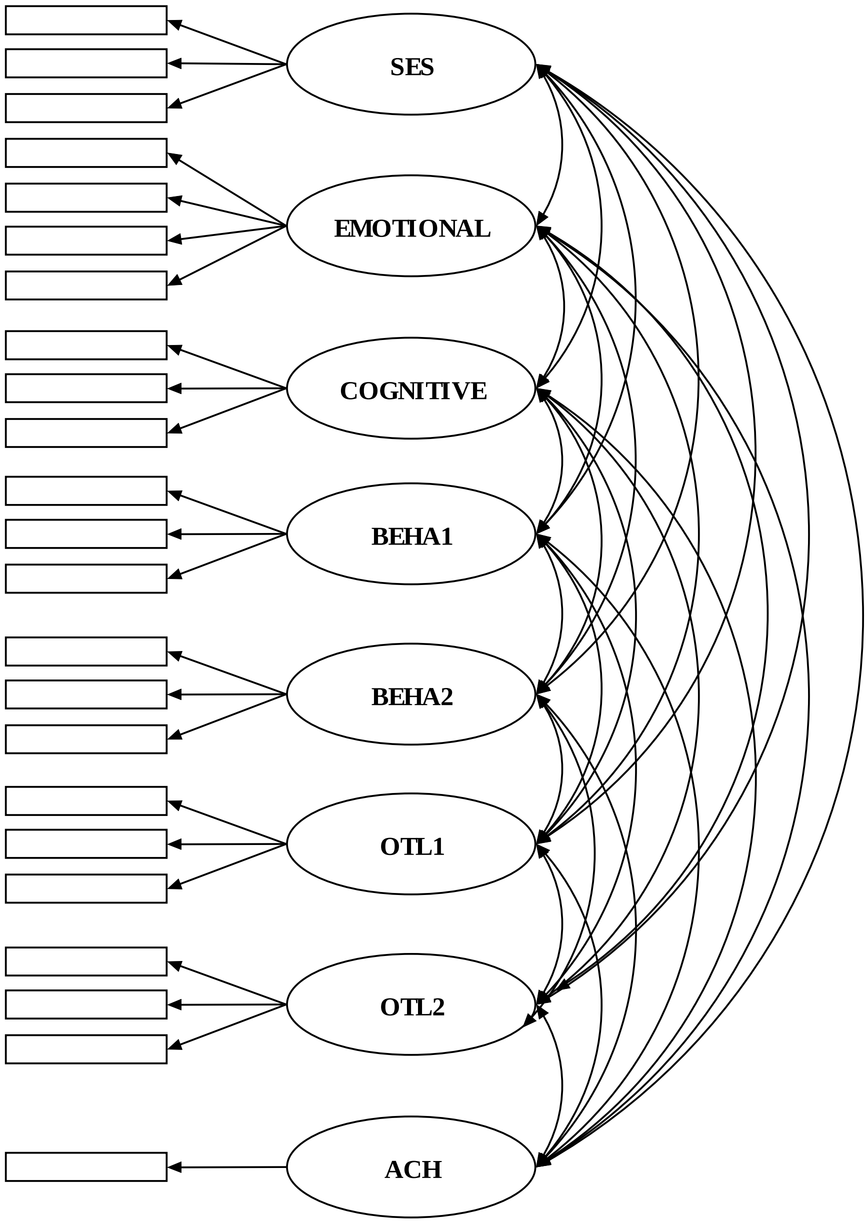

Figure 2. Full measurement model

Table 12

Overall Item Correlation Matrix

	ACH	SES1	SES2	SES3	ACT1	ACT2	ACT3	ACT4	ACT5	ACT6	TACT1	TACT2
ACH	1.00											
SES1	0.39	1.00										
SES2	0.27	0.36	1.00									
SES3	0.32	0.40	0.59	1.00								
ACT1	0.07	0.10	0.14	0.13	1.00							
ACT2	0.05	0.09	0.12	0.12	0.64	1.00						
ACT3	0.02	0.07	0.10	0.10	0.62	0.61	1.00					
ACT4	-0.09	-0.03	0.00	0.01	0.39	0.31	0.39	1.00				
ACT5	-0.04	0.02	0.01	0.03	0.31	0.25	0.32	0.53	1.00			
ACT6	-0.09	-0.02	0.01	0.01	0.32	0.28	0.31	0.40	0.40	1.00		
TACT1	0.08	0.10	0.06	0.07	0.16	0.14	0.15	0.07	0.06	0.04	1.00	
TACT2	0.04	0.06	0.02	0.02	0.15	0.14	0.13	0.08	0.05	0.06	0.53	1.00

Table 12 (Continued)

	ATCT3	TACT4	TACT5	TACT6	PSY1	PSY2	PSY3	PSY4	PSY5	PSY6	PSY7	ATCT3
ATCT3	0.08	0.10	0.09	0.11	0.30	0.31	0.24	0.06	0.03	0.07	0.55	0.49
TACT4	-0.09	-0.06	-0.04	-0.06	-0.01	-0.01	-0.02	0.04	0.03	0.02	0.31	0.38
TACT5	-0.02	-0.02	0.00	-0.01	0.08	0.08	0.08	0.07	0.05	0.03	0.38	0.44
TACT6	0.00	0.02	0.01	0.01	0.04	0.04	0.04	0.03	0.05	0.00	0.36	0.28
PSY1	0.16	0.17	0.14	0.14	0.14	0.12	0.12	0.16	0.24	0.14	0.03	0.03
PSY2	0.13	0.13	0.10	0.09	0.15	0.14	0.15	0.17	0.25	0.13	0.03	0.02
PSY3	0.24	0.20	0.17	0.18	0.14	0.15	0.11	0.11	0.19	0.11	0.03	0.04
PSY4	0.25	0.19	0.14	0.16	0.03	0.04	0.02	-0.02	0.07	0.01	0.04	0.03
PSY5	0.19	0.15	0.13	0.13	0.14	0.12	0.14	0.12	0.20	0.11	0.04	0.03
PSY6	0.11	0.13	0.10	0.10	0.19	0.15	0.17	0.25	0.34	0.21	0.03	0.03
PSY7	0.18	0.15	0.13	0.12	0.17	0.12	0.11	0.19	0.24	0.15	0.03	0.04

Table 12 (Continued)

	ATCT3	TACT4	TACT5	TACT6	PSY1	PSY2	PSY3	PSY4	PSY5	PSY6	PSY7
ATCT3	1.00										
TACT4	0.24	1.00									
TACT5	0.38	0.49	1.00								
TACT6	0.36	0.36	0.50	1.00							
PSY1	0.03	-0.02	0.01	0.01	1.00						
PSY2	0.02	-0.01	0.01	0.01	0.72	1.00					
PSY3	0.02	-0.02	-0.01	0.00	0.59	0.62	1.00				
PSY4	0.00	0.01	0.01	0.03	0.39	0.41	0.60	1.00			
PSY5	0.02	-0.02	0.01	-0.02	0.39	0.36	0.31	0.18	1.00		
PSY6	0.02	0.00	0.01	0.01	0.54	0.52	0.41	0.21	0.55	1.00	
PSY7	0.03	0.02	0.03	0.01	0.54	0.51	0.41	0.28	0.59	0.59	1.00

After creating new composites by taking average of highly correlated indicators, the full hypothesized measurement model was tested (Figure 2). Overall correlation matrix of all scales is listed at Table 12. For each latent variable, the first indicator was set to a value of 1.0, and thus *t* scores were not computed. The fit indices for the full measurement model were high, indicating a well-fitting model in which data fit well to the hypothesized model, $\chi^2(263, N=4107) = 2294.65, p<.05$. The goodness-of-fit index (GFI) is .96, and the adjusted goodness-of-fit (AGFI) index is .94. The comparative fit index (CFI) was 97. The root-mean-square error of approximation (RMSEA) is .043 and the standardized root mean square error of approximation (SRMR) was also less than .036. Overall, these fit indices indicate a theoretically sound model that explained the data well.

Table 13

Standardized Loading, Reliability, and Validity of the Final Measurement Model

Label	Construct and indicators	Standardized loading	*t*	Reliability	Variance extracted estimate
	Family Background (SES)			0.72	0.47
SES1	Number of Books In Home	0.54	–	0.29	
SES2	mother's education	0.72	28.52	0.52	
SES3	father's education	0.78	28.54	0.61	
	Emotional Engagement (Interest/Efficacy in Science)			0.82	0.54
PSY1	I would like to take more science in school.	0.84	–	0.71	
PSY2	I enjoy learning science.	0.85	58.32	0.72	
PSY3	I usually do well in science. + I learn things quickly in science.	0.71	48.38	0.50	

PSY4	Science is more difficult for me than for many of my classmates. _R + Science is not one of my strengths. _R	0.46	28.85	0.21	
	Cognitive Engagement (Value of Science)			*0.80*	*0.57*
PSY5	I need to do well in science to get into the <university> of my choice.	0.69	–	0.48	
PSY6	I think learning science will help me in my daily life. + I need science to learn other school subjects.	0.78	42.12	0.61	
PSY7	I would like a job that involved using science. + I need to do well in science to get the job I want.	0.8	42.75	0.64	
	Behavioral Engagement I			0.82	0.62
ACT1	conduct an experiment or investigation	0.81	–	0.66	
ACT2	work in small groups on an experiment or investigation	0.78	48.84	0.61	
ACT3	write explanations about what was observed and why it happened	0.78	48.85	0.61	
	Behavioral Engagement II			0.71	0.45
ACT4	study the impact of technology on society	0.73	–	0.53	
ACT5	relate what you are learning in science to your daily life	0.71	36.35	0.50	
ACT6	present your work to the class	0.57	30.79	0.32	
	Behavioral Engagement III			0.55	0.3
ACT7	listen to the teacher give a lecture-style presentation	0.45	–	0.20	

ACT8	work problems on your own	0.62	19.03	0.38	
ACT9	begin your homework in class	0.55	18.58	0.30	
	Opportunity to Learn I			0.76	0.52
TACT1	ask them to formulate hypotheses or predictions to be tested	0.73	−	0.53	
TACT2	ask them to design or plan experiments or investigations	0.7	38.30	0.49	
TACT3	ask them to conduct experiments or investigations + ask them to work together in small groups on experiments or investigations + ask them to write explanations about what was observed and why it happened	0.74	39.95	0.55	
	Opportunity to Learn II			0.72	0.46
TACT4	ask them to put events or objects in order and give a reason for the organization + ask them to study the impact of technology on society	0.61	−	0.37	
TACT5	ask them to learn about the nature of science and inquiry	0.79	32.53	0.62	
TACT6	ask them to relate what they are learning in science to their daily lives	0.63	29.91	0.40	
	Science Achievement			1.00	1
ACH	National Science Rasch Score	1	−	1.00	

The standardized item loadings, item reliability, construct reliability and variance extracted estimate were evaluated according to Fornell and Larcker's (1981) criteria. A variance extracted of greater than 0.50 indicates that the validity of both the construct and the individual variable is high. For three of nine latent constructs (SES, behavioral engagement II and OTL II), the estimates were slightly lower than .50. But it is important to note that this variance extracted estimate test is very conservative. Given the significant factor loadings and high reliabilities, the constructs were retained in the final measurement model. But variance extracted estimate of the behavioral engagement III is only 0.30. That means there was 70% error variance in that latent construct in measurement model, so this unreliable latent construct was not included in future analyses.

Structural Models

Testing the hypotheses

Hypothesis 1: Science classroom activities (behavioral engagement) are related to science achievement.

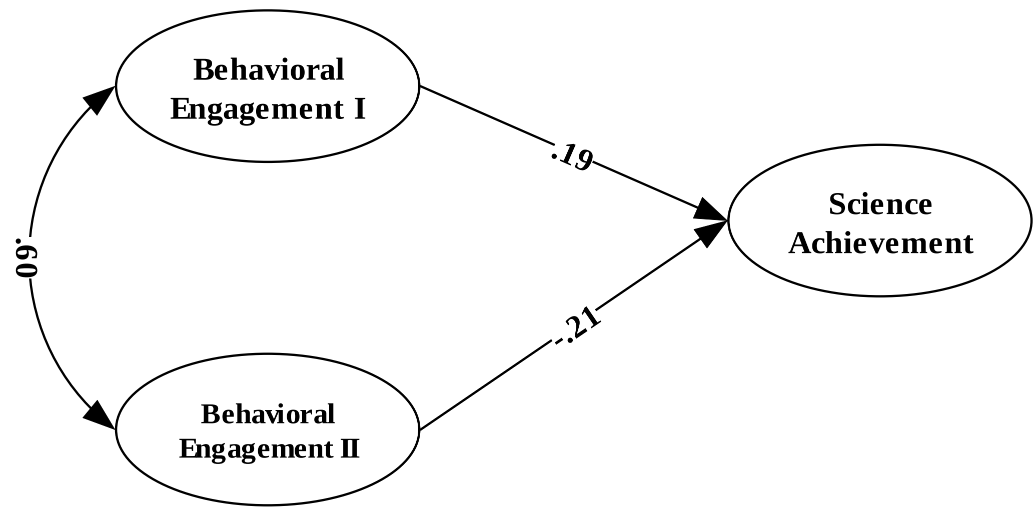

Figure 3. Structural model of behavioral engagement on science achievement (hypothesis 1)

The fit indices for the model were high, indicating a well-fitting model in which data fit well to the final model. The goodness-of-fit index (GFI) was .99, and the adjusted goodness-of-fit (AGFI) index was .98. The comparative fit index (CFI) was .99. The root- mean square error of approximation (RMSEA) was .05. Overall, these fit indices indicate a theoretically sound model that explained the data well. Figure 3 illuminates the effects of behavioral engagement on science achievement. Both of these two behavioral engagement factors were significantly related to science achievement but the directions were different. The first behavioral engagement factor, activities related to core scientific investigation (i.e., conduct experiment in the class) had a significant positive effect on science achievement score ($\gamma = 0.19$) while the other behavioral engagement factor which reflected the connection of science to society (i.e., related science learning to daily lives) negatively influenced science achievement score ($\gamma = -0.21$). One explanation can be that the more time students spent in thinking about science and daily lives, the less time he/she had to spend on the core learning activities, hence a negative effect on achievement. However, only a very small proportion of variance (R^2 = .03) was explained by this model. There were some other important factors related to science achievement which were not part of the model. In a second model, the teachers' instructional activities (OTL) were analyzed to assess their effect on science achievement score..

Hypothesis 2a: Teacher reports of instructional activities are related to students' science achievement.

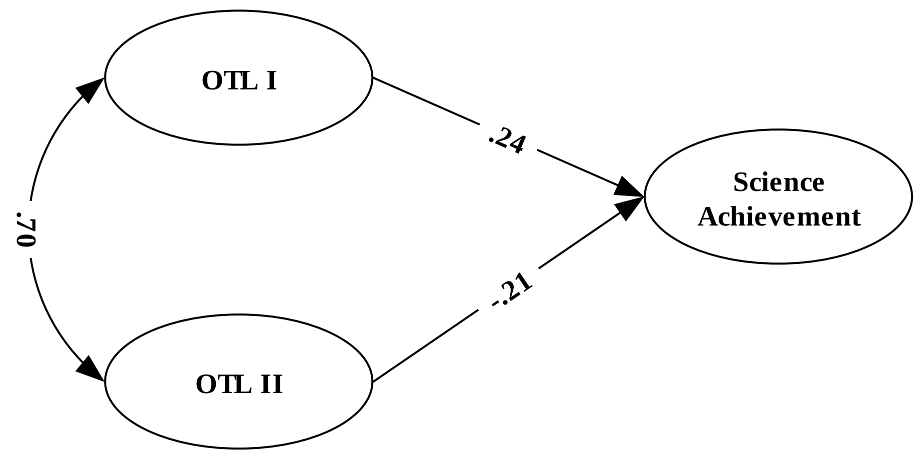

Figure 4. Structural model of opportunity to learn on science achievement (hypothesis 2a)

The model of effects of OTL (teachers' instructional activities) on student science achievement was also fitted well. The goodness-of-fit index (GFI) was .98, and the adjusted goodness-of-fit (AGFI) index was .95. The comparative fit index (CFI) was .98. The standardized root mean square error of approximation was .07. Similar to the effect of students' class activities, the effect of teachers core scientific activities (i.e., ask students design or plan experiments) on achievement was positive ($\gamma = 24$). But when teachers do more activities related to nature of science and connect science to society (i.e., ask them learn about nature of science an inquiry), the relationship is negative ($\gamma = -0.21$).The measures of this OTL factor included three indicators: teachers ask students to "learn about the nature of science and inquiry"; "related learning in science to daily lives" and the composite of the following two items: teachers ask students to "put event or objects in order and give a reason for the organization" and "study the impact of technology on society". The Pearson correlation coefficients of these indicators and science achievement were: -.02, .00, and -.09 respectively. It is clear that these items had no or very small negative association with science achievement. This implies that the student's science achievement score is not associated to these kinds of activities teachers asked students to do in class. In other words, the students' achievement score tended to be lower when teachers frequently engaged in many diverse but less focused activities such as showing students the impact of science on daily life (Figure 4). Another interpretation of this negative effect is the possible mismatch between the content of the course and the focus of the achievement test. The achievement test was probably designed to measure content knowledge of science and did not emphasize knowledge of the application of science. In sum, this result was unexpected and we can only speculate on the reasons for it.

Table 14

Standardized Direct, Indirect, and Total Effects of OTL and Behavioral Engagement on Science Achievement (Hypothesis 2b)

	Students' science engagement						Science achievement		
	Behavioral engagement I			Behavioral engagement II					
	direct	indirect	total	direct	indirect	total	direct	indirect	total
Exogenous variables									
OTL I	.60*		.60*	.18*		.18*		.05*	.05*
OTL II	-.34*		-.34*					-.04*	-.04*
Endogenous variables									
Behavioral engagement I							.14*		.14*
Behavioral engagement II							-.16*		-.16*

Note: *, $p<.05$

Hypothesis 2b: Teachers' reported factors are related to students' reported factors, and, in turn, affect student science achievement. Students' engagement in learning mediates the effect of OTL factors on achievement.

The fit indices for the model 2b are all in acceptable range (GFI=.94, CFI= .93, AGFI= .91, RMSEA=.08, and SRMR = .09). These fit indices indicate a theoretically sound model that explained the data well.

This model tests the hypothesis that students' classroom behavioral engagement in science class activities acts as a mediator between OTL and science achievement. There was 4% variance of science achievement explained by this model. All path coefficients on Figure 5 and Table 14 are significant. Teacher's instruction when it emphasized the methods of scientific investigation such as conducting experiment etc. had significant positive effects on both factors of students' engagement in science classroom. The effect on students' scientific activities ($\gamma = .60$) was much stronger than the effect on students' activities related to science's connection to society ($\gamma = .18$). The link between teacher's instructional emphases on science's relation to society (connection to society) to students' engagement in scientific activities was negative ($\gamma = -.34$). The interpretation is that if the teacher spends more time in doing the activities that focus on the relationship of science to society, students have less time to do core science activities such as conducting experiments and writing scientific explanations. No significant effect of teacher's instructional emphasis on science's relation to society on the second factor of students' engagement was found. Table 15 presents the correlations of the same questions that both students and teachers were asked. It was surprising that students and teachers' responses to science class activities were not highly correlated although the correlation coefficients were significant due to the large sample size. The correlation coefficients of teachers' and students' response to "how often study the impact of technology on society" and "how often relate learning in science to your daily life" were .04 and .06

respectively. The weak correlations confirmed that there was no significant effect of the second OTL factor on the students' second behavioral engagement factor.

Table 15

Correlations between Students and Teachers' Responses

	Teachers ask student to				
	conduct experiments or investigations	work together in small groups on experiments or investigations	write explanations about what was observed and why it happened	study the impact of technology on society	relate what they are learning in science to their daily lives
Students					
conduct an experiment or investigation	0.29**				
work in small groups on an experiment or investigation		0.26**			
write explanations about what was observed and why it happened			0.17**		
study the impact of technology on society				0.04**	
relate what you are learning in science to your daily life					0.06**

Note: **, *p*<.01

Science achievement was directly influenced by both classroom engagement factors which were students' reports of what they do in the class and how frequently. The effect of engagement in classroom activities that were related to methods of science and scientific inquiry was positive ($\beta = .14$) while the effect of the second factor, engagement in activities related to the relationship of science to society, was negative

($\beta = -.16$). That means when students spend class time in thinking the importance of science and its relationship to society and students report studying the relationship of science to society, its effect on student achievement is negative. The overall item correlation matrix (Table 12) showed negative correlations between science achievement and all three original indicators. These are students' response to how often do you "study the impact of technology on society"; "relate what you are learning in science to your daily life"; and "present your work to the class". That means these three specified activities were negatively associated with student' science achievement. It is possible that such emphasis cuts down the time on other classroom activities which may have a more direct and strong relationship to science achievement. Another explanation is a mismatch between the second factor (student engagement in activities related to science and society) and the science achievement test. The science test focused mainly on content knowledge and not the implication of science for society. Overall the frequency of students' classroom activities and teachers' instructional activities turned out to have a complex relationship with science achievement.

There was no direct effect OTL on student science achievement, however, significant indirect effect had been found through behavioral engagement. In other words, all variance of science achievement was explained by student behavioral engagement in science class activities which supported the mediation hypothesis. The effect of second OTL factor (implication science to society) negatively affected science achievement. However, the indirect effect was weak (β=-.02) and only a small proportion (4%) of variance of science achievement explained by OTL and behavioral engagement factors, there were other more important factors contributing to science achievement.

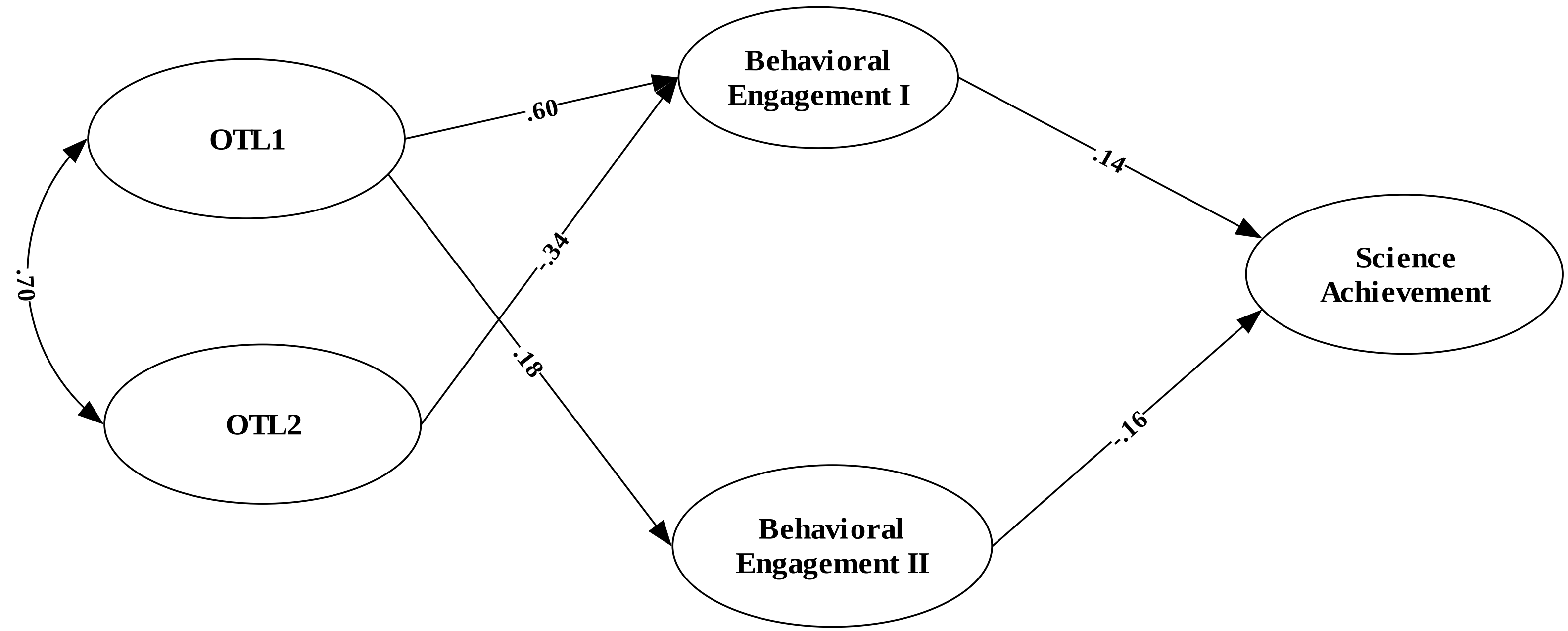

Figure 5. Structural model of OTL and behavioral engagement on science achievement (hypothesis 2b)

Hypothesis 3: Science emotional and cognitive engagement (interest/efficacy in science and future value of science) positively influence science achievement, controlling for family background.

Even though both hypotheses 1 and 2 were supported and showed significant results, only 4% of total variance was explained in science achievement, indicating that there were other factors that explain variance in science achievement. In a follow up model we included individual level variables to further understand differences in science achievement. Factors that reflected students' affective and cognitive engagement in science learning were included in the model. Student family background was included to control for socio-economic status. It was initially hypothesized that there was no direct effect of SES on science achievement. The fit indices showed the model to have a poor-fit (GFI=,89, CFI=.89, AGFI=.81, RMSEA=.13, SRMR=.17). Modification indices suggested add a direct path from SES to science achievement. Thus, certain changes were made in the model specification. (Table 16 and Table 17)

Table 16

Modification Indices Suggest to Add (Hypothesis3)

	Decrease in Chi-Square	New Estimate
Emotional←Cognitive	1221.9	0.70
ACH ←SES	506.2	4.42

Table 17
Fit Indices and Chi-Square Difference Test (Hypothesis 3)

	χ^2	*df*	$\Delta\chi^2$	Δdf	CFI	GFI	AGFI	RMSEA	SRMR
Model A	2804.42	39			.89	.89	.81	.13	.17
Model B (Emo← Cognitive)	1222.29	38	1582.13**	1	.96	.95	.91	.087	.08
Model C (ACH ←SES)	680.35	37	541.94**	1	.98	.97	.95	.065	.05

Note: **, p<.01

The fit indices for model C were high, indicating a well-fitting model in which data fit well to the final model. The Goodness-of-Fit Index (GFI) was .97, and the Adjusted Goodness-of-Fit Index (AGFI) was .95. The Comparative Fit Index (CFI) was .98. The Root-Mean-Square Error of Approximation (RMSEA) was .067. Overall, these fit indices indicate a theoretically sound model that explained the data well. The model explained 20% of variance in science achievement. The model showed that science emotional and cognitive engagement factors were important predictors of science achievement.

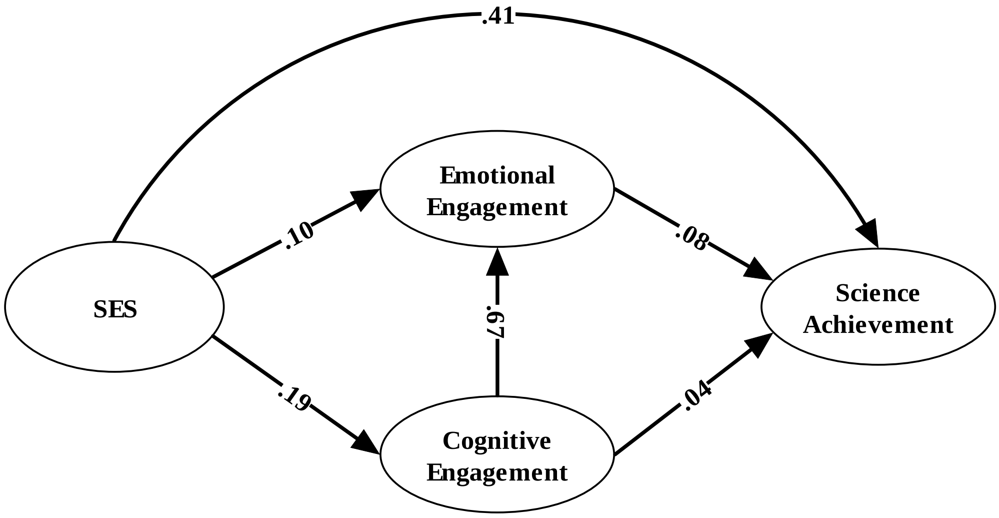

Figure 6. Structural model of emotional and cognitive engagement on science achievement (hypothesis 3)

Table 18

Standardized Direct, Indirect, and Total Effects of Emotional and Cognitive Engagement and SES on Science Achievement (Hypothesis 3)

	Students' science psychological engagement						Science achievement		
	Emotional engagement			Cognitive engagement					
	direct	indirect	total	direct	Indirect	total	direct	indirect	total
Exogenous variables									
SES	.10*	.13*	.23*	.19*		.19*	.41*	.03*	.44*
Endogenous variables									
Emotional engagement							.08*		.08*
Cognitive engagement	.67*		.67*				.04*	.05*	.09*

Note: *, *p*<.05

Figure 6 elaborated all the effects between latent variables. Because the model was considered adequate the direct and indirect effects for significance and magnitude were interpreted in Table 18. Family background (SES) had strongly positive direct effects on both emotional and cognitive engagement and on science achievement. Students from higher SES family background had positive attitudes towards science and achieved higher on standardized science test. Emotional engagement (interest and efficacy in science) and cognitive engagement (value of science) are significant positive influences on science achievement. That means students who like science (such as enjoy learning science) and think science is important (such as need science to get a job) usually have better science achievement. There is a significant relationship between emotional and cognitive engagement. The effect of valuing science on interest in science was significant ($\beta = .67$). That means, future value of science not only had direct effect, but also had indirect effect on science achievement. The indirect effect was strong. Almost half the effect of valuing science was indirect which was through the mediator, interest in science factor (emotional/affective engagement). It could be explained that the students who believe science is important also have higher interest in learning science, which, in turn, increased their science achievement.

Final structural model

After testing above three hypotheses, we had an initial understanding of the relationships between OTL, behavioral, psychological engagement in learning, and science achievement. To further investigate the affective factors and class room behavior factors together. The final model was included teachers' pedagogical factors, students' behavioral and psychological factors and science achievement (Figure 7).

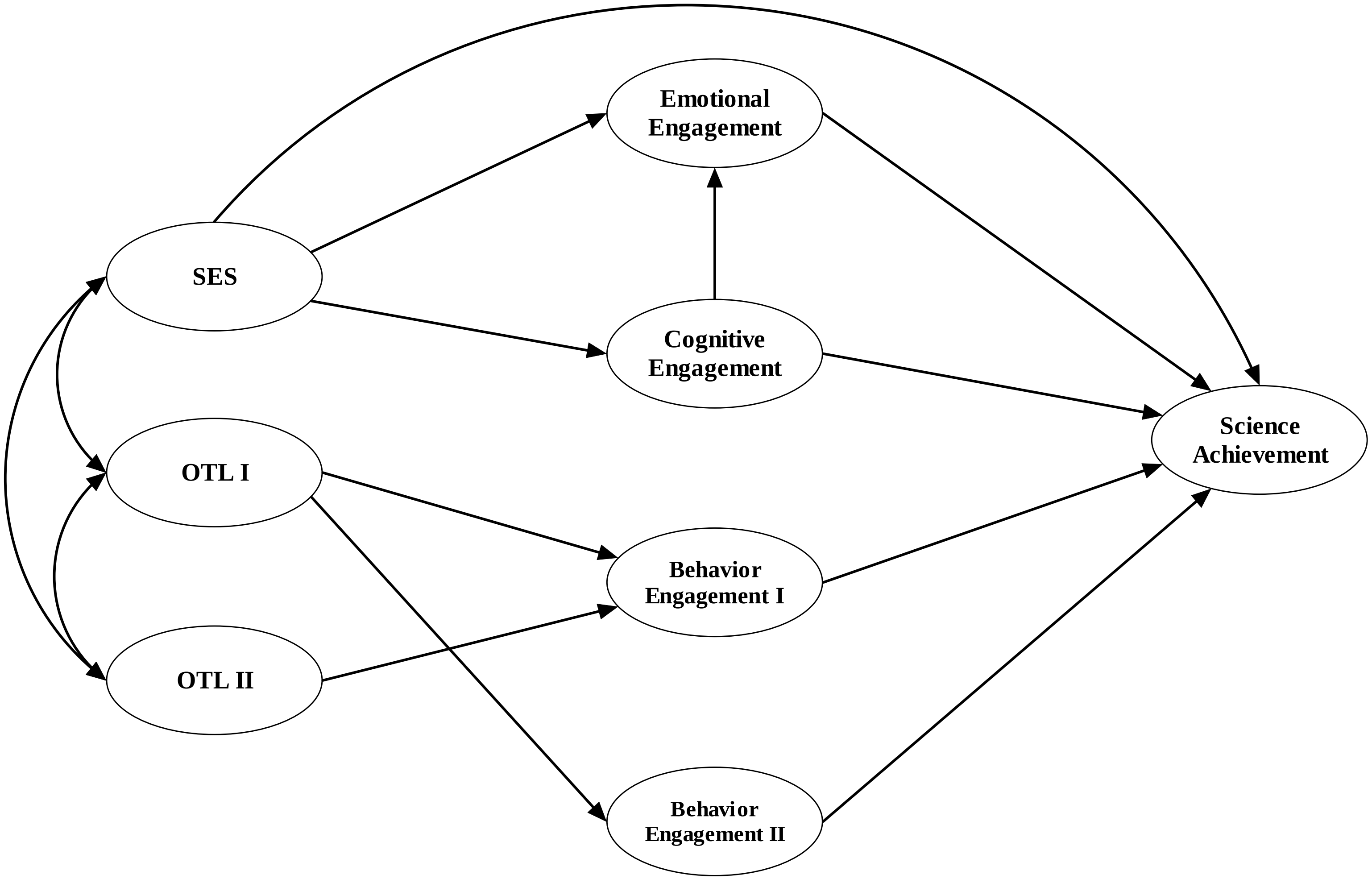

Figure 7. Hypothesized final SEM model

Table 19

Modification Indices Suggest to Add (Final Model)

	Decrease in Chi-Square	New Estimate
BEHA1←BEHA2	850.2	.56
COGNITIVE←BEHA2	428.9	.38
BEHA1←SES	35.9	.13

Table 20

Fit Indices and Chi-Square Difference Test (Final Model)

	χ^2	df	$\Delta\chi^2$	Δdf	CFI	GFI	AGFI	RMSEA	SRMR
Model A	3254.20	215			.95	.94	.92	.059	.085
Model B (BEHA1← BEHA2)	2314.13	214	940.07**	1	.96	.95	.94	.049	.068
Model C COGNITIVE ←BEHA2)	1894.99	213	419.14**	1	.97	.96	.95	.044	.038
Model D (BEHA1← SES)	1852.06	212	42.93**	1	.97	.96	.95	.043	.036

Note: **, $p<.01$

The fit indices of the hypothesized model (A) were good, but could be furthered improved by minor respecification of the model (GFI and AGFI <.95). When checked the suggested modification indices, three paths were strongly suggested in the model (Table 19). One was the link from the second behavioral engagement factor (connection to society) to the first behavioral engagement factor (scientific investigation activities). The other was from behavioral engagement factor (connection to society) to cognitive engagement (valuing science). The third was the link from SES to behavioral engagement I (in class scientific activities). Chi-square difference tests showed that the model was significantly improved when these three paths were added step by step (Table 20). Figure 8 elaborates the magnitude and directions of the coefficients in the final model.

Science Achievement. Science Rasch score was directly influenced by all engagement factors and SES and indirectly affected by OTL, SES, and cognitive engagement (valuing science). The strongest direct effect on science achievement was from SES. Students' cognitive engagement in learning science was the next highest positive effect. Those students who want to learn science, value science and have aspiration to continue in science area are likely to have higher achievement in science. The interest and efficacy in science also had a significant positive effect on science achievement. The effect of students' engagement in classroom activities that promoted understanding of science's impact on society was negative. Approximately 24% of total variance in science achievement was explained by the overall model. Standardized direct, indirect, and total effects were shown in Table 21. The table was organized according to effects on endogenous variables, starting with science emotional and cognitive engagement factors, then two behavioral engagement factors, and science achievement. Detailed explanations of the final model are summarized as below.

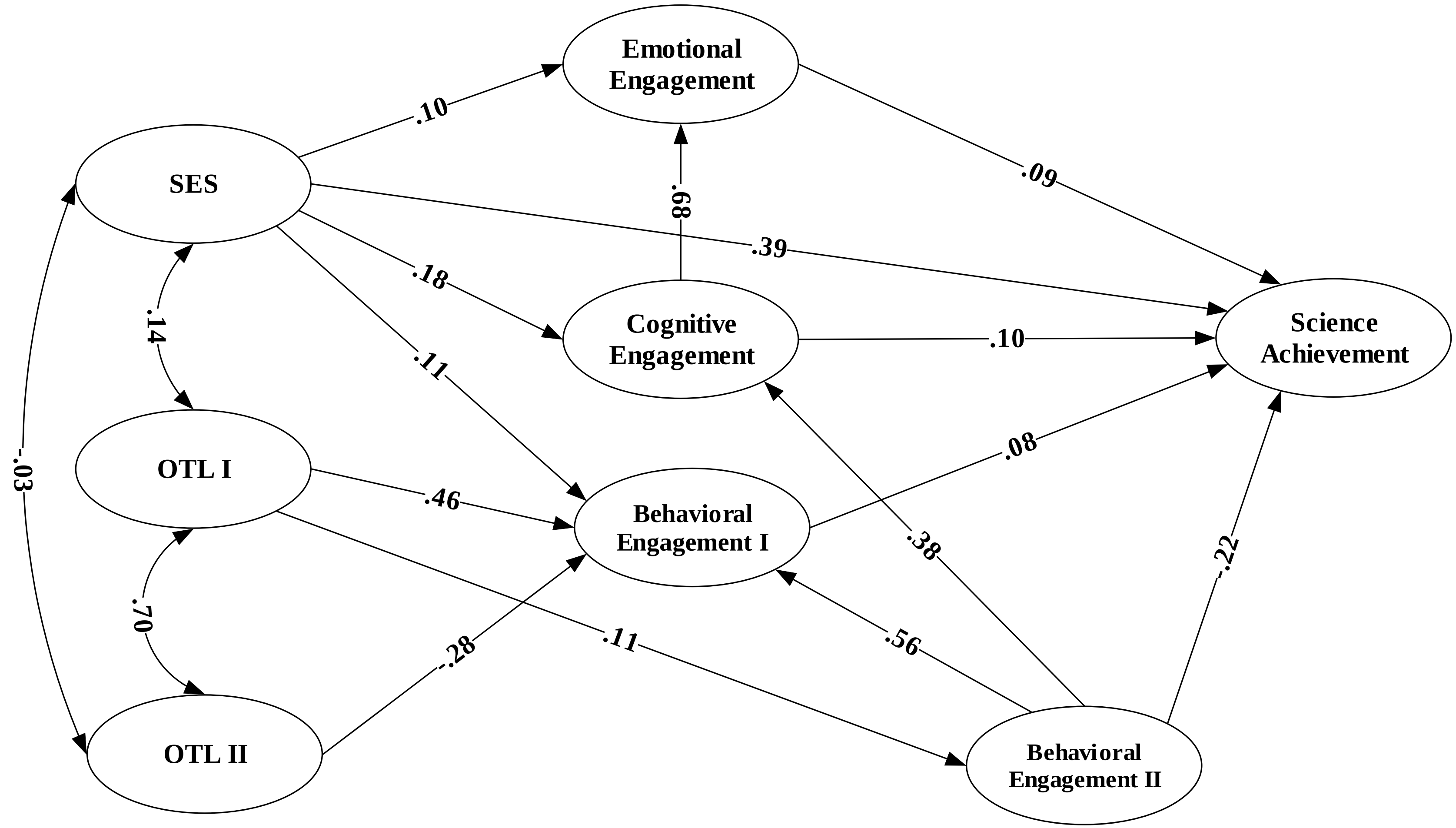

Figure 8. Final structural equation model

Table 21

Standardized Direct, Indirect and Total Effect (Final Model)

	Psychological						Science behavioral engagement						Science achievement		
	Emotional engagement			Cognitive engagement			Behavioral engagement I			Behavioral engagement II					
	direct	indirect	total	direct	indirect	total	direct	indirect	total	direct	indirect	total	direct	indirect	total
Exogenous variables															
SES	.10*	.12*	.22*	.18*		.18*	.11*		.11*				.39*	.05*	.44*
OTL1		.03*	.03*		.04*	.04*	.46*	.06*	.52*	.11*		.11*		.02*	.02*
OTL2							-.28*		-.28*					-.02*	-.02*
Endogenous variables															
Emotional													.09*		.09*
Cognitive	.68*		.68*										.10*	.06*	.17*
Behavioral I													.08*		.08*
Behavioral II		.26*	.26*	.38*		.38*	.56*		.56*				-.22*	.11*	-.12*

Note: *, *p*<.05

Effects of Socio-economic Status: The control variable in the study, SES, had large direct and indirect effects on the engagement variables and science achievement. The direct effect on science achievement was the strongest (γ=.39). Followed by the effect on cognitive engagement in valuing science (γ=.18). SES also had significant direct effect on emotional engagement (γ=.10) and behavioral engagement I (scientific investigation activities) (γ=.11). Two significant indirect effects of SES on endogenous variables were found. They were indirect effects on emotional engagement (γ=.12) and science achievement (γ=.05). It is evident that socioeconomic status exerts a significant role in students' learning behaviors and achievement. This study confirms the important role of family background in educational outcomes.

Effects of Opportunity to Learn (OTL) as measured by the teachers' reported classroom instructional activities: OTL factors had direct effects on student behavioral engagement. What teachers do in the classroom determines the activities of students in the classroom. The first OTL factor which means teachers emphasized instruction in core scientific activities such as conducting experiments had very strong direct effects on students' scientific investigation activities (γ=.46). In addition, this kind of effect directly influenced student behavioral engagement in connecting science to society (γ=.11). Slight but significant indirect effects of the first OTL factor were found on emotional engagement (γ=.03), cognitive engagement (γ=.04), behavioral engagement I in student scientific investigation activities (γ=.06), and science achievement (γ=.02). It is evident that when teachers focused instruction on learning about scientific methods it promoted student behavioral engagement in scientific investigation activities. Similarly when the teacher focused the instruction on the impact of science on society, the students also focused on similar activities and such emphasis directly affected scientific investigation activities. The effect was negative (γ=-.28). In turn, it had a slightly negative effect on science achievement (γ=-.02).

Effects of Emotional Engagement. Emotional engagement, as measured by the interest/efficacy in science was an important mediator between OTL and science achievement; it was influenced by SES, OTL, cognitive and behavioral engagement. It had a direct positive effect on science achievement ($\beta = .09$). The more interested students were in science, the higher they achieved in science.

Effects of Cognitive Engagement. Cognitive engagement had significant direct effect on emotional engagement ($\beta = .68$) and science achievement ($\beta = .10$). It also had an indirect effect on science achievement through the emotional engagement ($\beta = .06$). The total effect of cognitive engagement on achievement was .17. Students who value science and want to pursue science in college are more likely to have higher achievement in science.

Effects of Behavioral Engagement. The first behavioral engagement factor (scientific investigation activities) significantly influenced the outcome—science achievement ($\beta = .08$). The engagement of students in learning core science activities of running experiments and writing up results had a significant positive effect on achievement. The effect of the second behavioral engagement factor (connecting science to society) was complex. It had directly negative effect on science achievement ($\beta = -.22$). But it positively affected emotional, cognitive engagement and behavioral engagement in scientific investigation activities. So it also had positive indirect effect on science achievement ($\beta = .11$). The total effect was still negative ($\beta = -.12$).

Hierarchical Linear Modeling

The Unconditional Model

The use of the hierarchical linear model involved the single cross-section of data with a three-level structure consisting of students (Level 1) nested within classes (Level 2) nested within schools (Level 3). The fully unconditional model with no predictor variables was specified. The outcome measure, science achievement, was free to vary across three different levels of analysis: student, class and school. This model is described below.

Student-level model

Science Achievement for each student was estimated as a function of the class average plus random error:

$$Ach_{ijk} = \pi_{0jk} + e_{ijk}$$

where

Ach_{ijk} represents the Science Achievement of 8^{th} grade student *i* in class *j* and school *k* in 2003.

π_{0jk} represents the class mean Science achievement of class *j* in school *k*.

e_{ijk} represents the random error of student *i* in class *j* and school *k*

The notations, *i*, *j*, and *k* denote student, classes, and schools where there are

$i = 1, 2, 3, \ldots, n_{jk}$ students within class *j* in school *k*.

$j = 1,2, \ldots J_k$ classes within school *k*,

$k = 1, \ldots, K$ schools.

Class-level model

Science achievement classroom mean varies as a function of the school mean plus random error:

$\pi_{0jk} = \beta_{00k} + r_{0jk}$

where

β_{00k} represents the mean Science achievement in school *k*.

r_{0jk} represents the random error of class *j* within school *k*

School-level model.

Science school mean achievement varies randomly around a grand mean for all schools.

$\beta_{00k} = \gamma_{000} + \mu_{00k}$

where

γ_{000} represents the grand mean Science achievement for all schools.

μ_{00k} represents the random school effect, the deviation of school *k's* mean from the grand mean.

Table 22 presents the results of the fixed effects model. The results indicates that the grand-mean for science achievement is 150.29 (t=277.23, p<.05). In terms of the variance partitioning, the estimates of the variance components at student level Var(e_{ijk}) =60.20 is the variance between students within classes, Var (r_{0jk}) =24.08 (χ^2=782.85, p<.05) is the variance between classes within schools, and Var(u_{00k}) =18.97 (χ^2= 273.58, p<.05) is the variance between schools. These significant random effects suggest there is large variation in class and school level. The estimates of the percentage of variance in students' science achievement at each level show that 58.31% of the variability in science achievement scores was over measurements within students (over level-1 units); 23.32% of the variability in the outcome was at class level (level-2); while 18.37% of the variability was at school level (level-3). Because more than 40% of the variance accounted was at the class and school level, a three level HLM analysis was justified.

Table 22

Three-Level Analysis of Unconditional Model

Fixed effects		S.E.	*t* Ratio	
Mean science achievement, γ_{000}	150.29	0.54	277.23	
Random effects	Variance component		Chi-square	*p* Value
Students (level 1), e_{ijk}	60.20	7.76		.00
Classes (level 2), r_{0jk}	24.08	4.90	782.85	.00
Schools (level 3), u_{00k}	18.97	4.36	273.58	.00
Variance decomposition (percentage by level)				
Level 1	58.31%			
Level 2	23.32%			
Level 3	18.37%			

Conditional Model

An explanatory model that allows estimation of the effect of engagement on individual science leaning is considered. Because only one to three classes per school were selected in the data which was too small to allow estimating a model with multiple random slopes at level two, three separate research questions were posed:

1. How does students' emotional engagement measured by interest/efficacy in science affect their science achievement controlling for content coverage,

teachers' quality, school enrichment/remedial program, and school SES? Is there any evidence that these effects vary by class and school level?

2. How does students' cognitive engagement measured by "value of science learning for future" affect their science achievement controlling for content coverage, teachers' quality, school enrichment/remedial program, and school SES? Is there any evidence that these effects vary by class and school level?
3. Does students' behavioral engagement in science classroom activities affect on their science achievement controlling for content coverage, teachers' quality, school enrichment/remedial program, and school SES? Is there any evidence that these effects vary by class and school level?

Predictors in level-1 are EMOTIONAL ENGAGEMENT in science in research question one; COGNITIVE ENGAGEMENT in science in research question two; and BEHAVIORAL ENGAGEMENT in class activities in research questions three. They are measured by the average of observed items. In this part of the analysis we used class and school level factors that provide differential opportunities to learn to the students in different classes and schools. The OTL factors used in level-2 are the same in all three research questions. These are LICENSE (science teacher has full teaching license or certificate) and TOPICS (the percentage of overall topics taught in class as listed in TIMSS questionnaire). Factors in level-3 model are REDUCED/FREE LUNCH (the percentage of school reduced/free lunch); and ENRICHMENT/REMEDIAL (school offers science enrichment or remedial program) and they are the same in all three research questions (See Table 23 & Table 24 for descriptive statistics). All HLM models were weighted using HOUWGT at the student level.

Table 23

Descriptive Statistics in HLM Analyses

		N	Percentage (categorical)	Mean	S.D.
Student level					
Science Rasch Score		8544		149.59	10.17
Emotional Engagement		8183		2.91	.76
Cognitive Engagement		8354		2.88	.77
Behavioral Engagement I		8275		2.85	.83
Behavioral Engagement II		8257		2.43	.78
Class level					
LICENSE	Yes	328	88.65%		
	No	42	11.35%		
	Total	370			
TOPIC Coverage		385		79.05	18.13
School level					
Enrichment/ Remedial	Yes	73	44.79%		
	No	90	55.21%		
	Total	163			
Free/Reduced Lunch		163		37.75	26.64

Note: Sample size is unweighted

Table 24

Correlation Matrix for Student Level (Level-1) Predictors

		1	2	3	4	5
1	Science Rasch Score	1				
2	Emotional Engagement	0.28**	1			
3	Cognitive Engagement	0.17**	0.51**	1		
4	Behavioral Engagement I	0.07**	0.16**	0.21**	1	
5	Behavioral Engagement II	-0.09**	0.16**	0.32**	0.49**	1

Note: **, $p<.05$

Research question one:

Table 25

Taxonomy of HLM Model for Change (Research Question 1)

Model 1	Level-1 model	Level-2 model	Level-3 model
1-A	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}(EMOTIONAL) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + r_{0jk}$	$\beta_{00k} = \gamma_{000} + u_{00k}$
		$\pi_{1jk} = \beta_{10k} + r_{1jk}$	$\beta_{10k} = \gamma_{100}$
1-B	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}(EMOTIONAL) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01k}(LICENSE) + \beta_{02k}(TOPICS) + r_{0jk}$	$\beta_{00k} = \gamma_{000} + u_{00k}$
		$\pi_{1jk} = \beta_{10k} + \beta_{11k}(TOPICS) + \beta_{12k}(LICENSE) + r_{1jk}$	$\beta_{01k} = \gamma_{010}$
			$\beta_{02k} = \gamma_{020}$
			$\beta_{10k} = \gamma_{100}$
			$\beta_{11k} = \gamma_{110}$
			$\beta_{12k} = \gamma_{120}$
1-C	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}(EMOTIONAL) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01k}(LICENSE) + \beta_{02k}(TOPICS) + r_{0jk}$	$\beta_{00k} = \gamma_{000} + u_{00k}$
		$\pi_{1jk} = \beta_{10k} + \beta_{11k}(TOPICS) + r_{1jk}$	$\beta_{01k} = \gamma_{010}$
			$\beta_{02k} = \gamma_{020}$
			$\beta_{10k} = \gamma_{100}$
			$\beta_{11k} = \gamma_{110}$
1-D	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}(EMOTIONAL) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01k}(LICENSE) + \beta_{02k}(TOPICS) + r_{0jk}$	$\beta_{00k} = \gamma_{000} + \gamma_{001}(LUNCH) +$
		$\pi_{1jk} = \beta_{10k} + \beta_{11k}(TOPICS) + r_{1jk}$	$\gamma_{002}(ENRICH / REMEDIAL) + u_{00k}$
			$\beta_{01k} = \gamma_{010}$
			$\beta_{02k} = \gamma_{020}$
			$\beta_{10k} = \gamma_{100}$
			$\beta_{11k} = \gamma_{110}$

Beginning with the lower level, the final model is derived by four steps following Singe and Willett's (2003) model specification. Deviance difference tests were applied to determine the models (Table 25).

Model 1-A presents a random-coefficients regression model. That is the emotional engagement as measured by interest/efficacy in science as level 1 predictor and both intercept π_{0jk} (class means science achievement) and slope π_{1jk} (effect of emotional engagement in science) vary randomly. There are no class and school level predictors. Table 26 elaborates that both the level 2 variance are significant at 0.05 level (Var(r_{0jk})=19.44, p<.01; Var(r_{1jk})=2.16, p<.05). There is evidence that class mean science achievement and the effect of emotional engagement on science achievement are different by class.

Model 1-B presents an "Intercepts- and Slopes-as Outcomes" model. The class level factors (science teacher has full license and overall taught topics) were added to predict the class-mean achievement (π_{0jk}) and the effect of emotional engagement on science achievement (π_{1jk}). We found that the interaction effect of "teacher has full science license" and "emotional engagement" on science achievement is not significant (γ_{120}=0.44, p>.05).

Model 1-C is a parsimonious model compared to Model B. The insignificant interaction effect of "teacher has full science license" and "emotional engagement" is deleted from model B. The deviance difference test shows $\Delta\chi^2(1)$= 30416.05-30415.66=0.39 (p>.05), which means that there is no significant difference between these two models. Model C is the retained model due to parsimony in future steps.

Model 1-D adds level 3 (school level) predictors: "percentage of free/reduced lunch" and "school offers enrichment/remedial science classes". Both of these two variables significantly predict the school-mean science achievement. No other random effects have been found besides u_{00k}. In other words, the class level effect of "science teacher has full license" and "overall taught topics" and mean effect of "emotional engagement" in science on achievement are the same across all schools.

Model 1-D is adopted as the final model to answer the first research question.

Level-1 model

$$Ach_{ijk} = \pi_{0jk} + \pi_{1jk} \text{ (EMOTIONAL ENGAGEMENT)} + e_{ijk} \quad [4.4]$$

Where

Ach_{ijk} represents the Science Achievement of each student *i* in class *j* and school *k*.

π_{0jk} represents the class mean Science achievement of class *j* in school *k*.

π_{1jk} represents the slope of students' emotional engagement in science on science achievement of class *j* in school *k*.

e_{ijk} represents the random error of student *i* in class *j* and school *k*

The level-2 model represented the variability in class mean achievement π_{0ij} and slope of emotional engagement, π_{1ij}, between classes within schools. The effects of class level OTL variables were represented in this level:

$$\pi_{0jk} = \beta_{00k} + \beta_{01k} \text{ (LICENSE)} + \beta_{02k} \text{ (TOPICS)} + r_{0jk} \quad [4.5]$$

$$\pi_{1jk} = \beta_{10k} + \beta_{11k} \text{ (TOPICS)} + r_{1jk} \quad [4.6]$$

where

β_{00k} represents the school mean science achievement in school *k*.

β_{01k} represents the effect of "teacher has full science teaching license or not" on class mean science achievement in school *k*.

β_{02k} represents the effect of "the percentage of overall taught science topics" on class mean science achievement in school *k*.

β_{10k} represents the mean effect of "students' emotional engagement in science" on science achievement in school *k*.

β_{11k} represents the effect of "the percentage of overall taught science topics" on the slope of "students' emotional engagement in science" in school *k*.

r_{0jk} represents the random error of class *j* within school *k*

r_{0jk} represents the random error of emotional engagement slope within class *j* in school *k*

Equations [4.5] explained the model that LICENSE (a dummy variable indicating teacher has full science teaching license) and TOPICS (the percentage of overall taught science topics) were related to class mean science achievement. Equation [4.6] specified that the effect of emotional engagement in science on science achievement was a function of TOPICS (the percentage of overall taught science topics) and differs among class within each school *k*. In other words, there is an interaction effect of emotional engagement and topic coverage on student science achievement.

The level-3 model represents the variability among schools in the five β coefficients. We found that school offers enrichment or remedial class and school rate of reduced/free lunch predict school mean science achievement. But school level factors can not predict the effect of students' emotional engagement in science (β_{10k}) and its intercept (β_{11k}). Also, the effect of teachers' license (β_{01k}) and taught topics (β_{02k}) is constant across all schools. Thus, we pose the following level-3 model:

$$\beta_{00k} = \gamma_{000} + \gamma_{001}\text{ (REDUCED LUNCH)) } + \gamma_{002}\text{(ENRICHMENT/REMEDIAL)} + \mu_{00k} \quad [4.7]$$

$$\beta_{01k} = \gamma_{010} \quad [4.8]$$

$$\beta_{02k} = \gamma_{020} \quad [4.9]$$

$$\beta_{10k} = \gamma_{100} \quad [4.10]$$

$$\beta_{11k} = \gamma_{110}. \quad [4.11]$$

For school level effects, the γ_{000} coefficient represents in this application the predicted school mean science achievement for a school which does not offer any enrichment or remedial science classes (ENRICHMENT/REMEDIAL = 0) and is an "affluent school" (FREE/REDUCED LUNCH = 0%). For such a school, the predicted mean science achievement is 133.19. For each 10% increment in free/reduced lunch rate, the expected achievement is reduced by 1.2 points (γ_{001} =-.12, p<0.5). For those schools which offer science enrichment or remedial class, the school means are 1.47 points higher than others (γ_{002} =1.47, p<.10).

For class level, at the class mean science achievement point, the gap of teacher has full license or not, γ_{010} is identical with β_{01k} , is 2.65 (p<.05) because the random effect is not significant. This means that teacher who has no full license or certificate (LICENSE=0) start out 2.65 points of class mean behind those teachers who have full license or certificate. γ_{020} , which is identical with β_{02k} , is .11(p<.01). That shows the effect of "the percentage of overall taught science topics" on class mean science achievement. For each 10% increment in taught topics, the expected class mean of science achievement is increased by 1.1 points.

The predicted mean effect of "students' emotional engagement in science" on science achievement γ_{100} , which is identical with β_{10k} , is 5.84 (p<.01) because the random component is not significant. In other words, the effect of emotional engagement in science on science achievement can be predicted as mean effect (β_{10k}) plus a function of TOPICS (percentage of taught science topics) ($\beta_{11k} = \gamma_{11k}$ =-.03, p<.01) and a random error. That means there is an interaction effect of emotional engagement and "percentage of taught science topics". This interaction effect is slightly negative, which can be explained as the more topics covered in science class, the less effect of emotional engagement on achievement. For a student whose class covered 100% of content topics, the effect of emotional engagement in science on science achievement was $\pi_{1jk} = \beta_{10k} + \beta_{11k}$ (TOPICS) = 5.81- .03(100) = 2.81. For an average student whose class covered 0% of topics, the effect of emotional engagement in science on science achievement $\pi_{1jk} = \beta_{10k} + \beta_{11k}$ (TOPICS) = 5.81- .03(0) = 5.81. Hence emotional engagement in science is a significant positive predictor on science achievement. The more interested the students are in science, the higher the achievement score they received. Figure 9 further elaborates the interaction effect of emotional engagement in science and "percentage of taught topics". When a student is not interested in science and does not like science (disagree a lot or disagree a little with science is interesting), the relationship between TOPICS (percentage of taught topics) and science achievement is positive. When students agree with science is interesting, there was a slightly negative relationship between TOPICS (percentage of taught topics) and science achievement. It can be seen that coverage of science topics was more important for those students who were less interested in science.

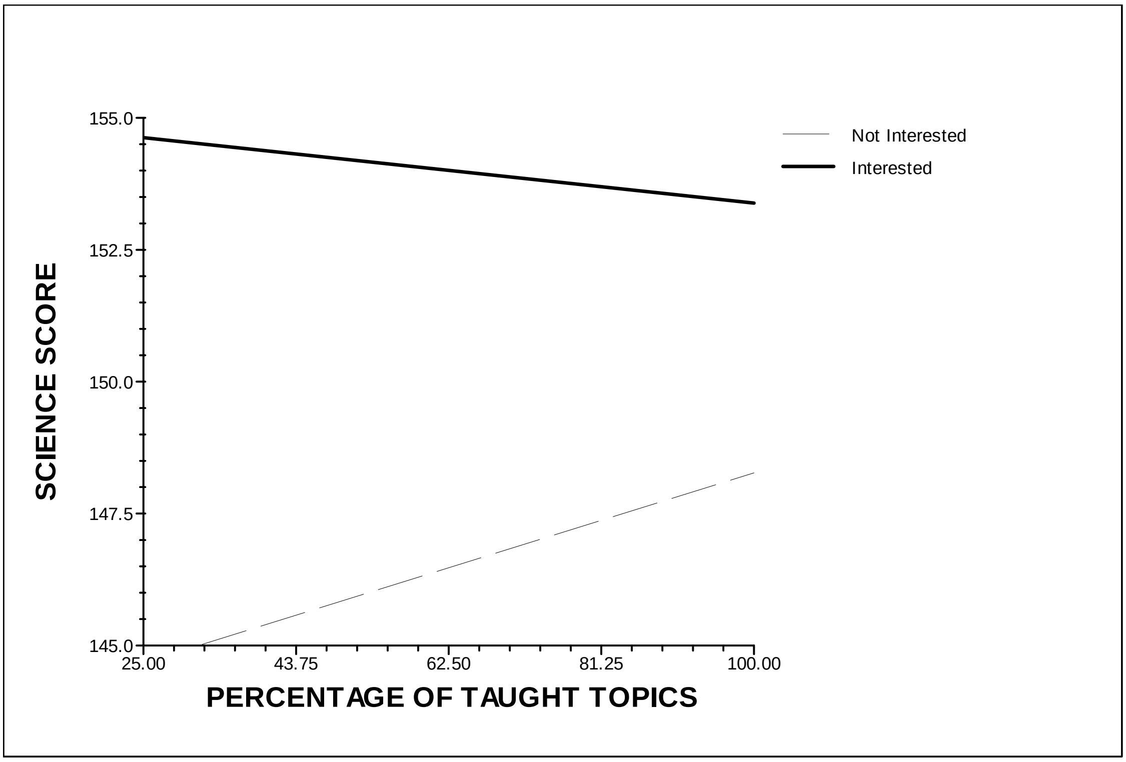

Figure 9. Effect of content coverage on science achievement

The variance components shown in Table 26 present estimated variance and related χ^2 statistics from the three-level model decomposition. Level 1 (student level) variance Var(e_{ijk}) = 54.04, which indicates student's science achievement is different within class. Level 2 (class level) variance Var(r_{0jk})=17.45 (p<.01) and Var(r_{1jk})=1.92 (p<.10), which indicates that class mean science achievement scores are varied within school; the effect of emotional engagement on science achievement is also varied between classes within school. Level 3 (school level) variance Var(u_{00k}) =5.26 (p<.01).

That implies residual variance still remains to be explained in school mean science achievement.

Research question two

To answer the second question, a three-level analysis was conducted by following five steps (Table 27).

Model 2-A presents a random-coefficients regression model with level 1 predictor "COGNITIVE ENGAGEMENT" which is measured by the value of science learning for future. Both intercept π_{0jk} (class mean science achievement) and slope π_{1jk} (effect of cognitive engagement in valuing science) vary randomly. There are no class and school level predictors in this model. Results in Table 28 show that both the level 2 variance are significant at 0.05 level (Var(r_{0jk})=20.45, p<.01; Var(r_{0jk})=1.58, *p*<.05). There is evidence that class mean science achievement and the effect of cognitive engagement on science achievement are different by class within school.

Model 2- B presents an "Intercepts- and Slopes-as Outcomes" model. Class level factors are added (science teacher has full license and overall taught topics) to predict the class mean achievement (π_{0jk}) and the effect of cognitive engagement on science achievement (π_{1jk}). We found that both effects on "COGNITIVE ENGAGEMENT" slope are not significant (γ_{110} =-.66, *p*>.05 and γ_{120} =-.01, *p*>.05). Furthermore, deviance difference test doesn't show significant improvement when all four parameters are added simultaneously ($\Delta\chi^2(4)$ =8.83, *p*>.05).

Table 26

Fixed and Random Effects of Emotional Engagement (Research Question 1)

		Parameter	Model 1-A	Model 1-B	Model 1-C	Model 1-D
Fixed Effects						
Class Mean Of Science Achievement (Intercept)	Intercept	γ_{000}	140.95** (0.71)	128.30** (3.73)	127.01** (3.61)	133.19** (3.56)
	Free/Reduced lunch	γ_{001}				-0.12** (0.01)
	Enrichment/Remedial	γ_{002}				1.47^{+} (0.79)
	License	γ_{010}		2.35 (1.47)	3.47** (1.22)	2.65* (1.31)
	Topics	γ_{020}		0.13** (0.04)	0.13** (0.04)	0.11** (0.04)
Effects of Emotional Engagement (Slope)	Intercept	γ_{100}	3.17** (0.21)	5.35** (1.08)	5.86** (0.94)	5.81** (0.92)
	Topics	γ_{110}		-0.03** (0.01)	-0.03** (0.01)	-0.03** (0.01)
	License	γ_{120}		0.44 (0.53)		

Table 26 (Continued)

		Parameter	Model A	Model B	Model C	Model D
Variance Components						
Level 1	Students variation	e_{ijk}	54.05 (7.35)	53.98 (7.35)	53.99 (7.35)	54.04 (7.35)
Level 2	Class mean variation	r_{0jk}	19.44** (4.41)	17.65** (4.20)	17.62** (4.20)	17.45** (4.12)
	Emotional engagement slope variation	r_{1jk}	2.16* (1.47)	2.00^{+} (1.41)	2.00^{+} (1.42)	1.92^{+} (1.38)
Level 3	School mean variation	u_{00k}	17.19** (4.15)	15.75** (3.97)	15.78** (3.97)	5.26** (2.29)
	deviance		30430.01	30415.66	30416.05	30357.31
	# of estimated parameters		7	11	10	12
	Δ deviance			14.35**	0.39	58.74**
	Δ *df*			4	1	2

Note: **, $p<.01$
*, $p<.05$
$^{+}$, $p<.10$

Table 27

Taxonomy of HLM Model for Change (Research Question 2)

Model 2	Level-1 model	Level-2 model	Level-3 model
2-A	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}\left(COGNITIVE\right) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + r_{0jk}$	$\beta_{00k} = \gamma_{000} + u_{00k}$
		$\pi_{1jk} = \beta_{10k} + r_{1jk}$	$\beta_{10k} = \gamma_{100}$
2-B	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}\left(COGNITIVE\right) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01k}\left(LICENSE\right) + \beta_{02k}\left(TOPICS\right) + r_{0jk}$	$\beta_{00k} = \gamma_{000} + u_{00k}$
		$\pi_{1jk} = \beta_{10k} + \beta_{11k}\left(LICENSE\right) + \beta_{12k}\left(TOPICS\right) + r_{1jk}$	$\beta_{01k} = \gamma_{010}$
			$\beta_{02k} = \gamma_{020}$
			$\beta_{10k} = \gamma_{100}$
			$\beta_{11k} = \gamma_{110}$
			$\beta_{12k} = \gamma_{120}$
2-C	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}\left(COGNITIVE\right) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01k}\left(LICENSE\right) + \beta_{02k}\left(TOPICS\right) + r_{0jk}$	$\beta_{00k} = \gamma_{000} + u_{00k}$
		$\pi_{1jk} = \beta_{10k} + r_{1jk}$	$\beta_{01k} = \gamma_{010}$
			$\beta_{02k} = \gamma_{020}$
			$\beta_{10k} = \gamma_{100}$
2-D	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}\left(COGNITIVE\right) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01k}\left(LICENSE\right) + r_{0jk}$	$\beta_{00k} = \gamma_{000} + u_{00k}$
		$\pi_{1jk} = \beta_{10k} + r_{1jk}$	$\beta_{01k} = \gamma_{010}$
			$\beta_{10k} = \gamma_{100}$
2-E	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}\left(COGNITIVE\right) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01k}\left(LICENSE\right) + r_{0jk}$	$\beta_{00k} = \gamma_{000} + \gamma_{001}\left(LUNCH\right) +$
		$\pi_{1jk} = \beta_{10k} + r_{1jk}$	$\gamma_{002}\left(ENRICH / REMEDIAL\right) + u_{00k}$
			$\beta_{01k} = \gamma_{010}$
			$\beta_{10k} = \gamma_{100}$

Model 2-C is a parsimonious model than Model 2-B. Insignificant effects of "teacher has full science license" and "overall topic coverage" on cognitive engagement slope were deleted from model 2-B. Deviance difference test shows $\Delta\chi^2(2)$ = 30598.16-30595.74 = 2.42 (p>.05), which shows no significant difference between these two models. The effect of "topic coverage" is not significant on the class mean science achievement ($\gamma_{020} = .04$, p>.05).

Model 2-D ignores the insignificant effect of "topic coverage" on school mean science achievement from Model 2-C. Deviance difference test shows $\Delta\chi^2(1)$ = 2.41 (p>.05), which is no significant shrinkage. Deviance difference between Model A and Model D ($\Delta\chi^2(1)$= 30604.57 - 30600.57 = 4.00 shows significant improvement when class-level OTL predictor "LICENSE" is added.

Model 2-E adds level 3 (school level) predictors: "percentage of free/reduced lunch" and "school offers enrichment/remedial science classes". Both of these two variables significantly predict the school mean science achievement. No other random effects have been found besides u_{00k} . In other words, the class level effect of "science teacher has full license" and mean effect of "cognitive engagement" in science on achievement are the same across all schools. Detailed explanations of model 2-E are shown as following:

level-1 model :

$$\text{Ach}_{ijk} = \pi_{0jk} + \pi_{1jk}\,(\text{COGNITIVE ENGAGEMENT}) + e_{ijk} \qquad [4.12]$$

Where

Ach_{ijk} represents the Science Achievement of each student i in class j and school k.

π_{0jk} represents the class mean Science achievement of class j in school k.

π_{1jk} represents the effect of valuing science on science achievement of class *j* in school *k*.

e_{ijk} represents the random error of student *i* in class *j* and school *k*

The level-2 model represents that class mean achievement as a function of "science teacher has a license" and a random error. Furthermore, there is variability in cognitive engagement slope, π_{1ij} , between classes within schools.

$$\pi_{0jk} = \beta_{00k} + \beta_{01k} \text{ (LICENSE)} + r_{0jk} \qquad [4.13]$$

$$\pi_{1jk} = \beta_{10k} + r_{1jk} \qquad [4.14]$$

where

β_{00k} represents the mean school science achievement in school *k*.

β_{01k} represents the effect of "teacher has full science teaching license or not" on class mean science achievement in school *k*.

β_{10k} represents the mean effect of valuing science on science achievement in school *k*.

r_{0jk} represents the random error of class mean of class *j* within school *k*

r_{1jk} represents the random error of mean effect of valuing science of class *j* within school *k*

Equations [4.13] can be explained that LICENSE (a dummy variable indicating teacher has full science teaching license) was related to class mean science achievement. Equation [4.14] specified that the effect of students' cognitive engagement on science achievement varied by class within each school *k*.

The level-3 model represents the variability among school in the three β coefficients, the same as in research question 1. School offers enrichment or remedial classes and school rate of reduced/free lunch predict school mean science achievement. But school level mean effect of cognitive engagement in science (β_{01k}) is constant across schools. Also, the slope for the teachers' license variable (β_{10k}) is constant across all schools too:

$$\beta_{00k} = \gamma_{000} + \gamma_{001} \text{ (REDUCED LUNCH))} + \gamma_{002}\text{(ENRICHMENT/REMEDIAL} + \mu_{00k} \quad [4.15]$$

$$\beta_{01k} = \gamma_{010} \quad [4.16]$$

$$\beta_{10k} = \gamma_{100} \quad [4.17]$$

All fixed effects are presented in Table 28. School level effects on science achievement are similar as in research question 1. Since it only interprets the intercept (mean science achievement), the γ_{000} coefficient represents in this application predicted school mean science achievement for a school which does not offer any enrichment or remedial science classes (ENRICHMENT/REMEDIAL = 0) and is an "affluent school" (FREE/REDUCED LUNCH = 0%). For such a school, the predicted science achievement is 146.87. For each 10% increment in free/reduced lunch rate, the expected initial status is reduced by 1.2 points (γ_{001} =-.12, p<0.5). For those schools that offer science enrichment or remedial class, their school mean are 1.61 (γ_{002} =1.61, p<0.5) points higher than others.

For the class level effect, at the class mean science achievement, the gap of teacher has full license or not, γ_{010} which is identical with β_{01k} , is 2.72 (p<.05) because there is no significant random effect. This means that teacher who has no full license or certificate (LICENSE=0) start out 2.72 points of class mean behind those teachers who have full license or certificate.

The predicted mean effect of students' cognitive engagement in science on science achievement β_{10k}, which is identical with γ_{100}, is 1.66. Hence the slope of cognitive engagement at student level is 1.66 which is smaller than the effect of emotional engagement but it is still a strong positive predictor on science achievement.

The variance components in Table 28 present estimated variance and related χ^2 statistics from the three-level decomposition. Level 1 (student level) variance Var(e_{ijk}) = 56.53, which indicates student's science achievement is different within classes. Level 2 (class level) variance Var(r_{0jk}) =21.18 (p<.01) and Var(r_{1jk}) =1.58 (p<.01), which indicates that class-mean science achievement scores varied between classes within school; the effect of cognitive engagement on science achievement is also varied between classes within school. Level 3 (school level) variance Var(u_{00k}) =5.86 (p<.01). That implies residual variance still remains to be explained in school mean science achievement.

Table 28

Fixed and Random Effects of Cognitive Engagement (Research Question 2)

		Parameter	Model 2-A	Model 2-B	Model 2-C	Model 2-D	Model 2-E
Fixed Effects							
Class Mean Of Science Achievement (Intercept)	Intercept	γ_{000}	145.68** (0.66)	134.57** (4.15)	138.73** (2.83)	142.72** (1.32)	146.87** (1.88)
	Free/Reduced Lunch	γ_{001}					-0.12** (0.02)
	Enrichment/Remedial	γ_{002}					1.61* (0.83)
	License	γ_{010}		5.28* (2.26)	3.73** (1.35)	3.27* (1.32)	2.72+ (1.47)
	Topics	γ_{020}		0.08 (0.04)	0.04 (0.03)		
Effect of Cognitive Engagement Slope (Slope)	Intercept	γ_{100}	1.65** (0.17)	3.39** (1.10)	1.65** (0.17)	1.65** (0.17)	1.66** (0.17)
	License	γ_{110}		-0.66 (0.69)			
	Topics	γ_{120}		-0.01 (0.01)			

Table 28 (Continued)

		Parameter	Model 2-A	Model 2-B	Model 2-C	Model 2-D	Model 2-E
Variance Components							
Level 1	Students variation	e_{ijk}	56.51 (7.52)	56.50 (7.52)	56.51 (7.52)	56.52 (7.52)	56.53 (7.52)
Level 2	Class mean variation	r_{0jk}	20.45** (4.52)	19.19** (4.38)	19.38** (4.40)	19.57** (4.42)	21.18** (4.60)
	Cognitive engagement slope variation	r_{1jk}	1.58* (1.26)	1.50* (1.22)	1.56* (1.25)	1.54* (1.24)	1.58* (1.26)
Level 3	School mean variation	u_{00k}	18.02** (4.25)	16.00** (4.00)	16.03** (4.00)	17.05** (4.13)	5.86** (2.42)
	Deviance		30604.57	30595.74	30598.16	30600.57	30545.27
	# Of Estimated Parameters		7	11	9	8	10
	Δ deviance			8.83**	2.42	2.41	55.3**
	Δ *df*			4	2	1	2

Note: **, *p*<.01
*, *p*<.05
+, *p*<.10

Table 29

Taxonomy of HLM Model for Change (Research Question 3)

Model 3	Level-1 model	Level-2 model	Level-3 model
3-A	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}(BEHA1)$ $\pi_{2jk}(BEHA2) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + r_{0jk}$ $\pi_{1jk} = \beta_{10k}$ $\pi_{2jk} = \beta_{20k}$	$\beta_{00k} = \gamma_{000} + u_{00k}$ $\beta_{10k} = \gamma_{100}$ $\beta_{20k} = \gamma_{200}$
3-B	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}(BEHA1)$ $\pi_{2jk}(BEHA2) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + r_{0jk}$ $\pi_{1jk} = \beta_{10k} + r_{1jk}$ $\pi_{2jk} = \beta_{20k} + r_{2jk}$	$\beta_{00k} = \gamma_{000} + u_{00k}$ $\beta_{10k} = \gamma_{100}$ $\beta_{20k} = \gamma_{200}$
3-C	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}(BEHA1)$ $\pi_{2jk}(BEHA2) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01k}(LICENSE) + \beta_{02k}(TOPICS) + r_{0jk}$ $\pi_{1jk} = \beta_{10k} + \beta_{11k}(LICENSE) + \beta_{12k}(TOPICS) + r_{1jk}$ $\pi_{2jk} = \beta_{20k} + \beta_{21k}(LICENSE) + \beta_{22k}(TOPICS) + r_{2jk}$	$\beta_{00k} = \gamma_{000} + u_{00k}$ $\beta_{01k} = \gamma_{010}$ $\beta_{02k} = \gamma_{020}$ $\beta_{10k} = \gamma_{100}$ $\beta_{11k} = \gamma_{110}$ $\beta_{12k} = \gamma_{120}$ $\beta_{20k} = \gamma_{200}$ $\beta_{21k} = \gamma_{210}$ $\beta_{22k} = \gamma_{220}$

3-D	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}(BEHA1)$ $\pi_{2jk}(BEHA2) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01k}(LICENSE) + \beta_{02k}(TOPICS) + r_{0jk}$ $\pi_{1jk} = \beta_{10k} + r_{1jk}$ $\pi_{2jk} = \beta_{20k} + r_{2jk}$	$\beta_{00k} = \gamma_{000} + u_{00k}$ $\beta_{01k} = \gamma_{010}$ $\beta_{02k} = \gamma_{020}$ $\beta_{10k} = \gamma_{100}$ $\beta_{20k} = \gamma_{200}$
3-E	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}(BEHA1)$ $\pi_{2jk}(BEHA2) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01k}(LICENSE) + r_{0jk}$ $\pi_{1jk} = \beta_{10k} + r_{1jk}$ $\pi_{2jk} = \beta_{20k} + r_{2jk}$	$\beta_{00k} = \gamma_{000} + u_{00k}$ $\beta_{01k} = \gamma_{010}$ $\beta_{10k} = \gamma_{100}$ $\beta_{20k} = \gamma_{200}$
3-F	$Ach_{ijk} = \pi_{0jk} + \pi_{1jk}(BEHA1)$ $\pi_{2jk}(BEHA2) + e_{ijk}$	$\pi_{0jk} = \beta_{00k} + \beta_{01k}(LICENSE) + r_{0jk}$ $\pi_{1jk} = \beta_{10k} + r_{1jk}$ $\pi_{2jk} = \beta_{20k} + r_{2jk}$	$\beta_{00k} = \gamma_{000} + \gamma_{001}(LUNCH) +$ $\gamma_{002}(ENRICH / REMEDIAL) + u_{00k}$ $\beta_{01k} = \gamma_{010}$ $\beta_{10k} = \gamma_{100}$ $\beta_{20k} = \gamma_{200}$

Table 30

Fixed and Random Effects of Behavioral Engagement (Research Question 3)

		Parameter	Model 3-A	Model 3-B	Model 3-C	Model 3-D	Model 3-E	Model 3-F
Fixed Effects								
Mean Science Achievement	Intercept	γ_{000}	151.27** (0.80)	151.27** (0.83)	142.55** (3.99)	145.26** (2.89)	148.19** (1.42)	152.60** (1.87)
	Free/Reduced Lunch	γ_{001}						-0.13** (0.02)
	Enrichment/ Remedial	γ_{002}						1.82* (0.84)
	License	γ_{010}			3.84* (1.97)	3.73** (1.41)	3.40* (1.37)	2.90* (1.47)
	Topics	γ_{020}			0.07 (0.04)	0.03 (0.03)		
Effect Of Behavioral Engagement I (Slope)	Intercept	γ_{100}	0.02 (0.21)	0.02 (0.22)	0.24 (1.33)	0.01 (0.22)	0.02 (0.22)	-0.001 (0.22)
	License	γ_{110}			0.44 (0.66)			
	Topics	γ_{120}			-0.01 (0.01)			
Effect Of Behavioral Engagement II (Slope)	Intercept	γ_{200}	-0.40* (0.19)	-0.42* (0.19)	0.34 (1.32)	-0.41* (0.19)	-0.41* (0.19)	-0.42* (0.19)
	License	γ_{210}			-0.56 (0.67)			
	Topics	γ_{220}			0 (0.01)			

Table 30 (Continued)

		Parameter	Model 3-A	Model 3-B	Model 3-C	Model 3-D	Model 3-E	Model 3-F
Variance Components								
Level 1	Students variation	e_{ijk}	58.64 (7.66)	57.43 (7.58)	57.43 (7.58)	57.43 (7.58)	57.43 (7.58)	57.40 (7.58)
Level 2	Class mean variation	r_{0jk}	23.85** (4.88)	50.51** (7.11)	49.86** (7.06)	50.27** (7.09)	50.65** (7.12)	45.38** (6.74)
	Behavioral engagement I slope variation	r_{1jk}		2.38** (1.54)	2.34** (1.53)	2.36** (1.53)	2.38** (1.54)	2.26** (1.50)
	Behavioral engagement II slope variation	r_{2jk}		1.97* (1.40)	1.97* (1.40)	1.98* (1.41)	1.98* (1.41)	2.15* (1.47)
Level 3	School mean variation	u_{00k}	18.04** (4.25)	18.28** (4.28)	18.53** (4.07)	16.56** (4.07)	17.29** (4.16)	5.31** (2.30)
	Deviance		30635.35	30678.42	30671.93	30673.22	30674.45	30617.18
	# Of Estimated Parameters		6	11	17	13	12	14
	Δ deviance			43.07**	6.49	1.29	1.23	57.27**
	Δ *df*			5	6	4	1	2

Note: **, p<.01
*, p<.05
+, p<.10

Research question three

Table 29 shows the six steps used to derive the final model of the effect of behavioral engagement. Model 3-A presents a simple three-level HLM with two behavioral engagement factors as level 1 predictors. There are no random effects of slopes. It can be seen that behavioral engagement I (students' engagement in scientific investigation activities) is not significant at .05 level. However, deviance difference test shows significant change ($\chi^2(1)$ =30641.37 - 30635.35 = 6.02, p<.05). This predictor remains in this model.

Model 3-B presents a random-coefficients regression model with two behavioral engagement predictors. Both intercept π_{0jk} (class mean science achievement) and slopes π_{1jk} and π_{2jk} (effect of two behavioral engagement factors) vary randomly. There are no class and school level predictors in this model. Table 30 elaborates that the three level 2 variance are significant at 0.05 level (r_{0jk} =50.51, p<.01; r_{0jk} =2.38, p<.01; r_{1jk} =1.97, p<.05). There is evidence that class mean science achievement and the effect of behavioral engagement in students' class activities on science achievement are different by class within school.

Model 3-C presents an "Intercepts- and Slopes-as Outcomes" model. Class level factors are added (science teacher has full license and overall taught topics) to predict the class mean achievement (π_{0jk}) and the effect of two behavioral engagement factors on science achievement (π_{1jk} and π_{2jk}). We found that the both effects on "BEHAVIORAL ENGAGEMENT I" slope are not significant (γ_{110} =.44, p>.05 and γ_{120} =-.01, p>.05). Similarly, no significant effects were found on "BEHAVIORAL ENGAGEMENT II" slope (γ_{210}=-.56, *p*>.05 and γ_{220}=.00, *p*>.05). Furthermore, deviance difference test doesn't show significantly improvement when all four parameters are added simultaneously ($\Delta\chi^2(6)$=6.49, *p*>.05).

Model 3-D is a parsimonious model than Model 3-C. Insignificant effects of "teacher has full science license" and "overall topic coverage" on two behavioral engagement slopes were deleted from model 3-C. Deviance difference test shows $\Delta\chi^2(4)$= 1.29 (*p*>.05), which means no significant difference between these two models. The effect of "topic coverage" is not significant on the class mean science achievement ($\gamma_{020} = .03$, *p*>.05).

Model 3-E ignores the insignificant effect of "topic coverage" on school mean science achievement from Model 3-D. Deviance difference test shows $\Delta\chi^2(1)$ = 1.23 (*p*>.05), which is no significant shrinkage. Deviance difference between Model 3-B and Model 3-E ($\Delta\chi^2(1)$= 30678.42 - 30674.45= 3.97, *p*<.05) shows significant improvement when class-level OTL predictor "LICENSE" is added.

Model 3-F adds level 3 (school level) predictors: "percentage of free/reduced lunch" and "school offers enrichment/remedial science classes". Both of these two variables significantly predict the school-mean science achievement. No other random effects have been found besides u_{00k} . In other words, the class level effect of "science teacher has full license" and mean effect of two "behavioral engagement" factors on achievement are the same across all schools. Detailed descriptions as follows:

level-1 model:

$$\text{Ach}_{ijk} = \pi_{0jk} + \pi_{1jk}\ (\text{BEHAVIOR1}) + \pi_{2jk}\ (\text{BEHAVIOR2}) + e_{ijk} \qquad [4.18]$$

where

Ach_{ijk} represents the Science Achievement of each student i in class j and school k.

π_{0jk} represents the class mean Science achievement of class j in school k.

π_{1jk} represents the effect of behavioral engagement I in scientific investigation such as conduct experiment on science achievement of class j in school k.

π_{2jk} represents the effect of behavioral engagement II in innovative activities such as relate science to society on science achievement of class j in school k.

e_{ijk} represents the random error of student i in class j and school k

The level-2 model represents the variation in behavioral engagement parameters, π_{1ij}, between classes within schools. The only class level OTL variable remaining in this model is LICENSE. And all three random effects at class level are significant.

$$\pi_{0jk} = \beta_{00k} + \beta_{01k} \text{(LICENSE)} + r_{0jk} \quad [4.19]$$

$$\pi_{1jk} = \beta_{10k} + r_{1jk} \quad [4.20]$$

$$\pi_{2jk} = \beta_{20k} + r_{2jk} \quad [4.21]$$

where

β_{00k} represents the mean school science achievement in school k.

β_{01k} represents the effect of "LICENSE" on class mean science achievement in school k.

β_{10k} represents the mean effect of behavioral engagement I on science achievement in school k.

β_{20k} represents the mean effect of behavioral engagement II on science achievement in school k.

r_{0jk} represents the random error of class mean of class j within school k

r_{1jk} represents the random error of mean effect of behavioral engagement I of class *j* within school *k*

r_{2jk} represents the random error of mean effect of behavioral engagement II of class *j* within school *k*

Equations [4.19] can be explained that LICENSE was related to class mean science achievement. Equation [4.20] and [4.21] specified that the effects of student behavioral engagement in class activities varied by class within each school *k*.

The level-3 model represents the variability among schools in the four β coefficients. The results are similar as in the first two questions: Offering enrichment or remedial class and school rate of reduced/free lunch predict school mean science achievement. But school level mean effect of teacher has full license (β_{01k}) is constant across schools. Also, the effect of two behavioral engagement factors (β_{10k} , β_{20k}) were constant across all schools:

$$\beta_{00k} = \gamma_{000} + \gamma_{001}\text{ (REDUCED LUNCH)) } + \gamma_{002}\text{(ENRICHMENT/REMEDIAL } + \mu_{00k} \quad [4.22]$$

$$\beta_{01k} = \gamma_{010} \quad [4.23]$$

$$\beta_{10k} = \gamma_{100} \quad [4.24]$$

$$\beta_{20k} = \gamma_{200} \quad [4.25]$$

School level effects on science achievement are similar as in research question 1 and 2. Since it only interprets the intercept (mean science achievement), the γ_{000} coefficient represents in this application predicted school mean science achievement for a school which does not offer any enrichment or remedial science classes (ENRICHMENT/REMEDIAL = 0) and is an “affluent school” (FREE/REDUCED LUNCH = 0%). For such a school, the predicted science achievement is 152.60. For each 10% increment in free/reduced lunch rate, the expected initial status is reduced by

1.3 points (γ_{001}=-.13, p<0.1). For those schools that offer science enrichment or remedial class, their school mean are 1.82 points higher than others (γ_{002}=1.82, *p*<0.5).

For class level effect, at the class mean science achievement, the gap of teacher has full license or not, γ_{010} which is identical with β_{01k}, is 2.90 (*p*<.05) because there is no significant random effect. This means that for a teacher who has no full license or certificate (LICENSE=0) the class mean starts out 2.90 points behind for classes where teachers have full license or certificate. The class level effect of "science teacher has full license" on achievement is the same across all schools.

The predicted mean effect of two behavioral engagement factors on school mean science achievement β_{10k} and β_{20k}, are identical with γ_{100} and γ_{200}. That means the fixed effects of two behavioral engagement factors were - .001 (*p*>.05) and -.42 (*p*<.05) respectively. The mean effects of two "behavioral engagement" factors on school mean achievement are constant across all schools. The effect of behavioral engagement I (In-class regular scientific activities) was almost 0. The effect of engagement II was negative and much stronger. That means when students spend class time in thinking the importance of science and its relationship to society and students report studying the relationship of science to society, its effect on student achievement is negative. This effect was similar in SEM analysis. A discussion of this interesting phenomenon is included in SEM model. It is possible that such emphasis cuts down the time on other classroom activities which may have positive relationship to science achievement. Another explanation is a mismatch between the second factor (student engagement in innovative activities related to the role of science in society) and the science achievement test. The science test focused mainly on content knowledge and not the implications of science for society.

Estimated variance and related χ^2 statistics from the three-level decomposition are presented in Table 30. All three level intercept variance were significant (σ_e^2 = 57.40; τ_π =45.38, p<.01; τ_β =5.31, p<.01). That means students science achievement scores were varied by class and in turn, varied by school. The class level variance of effect of two behavioral engagement factors (Var(r_{1jk}) and Var(r_{2jk})) are 2.26 and 2.15 respectively which are significant. That means the effect of behavioral engagement in class activities were varied by class; different classes had different magnitudes of effects. The correlation coefficient between r_{1jk} and r_{2jk} is -.55 (Cov(r_{1jk} , r_{2jk}) =-1.21), so it can be concluded that two behavioral engagement factors are highly related.

CHAPTER FIVE

SUMMARY, CONCLUSION, AND DISCUSSIONS

The purpose of this study was to test a model of the mediating effects of student science engagement on the relationship between the opportunity to learn (OTL) and the students' science achievement. This study used two different methods (structural equation modeling (SEM) and hierarchical linear modeling (HLM) to analyze the relationships among OTL, engagement, and science achievement. This chapter presents a summary as well as a discussion of the findings. Limitations and recommendations for future research conclude the chapter.

Summary of Research Findings

Findings from Structural Equation Models

Two-step approach was used in SEM analysis, first the measurement model was tested; the reliability and validity of observed variables were examined, as well as the relationships between observed and latent variables. Seven latent constructs had been found in this study. They were: family socioeconomic status, two opportunity to learn factors measured by teaching instruction, two students' behavioral engagement factors measured by students' class activities, students' emotional engagement, and students' cognitive engagement. And then structural model was tested; the relationships among the latent variables were interpreted. Following is a summary of the results of the structural relationships in the model.

Effects of opportunity to learn (OTL) factors

In this study we hypothesized that teachers' instructional practices provide students an opportunity to learn. For example, if the teacher focuses on the activities that are related to conducting experiments, writing results etc, she/he is providing students opportunities to learn scientific methods. Otherwise, if the teacher focused on activities that are related learning into society, she/he is providing students opportunities to learn the nature and implications of science. Thus, the teachers' account of classroom activities formed the two OTL factors considered in the study: the first was the focus on teaching scientific method; the other was teaching science in relation to society. The results of the study indicated that OTL is significantly related to students' behavioral engagement. The teacher's instruction, when it emphasized the methods of scientific investigation, such as asking students conduct experiments etc., had a significant positive effect on students' engagement in science classroom activities. The effect on scientific activities was much stronger than the effect on activities related to science's connection to society. The link between the teacher's instructional emphasis on science's relation to society with the students' engagement in scientific activities was negative. It means that, if the teacher spends time on activities that focus on the relationship of science to society, students have less time to do core science activities, such as conducting experiments and writing scientific explanations.

No significant direct effects were found from OTL factors to students' emotional engagement, cognitive engagement and science achievement. However, slight but significant indirect relationships of OTL to emotional, cognitive engagement, and science achievement were discovered through behavioral engagement. When teachers emphasize scientific investigation in experiments, it slightly encourages the students' emotional and cognitive engagement, and in turn increases the students' science achievement. When the teachers' instructional activities emphasize science's relation to society, students tend to have a low achievement. In sum, the first OTL factor had direct positive effects on behavioral engagement and small indirect but significant positive effects on students' emotional and cognitive engagement in science learning. The second OTL factor had negative effect on students' behavioral engagement in scientific investigation activities. Both OTL factors had minor indirect but significant effects on science achievement. The first OTL factor was positive, while the second one was negative.

Effects of socioeconomic status (SES)

SES played an important role in both science engagement and science achievement. This finding confirms the earlier findings about the strong effect of SES on students' engagement and learning in science (Lee & Luykx, 2007). Students, who are from a high SES family, are more likely to have a strong interest in science as well as feel more efficacy in regard to science learning. High SES students are also more likely to think that science is important for their future and they express interest in pursuing scientific careers. No doubt, these students have higher science achievement. The relationship between SES and the students' behavioral engagement in scientific investigation is positive. Students from a high SES family are more likely to engage in conducting experiments or writing explanations. A relationship between SES and student engagement in connection to society was not found. Thus, SES had the strongest direct and indirect effects on science achievement mediated through the emotional and cognitive engagement of the students in science. These results are not surprising and have been found in many studies of science achievement. What the study contributes is the evidence in support of the mechanism of the effect of SES on achievement. A part of the effect of SES is through students' engagement.

Effects of engagement on science achievement

The study also considered the role of student engagement in science achievement. Three engagement factors were included in the model: emotional or affective engagement, cognitive engagement and behavioral engagement. To some extent the engagement factors acted as the mediators between the classroom learning opportunities and science achievement, as was hypothesized. Both cognitive and behavioral engagement had direct and indirect effects on the science achievement while the effect of emotional engagement was direct only. The students' emotional and cognitive engagement had a direct positive effect on science achievement, and these two kinds of engagement were strongly related. Students, who were interested in science and considered it important tended to also value it for their future education and occupational aspirations. In turn, students who had emotional and cognitive engagement in science learning were higher achievers in science. Realizing the importance of science for future education would promote student interest in the subject. The effect of behavioral engagement on science achievement was complex. When students engaged in science experiments such as designing or conducting experiments, they had higher achievement. However, when students studied implications of science to society in class, their science achievement was lower. Deep explanations are investigated in discussion part of this chapter. Overall, emotional and cognitive engagement was important for science achievement. The model explained about 24% variance in science achievement.

Findings from HLM analyses

HLM analyses were considered for the study due to the nested structure of the data. HLM analysis reexamined the effects of emotional, cognitive and behavioral engagement on science achievement, taking class and school membership into account. About 23% of the total variance in science achievement was explained by class level and 18% by school level variables. That indicates that there is a large variation in students' science achievement among classes and schools.

<u>Student level analysis</u>

Emotional and cognitive engagement had significant positive effects on science achievement. In general, liking science and realizing its importance lead to high science achievement. The effects of emotional and cognitive engagement differ between classes within a school. So the effects of engagement for certain classes seemed to stronger while some other classes seemed to weaker.

The effects of the two behavioral engagement factors on science achievement were different. There is a slightly positive effect of behavioral engagement of students' scientific investigation activities, such as conducting experiments, on achievement when school OTL variables are not taken into account. After adding the class and school level predictors, the effect of engagement in student scientific investigation activities shrank to 0. Collinearity can explain this shrinkage because this engagement factor was correlated with school SES.

Similar to the results of SEM, the effect of the second behavioral engagement factor was negative, too. Students who spent much time learning the relation of science to society did not do well on the science test. We can only speculate on the reasons for such a result. These are discussed in the discussion part of the chapter.

Class/teacher level analysis

We considered two different OTL factors for this part of the analysis: teacher license and topics covered in class. The two class level OTL variables, the science teacher's full license, and the topic coverage, significantly influenced the class mean science achievement. The effect of a science teacher having a full license is strong. Teachers who had a full teaching science license or certificate got higher class mean achievement than those who have not. This is a significant policy relevant finding, given that many teachers who teach science are not licensed to teach science. There is not only a shortage of science teachers nationwide but it is also becoming a serious problem to recruit and retain science teachers. Given these conditions it is likely that number of teachers who teach science without having a license is likely to increase.

The second teachers' level factor considered was the percentage of topics covered by the teachers. Topic coverage was found to be related to science achievement. There was also an interaction effect of topic coverage and student emotional engagement. For those students who are not interested in science, the relationship between topic coverage and science achievement is positive. It indicated that when the students were low on emotional engagement, the more topic coverage resulted in higher achievement. But for more interested students, the larger coverage of topics was not effective in increasing achievement. On the contrary, for those students who are interested in science, the relationship between topic coverage and science achievement was slightly negative. We can only speculate about the reasons for these results. It is possible that more interested students need deeper engagement in the topics and not a surface level instruction on many topics. It also reminds us of the criticism that some researchers based on their analysis of international data have leveled against science teaching in U.S. They have argued that the curriculum in U.S. is a mile long and an inch deep (Schmidt, McKnight, & Raizen, 1997).

School level analysis

For the school level predictors we used two factors: SES as measured by the free/reduced lunch, and whether the school offered enrichment or remedial programs in science. Both of these factors were significant predictors of science achievement of a school. As expected, school SES measured by free/reduced lunch rates and schools that offer science enrichment or remedial programs significantly affected the school mean achievement. Lower percentages of free/reduced lunches signify “affluent schools”. Such schools had higher mean science achievement. Schools that offer science enrichment or remedial programs also had a higher science achievement than those with no enrichment/remedial programs. These results confirm that high SES schools and schools with more academic options are likely to have higher achievement.

Discussion and Implications

The results of the study are both theoretically and practically significant. These results confirm the findings of the earlier studies and show that the students’ family background and students’ interest and self-efficacy play a dominant role in their achievement and aspirations in science. The study also confirmed that students’ engagement factors do mediate the relationship of classroom teaching and science achievement. The classroom teaching's effect is through students' classroom activities. Yet the relationship between teachers' instruction activities and students' science achievement in the final SEM model was weak and indirect. A possible explanation for the weak relationship is that teachers' instruction and students' class activities share the same variance in the students' science achievement. In other words, the effects of teacher OTL measured by teachers' instruction on students' achievement also can be explained by students' class activities (behavioral engagement). The direct effect of student activities is stronger than teacher instruction on the student science achievement, because "the instruction some students receive might be more effective than what other students receive" (Le et al., 2006 p.77). In other words, students' behavioral engagement in science class activities is related to the teachers' instruction and, also, is a psychological investment in learning. The effect of OTL in teachers' instruction on student science achievement is mediated by student behavioral engagement. This finding is consistent with the results of some earlier studies. For example, Le et al (2006) found an insignificant relationship between reform-oriented instruction in science and student achievement in a three-year longitudinal study. Furthermore, this study found an indirect effect of the teachers' instruction on achievement. When teachers ask students to do more science activities, such as conducting experiments, students are more engaged in these scientific investigations, which positively affect their science achievement. On the contrary, when teachers ask students to think more about science as it relates to society, students are less engaged in regular science class activities such as conduct an experiment or investigation. As a result, the effect of this OTL factor (instruction in connection of science to society) on student achievement becomes negative. For increasing students' achievement in science it is important to focus instruction on science content and hands-on learning activities.

Two student behavioral engagement factors were considered: one was scientific investigation, such as conducting experiments or writing an experiment's explanation; the other was activities that focus on science's relation to society. The first factor positively affected student science achievement while the latter one had a negative effect. Gamoran and Nystrand (1992) explained that the class discourse is related to the teacher's quality of instruction. When students engage in substantive thinking and conversation on one topic, the topic must be covered in appropriate detail, and teachers have to expand one lesson to several classes, which may lead to less time on other activities. Moreover, they found that engagement in class discourse is different between high and low ability students. Engagement in discourse "had positive effects in honors classes and negative effects in remedial classes." Experienced teachers of honors classes more often focused on pertinent questions relating to tests, thus students achieved higher on exams. To promote students' achievement, teachers should avoid too many activities that do not form the core of science learning. The negative effect of the instructions in connection of science to society could also be related to the mismatch between teaching curriculum and test content. This is an issue which is extensively debated because of the rise in high stakes testing.

It is important to note that although the effect of OTL in instruction in connection of science to society and the related behavioral engagement factor II both had negative effects on science achievement, the effect of second behavioral engagement on cognitive engagement was positive. So it had some interesting and inconsistent effects. Clearly more research is needed to investigate the direct and indirect effect of science instruction that includes discourse on the nature of science and its relationship to society. We should not be too quick in discarding such discourse because it does stimulate cognitive engagement and makes students aware of the value of science. Secondly, it is important for teachers and curriculum specialists of science to understand and examine the proper balance of various classroom activities and topics to be covered. Because class engagement in discussion of science in relation to society also encourages students to value science and, in turn, student achievement, the frequency of these kinds of activities needs to be further investigated. Teachers should ponder the right mix of students' behavioral engagement in class scientific investigation and connection to society activities, and try to find a balance of these two factors in order to promote the students' science interest and achievement. Furthermore, it is important to create assessment that aligns well with teaching and curriculum. This is an area of continued debate and interest in increasingly intense testing environment

Approximately 24% of the variance of science achievement is explained by the overall SEM model. Only 4% was accounted for by teacher instruction and student class behavioral engagement. A large amount of variance (20%) is from students' emotional, cognitive engagement and SES. This finding confirmed Snow's (1992, 1994, 1996) aptitude theory that cognitive and conative-affective processes affect achievement. It suggests that it is important for teachers to encourage student interest in science and enhance their efficacy in science learning. Emotional and cognitive engagements were significantly related to SES which was, no doubt, a predominant predictor of student achievement. The reason is that parents of a higher socioeconomic status have more money and resources to provide better opportunities for their children's learning (e.g., enrolling in a high-performing school, attending extra lessons, or looking for personal mentors).

Another interesting result found in this research is there was an interaction between emotional engagement and topic coverage. For those students who had high emotional engagement in science learning, the effect on science achievement was slightly negative. It meant that for students with high emotional engagement the more the topics covered in class, the lower their achievement. It is important to note this effect is very small. One can only speculate about the reasons for such an effect. It is possible that these students want a deeper learning of the course material and are not satisfied with larger coverage of topics. These high-performing students may not be satisfied with the topics they learned in class; they may seek extra enrichment lessons, so the effect of class topic coverage may not quite relate to their achievement. On the other hand, for those students who are low on emotional engagement, the effect of topic coverage was positive. The more topics covered in class, the higher their achievement. The magnitude of the negative effect for high-performing students is slight while the positive effective for low-performing students is strong. In any case, the reason for these complex effects needs to be investigated further. It also brings to mind the more general critique of U.S. science curriculum that compared to curriculum in other developed countries teachers here cover many more topics while in other countries teachers cover fewer topics but in greater depth. This finding suggests that teachers and curriculum specialists need to be thoughtful about the topics covered in a course and should strive for an optimal balance between breadth and depth in knowledge. Reformed science instruction, with student-centered pedagogy and proper balance of in-depth coverage of topics may be more suitable for all students.

Overall, the findings of the study have implications for teachers, educators and policy makers. The results of the study show important effects of engagement and classroom opportunities to learn on student achievement. Parents, science teachers and school administrators should work to enhance the students' engagement in science learning. Clearly, affective factors such as interest in science and students' self-efficacy in science are individual level variables that are highly correlated with science achievement.

Although the individual characteristics of students such as family socioeconomic status, emotional and cognitive engagement in learning had large effects on achievement, the most policy relevant findings of the study are about the effects of the school level factors. The results of the study suggest that schools that offer science enrichment or remedial programs were likely to have higher overall science achievement. Thus, expanded academic opportunities to meet the needs of both high achieving and low achieving students promote science achievement. Another important and policy related finding was that teachers with science teaching license had higher overall class mean for science achievement. Teacher quality and preparation are important ingredients for improvement of science education. This issue will become increasingly salient as the teacher shortages in science and mathematics make it difficult to recruit and retain licensed science teachers, especially in resource poor schools. These schools are also more likely to have high risk student populations. This will exacerbate the poor achievement of students who come from lower SES communities to poor resource schools with fewer licensed teachers. One recommendation is the on-going professional development opportunities for teachers. Federal and state agencies should assist school districts in training their teachers and staff. Teachers not only need to learn how to effectively teach the diverse students and find optimal scientific instruction methods, but also need to encourage student engagement in science to promote their achievement.

Contribution of the Study

Conceptual Contribution

The main purpose of this study was to examine the relationship of opportunity to learn, science engagement, and science achievement. Empirical research on engagement is inconsistent on its dimensions as well as its relationship to teaching practices and achievement. The main contribution of this study is an expanded framework to understand and examine differences in science achievement. The study included many complex variables and relationships in its model of science achievement. The results of the study present a complex picture of how teachers' practices are related to what students do and how these different types of engagement in learning lead to achievement. The conceptual model in this study expanded Greenwood's (1996) performance-based models of instruction presenting engagement as being emotional, cognitive, and behavioral, and adding students' background information as an exogenous variable. In addition, class and school level variables were also included to model the complex nested structures of the setting in which learning takes place. The study also provides ideas for future research.

Methodological Contribution

The main methodological contribution of this study is the use of both structural equation modeling (SEM) and hierarchical linear modeling (HLM). The use of SEM not only enabled the specification of relations between observed variables and latent constructs, but also the relationship among latent constructs. The mediating effects of engagement and the correlations between exogenous variables could not be assessed without the use of this methodology. Moreover, SEM provides the fit indices to assess how well the model fits the data. Such fit indices make it possible to evaluate the adequacy of the theoretical model in explaining the data. HLM allows variance in outcome variables to be analyzed at multiple hierarchical levels. The three-level hierarchical linear model used in this study explained the schools and teachers' variability. Both methodologies in conjunction provided a much better insight into the complexity of science achievement and provided a deeper understanding of why there is such a large variability in science learning outcomes.

Limitations of the Study

While this study makes conceptual and methodology contribution to research in science education, like any research study it has some limitations. First, this study is based on secondary dataset. Although secondary dataset has several advantages such as timeliness, no cost, large sample size and generalizability of findings, it does not provide indicators that completely fit the research question. TIMSS, the dataset used in this study provides extensive background information to address concerns about the quantity, quality, and content of instruction. However, "secondary studies must strategically select and work with the data from those studies, such as being aware of the constraint of the data while interpreting the results, interpreting data combining several items to create appropriate indicators" (Li et. al., 2006. p.305). For example, one dimension of OTL, teachers' instruction, was measured by the frequency of teachers' use of certain methods. The frequency here can not represent the quality of instruction. In addition, there are no existing variables in TIMSS to measure the content emphasis and quality of instructional delivery. This could be one reason why this study did not find strong effect of OTL on science achievement and the effect size is very small. Clearly the OTL measures in TIMSS data were limited. The same is true of engagement measures; engagement is a multidimensional construct which includes a broad perspective. This study used students' interest in science, value of science and science class activities to measure emotional, cognitive, and behavioral engagement respectively. There can be many more measures of these dimensions of engagement. Behavioral engagement should be of utmost importance to achievement but the findings of the study did not support it. It is most likely that these classroom activities are not good measures of behavioral engagement. Another speculation for poor results of science behavioral engagement and especially the negative effect of engagement factor that included indicators where students relate science learning to daily lives or present work to the class: Does this reflect good teaching or a classroom where teacher has relinquished the teaching role?

Another limitation of the study included reliance on self-reported data. There is measurement error in such data. Science engagement construct should be further studied so better measures of both in-class and out of class engagement in science learning can be developed. Despite these limitations the study contributes to better understanding of science learning and has implications for practice.

Forty one percent of variance in students' science achievement was explained by class/teacher and school where 23% was from class/teacher level. The large amount of variance between classes indicates that science teacher quality does matter and has a strong effect on the students' science achievement. However, there is a limitation of this dataset design that the number of classes per school is too small to estimate a model with multiple random slopes at level-2 (class/teacher level). In TIMSS 2003, only 1 to 3 classes were selected in sampled schools. One suggestion to the survey designers is that more classes per school are needed in order to investigate the variability of the classes.

Directions for Future Research

This study raises some questions for further research. First, the effects of students' behavioral engagement in classroom activities were small and in case of the second factor unexpected. Engagement in regular scientific experiment or investigation positively affects science achievement while the effect of engagement in learning the connection of science to society is negative. However, these two engagement factors are positively correlated. We can only speculate about these results. These findings certainly provide rationale for more research, based on new data. Gamoran and Nystrand (1992) have indicated that behavioral engagement is related to students' ability and teachers' quality. This study did not address these issues but these ideas should be further studied to provide solid empirical results. Furthermore, more demographic variables such as ethnicity should be included in understanding differences in achievement. More qualitative and mixed methods research is needed for deeper knowledge of classroom practices. These observational methods offer the most promise for furthering understanding of class room instructional practices and study of engagement.

In terms of methodology this study applied structural equation modeling (SEM) and hierarchical linear model (HLM) separately. An advance method, multilevel structural equation modeling (MSEM) has been proposed as an extension to structural equation modeling for analyzing data with nested structure, but few applications have been used in educational research. Future researchers could reexamine the results of this study by the advanced statistical methods. Certainly the study points to the need for more research using newer methods and deeper conceptualization of the study's constructs.

Conclusions

This study corroborated the earlier results by researchers and provided further evidence in support of the relationship among opportunity to learn, science engagement, and science achievement. It provided evidence that students' science achievement and the effects of various types of engagement are varied across classes and schools due to different opportunities and conditions. Our results further support that including the family background and class room opportunities improves the explanation of individual differences in science achievement. Emotional and cognitive engagement explained a large variance in science learning. The study has implications for practice both at individual and school level. Enhancing engagement in learning and increasing the opportunities to learn science for all students will have positive effect on science achievement.

References

Alkin, M., Doby, W., & Lindheim, E. (1990). *The schools program evaluation reports case studies: 1988-89 updata*. Los Angeles: LAUSD.

American Association for the Advancement of Science Washington DC. (1989). *Science for all Americans: Summary, project 2061*: Andrew W. Mellon Foundation, New York, NY. Carnegie Corp. of New York, NY.

Anderman, E. M. (2002). School effects on psychological outcomes during adolescence. *Journal of Educational Psychology, 94*, 795-809.

Anderson, B. J. (1990). Minorities and mathematics: The new frontier and challenges of the nineties. *Journal of Negro Education, 59*(3), 260-272.

Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and recommended two-Step approach. *Psychological Bulletin, 103*, 411-423.

Angell, C., Kjærnsli, M., & Lie, S. (2006). Curricular effects in patterns of student responses to TIMSS science items. In S. J. Howie & T. Plomp (Eds.), *Contexts of learning mathematics and science* (pp. 277-290). London: Routledge.

Bagozzi, R. P. (1991). Further thoughts on the validity of measures of elation, gladness, and joy. *Journal of Personality and Social Psychology, 61*, 98-104.

Baloglu, M. (2000). *A brief introduction to structural equation modeling techniques: Theory and application*. (ERIC Document Reproduction Service No. ED 438 300).

Bandura, A. (1994). Self-efficacy. In V. S. Ramachaudran (Ed.), *Encyclopedia of human behavior* (Vol. 4, pp. 71-81). New York: Academic Press. (Reprinted in H. Friedman [Ed.], *Encyclopedia of mental health*. San Diego: Academic Press, 1998).

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Berends, M. (1995). Educational stratification and students' social bonding to school. *British Journal of Sociology of Education, 16*(3), 327-352.

Bollen, K. A. (1989). *Structural equations with latent variables*. New York: John Wiley & Sons.

Brophy, J., & Good, T. (1986). Teacher behavior and student achievement. In M. Writtrock (Ed.), *Handbook of research on teaching*. New York: MacMillan Publishing.

Butler, M. B. (1999). Factors associated with students' intentions to engage in science learning activities. *Journal of Research in Science Teaching, 36*(4), 455-473.

Byrne, B. M. (1998). *Structural equation modeling with LISREL, PRELIS, and SIMPLIS: Basic concepts, applications, and programming*. Multivariate applications book series. Mahwah, N.J.: L. Erlbaum Associates.

Chang, M., Singh, K., & Mo, Y. (2007). Science engagement and science achievement: Longitudinal models using NELS data. *Educational Research and Evaluation, 13*(4), 349-371.

Chapman, E. (2003). Alternative approaches to assessing student engagement rates. *Practical Assessment, Research & Evaluation*, 8(13). Retrieved from http://PAREonline.net/getvn.asp?v=8&n=13.

Connell, J. P. (1994). Educational risk and resilience in African-American youth: Context, self, action, and outcomes in school. *Child Development, 65*(2), 493-506.

Connell, J. P., & Wellborn, J. G. (1991). Competence, autonomy, and relatedness: A motivational analysis of self-system processes. In M. R. Gunnar & L. A. Sroufe (Eds.), *Minnesota Symposium on Child Psychology* (Vol. 23, pp. 43-77). Hillsdale, N.J.: Erlbaum.

Cooper, H., Robinson, J. C., & Patall, E. A. (2006). Does homework improve academic achievement? A Synthesis of research, 1987-2003. *Review of Educational Research, 76*(1), 1-62.

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

Duschl, R. A., Shouse, A. W., & Schweingruber, H. A. (2007). What research says about K-8 science learning and teaching. *Principal, 87*(2), 16-22.

Eccles, J. S., Wigfield, A., & Schiefele, U. (1997). Motivation to succeed. In W. Damon & N. Eisenberg (Eds.), *Handbook of child psychology* (5th ed., Vol. 3, pp. 1017-1095). New York.

Feltz, D. L., & Riessinger, C. A. (1990). Self-efficacy; endurance; man; woman; imagery; feedback. *Journal of Sport & Exercise Psychology, 12*(2), 132 - 143.

Finn, J. D. (1989). Withdrawing from school. *Review of Educational Research, 58*, 117-142.

Finn, J. D. (1993). *School engagement and students at risk.* Washington, DC: National Center for Education Statistics.

Finn, J. D., & Rock, D. A. (1997). Academic success among students at risk for school failure. *Journal of Applied Psychology, 82*, 221–234.

Finn, J. D., & Voelkl, K. E. (1993). School characteristics related to student engagement. *Journal of Negro Education, 62*(3), 249-268.

Finn, J. D., Pannozzo, G. M., & Voelkl, K. E. (1995). Disruptive and inattentive withdrawn behavior and achievement among fourth graders. *Elementary School Journal 95*, 421-454.

Floden, R. E. (1981). Responses to curriculum pressures: A policy-capturing study of teacher decisions about content. *Journal of Educational Psychology, 73*(2), 129-141.

Fornell, C., & Larcker, D. (1981). Structural equation models with unobservable variables and measurement error. *Journal of Marketing Research, 18*(1), 39-50.

Fredricks, J. A., Blumenfeld, P. C., & Paris, A. H. (2004). School engagement: Potential of the concept, state of the evidence. *Review of Educational Research, 74*(1), 59-109.

Freedman, M. P. (1997). Relationships among laboratory instruction, attitudes toward science, and achievement in science knowledge. *Journal of Research in Science Teaching, 34*, 343-357.

Furlong, M. J., Whiple, A. D., St. Jean, G., Simental, J., Soliz, A., & Punthuna, S. (2003). Multiple contexts of school engagement: Moving toward a unifying framework for educational research and practice. *The California School Psychologist, 8*, 99-113.

Gamoran, A., & Nystrand, M. (1992). Taking students seriously. In F. M. Newmann (Ed.), *Student engagement and achievement in American secondary schools* (pp. 11-61). New York: Teachers College Press.

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford, CA: Stanford University Press.

Greenwood, C. R. (1996). The case for performance-based instructional models. *School Psychology Quarterly, 11*(4), 283-296.

Greenwood, C. R., Horton, B. T., & Utley, S. A. (2002). Academic engagement: Current perspectives on research and practice. *School Psychology Review, 31*(3), 328-349.

Guiton, G., & Oakes, J. (1995). Opportunity to learn and conceptions of educational equality. *Educational Evaluation and Policy Analysis, 17*(3), 323-336.

Hatcher, L. (1994). *A step-by-step approach to using the SAS system for factor analysis and structural equation modeling*. Cary, NC: SAS Institute.

Hayes, M. T., & Deyhle, D. (2001). Constructing difference: A comparative study of elementary science curriculum differentiation. *Science Education, 85*, 239-262.

House, J. D. (2004). Cognitive-motivational characteristics and science achievement of adolescent students: Results from the TIMSS 1995 and TIMSS 1999 assessments. *International Journal of Instructional Media, 31*(4), 411.

Hudley, C., Daoud, A., Hershberg, R., Wright-Castro, R., & Polanco, T. (2002, April). *Factors supporting school engagement and achievement among adolescents*. Paper presented at the Annual Meeting of the American Educational Research Association, New Orleans, LA.

Jensen, A. R. (1998). The suppressed relationship between IQ and the reaction time slope parameter of the hick function. *Intelligence*, 26(1), 43-52.

Jimerson, S. R., Campos, E., & Greif, J. L. (2003). Toward an understanding of definitions and measures of school engagement and related terms. *The California School Psychologist, 8*, 7-27.

Johnson, M. K., Crosnoe, R., & Elder, G. H., Jr. (2001). Students' attachment and academic engagement: The role of race and ethnicity. *Sociology of Education, 74*(4), 318-340.

Jöreskog, K. G., & Sörbom, D. (2006). LISREL 8.8. Chicago: Scientific Software International.

Ladd, G. W. (1990). Having friends, keeping friends, making friends, and being liked by peers in the classroom: Predictors of children's early school adjustment? *Child Development, 61*(4), 1081-1100.

Lau, S., & Roeser, R. W. (2002). Cognitive abilities and motivational processes in high school students' situational engagement and achievement in science. *Educational Assessment, 8*(2), 139-161.

Le, V. N., Stecher, B. M., Lockwood, J. R., Hamilton, L. S., Robyn, A., Williams, V. L., et al. (2006). Improving mathematics and science education: A longitudinal investigation of the relationship between reform-oriented instruction and student achievement. *RAND Corporation.*

Lee, O., & Luykx, A. (2006). *Science education and student diversity – Synthesis and research agenda*. New York: Cambridge University Press.

Lee, V. E. (2000). Using hierarchical linear modeling to study social contexts: The case of school effects. *Educational Psychologist, 35*(2), 125-141.

Leinhardt, G., & Seewald, A. M. (1981). Overlap: What's tested, what's taught? *Journal of Educational Measurement, 18*(2), 85-96.

Li, M., Ruiz-Primo, M. A., & Shavelson, R. J. (2006). Towards a science achievement framework: The case of TIMSS 1999. In S. Howie & T. Plomp (Eds.), *Contexts of learning mathematics and science: Lessons learned from TIMSS*. London: Routledge.

Manning, M. L. (1998). Gender differences in young adolescents' mathematics and science achievement. *Childhood Education, 74*(3), 168-171.

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

Martin, M. O. (Ed.). (2005). *TIMSS 2003 user guide for the international database*. Chestnut Hill: MA: TIMSS & PIRLS International Study Center, Boston College.

McDonnell, L. M. (1990). *Discovering what schools really teach: Designing improved coursework indicators. Joint Report JR-02*: Office of Educational Research and Improvement (ED), Washington, DC.

Mere, K., Reiska, P., & Smith, T. M. (2006). Impact of SES on Estonian students' science achievement across different cognitive domains. *Prospects: Quarterly Review of Comparative Education, 36*(4), 497-516.

Mo, Y., & Singh, K. (2008a). Parents' relationships and involvement: Effects on students' school engagement and performance. *RMLE Online: Research in Middle Level Education, 31*(10), 1-11.

Mo, Y., & Singh, K. (2008b, February). *The role of classroom instruction in science learning: Evidence from TIMSS 2003 data*. Paper presented at the Annual Meeting of Eastern Educational Research Association, Hilton Head, SC.

National Academy of Sciences - National Research Council Washington DC. (1996). *National science education standards*: Department of Education, Washington, DC. National Aeronautics and Space Administration, Washington, DC. National Science Foundation, Washington, DC. National Institutes of Health (DHEW), Bethesda, MD.

National Center for School Engagement. (2006). *Quantifying school engagement: Research report*. Colorado Foundation for Families and Children, Denver, CO

National Science Foundation Washington DC. National Science Board. (2008). Science and engineering indicators, 2008. Volume 1. NSB-08-01. *National Science Foundation.*

Newcomb, M. D., Abbott, R. D., Catalano, R. F., Hawkins, J. D., Battin-Pearson, S., & Hill, K. (2002). Mediational and deviance theories of late high school failure: Process roles of structural strains, academic competence, and general versus specific problem behavior. *Journal of Counseling Psychology, 49*(2), 172-186.

Newmann, F. M., (1992). Higher-order thinking and prospects for classroom thoughtfulness. In F. M. Newmann (Ed.), *Student engagement and achievement in American secondary schools* (pp. 62-91). New York: Teachers College Press.

Newmann, F. M., Wehlage, G. G., & Lamborn, S. D. (1992). The significance and sources of student engagement. In F. M. Newmann (Ed.), *Student engagement and achievement in American secondary schools* (pp. 11-39). New York: Teachers College Press.

Oakes, J., Gamoran, A., & Page, R. (1992). Curriculum differentiation: Opportunities, outcomes, and meanings. In P. W. Jackson (Ed.), *Handbook of research on* *curriculum: A project of the American Educational Research Association*. New York: Macmillan.

Papanastasiou, E., & Zembylas, M. (2004). Differential effects of science attitudes and science achievement in Australia, Cyprus, and the USA. *International Journal of Science Education, 26*, 259-280.

Pintrich, P. R., & De Groot, E. (1990). Motivated and self-regulated learning components of academic performance. *Journal of Educational Psychology, 82*, 33-40.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: applications and data analysis methods.* (2nd ed. ed.). Thousand Oaks: Sage Publications.

Rumberger, R. W., & Thomas, S. L. (2000). The distribution of dropout and turnover rates among urban and suburban high schools. *Sociology of Education, 73*(1), 39-67.

Schlechty, P. C. (2001). *Shaking up the school house: How to support and sustain educational innovation.* San Francisco: Jossey-Bass.

Schmidt, W. H., McKnight, C. C., & Raizen, S. (1997). *A splintered vision: An investigation of U.S. science and mathematics education*. Dordrecht, N.L.: Kluwer.

Schumacker, R. E., & Lomax, R. G. (1996). *A beginner's guide to structural equation modeling*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Schumacker, R. E., & Lomax, R. G. (2004). *A beginner' guide to structural equation modeling (2nd ed.)*. Mahwah, N.J.: Lawrence Erlbaum Associates.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: modeling change and event occurrence*. Oxford: Oxford University Press.

Snow, R. E. (1989). Cognitive-conative aptitude interactions in learning. In R. Kanfer, P. L. Ackerman & R. Cudeck (Eds.), *Abilities, motivation and methodology* (pp. 435-474). Hillsdale, N.J.: Lawrence Erlbaum Associates.

Snow, R. E. (1992). Aptitude theory: Yesterday, today and tomorrow. *Educational Psychologist, 27*, 5-32.

Snow, R. E. (1994). Abilities in academic tasks. In R. J. Sternberg & R. K. Wagner (Eds.), *Mind in context: Interactionist perspectives on human intelligence* (pp. 3-37). New York: Cambride University Press.

Snow, R. E., Corno, L., & Jackson, D. N. (1996). Individual differences in affective and conative functions. In D. C. Berliner & R. C. Calfee (Eds.), *Handbook of educational psychology* (pp. 243-308). New York: Simon & Schuster.

Sofroniou, N., & Kellaghan, T. (2004). The utility of third international mathematics and science study scales in predicting students' state examination performance. *Journal of Educational Measurement, 41*(4), 311-329.

Stedman, L. C. (1994). The sandia report and U.S. achievement: An assessment. *Journal of Educational Research, 87*(3), 133-147.

Stevens, F. I. (1993). Applying an opportunity-to-Learn conceptual framework to the investigation of the effects of teaching practices via secondary analyses of multiple-case-study summary data. *Journal of Negro Education, 62*(3), 232-248.

Stevenson, H. W., & Stigler, J. W. (1992). *The learning gap: Why our schools are failing and what we can learn from Japanese and Chinese education*. New York: Summit Books.

Venezia, A., & Maxwell-Jolly, J. (2007). *The unequal opportunity to learn in California's schools: Crafting standards to track quality* (Working Paper 07-2). PACE, Policy Analysis for California Education, University of California, Berkeley.
Wang, J. (1998). Opportunity to learn: The impacts and policy implications. *Educational Evaluation and Policy Analysis*, 20(3), 137-156.
Winfield, L. F. (1993). Investigating test content and curriculum content overlap to assess opportunity to learn. *Journal of Negro Education, 62*(3), 288-310.
Winfield, L. F., & Woodard, M. D. (1994). Assessment, equity, and diversity in reforming America's schools. *Educational Policy, 8*(1), 3-27.
Wyatt, T. (1991). The concept and measurement of opportunity to learn in the international context. *Draft for publication in the University of Sydney Monograph series*.
Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 21*, 3-17.