# Using TIMSS 2007 Data to Examine STEM School Effectiveness in an International Context

Author: Gabrielle Stanco



Boston College Electronic Thesis or Dissertation, 2012

BOSTON COLLEGE
Lynch School of Education

Department of
Educational Research, Measurement, and Evaluation

# USING TIMSS 2007 DATA TO EXAMINE STEM SCHOOL EFFECTIVENESS IN AN INTERNATIONAL CONTEXT

Dissertation
by

GABRIELLE STANCO

submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy

May 2012

# Abstract

Using TIMSS 2007 Data to Examine STEM School Effectiveness Factors in an International Context

Dissertation by Gabrielle Stanco

Advisor: Ina V.S. Mullis, Ph.D.

Because results from TIMSS 2007 showed a gap in mathematics and science achievement between students in the United States and those in the top–performing countries, TIMSS 2007 data were used to investigate how school effectiveness factors known to be strongly associated with higher STEM achievement operated in the United States compared to Chinese Taipei, the Czech Republic, Singapore, and Slovenia. In each of the five countries, multilevel modeling was used to examine STEM achievement in relation to 11 school effectiveness factors associated with school resources, fidelity of curriculum implementation, and school climate, controlling for student home resources. A secondary purpose of this dissertation research was to help the TIMSS & PIRLS International Study Center prepare for multilevel modeling planned for the TIMSS and PIRLS 2011 data.

Findings from this research showed that across the five countries, there were differences in how important school effectiveness factors operated. Teacher preparation, teaching the curriculum, and using instructional strategies involving reasoning and inquiry all were important school characteristics related to STEM achievement in some countries. A school environment conducive to learning emerged as being strongly

associated with high STEM achievement in three of the countries, including the United States. Both absence of discipline and attendance problems as well as a school climate supportive of academic success were important predictors of student STEM achievement. This dissertation research also showed the potential of using TIMSS data as a basis for conducting school effectiveness analyses across different country contexts.

# Acknowledgements

During the process of this dissertation study, I found myself thinking of the idiom, “It takes a village to raise a child.” Only in this case, it took a village to complete a dissertation. First and foremost, I would like to thank my dissertation advisor, Dr. Ina Mullis, and my readers, Dr. Michael Martin and Dr. Katherine McNeill. Dr. Mullis provided guidance and support throughout this process by keeping me on a timeline and continuously encouraging me to improve my research through constructive criticism of the many drafts submitted. Her thoughtful and detailed feedback helped to shape this dissertation into a work that I could be proud of. Dr. Martin gave invaluable advice on both the theoretical and technical aspects of this research. He helped me to keep thinking of the big picture and how all of the pieces come together to tell a story, guiding my work into a cohesive narrative. Dr. McNeill provided her time and expertise, helping in particular to develop the theoretical framework.

I am also grateful to my colleagues at the TIMSS and PIRLS International Study Center for providing a supportive and inspiring environment in which to work. Katie Trong Drucker supported me throughout the dissertation process through good advice, an ear to listen, and tasty treats. Alka Arora and Corinna Preuschoff were great sounding boards for methodological puzzles and Victoria Centurino and Karen Lam listened to my weekly progress updates and helped me stay inspired.

My family was also an important support system throughout the dissertation process. My mother, Michele Stanco, whose lifetime teaching career inspired me to pursue a degree in education, always kept me grounded with news from the classroom about what my results might mean. My sisters, Natalie and Hilary Stanco answered my endless phone calls with encouragement, support, and care and never tired of hearing about my work.

And last, but not least, I would like to thank my boyfriend, Matt Presley. From the beginning of the process, he has been supportive and encouraging, helping me to stay motivated and good-humored through the late nights and bleary-eyed revisions with dinner, shoulder rubs, and constant reassurance. He listened to my concerns about this variable or that slope with endless patience and helped me to maintain a balance between work and life.

# Table of Contents

# List of Tables

# List of Figures

# Chapter 1

## *Purpose of the Dissertation Research*

### Introduction

This dissertation used Trends in International Mathematics and Science Study (TIMSS) 2007 data to examine characteristics of schools in countries with high academic achievement in mathematics and science at the eighth grade. More specifically, a school effectiveness analytic approach was used to examine school factors related to student success in science, technology, engineering, and mathematics (STEM). Since schools are the primary institution for learning in almost all countries, investigating the characteristics of schools in high–performing countries can provide valuable information to education policymakers. In particular, analyzing the characteristics of schools whose students have high achievement in STEM subjects will highlight important school policies, practices, and resources that could be used to improve education.

School effectiveness analyses seek to improve educational practice by studying what makes for a successful school beyond having a student body from advantaged socioeconomic backgrounds. That is, what factors under school control can be manipulated to improve achievement. From an analytic perspective, school effectiveness studies make use of multilevel modeling in order to analyze the relationships between school factors and achievement after controlling for student characteristics. This dissertation approached the school effectiveness investigation by capitalizing on the

TIMSS design, which assesses the same students in mathematics and science. This made it possible to create an overall measure of achievement that is a good indicator of science, technology, engineering, and mathematics (STEM). Even though TIMSS was not designed to measure STEM directly, the combined mathematics and science items together provided a good measure of STEM achievement. The research focused on how major school effectiveness factors related to teaching and learning mathematics and science operated in the United States as compared to Chinese Taipei, the Czech Republic, Singapore, and Slovenia.

## Description of the Problem

Results from TIMSS 2007 show a gap in mathematics and science achievement between U.S. students and those in the top–performing countries, particularly at the advanced level. In mathematics, only 6 percent of U.S. eighth grade students scored at or above the advanced international benchmark in mathematics (Mullis, Martin, & Foy, 2008). Seven countries had higher percentages of students performing at the advanced level in mathematics at the eighth grade compared to the United States (Gonzales et al., 2008). Similar to the percentages in mathematics, only 10 percent of U.S. eighth grade students scored at or above the advanced international benchmark in science (Martin, Mullis, & Foy, 2008). Outperforming the United States, six countries at grade 8 had greater percentages of students reaching the advanced level in science (Gonzales et al., 2008).

National assessment results in the United States support the findings of mediocre achievement in STEM subjects. Most recently, results from the National Assessment of

Educational Progress (NAEP) revealed that only 30 percent of eighth grade students were at or above the *proficient* level in science (National Center for Education Statistics, 2011). Results from the NAEP mathematics assessment were similar—34 percent of eighth grade students were at or above the *proficient* level in mathematics (National Center for Education Statistics, 2009).

Particularly relevant in a globalized economy, achievement in science, technology, engineering, and mathematics (STEM) is a strong gauge of a country's future economic productivity (Hanushek, Jamison, Jamison, & Woessmann, 2008). Recent research suggests that human capital has a direct influence on long-term economic development. For example, differences in cognitive reasoning in mathematics, science, and reading explain the majority of differences in economic growth rates across OECD countries (Organisation for Economic Co-operation and Development, 2010a). In the United States, preparing students for work in STEM areas is critical, since the Bureau of Labor Statistics (2007) projects that jobs for STEM occupations will grow 22 percent between 2004 and 2014. It is therefore important that the education system prepare enough students to fulfill the workforce needs in these fields or risk falling behind more scientifically advanced countries in the global marketplace.

Previous TIMSS research indicates that top–achieving countries in mathematics and science have a rigorous curriculum (Mullis & Martin, 2007). In particular, these countries tend to have established standards that clearly define curricular intentions, are specific about the content involved and what is expected of students, and are coherent across grade levels (Schmidt, McKnight, Valverde, Houang, & Wiley, 1997; Schmidt, Raizen, Britton, Bianchi, & Wolfe, 1997). However, adopting a rigorous curriculum is not enough for schools to support high percentages of students reaching advanced levels of STEM achievement. The teacher and school need to implement the curriculum with fidelity such that students learn the material and support learning through a positive school climate. Using a school effectiveness analytic approach and employing multilevel modeling, this dissertation focused on what could be learned about how school factors related to success in STEM achievement operate in the TIMSS 2007 top–performing countries.

## IEA's Trends in International Mathematics and Science Study (TIMSS)

IEA (International Association for the Evaluation of Educational Achievement) pioneered international assessments in the 1950s as a way for countries to learn from each other (Husen, 1973). TIMSS together with PIRLS (Progress in International Reading Literacy Study) comprise IEA's core cycle of international studies, which are directed by the TIMSS & PIRLS International Study Center at Boston College. To measure trends in mathematics and science achievement, TIMSS has been assessing fourth and eighth grade students every four years since 1995. TIMSS 2007 was the fourth cycle in the TIMSS study and had 50 countries and 7 benchmarking participants at the eighth grade. As the most current TIMSS data available for public use, TIMSS 2007 data is particularly appropriate for this dissertation because it assesses nationally representative samples of students, allowing for generalizations at the country level as well as cross-country comparisons. The TIMSS design, which assesses the same eighth grade students in mathematics and science, makes it possible to create an overall STEM achievement measure encompassing both the mathematics and science items. In addition, TIMSS collects and reports extensive information about participating countries' education systems in terms of their organization, curricula, and instructional practices in order to improve education policy-making.

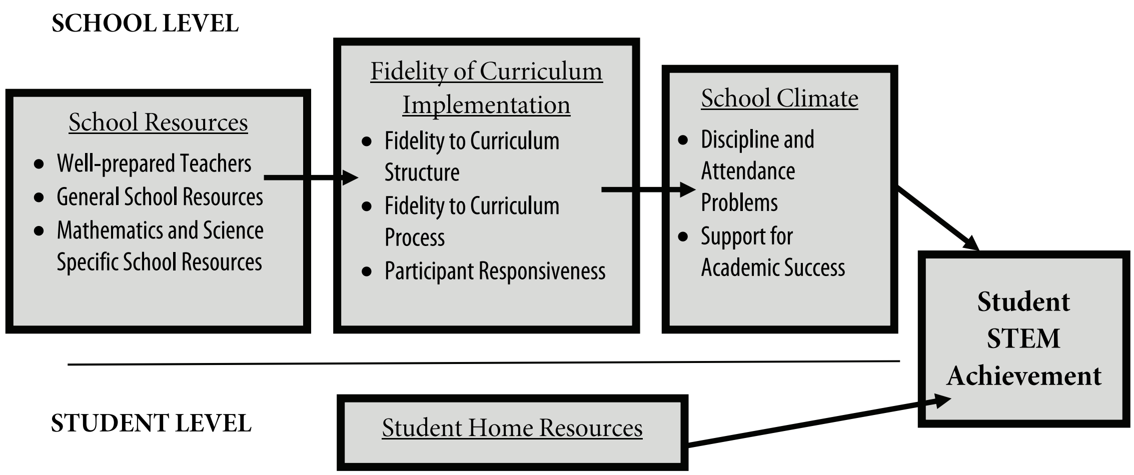
## Purpose of the Study

Differences between schools in the United States and four high–achieving TIMSS 2007 countries were examined to predict the extent to which improving various school characteristics would increase STEM achievement in the United States. That is, relative school effectiveness factors were examined in the United States compared with Chinese Taipei, Singapore, the Czech Republic, and Slovenia using multilevel modeling. Chinese Taipei and Singapore were selected to represent the group of Asian countries that performed exceptionally well in both mathematics and science at the eighth grade, and the Czech Republic and Slovenia were selected to represent the Central European countries that performed especially well in science. The two European countries taught biology, chemistry, physics, and geography as separate subjects at the eighth grade instead of one integrated science course as was done in the other three countries in this study. Comparing countries with structural differences in how instruction is delivered provides information about effective STEM policies in different contexts.

At the eighth grade, Chinese Taipei and Singapore had higher average mathematics and science achievement than all other countries participating in TIMSS 2007 along with Korea. Chinese Taipei's average achievement score was 598 in mathematics and 561 in science and Singapore's average achievement scores were 593 and 567 in mathematics and science, respectively. The Czech Republic and Slovenia also performed admirably well in science. Their achievement scores were 539 and 538, respectively. Essentially, they were only outperformed by Asian countries in eighth grade science. The United States' average achievement score for eighth grade mathematics was 508 and their average achievement score for eighth grade science was 520.

School effectiveness in this dissertation was examined using the conceptual model for a country to attain high STEM achievement shown in Figure 1.1. The model recognizes that student achievement is a multilevel phenomenon of students and schools, meaning the model distinguishes between levels of education. At the student level, students' STEM achievement is influenced by individual students' home resources. At the school level, student STEM achievement is influenced by the characteristics of the school, including school resources, fidelity of curriculum implementation, and school climate. School resources relate to human and physical resources a school has available and in this conceptual model, include well-prepared teachers, general school resources, and mathematics and science specific school resources. Fidelity of curriculum implementation comprises fidelity to curriculum structure (i.e., teaching the topics in the intended curriculum and duration of instruction), fidelity to curriculum process (i.e., effective instructional practices), and participant responsiveness (i.e., student attitudes toward learning), which underlies both of the other curriculum factors. School climate represents both the negative aspects of school climate, such as discipline and attendance problems and the positive aspects, such as support for academic success.

Figure 1.1. Conceptual Model for a Country to Attain High Achievement in Science, Technology, Engineering, and Mathematics (STEM)




While high STEM achievement involves numerous factors, this dissertation focused on those factors that may be directly manipulated by policy changes. When conducting school effectiveness research, it is important to take into account existing contexts, primarily students' home background since school effectiveness is confounded with other factors related to student achievement. By focusing on what can actually be improved, policymakers can identify the essential elements that can be changed at the school level, since the existing student population cannot be directly manipulated. Although students' home environment characteristics can never be controlled for completely, a student–level home resources factor was included in these analyses as a covariate in the multilevel models to help isolate the relationship between important school factors and STEM achievement.

In preceding analyses using TIMSS data (Mullis, Martin, Gonzalez, & Chrostowski, 2004; Mullis, Martin, and Foy, 2008), certain school factors associated with higher mathematics and science achievement emerge. First, effective schools appear to be equipped with well–prepared teachers who feel ready to teach nearly all the TIMSS topics based on high education levels and majors in the field of study being taught. Second, well–resourced schools and classrooms are associated with higher STEM achievement as indicated by factors such as principals reporting that resource shortages do not adversely affect instruction and that schools are equipped with resources specific to teaching STEM topics, such as science laboratories. Third, higher–performing schools tend to actually teach the intended mathematics and science curriculum, use effective instructional strategies to do so, and have students that respond to this instruction through positive attitudes toward learning mathematics and science. Finally, school climates where there are few discipline and attendance problems and that support academic success are associated with higher achievement. This dissertation examined how these important factors operate in the top-performing TIMSS 2007 countries compared to the United States.

This dissertation research had the added purpose of exploring school effectiveness in a subset of top-performing TIMSS countries before school effectiveness analyses are conducted for all participating countries in TIMSS and PIRLS 2011. In 2011, for the first time, TIMSS and PIRLS were administered on the same schedule, allowing countries to assess the same fourth grade students in mathematics, science, and reading. In order to capitalize on this opportunity for examining the relative school effectiveness in reading, mathematics, and science, school effectiveness models for each country will be built. This dissertation study provides a “dress rehearsal” for that large scale effort.

## Research Questions

The overarching research question investigated in this dissertation was as follows: How do school effectiveness factors related to teaching and learning STEM subjects operate together in the United States compared to top-achieving countries in TIMSS 2007?

To answer this question, the following more specific questions were addressed:

1. How do school resources (well-prepared teachers, general school resources, and resources specific to mathematics and science instruction) add to the explanation of STEM school effectiveness in each country?
2. How does fidelity of curriculum implementation (fidelity to curriculum structure, fidelity to curriculum process, and participant responsiveness) add to the explanation of STEM school effectiveness in each country?
3. How does school climate (discipline and attendance problems and support for academic success) add to the explanation of STEM school effectiveness in each country?
4. How does the combination of school resources, fidelity of curriculum implementation, and school climate add to the explanation of STEM school effectiveness in each country?

## Importance of the Study

With the goal of recent education policies being to help set high expectations for student achievement, investigating what makes schools in high–performing TIMSS countries effective could provide valuable information to policymakers about how to improve schools. For example, the $4.35 billion Race to the Top grant program has been directed to ensure a competitive preference to states that commit to improve STEM education, among other goals (U.S. Department of Education, 2009). Also, President Obama has prioritized the improvement of STEM education through the Educate to Innovate campaign, requesting $206 million dollars for increasing effective teaching and learning in STEM subjects (U.S. Department of Education, 2011).

Funding for STEM education is an important step toward improving student achievement in STEM fields in order to meet the market needs of the 21st century. However, further research is needed to target spending in order to most efficiently improve how schools implement such programs. The U.S. Department of Education (2007) notes "there is a general dearth of evidence of effective practices and activities in STEM education" (p. 3). Results from this dissertation study could provide guidance about how curriculum implementation and school climate factors operate in relation to STEM achievement.

Many studies have investigated school effectiveness in general. However, few have specifically examined the school factors relating to success in STEM achievement (Teddlie & Reynolds, 2000). Other school effectiveness studies have been criticized for focusing on only one curricular subject at a time, usually language or mathematics (Creemers & Kyriakides, 2006). Therefore, the outcome measure in this dissertation was a composite STEM achievement measure computed by adding together students' TIMSS mathematics and TIMSS science achievement scores, reflecting student performance across the multiple components of both subjects. Previous research indicates that combining student-level scores from different subject areas provides a more consistent and reliable measure of school effectiveness (Crone, Lang, Teddlie, & Franklin, 1995). The STEM measure is a very robust achievement measure based on 429 items at the eighth grade (224 multiple-choice items and 205 constructed-response items).

The extensive TIMSS background information about teaching and learning in mathematics and science at the eighth grade provided broad information to examine school effectiveness across STEM curriculum areas. TIMSS 2007 contains background data from four types of questionnaires—curriculum, school, teacher, and student. The curriculum questionnaire provides country-level data about the structure and content of the intended curriculum in mathematics and science. The school, teacher, and student questionnaires provide data relating to the mathematics and science content taught in classrooms, the instructional approaches used, school and classroom resources, teacher preparation, and student experiences and attitudes toward mathematics and science (Erberber, Arora, & Preuschoff, 2008). These data constitute a rich resource for examining predictors of school effectiveness across the different aspects of the education system.

Using the rich background information available from TIMSS 2007, this dissertation explored school resources, curriculum implementation, and school climate factors related to effective STEM education. A 2005 survey conducted by the United States Government Accountability Office (2005) noted there were 207 federal education programs designed to increase the number of students studying in STEM fields and to improve the quality of STEM education. However, the programs targeted elementary and middle school students the least frequently, with high school, college, graduate, and postgraduate students being the target groups for the majority of programs. Examining what STEM-related school practices are effective internationally at the eighth grade provides information currently lacking for the middle school level. Investigating the effective school practices for increasing eighth grade STEM education could provide targeted information to policymakers and educators about how best to improve STEM teaching and learning in middle schools.

Finally, while school education research has focused mostly on national contexts, there is much to be learned from a cross–national perspective about how to improve schools (Teddlie & Reynolds, 2000). Hanushek & Woessmann (2010) identify several advantages of cross-country comparative approaches over national approaches to examining educational effectiveness. One advantage is that international studies can examine differences in system-wide processes, such as the use of a national high-school graduation exam, that are hard to examine within one country, since there is no variation. International comparative studies can also provide information about whether any particular relationship found is specific to one country context or may be generalized to a broader context. Related to the contextual issue, results from an international study may be used to determine if a particular relationship differs across contexts because of an interaction variable, not identifiable when looking at only one country's results. Thus, results from an international context focusing on school effectiveness in high–achieving TIMSS countries highlights what school factors might be improved across a variety of contexts to attain the high achievement in STEM areas necessary for healthy economic growth.

Considering what leads to a country attaining high achievement in science, technology, engineering, and mathematics involves a number of complex factors. This dissertation focused on the school resources, fidelity of curriculum implementation, and school climate factors that are most important to consider after the adoption of a rigorous STEM curriculum and controlling for students' home resources.

# Chapter 2

## *Review of the Literature*

### Introduction

This dissertation examined school effectiveness in STEM achievement in the United States compared to four top-achieving countries in TIMSS 2007—Chinese Taipei, Singapore, the Czech Republic, and Slovenia. The first section of the chapter provides an overview of the school effectiveness model, including the use of multilevel modeling and gives a brief history of international assessments. The second section reviews literature about student home resources related to high achievement in STEM subjects, including educational resources in the home and parental education level. The last section reviews the literature related to school resources for learning STEM subjects, fidelity of curriculum implementation, and a school climate that supports academic success.

### School Effectiveness Research

School effectiveness analyses seek to improve educational practice by studying what makes for a successful school. For more than 40 years, researchers have been examining different aspects of school effectiveness in order to improve educational outcomes for students (Sammons, 2007). While definitions of school effectiveness vary, most researchers agree that, when comparing schools with similar student populations, an effective school is one that "adds extra value" to student achievement (Sammons, 2007, p. 13). That is, the characteristics students have when entering school have been shown to have a strong relationship with achievement and should be explicitly controlled in the analysis model in order to better isolate the effects of a school. An effective school has the capacity to improve student achievement despite the characteristics of the entering student body.

School effectiveness research distinguishes itself from other strands of educational effectiveness research, such as economically oriented studies and instructional effectiveness studies by focusing on the importance of school difference (Scheerens & Bosker, 1997). Research investigating school difference has provided important contributions to the field of education by focusing on the aspects of education that have an influence on achievement and using results to suggest improvements and shape reform policies. Teddlie (2010) describes four aspects of the school effectiveness research tradition that have influenced how educational research is conducted today.

- The school effectiveness tradition approaches educational research from the perspective that schools matter in the education of students and that differences in schools lead to differences in students' education.
- Effectiveness research acknowledges that school differences present challenges to educators, but promotes the idea that schools should take responsibility for the education of their students and by using results of effectiveness research, overcome challenges.
- School effectiveness research uses the school as the major unit of change in educational reform.
- School effectiveness research recognizes that schools vary according to context and change and evolve over time.

This dissertation approaches the school effectiveness investigation by considering what school factors or improvements would enable a country to attain high achievement in science, technology, engineering, and mathematics. While STEM achievement involves numerous factors, this dissertation will focus on those factors that may be directly manipulated by policy changes. Luyten, Visscher, and Witziers (2005) define school effectiveness research as investigating performance differences within and between schools and examining the "malleable factors that enhance school performance" (p. 249). If research can help identify how the manipulable characteristics of an effective school operate, underperforming schools might be helped by adopting these best practices as far as is possible.

### *Use of Multilevel Modeling in School Effectiveness Research*

The origins of school effectiveness research can be traced back to the mid-1960s in the United States when most educational research involved investigating the relationship between inputs (human and physical resources) and outputs (student outcomes) (Teddlie & Reynolds, 2000). The major research question during this early stage of school effectiveness research was "Do schools have measurable impacts on student achievement?" (Rumberger & Palardy, 2004). A school was defined by its material resources and differences in student achievement were attributed to unequal opportunities in terms of school environments (Scheerens & Bosker, 1997).

However, according to Teddlie and Reynolds (2000), educational researchers were criticized for not measuring the educational processes within schools. Not accounting for school and classroom processes caused much of the differences researchers found between schools to be attributed to student background characteristics rather than educational practices. In addition, a lack of sophisticated methodology prevented researchers from making fair comparisons between schools to assess the unique contribution of an individual school on student achievement (Creemers, Kyriakides, & Sammons, 2010).

In order to address the conceptual and methodological critiques of school effectiveness research, more contextual factors and sophisticated methodologies, such as the development of psychological scales and more sensitive outcome measures were applied to school effectiveness research in the 1980s (Townsend, 2007). The most notable methodological advance was the development of multilevel statistical modeling to more accurately estimate the effects of factors at different levels in the education system (i.e., the student level, classroom level, and school level) on student outcomes (Rumberger & Palardy, 2004). Advances in technology also made computer programs that could accommodate multilevel modeling more widely available (Teddlie & Reynolds, 2000).

Multilevel statistical models are used for clustered data. In educational research, because students in schools share the context of the classroom or school, they are considered to be clustered (Raudenbush & Bryk, 2002). This makes students in one classroom or school more similar to each other than they are to students in other classrooms or schools. The students' shared context leads to several statistical challenges for using Ordinary Least Squares regression, including incorrect standard error calculations, heterogeneity of regression slopes, and aggregation bias (Bickel, 2007).

That is, standard errors are smaller than they should be, the relationship between the predictor variables and the outcome measure is incorrectly modeled as constant across groups, and individual-level data is aggregated to the group level, causing loss of information. These statistical errors can result in an inflated Type I error rate, indicating a greater likelihood of finding significance than is appropriate.

Using multilevel models to analyze nested data helps correct for these statistical errors. First, multilevel modeling computes a separate regression equation for each level in the data, which results in correct standard errors being estimated (Bickel, 2007). This also allows regression effects to be calculated at more than one level, so researchers may assess whether or not student-level variables vary across schools. Second, multilevel modeling addresses the heterogeneity of regression slopes issue by allowing the relationship between the predictor variables and outcome measure to vary randomly within each level, which provides a more precise estimate of the relationship between factors. Third, multilevel modeling resolves aggregation bias by building a regression equation for each level of the data, which describes how the variables at one level influence the variables at the other levels.

Beyond the statistical aspects, multilevel models can provide substantive help when investigating a research problem with multiple unit levels by producing statistics that can be evaluated at each level. Bickel (2007) notes, "the possibility of individual level-effects *and* contextual effects in the same analysis is one compelling reason why multilevel modeling has become so conspicuous in the study of student achievement" (p. 3). With regard to school effectiveness research, multilevel modeling allows for predicting student achievement based on school-level characteristics, while controlling for the effects of student background characteristics at the student-level. Controlling for student background removes any variance related to student characteristics from the analysis, focusing the relationship between school effectiveness factors and student achievement.

### *International Assessments*

Parallel to the school effectiveness research tradition developing in the United States from the 1960s onward, international assessments were emerging, namely IEA's First International Mathematics Study (FIMS). FIMS was notable in the field of educational research because it provided detailed information about the inputs of education systems, such as the mathematics curriculum, teacher education and preparation, and student attitudes and home background that were related to mathematics achievement. IEA published results from FIMS in 1967, including data on mathematics achievement as well as contexts of student achievement for 12 countries (Husen, 1996). FIMS focused on a country's policies and practices as input variables and student achievement as the output variable.

FIMS was followed by the First International Science Study (FISS) conducted between 1971 and 1972 in 19 countries as part of IEA's Six Subject Survey that also included reading comprehension, civic education, literature, English as a foreign language, and French as a foreign language (Husen, 1996). FISS was conducted during a time when many countries were undergoing curriculum reform in science education with the goal to provide data that would improve science education (Murphy, 1996). It followed the same principles as FIMS, providing information about inputs and outputs of educational systems.

Building on the success of FIMS and FISS, IEA administered the Second International Mathematics Study (SIMS) and the Second International Science Study (SISS) in the early 1980s. Important in the field of educational research, SIMS and SISS reported extensive information about the ongoing processes within schools across countries along with the inputs and outputs to better contextualize mathematics and science achievement results (Husen, 1996; Murphy, 1996). SIMS and SISS were the first IEA assessments to use a three-level curriculum model—the *intended curriculum* describes the national goals a country has planned for its students, the *implemented curriculum* represents what is actually taught to students, and the *attained curriculum* represented by the assessment results, refers to what the student has learned or experienced as a result of being in school (Murphy, 1996).

IEA's Third International Mathematics and Science Study (TIMSS) conducted in 1995 was the start of a new era in international assessments. Bringing together mathematics and science in a single study, TIMSS tested students at three levels of the education system (the end of primary schooling, middle or lower secondary schooling, and final grade) and gathered extensive data on countries' mathematics and science curricula as well as home, school, and classroom contexts for learning (Mullis & Martin, 2007). TIMSS contributed methodologically to the international research tradition by designing assessments to represent a broader range of content and using complex psychometrics to analyze results.

TIMSS 1999 or TIMSS-Repeat (as it was known at the time), administered in 35 countries (Mullis & Martin, 2007), was designed as a replication of TIMSS 1995 at the eighth grade only, allowing those countries that participated at the eighth grade in 1995 a measure of trend in student performance from 1995 to 1999. For countries that participated at the fourth grade in 1995, TIMSS 1999 provided the opportunity to reevaluate the population of students originally assessed as fourth graders but now in the eighth grade four years later.

Based on the success of TIMSS 1995 and TIMSS 1999, IEA committed to implementing TIMSS as a trend study of student achievement in mathematics and science at the fourth and eighth grades (Mullis & Martin, 2007). Renaming it the Trends in International Mathematics and Science Study, IEA has administered TIMSS every four years, most recently in 2011. The results of the TIMSS studies continue to provide important information about school effectiveness factors over time in education systems worldwide.

The Progress in International Reading Literacy Study (PIRLS) was developed by IEA and the International Study Center at Boston College to be the corresponding reading assessment to TIMSS (Mullis, Martin, Gonzalez, & Kennedy, 2003). PIRLS assesses students internationally at the fourth grade on a range of reading comprehension strategies relating to literary and informational purposes. PIRLS is conducted every five years, including assessments in 2001, 2006, and soon 2011.

Kellaghan (1996) notes four ways IEA studies improved educational research and the ways policymakers could use the results of such research. First, IEA studies provided data on the inputs, processes, and outputs of educational systems so that policymakers could have a broader view of the context of achievement. Second, the studies provided comparative information at an international level about student achievement, and also about the different factors involved in achievement, such as the organization of the education system or curriculum content. Third, IEA capitalized on the variability of education systems worldwide, broadening the effectiveness factors that could be studied to cross-country conditions instead of national conditions. Fourth, the quantitative methods employed by IEA increased the quality of comparative education studies over other similar studies being conducted at the time.

Another international assessment that emerged in the late 1990s was the Program for International Student Assessment (PISA). PISA is an international assessment of 15-year old students' knowledge and skills in mathematics, science, and reading. It provides data on student performance in reading, mathematics, and science literacy as well as contextual information regarding students' home backgrounds by asking students and national contexts (Organisation of Economic Co-operation and Development, 2011). PISA has been conducted four times in 2000, 2003, 2006, and 2009.

Current school effectiveness research has capitalized on the data provided by international assessments to investigate the various factors that are working in education across a variety of international contexts. For example, Martin, Mullis, Gregory, Hoyle, and Shen (2000) conducted a school effectiveness analysis using TIMSS 1995 data. The authors identified characteristics that distinguished low- from high-achieving schools and then examined those factors using hierarchical linear modeling to control for student home background. They conducted analyses in two sets of countries to examine factors related to mathematics achievement (18 countries) and science achievement (14 countries). Although the factors investigated did not have a consistent relationship to mathematics and science achievement across all countries, the following factors were found to relate to school achievement: school size and location, school climate, student attitudes toward mathematics and science, instructional activities, doing homework, teacher characteristics, and student aspirations for further education.

Other school effectiveness research focuses on the top–performing school systems identified by TIMSS, PIRLS, and PISA to examine what factors make these education systems so successful. McKinsey & Co. examined approximately 20 of the world's best performing school systems and identified three major factors contributing to student success across the systems under investigation: 1) getting high quality people to become teachers, 2) developing those high quality people into effective teachers, and 3) ensuring the education system delivers the best instruction for each student (Barber & Mourshed, 2007). More recently, McKinsey & Co.  examined continually improving school systems around the world to determine which system–wide interventions were associated with increased student performance. After linking achievement scales across a range of assessments, including TIMSS, PIRLS, PISA, and NAEP, results indicated successful interventions are systematic, contextualized, and sustained through collaborative practices (Mourshed, Chijioke, & Barber, 2010). Further, research conducted by the Organisation for Economic Co–operation and Development (2010c) has focused on the "strong performers and successful reformers" in education, mostly in terms of system–

wide factors that contribute to high educational performance, such as examination and instructional systems, teacher quality, and school finance.

## Research on Factors Associated with STEM Achievement

School effectiveness research has identified many different factors that account for differences in school achievement, ranging from family background to community characteristics to educational experiences (Sammons, 2007). The school effectiveness research in this dissertation focused on school resources, fidelity of curriculum implementation, and school climate factors that differentiate between more and less effective schools. Hanushek & Woessmann (2010) define school resources as teacher and material resources, such as expenditure per student and availability of instructional materials. Teddlie (2010) identifies overarching effective school practices related to curriculum implementation and school climate that have emerged from the school effectiveness literature over the years. These factors include: effective instruction, focus on learning (i.e., a viable curriculum and opportunity to learn), positive school culture, high expectations for students and staff, parental involvement, and emphasizing student responsibilities.

Within the conceptual framework of examining school resources, fidelity of curriculum implementation, and school climate, the school effectiveness factors reviewed in this section concentrate on those that are included in the TIMSS 2007 database. That is, the literature reviewed describes factors related to student and school effectiveness good for studying school differences in STEM achievement and that also are available in the TIMSS 2007 database.

Because controlling for student background is central to school effectiveness research, this section first presents research on student home resources, including educational resources in the home and parental education level. Next, there is a discussion of school resource factors, including teacher preparation, general school resources, and mathematics and science specific school resources for learning STEM subjects. Following that is a description of research relating to fidelity of curriculum implementation factors, and finally, school climate.

### *Student Home Resources*

In school effectiveness studies, it is important to measure students' family background since school effectiveness results are confounded with other factors related to student achievement if these factors are not accounted for. Buchmann (2002) discusses several reasons why controlling for family influences is important, especially in international comparative studies of education. First, he contends that controlling for family background allows researchers to examine the role of the school in improving student learning and achieving equality of educational opportunities for students coming from diverse backgrounds. Second, he promotes researching family background since it leads to increased knowledge about how a student's home context affects his or her ability to learn and academic achievement in school. This information can help policymakers address inequities in background by providing extra services and resources for students who need extra help. Third, particularly important for comparative studies, he purports that investigating family background allows researchers to measure the distribution of achievement across countries. That is, if the student population across countries is very different, this could lead to differences in academic achievement that are not solely related to the quality of education in each country. This information could provide more context to policymakers who could adjust goals for their education systems.

## Home Educational Resources

To provide a measure of a student's home background, educational research makes use of a range of indicators. Since it is often difficult to measure directly, much of the current research uses a composite of variables, including resources in the home to provide context for academic achievement (Sirin, 2005). A recent meta-analysis of educational research examining the relationship between socioeconomic status and academic achievement revealed that many educational research studies include books in the home, computers, and study aids or a study room as indicators of the social and economic status of a household (Sirin, 2005).

For example, Marks, Cresswell, and Ainley (2006) investigated the relationship between home factors and reading, mathematics, and science achievement, by regressing PISA 2000 achievement data on socioeconomic background with or without controlling for material and cultural resources. Results indicated that material resources in the home (income measures, such as a dishwasher, internet connection, car, etc.) and cultural resources in the home (books in the home and cultural possessions, such as classic literature) accounted for between one fifth and one half of the influence of (variance related to) socioeconomic background on mathematics and science achievement across countries.

Similarly, Shelley and Su (2011) used hierarchical linear modeling to investigate the effects of student- and school-level variables on the mathematics achievement of 15-year old students in the United States participating in PISA 2003. Results indicated that a composite variable of parents' highest occupational level and PISA's estimate of household possessions had a positive relationship with students' mathematics achievement (Shelley & Su, 2011). PISA's estimate of household possessions included 13 different household items from educational software and a dictionary to a dishwasher (Organisation for Economic Co-operation and Development, 2005).

International comparative research conducted by Martins and Veiga (2010) supports the research previously described, suggesting an association between home educational resources and academic achievement. The researchers decomposed socioeconomic-related inequalities in the students' mathematics achievement in 15 European countries using hierarchical linear modeling of PISA 2003 data. They found that the most consistent socioeconomic variable associated with inequalities across countries was the cultural environment of the home as measured by PISA's estimate of household possessions and books in the home. Having fewer home possessions was associated with lower mathematics performance in several countries. Similarly, the average mathematics achievement of students possessing more books was higher than students with fewer books in every country.

The research previously described indicates there is a relationship between home educational resources and mathematics and science education on average internationally. Examining the countries under study in this dissertation—Chinese Taipei, Singapore, Slovenia, the Czech Republic, and the United States, however, indicates the context of this relationship differs somewhat between countries. For example, differences in home educational background between Asian countries and the U.S. are examined by Wang (2004). Using TIMSS 1995 data, he examined if family background factors were equally related to middle school mathematics achievement in both Hong Kong (using the Chinese sample of students) and the United States. Results from a regression analysis indicated that the presence of study aids and more books at home were related to mathematics achievement in both countries. However, while the presence of study aids was an equally strong predictor of mathematics achievement in both countries, the number of books in the home was a stronger predictor of mathematics achievement in the U.S. compared to Hong Kong where typically there is little space in homes for books.

Tsui (2005) examined the relationship between family income, home environment, and mathematics achievement of eighth grade students in China and the United States. The mathematics scores of the Chinese sample (N=1,021) were from a second-semester final examination of Chinese eighth grade students from three districts. The mathematics scores of the American sample were from the eighth grade mathematics portion of the National Education Longitudinal Study (and the sample was restricted to students who were only children or had only one sibling for comparison purposes with the Chinese cultural context, yielding approximately 8,747 students). The home environment variable included owning educational materials, such as a daily newspaper, a regularly received magazine, an encyclopedia, a dictionary, a computer, a calculator, and at least 50 books. The researcher found that within each country, both Chinese and

American families with relatively higher family income had more educational materials in the home than did families with lower income. However, low-income Chinese families had more learning materials in the home than did low-income American families. Survey results indicated that education spending in China was the second largest expense after food and 55 percent of Chinese families surveyed hired tutors for their children, providing an indication of the emphasis Chinese parents place on education.

Researchers also have examined the relationship between student achievement and family background in Central European countries, such as the Czech Republic. Strakova (2007) used data from three international datasets (PIRLS, TIMSS, and PISA) to examine how the relationship between student achievement and family background changed over time from fourth to eighth to tenth grade in the Czech Republic compared to Canada, Finland, and Sweden. These four countries have similar student achievement and similar variation in socio-economic background but differ in terms of the differentiation of students in the school system (i.e., tracking). Specifically, the Czech Republic employs student tracking into different schools starting at age eight and continuing at ages eleven and thirteen. In contrast, Finland, Sweden, and Canada do not employ tracking in their education systems. As a measure of the background characteristics of the students, the researcher used a composite of students' highest parental education, the main job of the mother, and the main job of the father in PIRLS (fourth grade); students' highest parental education, number of books at home, and possession of a computer and a dictionary in TIMSS (eighth grade) and an international index of economic, social, and cultural status in PISA (tenth grade). Based on a hierarchical linear modeling analysis, results indicated that at the primary school level the relationship between a student's achievement and his or her socio-economic status was comparable in all four countries. However, in higher grades (once students had been tracked into different school types) the difference in achievement between students with higher socioeconomic status compared to lower socioeconomic status in the Czech Republic was greater than the difference in Finland, Sweden, or Canada. Thus while home educational resources are important factors related to mathematics achievement in many countries, the relationship may differ in each of the country contexts in this dissertation study.

For international educational studies, the number of books in the home sometimes is used by itself as a reliable factor representing the students' family background since it is readily comparable across countries (Hanushek & Woessmann, 2010). Schuetz, Ursprung, and Woessmann (2008) contend that books at home is straightforward for students to report since it involves simply counting (or estimating) the physical number of literary resources in the home. They also provide empirical support for using the number of books at home as a student background variable. Using the PIRLS 2001 data for six European countries, the researchers created interaction variables for each country, such as educational expenditure per student by number of books at home or country mean test score by number of books at home. The researchers then regressed household income on books at home and the country interaction variables. Results indicated that the association between household income and books at home does not vary significantly between countries, making it a stable cross-country predictor variable.

Research conducted by Woessmann (2004) used books in the home as a measure of home resources. Researchers conducted a regression analysis using TIMSS 1995 data to examine the effects of family background factors on student performance in the United States and Europe. Results indicated that books in the home were related to achievement across countries, even after controlling for parental education level.

TIMSS 2007 results on home educational resources show there is a relationship between eighth grade mathematics and science achievement and home educational resources internationally. At the eighth grade, TIMSS 2007 data indicate that students with a computer at home had higher average mathematics and science achievement than students without a computer (Mullis, Martin, & Foy, 2008; Martin, Mullis, & Foy, 2008). In particular, students with an Internet-connected computer had higher achievement than student that did not.

TIMSS 2007 asked about the number of books in the home and similar to other research cited, the data suggest that students from homes with abundant literacy resources have higher mathematics and science achievement than students from homes with fewer resources. Specifically, eighth grade students from homes with more than 100 books had higher average mathematics achievement than those from homes with fewer books (Mullis, Martin, & Foy, 2008). The same pattern was found for science, with students coming from homes with more than 100 books having higher than average science achievement than those coming from homes with fewer literary resources (Martin, Mullis, & Foy, 2008).

## Parental Education Level

Considerable research shows parental education has an important relationship with mathematics and science achievement. In a meta-analysis of educational research, Sirin (2005) identified parental education as the most commonly used factor in research examining family background and student achievement. He comments that parental education is often used in research because it is generally a stable variable, established when a student is young and tends to remain the same over time. Furthermore, the relationship between parental education and student achievement is moderately strong. In his meta-analysis, there was a moderate effect size detected (Cohen's $d$=0.30) between achievement and parental education across approximately 30 different studies.

Byrnes and Miller (2007) found parental education to be related to success in mathematics and science. The researchers used National Education Longitudinal Study data to investigate the interplay between different factors related to middle school mathematics and science achievement, including socio-economic family background, advanced coursework in mathematics and science, motivation to learn mathematics, middle school grade point average (GPA), race, and gender. The socio-economic family background variable comprised parental education, occupation, and income. Using hierarchical linear modeling and structural equation modeling, the researchers examined the interrelationship between the predictors, explaining the maximum amount of variance in student mathematics or science achievement. Results indicated that socio-economic family background composite variable and GPA were the strongest predictors of both mathematics and science achievement.

Because financial and familial effects are often confounded in educational research, Chevalier and Lanot (2002) conducted a study to examine the effect of changes in household income on whether or not students remain in school through compulsory education (age 16 in the education system of Great Britain) while holding constant family characteristics. Using longitudinal data from the National Child Development Study and the British Cohort Study, the authors examined the benefit of financial incentives to keep students enrolled in school through compulsory education after accounting for family characteristics, including parental education, father's socio-economic status, number of siblings, having natural parents or not, and race. Results suggested the effect of family income on the age students leave school is still a significant determinant, but that the effects are moderated by parental education. That is, despite financial incentives, parental education was an important determinant of how long a student remained in post compulsory education, indicating it has a powerful effect on students' school success.

From a comparative international perspective, research investigating parental education and achievement between countries compared to within countries reveals more details about the complex relationship. Martins and Veiga (2010) used hierarchical linear modeling to examine how inequalities in parents' education were related to students' PISA 2003 mathematics achievement in 15 European countries. Results suggest there is a positive relationship between average parental education and average mathematics achievement in each country. On average, differences in parents' education levels accounted for 25 percent of a country's total variance in educational inequality. However, results also suggested that parental education was more important within a country compared to between countries in terms of explaining socioeconomic-related variance in achievement scores. That is, the relationship between parental education and student achievement is moderated by a country's level of socio-economic inequality—the higher the inequality, the greater the relationship parental education level has with mathematics achievement.

The *TIMSS 2007 Mathematics Report* provides data relating to parental education. The data indicate that parental education levels vary widely between different countries. For example, the United States data indicated 44 percent of eighth grade students had at least one parent with a university degree, whereas only 20 percent of students in high-performing Chinese Taipei and Singapore had a parent with a university degree (Mullis, Martin, & Foy, 2008). Despite the variance across countries, higher levels of parental education were still associated with higher average mathematics and science achievement in almost all countries. On average across countries, the mathematics achievement of eighth grade students with university-educated parents was approximately 41 points greater than that of students whose parents completed upper-secondary school (e.g., high school in the United States).

### *School Resources*

The literature reviewed in this section provides information about the school resources factors specifically relevant to the conceptual model used in this dissertation and that are available in the TIMSS 2007 database. These include aspects of teacher preparation, such as degree and major thought to be related to content knowledge, as well as general and mathematics and science specific school resources related to learning STEM subjects, including computers and science laboratories.

## Teacher Preparation

Useful for school effectiveness research, differences in teacher quality could help explain differences in student achievement at the classroom or school level. The National Research Council's Committee on Indicators of Precollege Science and Mathematics Education identified various indicators that should be used to gather evidence about the condition of science and mathematics education (Murnane & Raizen, 1988). Teacher preparation, including level of education and major was an important committee recommendation to indicate the professional quality of mathematics and science teachers.

In particular for STEM education, teachers must have the capacity to teach in the mathematics and science disciplines since improving teaching quality is one of the most important ways to improve students' STEM learning outcomes (Fulton & Britton, 2011). Indeed, research has shown teachers' degree of success in delivering the curriculum, as measured by student performance, is related to their preparedness to teach, including their level of formal education completed and whether or not they hold a subject–specific academic degree (Darling-Hammond, 2000). A meta-analysis of studies examining teacher characteristics and student achievement indicates that teachers' level of formal education in conjunction with the subject area of teachers' major is related to student achievement (Wayne & Youngs, 2003).

Given the research indicating the importance of teachers' content knowledge as well as pedagogical skills, researchers suggest that high quality teacher preparation programs include learning in a particular content area related to the curriculum the teacher is expected to teach as well as learning pedagogical techniques, such as how to encourage students to think in different ways (Grossman, Schoenfeld, & Lee, 2005). However, many teachers in the U.S. are not adequately prepared to teach STEM subjects. A recent report examining successful schools for STEM education comments that many U.S. teachers are currently underprepared in terms of pedagogical preparation and content knowledge to teach STEM-related courses (Committee on Highly Successful Schools or Programs in K-12 STEM Education & National Research Council, 2011).

Especially for teaching at the secondary level, teachers' content coursework has been found to have a positive association with student achievement (Rice, 2003). Particularly in the mathematics and science content areas, high quality teacher preparation is essential for promoting greater student academic achievement. In a review of teacher quality research, Rice (2003) notes that when studies examining the relationship between teachers with advanced degrees and student achievement take into account the subject area of the advanced degree, there is generally a positive effect of subject-specific advanced degrees on student mathematics achievement, and to a lesser extent, science achievement.

Research conducted by Goldhaber and Brewer (2000) confirms that teachers' level of formal education together with their content focus is an important factor related to student STEM achievement. The researchers examined how teachers with higher level degrees in the subject matter they taught influenced their students' achievement. Using data from the National Education Longitudinal Study, the researchers investigated the test scores of students of teachers with different levels of formal education. After controlling for differences in student background, results indicated that twelfth grade students who had teachers with bachelor's or master's degrees in mathematics had higher mathematics test scores relative to those whose teachers with out-of-subject degrees. Results for science were not statistically significant.

Research conducted by Hill, Rowan, and Ball (2005) provides evidence that teachers' mathematics content knowledge is related to gains in students' mathematics achievement through instructional practices. Using a measure of mathematical knowledge for teaching, the authors gave teachers multiple-choice items representing mathematics skills used by teachers during instruction, such as "explaining terms and concepts to students", "using representations accurately in the classroom", and "providing students with examples of mathematical concepts, algorithms, or proofs" (Hill, Rowan, & Ball, 2005, p. 373). The researchers used linear mixed models to measure the association between teachers' mathematics knowledge and student achievement gains on the McGraw-Hill's Terra Nova Complete Battery. Two cohorts of student data were used—one cohort of students who entered the study in kindergarten and was followed through second grade and another that entered the study in third grade and was followed through fifth grade. Results indicated that teachers' knowledge for teaching mathematics was a significant predictor of student gains for both grade level cohorts.

Researchers also have investigated the relationship between state-wide policies relating to teacher education and student performance. For example, Darling-Hammond (2000) examined data from the Schools and Staffing Surveys as well as the 1990, 1992, 1994, and 1996 National Assessment of Educational Progress (NAEP) assessments in reading and mathematics. Results from regression analyses indicated the most consistent highly significant predictor at the state level of student achievement in reading and mathematics was the proportion of well-qualified teachers in a state. Well-qualified teachers were defined as those with full certification and a degree in the field to be taught. Results also indicated teachers' education levels (e.g., teachers with master's degrees) had positive relationships with educational outcomes, although were less powerful predictors than certification and major.

At the country-level, differences in teacher qualifications appear to be related to differences in a country's performance on international assessments, such as TIMSS. Wang, Coleman, Coley, and Phelps (2003) examined differences in teacher education and preparation in the United States compared to eight countries that outperformed the U.S. on TIMSS 1999, including Hong Kong and Singapore, which are particularly relevant to this dissertation study. The researchers developed a questionnaire to elicit information about each country's teacher education preparation process, including standards for entrance and exit into a teacher education program, characteristics of the education program for eighth grade mathematics and science teachers, and professional development requirements. Among other differences, researchers found that eighth grade students in the United States and Hong Kong were less likely than students in the other countries to have teachers with a mathematics or mathematics-education major or science or science-education major.

Teacher quality research focusing on the school level indicates there is an interaction between school-wide teacher quality and student demographics. Hill and Lubienksi (2007) conducted a study to determine if the teacher quality level was lower in schools with greater numbers of disadvantaged students. Using hierarchical linear modeling, they examined the mathematical knowledge of California teachers teaching Kindergarten through eighth grade using a series of multiple-choice items representing content and applied content knowledge of mathematics designed by the Study of Instructional Improvement and Learning Mathematics for Teaching project. Results revealed that there was a relationship between the population of students in a school and the level of teachers' mathematical knowledge within the school. That is, schools with larger proportions of low-income and minority students had teachers who had less mathematical knowledge for teaching, on average, than their counterparts at schools with lower proportions of low-income and minority students (Hill & Lubienski, 2007). Given that research previously cited suggests more mathematically knowledgeable teachers produce higher-achieving students in mathematics, these results suggest having lower school-wide teacher quality could have serious implications for the effectiveness of a school with an already disadvantaged student population.

Akiba, LeTendre, and Scribner (2007) investigated the interaction between teacher quality and student socioeconomic status from an international perspective. The researchers examined how the percentage of eighth grade students taught by high quality mathematics teachers differed across countries participating in TIMSS 2003 and what the opportunity gaps were between students from high and low socioeconomic backgrounds in terms of access to qualified teachers. High quality teachers were defined as those with full certification, majoring in mathematics or mathematics education, and with three or more years of teaching experience.  Results indicated that in countries where higher percentages of eighth grade students were taught by fully certified teachers, who majored in mathematics, with at least 3 years of teaching experience, the national average mathematics scores were significantly higher than in the other participating countries. However, researchers also identified an opportunity gap in many countries where students from lower socioeconomic backgrounds were being taught by less qualified teachers than students from comparatively higher socioeconomic backgrounds. All of the countries under study in this dissertation ranked high on this measure in terms of many students having unequal access to qualified teachers.

TIMSS 2007 found that teacher preparation varies across the countries under study in this dissertation. Across Chinese Taipei, Singapore, the Czech Republic, and the United States, the majority of eighth grade students had mathematics and science teachers with a university degree. In Slovenia though, the percentage of students who had teachers who had completed post-secondary education but not university (50% in mathematics and 52% in science) was higher than the percentage of students who had teachers who completed university but not a postgraduate degree (45% in mathematics and 44% in science) (Mullis, Martin, & Foy, 2008; Martin, Mullis, & Foy, 2008). The difference in percentages of teachers with and without university degrees is likely the result of teachers trained before and after the Declaration of Bologna enacted in 2006, requiring teachers to complete a university education program (Republic of Slovenia, Ministry of Education and Sport, 2011). Prior to the 2006 declaration, Slovenian teachers were trained through 4-year study programs in education.

Regarding teachers' educational emphasis in mathematics, TIMSS 2007 results indicate teachers' major area of study in their post-secondary education also varies across countries. In mathematics, 81 percent and 69 percent of students in Chinese Taipei and Singapore, respectively had mathematics teachers who had studied mathematics as their major area of study compared to 62 percent in the Czech Republic, 9 percent in Slovenia, and 42 percent in the United States. The percentages were similar for science, where 96 percent and 94 percent of students in Chinese Taipei and Singapore, respectively had science teachers who had studied science as their major area of study compared to 84 percent in the Czech Republic, 8 percent in Slovenia, and 57 percent in the United States.

## General and Mathematics and Science Specific School Resources

General as well as mathematics and science specific school resources are required to implement a rigorous curriculum so that students can attain high STEM achievement (Lee & Barro, 2001). General resources refer to school buildings and budget for general supplies as well as teachers and support staff. Mathematics and science specific school resources refer to computers and specialized instructional materials as well as laboratory equipment necessary for teaching mathematics and science. With current advances in technology, STEM curricula are beginning to proscribe more complex materials, especially in mathematics and science. For example, the 2014 NAEP Technological Literacy Assessment Framework calls for the assessment to be totally computer–based and includes items modeled after computer games, requiring students to manipulate various components presented to them (WestEd, 2011).Therefore, in order to successfully implement a STEM curriculum, teachers must have access to the technological and other resource components required to deliver such a curriculum (e.g., computer hardware, networking capabilities, etc.).

Research shows the overall availability of school resources, including general resources and mathematics and science specific resources has an influence on student achievement. For example, Schreiber (2002) examined the relationship between school resources and mean school advanced mathematics achievement using data from TIMSS 1995. In order to analyze this relationship, the researcher created a composite variable for school resources comprising 13 variables relating to principals' reports of the inadequacy of various resources. These resources ranged from general resources, including instructional materials, budget for supplies, school buildings, heating and lighting systems, and instructional space, to mathematics and science specific resources, such as computers, computer software, calculators, and audio-visual resources. Results from a hierarchical linear modeling analysis indicated there was a significant association between school resources and mean advanced mathematics achievement. Based on his findings, Schreiber (2002) concluded that schools that have fewer resources provide less student access to mathematics knowledge.

Research conducted by Archibald (2006) examines general and mathematics and science specific school resources in terms of schools' per pupil spending. The researcher used fiscal data from one U.S. state's school expenditures to examine four categories of school resources—instruction, instructional support, leadership, and operations.

Resources for instruction included funding for teachers, support staff, technology and software, trips, and instructional materials and supplies. Resources for instructional support included funding for libraries, extracurricular activities, student health services, curriculum development, staff development, and a variety of social work staff (e.g., therapists, psychologists, evaluators, etc.). Operations resources encompassed funding for transportation, food service, safety, utilities, and building maintenance. Finally, leadership resources were devoted to funding for principals and assistant principals in addition to other administrative and evaluation staff. Using a 3-level hierarchical linear model, the researcher controlled for student- and teacher-level factors at the first two levels in order to isolate school-level effects at the third level. Results indicated per-pupil spending in all categories combined was positively related to achievement in mathematics and reading.

Research sponsored by the UNESCO Institute for Statistics provides a detailed look at how resources for instructional staff relate to student achievement (Willms, 2006). Using hierarchical linear modeling, researchers examined the relationship between the student-to-staff teaching ratio and reading performance on PISA 2000. The student-to-staff teaching ratio represents the total number of students in the school over the number of full-time equivalent teaching staff in that school. Results indicated that performance declined when the student-to-staff teaching ratio was greater than 25. Since in many schools full time teaching staff included school administrators, librarians, and special education teachers, Willms (2006) notes the student-to-staff teaching ratio is "closely related to the costs associated with educating each student, as staff wages are by far the largest component of the overall costs of education" (p. 58).

Focusing on operational resources, Uline and Tschannen-Moran (2008) provide details about the relationship between general resources and mathematics and reading achievement, indicating that school resources may affect student achievement in both direct and indirect ways. The researchers explored student achievement in relation to quality of facilities, resource support, and school climate using bivariate correlational analysis. Results showed a significant relationship between the quality of school facilities and student achievement in mathematics and English. In addition, quality facilities were significantly positively related to school climate, indicating school climate plays a mediating role in the relationship between quality of school facilities and student achievement. That is, the researchers found school climate is less likely to be perceived as orderly and teachers are less likely to show enthusiasm for their jobs when a school has inadequate facilities.

STEM achievement also has been associated with students' access to resources particular to mathematics and science learning, such as specialized instructional materials as well as laboratory equipment. The National Research Council (2006) examined the current state of school science laboratories in the United States and what they contributed to science education. Examining effectiveness research, they concluded that while laboratories did not seem to be more or less effective than other forms of science instruction for general subject matter mastery, they did appear to be associated with improvements in students' scientific reasoning. As would be anticipated, they found an interaction between schools that have access to laboratory resources and the socioeconomic background of the student population. Schools with higher concentrations of poor students and minorities were less likely to have adequate laboratory facilities and often had lower budgets for laboratory equipment and supplies than other schools.

Aladejana & Aderibigbe (2007) used Nigerian data from the Science Laboratory Environment instrument to examine how five components of the science laboratory environment—student cohesiveness, open-endedness, integration, rule clarity, material environment, and total environment—were correlated with students' academic performance. Results indicated that the material environment (i.e., the degree to which laboratory equipment and materials were adequate) was strongly positively associated with students achievement in high school chemistry.

TIMSS 2007 asked principals about the degree to which shortages of the following general resources affected their school's capacity to provide instruction: instructional materials; budget for supplies; school buildings and grounds; heating/cooling and lighting systems; and instructional space (Mullis, Martin, & Foy, 2008; Martin, Mullis, & Foy, 2008). Principals also responded to questions about shortages affecting mathematics or science instruction in particular, including shortages related to science laboratory equipment and supplies (science only), computers for mathematics and science instruction, computer software for mathematics and science instruction, calculators, library materials relevant to mathematics and science instruction, and audio-visual resources. Mathematics and science achievement on average across countries at the eighth grade was positively associated with students in schools where principals reported little instructional impact due to resource shortages.

### *Fidelity of Curriculum Implementation*

The literature reviewed in this section provides information about fidelity of curriculum implementation. Originating from the program evaluation field, particularly health efficacy studies, “fidelity” refers to the degree to which components of a program are implemented as intended (O’Donell, 2008). Fidelity of implementation is beginning to be applied to educational research to evaluate curricular programs (e.g., McNeill, Pimentel, & Strauss, 2011; Merchlinsky & Hansen Grafton, 2007; Ysseldyke, 2003) and as such, has been identified by researchers as an important component of educational research, because it can provide information about processes happening inside the black box of the classroom (O'Donnell, 2008). Mowbray, Holter, Teague, and Bybee (2003) identify two aspects of fidelity that should be measured—structure and process. Fidelity to structure involves the framework for delivery, which includes how well the program adheres to the intended design and the duration of the delivery (i.e., length of time). Fidelity to process measures the quality of delivery. A third component, participant responsiveness, underlies both aspects of fidelity and reflects the extent to which participants are engaged in the content (O'Donnell, 2008).

More specifically, this dissertation examined fidelity in the context of curriculum implementation, beginning with the extent of curriculum coverage and instructional time spent on mathematics and science. The delivery process is considered to encompass the various instructional strategies and activities used in classrooms. Participant responsiveness was measured by student attitudes toward learning mathematics and science.

## Fidelity to Curriculum Structure

The National Research Council's Committee on Indicators of Precollege Science and Mathematics Education identified the degree of content coverage of the implemented curriculum as an important indicator that should be used to gather evidence about the condition of science and mathematics education (Murnane & Raizen, 1988). Specifically, they recommended data be gathered from classroom teachers about the content presented to the student. A more recent report from the National Research Council examined evaluations of the effectiveness of mathematics curricula and made suggestions for improving methods used to evaluate curricula (National Research Council, 2004). Based on their evaluation, they recommended that degree of implementation be included as an important predictor of curriculum effectiveness and be evaluated in terms of how well practitioners make appropriate use of materials and exercise judgment in their use.

While adherence to the curriculum is considered important for assessing fidelity, defining adherence is not always a straightforward process. Remillard (Summer 2005) cautions that the relationship between the teacher and curriculum is multifaceted, with the teacher adapting the curriculum as needed for instruction, so determining the degree of fidelity for evaluative purposes should be carefully considered. Merchlinsky and Hansen-Grafton (2007) provide one example of how researchers have begun to evaluate fidelity to curriculum structure involving both adherence to the intended design and the duration of delivery. In evaluating a six-week physical science curriculum unit designed for students in grades 5-8, the researchers conducted a special study to define fidelity. The final definition of fidelity of implementation developed was, “the extent to which a unit is put into practice as designed (that is, as it was intended by the unit developer)” (Merchlinsky & Hansen-Grafton, 2007, p. 7). The researchers operationalized adherence to structure as the number of topics teachers taught each day and the amount of time they spent teaching each unit.

Fidelity to curriculum structure also has been found to be associated with increased STEM achievement. A study conducted by Taylor, van Scotter, and Coulson (2007) observed the learning gains of students whose teachers implemented a biology science curriculum program with fidelity to structure (i.e., adherence to program design) compared with the learning gains of students whose teachers implemented the program with less fidelity. Analyzing data from ninth grade students’ biology pre- and posttests along with classroom observations to determine how the curriculum was implemented, the researchers found a strong relationship between fidelity of curriculum implementation and student learning gains. Teachers using instructional materials with medium or high levels of fidelity to the program showed the strongest association with student gains.

Researchers also have examined fidelity to curriculum structure in the context of mathematics instruction. A study conducted by Ysseldyke et al. (2003) examined the impact of adding a computerized curriculum-based mathematics instruction system to regular mathematics instruction on mathematics achievement of students in grades 3–5. Using an experimental design, student achievement on a standard mathematics test was measured for students in classrooms where the computerized system was implemented (treatment group) compared to students in classrooms where only the traditional mathematics curriculum was implemented (control group). The degree to which teachers in the treatment group implemented the program according to the program standards also was measured by examining the mean number of objectives mastered and mean number of practice problems attempted by the class.  Results indicated that students in the treatment group had higher achievement than students in the control group. Moreover, students in classrooms where teachers implemented the system with high fidelity had higher gains in mathematics achievement than students in classrooms where teachers implemented the system with lower (partial) fidelity.

Fidelity to curriculum structure in terms of duration, or instructional time spent on STEM subjects also has been found to have a positive relationship with achievement. For example, Lavy (2010) used PISA 2006 data to examine the relationship between instructional time spent on mathematics, science, and reading and student achievement in each of these content areas. The relationship between instructional time and achievement was investigated for the entire sample of PISA participants as well for participants stratified by level of development. Results from an OLS regression analysis indicated that instructional time was significantly related to student achievement in all subjects. Moreover, the relationship between instructional time and achievement was stronger for less developed countries.

TIMSS 2007 found that yearly instructional time spent on mathematics and science instruction varied across the countries under study in this dissertation. Teachers in Chinese Taipei, the United States, and the Czech Republic devoted more than the international average time to mathematics instruction (158 hours in Chinese Taipei, 148 hours in the United States, 128 hours in the Czech Republic, and 124 hours in Singapore). Teachers in Slovenia spent somewhat less than the international average time on mathematics instruction (113 hours). Results in science reflect the fact that the Czech Republic and Slovenia teach science as separate subjects. Teachers in the Czech Republic and Slovenia spent an average of 226 hours and 166 hours per year, respectively on science instruction. Teachers in the other countries devoted considerably less time to science instruction per year—145 hours in Chinese Taipei, 140 hours in Singapore, and 139 hours in the United States.

## Fidelity to Curriculum Process

Fidelity to process in the context of curriculum implementation relates to the quality of instruction, represented by the use of effective instructional practices. The U.S. National Council for Teachers of Mathematics (NCTM) has identified effective strategies for mathematics instruction as those centered around improving students' problem solving, reasoning and proof, communication, connections, and representations (National Council of Teachers of Mathematics, 2011). For science instruction, the National Research Council (NRC) identifies eight effective instructional practices in a K-12 STEM curriculum: 1) asking questions and defining problems, 2) developing and using models, 3) planning and carrying out investigations, 4) analyzing and interpreting data, 5) using mathematics, information and computer technology, and computational thinking, 6) constructing explanations and designing, 7) engaging in argument from evidence, and 8) obtaining, evaluating, and communicating information (National Research Council, 2011).

There are several studies that provide empirical support for the relationship between instructional practices like those proposed by the NCTM and NRC and student achievement in mathematics and science. For example, Lubienski (2006) examined fourth and eighth grade students' NAEP 2000 mathematics scores to determine which reform-oriented NCTM instructional practices correlated positively with mathematics achievement after controlling for race, socioeconomic status, and gender. Results from her hierarchical linear modeling analysis revealed that after controlling for confounding variables, collaborative problem solving and teacher knowledge of the NCTM Standards (i.e., teachers who were "very knowledgeable" about the standards) positively predicted fourth and eighth grade student achievement. Furthermore, emphasizing geometry; measurement; data analysis, statistics, and probability; and algebra and functions had a significant association with fourth grade student achievement. At the eighth grade, the same factors were positively associated with student achievement as well as calculator use, reasoning, and a de-emphasis of facts and skills.

In a similar analysis, Braun, Coley, Jia, and Trapani (2009) explored the relationship between teacher characteristics, instructional strategies, and science achievement as measured by the eighth grade NAEP assessment. Researchers used hierarchical linear modeling to examine the relationship between instructional practices and science achievement after controlling for differences in student demographics, home environment, and teacher characteristics. Results indicated teacher instructional strategies were associated with science achievement—some positively associated when used frequently, some negatively associated when used frequently, and some positively associated when used in moderation. Those instructional strategies associated with higher science achievement when used frequently included using a science textbook, doing hands-on activities in science, writing long answers to science tests and assignments, talking about measurements and results from hands-on activities, and working with others on a science activity or project. Pedagogical strategies associated with lower science scores when used frequently included students giving an oral report and using library resources. Lastly, there were seven strategies that were positively associated with higher science scores when used in moderation (what Braun et al. term the "Goldilocks" pattern). These included teachers doing a science demonstration and students taking a science test, discussing science in the news, reading a book or magazine about science, and preparing a written science report.

In particular for science instruction, research has examined the relationship between inquiry-based instructional strategies and science achievement. In a recent meta-analysis, researchers developed a framework for what constitutes inquiry-based instruction and examined research on the impact of inquiry science instruction on K-12 student outcomes (Minner, Levy, & Century, 2009). In the framework, the researchers defined inquiry science instruction as instruction having science content and student engagement with science content.  Moreover, they described inquiry instruction as involving five components—question, design, data, conclusion, and communication, and determined that students must interact with at least one of the components by being responsible for their own learning, using active thinking, or being motivated to learn. Student outcomes were measures of student understanding of science concepts, facts, and principles or student retention of that science understanding after a minimum of two weeks. The pattern of results across the 138 studies synthesized indicated using some level of inquiry-based instruction had a positive relationship with student understanding and retention of science content. In particular, instruction emphasizing active thinking and drawing conclusions from data or providing hands-on experiences with scientific phenomena were associated with increased likelihood of science understanding.

Other research has investigated middle school science inquiry instruction in particular. Fogleman, McNeill, and Krajcik (2011) evaluated how teachers implemented an inquiry-oriented middle school science curriculum and the resulting relationship with student science achievement. Student achievement was measured using gain scores from pre- and post-tests aligned to the curricular unit being evaluated. Hierarchical linear modeling results indicated that using the inquiry-oriented curriculum was associated with increased science achievement. How teachers enacted the curriculum also was important—teacher's choice of activity structure was significantly related to achievement. Specifically, students who did inquiry activities themselves had greater achievement gains than students whose teachers' did the activities as demonstrations.

Instructional practices also have been studied in the international context. Papanastasiou (2008) examined effective teaching in Cyprus by analyzing the mathematics results from TIMSS 1999. Using a residual regression analysis, he found two instructional strategies that accounted for school differences related to mathematics achievement. The first practice was described as "transmissional teaching"—when the teacher shows students how to solve problems or uses a blackboard or overhead projector during instruction, or when students copy notes from the board. The second practice was described as "active learning" and was operationalized as activities involving students working on mathematics projects, using situations from everyday life in solving math problems, working together in small groups, and trying to solve examples related to a new topic.

House (2005) examined cross-cultural variations in the relationships between instructional strategies and science achievement in high-performing Japan, Hong Kong, and Chinese Taipei using data from TIMSS 1999. Specifically, he examined how frequently teachers used the following strategies when introducing new science topics to their class: explained rules and definitions, discussed a problem related to everyday life, had students work together in small groups, had students relate what they knew to a new topic, had students look at the textbook while the teacher discussed it, and had students try to solve examples related to the new topic. In all three countries, having the teacher explain the rules and definitions of a new science topic and having students discuss a problem related to everyday life was associated with higher TIMSS 1999 science scores. Furthermore, in Hong Kong and Chinese Taipei, having the teacher ask students what they knew related to the new topic, looking at the textbook while the teacher discussed it, and solving an example related to the new topic also were strategies associated with higher science achievement.

TIMSS 2007 provides data regarding the frequency with which teachers had students do five mathematics content-related activities: practice adding, subtracting, multiplying, and dividing without using a calculator; work on fractions and decimals; write equations and functions to represent relationships; solve problems about geometric shapes, lines, and angles; and interpret data in tables, charts, or graphs. In all countries except Singapore, between 60 to 75 percent of eighth grade students were still practicing adding, subtracting, multiplying, and dividing in about half the mathematics lessons or more (Mullis, Martin, & Foy, 2008). Singapore had only 41 percent of students reporting they practiced these basic computations in about half the lessons or more.

TIMSS 2007 also reported data relating to teachers' focus on mathematics instructional activities related to problem-solving. Specifically, data were gathered to highlight differences between activities emphasizing memorizing how to work problems and working problems independently or explaining answers. At the eighth grade, memorizing formulas and procedures was not an instructional activity that many teachers reported doing frequently. Instead, the instructional focus in the two high-performing Asian countries as well as the United States appeared to be applying facts, concepts, and procedures to solve routine problems as well as explaining answers. In the Central European countries of the Czech Republic and Slovenia, the most frequent instructional activities were the same two as in the other countries in addition to relating what is being learned in mathematics to students' daily lives.

In science, TIMSS 2007 countries reported differences in the frequency with which eighth grade teachers engaged in scientific inquiry-related instructional activities

(Martin, Mullis, & Foy, 2008). These activities included making an observation and describing what was seen, giving an explanation about what was being studied, watching the teacher demonstrate an experiment or investigation, design or plan an experiment or investigation, conduct an experiment or investigation, work in small groups, and, relate what is being learned in science to daily life. The two Asian countries as well as the two central European countries in this dissertation reported eighth grade students having teachers who related what students were learning in science to their daily lives more than any other inquiry activity, on average. In the United States, equally high percentages of students had teachers reporting they gave explanations about something they were studying and related what students were learning in science to their daily lives.

## Participant Responsiveness

Participant responsiveness underlies both aspects of fidelity of curriculum implementation and reflects “the extent to which participants are engaged by and involved in the activities and content of the program” (O'Donnell, 2008, p. 34). Participant responsiveness can be gauged by student attitudes toward learning, which provide information about curriculum implementation since they may be considered a report from the student’s perspective on the effectiveness of a teacher’s instruction (McLaughlin et al., 2005). Students must attend to the subject matter being taught and the willingness of the student to participate in instruction is related to their motivation and attitudes toward the subject (McLaughlin et al., 2005). Therefore, student attitudes toward learning in general and toward school subjects in particular are important indicators of a students' motivation and potential academic success (Creemers & Kyriakides, 2008).

Hattie (2009) conducted a meta-analysis of student attitudes toward school and found that attitudes toward mathematics or science were related to mathematics and science achievement across 288 studies (with an effect size of d=0.36). He suggests that because it is a correlate of achievement, developing positive attitudes toward school is a desirable outcome of schooling.

Singh, Granville, and Dika (2002) used data from the National Education Longitudinal Study to examine the association between attitude, motivation, and academic engagement and eighth grade students' mathematics and science achievement using structural equation modeling. Student attitudes were measured by three variables indicating if a student looked forward to mathematics or science class, thought mathematics or science would be useful in the future, and if the student was bored in school. Motivation was represented by students' school and class attendance and frequency of coming to class prepared (e.g., with pencils, books, and homework). Academic time was measured through examining time spent on homework and time spent watching television (reverse-coded). Results indicated that attitude toward mathematics was directly related to mathematics achievement. Attitudes toward mathematics and science also had indirect effects on achievement through academic time and motivation. For example, students who had a positive attitude toward mathematics and science spent more time on science homework and less time watching television.

Similarly, more motivated students had a more positive attitude toward mathematics and science.

Ma and Xu (2004) contribute to the literature on student attitudes by exploring the causal direction between attitudes and achievement. The researchers used structural equation modeling to examine nationally representative data from the Longitudinal Study of American Youth for students in grades 6-12. Student attitudes were measured using students' responses to three items: "mathematics is useful in everyday problems", "mathematics helps a person think logically", and "I will use mathematics in many ways as an adult" (Ma & Xu, 2004, p. 261). Results indicated that prior achievement significantly predicted changes in attitude in later grades; but that the opposite was not true. That is, prior attitudes did not predict later achievement.

TIMSS 2007 results show a relationship between student attitudes toward learning and mathematics and science achievement. In TIMSS 2007, student responses to three statements about mathematics—"I enjoy learning mathematics", "Mathematics is boring" (reverse coded), and "I like mathematics" were combined to form the Index of Students' Positive Affect Toward Mathematics as measured by how much students agreed with the three statements on average (Mullis, Martin, & Foy, 2008). TIMSS also created a comparable index for science called the Index of Students' Positive Affect Toward Science (Martin, Mullis, & Foy, 2008). On average across countries, eighth grade students at the high level of each index had higher average mathematics or science achievement than students at the medium or low levels.

### School Climate

School climates that support academic success are associated with higher achievement, including those having few discipline and attendance problems, or having a more positive environment (Mullis, Martin, Ruddock, O'Sullivan, & Preuschoff, 2009). School environment has been noted in the school effectiveness literature as an important factor that differentiates more effective from less effective schools (Reynolds & Teddlie, 2000). However, there are many interpretations in the literature about what factors are involved in school climate. Across the different studies, common factors emerge related to positive school climate factors, such as teacher understanding of school goals and students' desire to do well in school (Johnson & Stevens, 2006; Reynolds & Teddlie, 2000). Negative school climate factors also emerge, including those related to discipline and attendance problems in a school, such as school safety issues, student conflicts, student absenteeism, and classroom disturbance (Austin & Bailey, 2008; Koth, Bradshaw, & Leaf, 2008; Lubienski, Lubienski, & Crane, 2008). Despite the varying definitions, there appears to be a consistent relationship between school climate and student achievement in many studies.

Reynolds and Teddlie (2000) note that the process of producing a positive school culture entails creating a shared vision, creating an orderly environment, and emphasizing positive reinforcement. Hoy, Tarter, and Hoy (2006) discuss school climate using a similar framework, but focus on a school's level of "academic optimism", a factor representing a school's press for academic achievement comprising the school's academic emphasis, collective efficacy, and degree of teacher-parent trust. Academic emphasis was operationalized as the degree to which the faculty stressed high achievement, and students worked hard, cooperated, and were respectful to high-achieving peers. Collective efficacy related to how strongly teachers felt the faculty and staff at the school could have positive effects on students. Finally, the degree of teacher-parent trust was measured by how much the faculty trusted in students and parents. Results indicated high academic optimism had a positive relationship with twelfth grade student achievement in a state-mandated 12th-grade test in mathematics, science, reading, social studies, and writing after controlling for demographic factors and prior achievement.

Other researchers define school climate by its positive and negative aspects. Lubienski, Lubienski, and Crane (2008) examined the role of school climate in student mathematics achievement by broadly defining school climate to include seven factors: teacher morale, conflicts among students, drugs or alcohol problems, parental involvement, parental volunteerism, parents talking about schoolwork with students at home, and student attendance. Results from a hierarchical linear modeling analysis revealed that parental involvement, talking about studies at home, teacher morale, conflicts among students (reverse coded), and student attendance were all positively related to eighth grade student achievement on the NAEP 2003 mathematics test.

Another perspective on the factors involved in school climate and their relationship to achievement is outlined in a report prepared for the California Department of Education Safe and Healthy Kids Program Office (Austin & Bailey, 2008). Researchers created scales of six school climate factors using California School Climate

Survey responses from teachers, administrators, and other staff teaching in California public schools—staff-student relationships, student behaviors that facilitate learning, school safety (including lack of violence and victimization), and substance use (low problem levels). The academic performance of the school was measured as low, medium, or high according to the state's Academic Performance Index (API). Comparing results of the school climate scales to the API scores of each school indicated that academic performance was positively associated with each of the school climate scales at the elementary and middle school level.

A multilevel study conducted by Koth, Bradshaw, and Leaf (2008) investigated classroom-and school-level factors accounting for perceptions of school climate, including the order and discipline level of the school and the extent of students' academic motivation to learn. At the school level, larger student enrollment (i.e., school size) and high faculty turnover rates (i.e., the percentage of faculty new to the school that year) were associated with lower perceptions of order and discipline as well as achievement motivation, after controlling for student- and classroom-level factors. At the classroom level, teachers new to the profession, large class sizes, and a concentration of students with behavior problems were associated with a school climate perceived as less safe and orderly compared to classrooms with more experienced teachers, smaller class size, and with fewer disruptive classmates.

Research indicates that the relationship between school climate and student achievement may be mediated by socioeconomic factors. Johnson and Stevens (2006) used a modified version of the School-Level Environment Questionnaire (SLEQ) to measure teachers' perceptions of school climate, including degree of teacher affiliation, atmosphere of innovation, teacher involvement in decision-making, adequacy of resources, and student support. Researchers used structural equation modeling to examine results from the SLEQ and fourth grade student achievement as measured by the Terra Nova Survey Plus standardized achievement test. Results indicated a positive relationship between teachers' perceptions of school climate and student achievement. However, the association between school climate and student achievement was mediated by school socioeconomic status. That is, there was a strong relationship between school climate and student achievement in schools in high socioeconomic areas, but a weaker relationship in schools in low socioeconomic areas.

Research conducted by the National Center for Education Statistics indicates large numbers of teachers left the profession in the early 1990s citing "student discipline problems" as one of the major reasons for leaving (Mayer, Mullens, & Moore, 2000, p. 14). The researchers note that reducing the numbers of well-prepared teachers in a school negatively affects a school's professional community, which has indirect effects on student achievement.

Individual components of a school climate also have been researched in relation to student achievement. For example, school safety issues have been identified as one of the important correlates of student achievement (Barton, 2003). Especially relevant for promoting higher student achievement are schools that are "safe, welcoming, stimulating and nurturing" and that encourage respectful interactions between individuals

(Committee on Highly Successful Schools or Programs in K-12 STEM Education & National Research Council, 2011, p. 24).

In particular, schools with high levels of student bullying do not constitute climates that are conducive to learning. For example, Rothon, Head, Klineberg, and Stansfeld (2010) used logistic regression to examine whether students who were bullied reached the UK national academic benchmark in English, mathematics, and science for their grade level or not. The researchers defined bullying as "a desire to hurt and the execution of a harmful action; it is characterized by repetition and either a physical or a psychological power imbalance" (Rothon et al., 2010). Results indicated that bullying had a strong impact on a student's odds of achieving the national academic benchmark for their age. That is, students who had been bullied within the last semester were approximately half as likely to achieve the benchmark as students who had not been bullied within the last semester.

Students must also have the opportunity to learn STEM topics, so problems relating to student and teacher absenteeism are strongly associated with lower achievement (Abadzi, 2007). For example, Singh, Granville, and Dika (2002) used data from the National Education Longitudinal Study to examine the association between students' school and class attendance and eighth grade students' mathematics and science achievement using structural equation modeling. Student responses to three survey items were used to create a composite variable of the level of students' school and class attendance—frequency of missing school, skipping class, and tardiness. Results indicated that level of students' school and class attendance was directly related to mathematics achievement.

TIMSS 2007 reported information about the relationship between attendance problems and student achievement across countries. School principals were asked three questions about how serious the following attendance problems were in their school: students arriving late, absenteeism (i.e., unjustified absences), and skipping class. TIMSS 2007 developed the Index of Good Attendance at School (GAS) based on a composite of the three items (Mullis, Martin, & Foy, 2008; Martin, Mullis, & Foy, 2008). Schools at the high level of the index reported that all three behaviors never occurred or were not a problem. Schools at the low level of the index reported that two or more of the behaviors were a serious problem or that one was serious and the other two were minor problems. Schools responding with all other response combinations were categorized at the medium level of the index. Results across countries at the eighth grade indicated that the high level of the GAS index was associated with higher mathematics and science achievement compared to the low level of the GAS index. That is, students attending schools where students arriving late, absenteeism, and skipping class were not serious problems had the highest mathematics and science achievement. Regarding the countries under study in this dissertation, Chinese Taipei had the second highest percentage of students in the high level of the GAS index internationally (52%), while Singapore, the Czech Republic, and Slovenia all had higher than the international average percentages of students at the high level of the index. The United States had only 15 percent of its students at the high level of the GAS, below the international average of 21 percent.

TIMSS 2007 results support previous research cited about the positive relationship between mathematics and science achievement and school climate from an international perspective. Reports from principals and teachers in TIMSS 2007 were used to develop two indicators of school climate, one for teachers' perceptions and one for principals' perceptions (Mullis, Martin, & Foy, 2008; Martin, Mullis, & Foy, 2008). The indicators were based on their responses (very high to very low) to a series of questions regarding: teachers' job satisfaction, teachers' understanding of the school's curricular goals, teachers' degree of success in implementing the school's curriculum, teachers' expectations for student achievement, parental support for student achievement, parental involvement in school activities, students' regard for school property, and students' desire to do well in school. Students were assigned to the high level of the Principals' Perception of School Climate (PPSC) index if they attended schools where the principal averaged high or very high on all aspects of school climate, and to the low level where the principal averaged low or very low. Students at the medium level had principals with other response combinations. Mathematics and science achievement was positively associated with principal's perceptions of school climate on average across countries.

TIMSS 2007 asked teachers the same eight questions as the principals regarding their perceptions of school climate and created the Index of Mathematics Teachers' Perception of School Climate (PPSC) and Index of Science Teachers' Perception of School Climate (PPSC). Although teachers had a less positive outlook on school climate than principals, average mathematics or science achievement was still positively related to teachers' perceptions of school climate at both fourth and eighth grades. That is, average achievement was higher among students at the high index level and lower among students at the low index level.

## Summary

This section reviewed the literature on school effectiveness factors related to STEM achievement within the conceptual framework of school resources, fidelity of curriculum implementation, and school climate factors. A number of home resources and school context factors were found to be associated with science, technology, engineering, and mathematics achievement. In relation to student home resources, higher levels of home educational resources and parental education were found to be associated with higher STEM achievement. At the school level, school resources, including teacher preparation as well as general and mathematics and science specific school resources were related to STEM achievement. Effective schools appear to be equipped with well–prepared teachers who feel ready to teach nearly all the TIMSS topics based on high education levels and majors in the field of study being taught. Further, well–resourced schools and classrooms were associated with higher STEM achievement as indicated by factors, such as principals reporting that resource shortages do not adversely affect instruction and that schools are equipped with resources specific to teaching STEM topics, such as science laboratories.

Fidelity of curriculum implementation factors, comprising fidelity to curriculum structure, fidelity to curriculum process, and participant responsiveness also were related to STEM achievement. Although many countries have established policies and standards about the intended curriculum in mathematics and science, how closely schools adhere to the curriculum in terms of the topics taught and time spent on instruction as well as the quality of the strategies they use to teach the curriculum are important factors in school effectiveness. Higher–performing schools tend to actually teach the intended mathematics and science curriculum and use effective instructional strategies to do so. Students need to have been taught the relevant topics in order to succeed in STEM achievement, including number, algebra, geometry, and data and chance, as well as biology, chemistry, physics, and earth science. Participant responsiveness, gauged by student attitudes toward learning as a measure of effective instruction also is associated with higher achievement.

Finally, school climate factors were related to STEM achievement. School climates that support academic success are associated with higher achievement, including those having few discipline and attendance problems, or having a more positive environment for student learning.

# Chapter 3

## *Methodology*

### Introduction

This dissertation research used multilevel modeling to examine how school resources, curriculum implementation, and school climate factors operated in Chinese Taipei, Singapore, the Czech Republic, and Slovenia compared to the United States. The data measuring the school factors was collected via the TIMSS 2007 questionnaires administered to students, teachers, and school administrators. The dependent variable was TIMSS 2007 eighth grade mathematics and science achievement data combined into a STEM outcome measure.

This chapter documents the methodological aspects of the study, beginning with an overview of the TIMSS 2007 assessment, including the target population and sampling design, achievement tests, assessment design, achievement scaling, contextual background data, and the TIMSS 2007 database used for analysis. The next section describes the process for developing school effectiveness factors contributing to success in STEM achievement as well as a student-level home resource control, including exploratory analysis, evaluating items for scaling, and scaling the items. The last section describes the main analysis approach to this research, which involved creating a series of multilevel statistical models to examine the relationship between the school effectiveness factors and STEM achievement, after controlling for student home resources.

## Description of TIMSS 2007

### *Target Population and Sampling Design*

The TIMSS 2007 target populations included all students enrolled in the fourth and eighth years of formal schooling, counting from the first year of primary school as defined by UNESCO's International Standard Classification of Education (ISCED) (Joncas, 2008). This dissertation used only the eighth grade data.

TIMSS 2007 used rigorous sampling of schools and students in order to achieve nationally representative samples for each country. A two-stage stratified cluster design was used, whereby first schools and then intact classrooms were sampled (Joncas, 2008). In the first stage, schools were sampled using the probability proportional-to-size (PPS) technique, which allows larger schools a greater chance of being selected for the sample. The second stage involved sampling one or more intact classes from the eighth grade in the sampled schools. Classrooms were selected with equal probabilities. Because Singapore has very large classes, the classes within schools also were sampled with a PPS method and the students within the classes were selected with equal probabilities, so only a subsample of students (19 students in each class) was selected to participate.

Sampling weights were calculated for each country that accounted for the probability of selection for each student, classroom, and school participating in TIMSS 2007 (Joncas, 2008). For each student, an overall sampling weight (TOTWGT) was calculated as the product of three component weights—1) the school weight, 2) the classroom-within-school weight, and 3) the student-within-classroom weight and represents the probability of that student being selected from within that classroom, within that school, and within that country. School weights were calculated to reflect the probability a sampled school was selected from among all the eligible schools in the country. The classroom-within-school weight reflects the probability the sampled classroom was selected from all the classrooms in a school at the target grade level (i.e., eighth grade). Because most of the participants used in this dissertation research sampled intact classrooms, the student-within-classroom weight for most countries was 1.0 because each student in the sampled class was certain to be selected. In Singapore, where students were further sampled within classrooms, the student weight reflects the probability that the student is sampled from within the classroom.

In analyzing the data for this dissertation study, the sampling weights provided in the *TIMSS 2007 International Database* (Foy & Olson, 2009) were used, since they properly account for the sample design, take the stratification into account, and include adjustments for non-response (Joncas, 2008). More specifically, the student sampling weight, TOTWGT, which sums to the national population size, was used in all analyses, although it was transformed in some analyses, where appropriate. In order for countries to contribute equally to parameter estimates in Rasch scales using teacher data, a transformed version of TOTWGT was used (SENWGT), so that the sample size summed to 500 in each country. The SENWGT also was adjusted for the number of mathematics or science teachers each student had so the teacher scores contributed equally to the number of students in each country. For the multilevel analyses, TOTWGT was normalized by the software so that the sum of the weights was equal to the student sample size in the data. The use of the weights in this dissertation ensured that all subgroups of the sample were properly represented in the estimation of population parameters.

Typically, approximately 150 schools were sampled in each participating country, which yielded a representative sample of approximately 4,500 students in each participant country (Joncas, 2008). Table 3.1 presents the total number of schools, classrooms, and students that participated in TIMSS 2007 at the eighth grade for the countries in this research.

Table 3.1. Number of Schools, Classrooms, and Students

| Country | Number of Schools | Number of Classrooms per School | | Number of Students |
|---|---|---|---|---|
| | | 1 Classroom | 2 Classrooms | |
| Chinese Taipei | 150 | 150 | -- | 4,046 |
| Czech Republic | 147 | 87 | 60 | 4,845 |
| Singapore | 164 | 33 | 131 | 4,599 |
| Slovenia | 148 | 47 | 101 | 4,043 |
| United States | 239 | 25 | 214 | 7,377 |

Chinese Taipei sampled 150 schools with one classroom per school at the eighth grade, yielding 4,046 students. The Czech Republic sampled 147 schools with one or two classrooms sampled in each school, resulting in a sample of 4,845 eighth grade students. Singapore sampled 164 schools and mostly two classrooms per school using PPS as previously mentioned, drawing 19 students from each class. This resulted in a student sample of 4,599. Slovenia sampled 148 schools and two classrooms per school when the school had at least 40 eighth grade students (one classroom otherwise), yielding 4,043 students. The United States sampled 239 schools and mostly two eighth grade classrooms per school, resulting in a student sample of 7,377.

### *Mathematics and Science Achievement Tests*

The achievement items included in the TIMSS 2007 Mathematics and TIMSS 2007 Science assessments were designed to measure content and cognitive dimensions of student understanding. As described in the *TIMSS 2007 Assessment Frameworks* (Mullis et al., 2005), the eighth grade mathematics content domains included *number*, *algebra*, *geometry*, and *data and chance*. The eighth grade content domains in science included *biology*, *chemistry*, *physics*, and *earth science*. The cognitive domains measured by the items were the same for mathematics and science—*knowing*, *applying*, and *reasoning*. *Knowing* items generally require students to recall facts and procedures, while *applying* items focus on students' ability to apply their knowledge and conceptual understanding to answer questions. *Reasoning* items require the student to solve non-routine problems and extend their knowledge to new situations. In the science assessment, items also assess scientific inquiry as an overarching assessment strand, overlapping the various content and cognitive domains.

The STEM achievement measure used in this dissertation is based on 429 items, approximately half multiple-choice format and half constructed-response. Multiple-choice items provide students with four response options, of which only one is correct, and are worth one score point each. Constructed-response items requiring students to provide their own answers are worth one or two score points, depending on the type of question and what it requires of students. At the eighth grade, the TIMSS 2007 mathematics assessment consisted of 117 Multiple-choice items and 98 constructed-response items (Ruddock, O'Sullivan, Arora, & Erberber, 2008). The TIMSS 2007 science assessment at the eighth grade consisted of 107 multiple-choice items and 107 constructed-response items.

### Assessment Design

Because of the wide content coverage in the TIMSS mathematics and science assessments, there are many more items on each assessment than any one student could answer in the allotted amount of testing time. Therefore, TIMSS routinely uses a matrix-sampling approach to allocate the pool of items across student achievement booklets for each subject, with each student answering the science and mathematics items in only one booklet. More specifically, items were grouped into “item blocks” by subject comprising approximately 10-15 items each and distributed across the test booklets. In TIMSS 2007, 14 student test booklets, each containing two science blocks and two mathematics blocks were assembled from 28 item blocks. Each item appears in two of the booklets so that student responses may be linked across booklets, using Item-Response Theory (IRT) scaling techniques.

### Achievement Scaling

In order to estimate student achievement scores, TIMSS 2007 made use of Item Response Theory (IRT) scaling in combination with conditioning and multiple imputation (“plausible values” methodology) as explained in the *TIMSS 2007 Technical Report* (Foy, Galia, & Li, 2008). TIMSS makes use of plausible values methodology to get accurate estimates of student achievement on the assessment as a whole despite the fact that each student only completed one test booklet. The plausible values approach uses all available data, including students' responses to the achievement items they were administered and all available background data, a process known as "conditioning" to estimate the characteristics of students' likely achievement distributions (Foy et al., 2008). Five imputed scores or plausible values are then generated from each student's likely achievement distribution.

The STEM outcome measure in this dissertation is a sum of students' TIMSS mathematics and TIMSS science achievement scores. In the TIMSS dataset there are five plausible values for each student on each assessment (mathematics and science), representing student achievement. The STEM measures were computed by adding each student's first plausible value from the mathematics assessment to his or her first plausible value for the science assessment, and so on, until each student had five STEM composite measures.

### *Contextual Background Data*

Data from the TIMSS 2007 student, teacher, and school questionnaires were used for these dissertation analyses. The *TIMSS 2007 Student Questionnaire* was given to the eighth grade students participating in TIMSS 2007 to collect information about their home and school environments and their experiences learning mathematics and science. Students provided information about themselves, such as their gender and age, as well as their home environment, including the language spoken at home and how many books there are in the home. Items also addressed students' motivation and attitudes toward learning mathematics and science (e.g., how much the student likes or values mathematics). The student questionnaire also covered characteristics of the school environment, including school climate, resources, and safety concerns. There was one version of the student questionnaire for countries where science is taught as a single integrated subject, including Chinese Taipei, Singapore, and the United States. For students in countries such as the Czech Republic and Slovenia where science is taught as separate subjects (i.e., biology, chemistry, physics, and earth science), the student questionnaire posed the same questions about attitudes toward science and instructional practices in science lessons as the general science version, but about each science subject area separately. There were 21 items on the *TIMSS 2007 Student Questionnaire* used for students in countries teaching general science and 33 items on the *TIMSS 2007 Student Questionnaire – Separate Science Subjects*.

The *TIMSS 2007 Teacher Questionnaire* collected information about the teachers of the students participating in the assessment. The teacher of the mathematics class sampled for the TIMSS 2007 assessment was asked to complete a questionnaire specific to teaching mathematics to that class, and each of the science teachers of students in that class were asked to complete a questionnaire specific to teaching science to those students. Both the mathematics teacher and science teacher questionnaires collected demographic data, asking for the teacher's age, years teaching, and level of education and major(s). The two teacher questionnaires also included parallel questions across subjects relating to instructional practices, content topic coverage, lesson preparation, homework, and assessment. Overall, the *TIMSS 2007 Mathematics Teacher Questionnaire* included 33 items and the *TIMSS 2007 Science Teacher Questionnaire* included 31 items.

The 22-item *TIMSS 2007 School Questionnaire* collected information from principals of each sampled school, including general school characteristics, such as the total enrollment, the percentage of students from economically disadvantaged backgrounds, and total instructional time. Issues surrounding school climate, such as levels of parental involvement, teacher job satisfaction, and student behavior also were addressed as were schoolwide policies on instructional practices, including ability grouping and enrichment services. The questionnaire also asked about any difficulties in recruiting mathematics and science teachers and the availability of school resources, such as science laboratories.

### The TIMSS 2007 International Database

The *TIMSS 2007 International Database* containing the students' mathematics and science achievement data together with the student, teacher, school, and curricular questionnaire data is available for public use (Foy & Olson, 2009). The *TIMSS 2007 User Guide* (Foy & Olson, 2009) provides information about the organization and content of the database, while the *TIMSS 2007 Technical Report* (Olson, Martin, & Mullis, 2008) describes the TIMSS 2007 assessment methods and procedures in detail. Both were used as important resources for this dissertation study.

In addition to the international versions of all the background questionnaires administered in TIMSS 2007, the database also provides details on national adaptations that were applied to the background questionnaires. That is, country-specific questions added by individual countries to suit their national context. Items used in this dissertation study were reviewed for any adaptations countries might have made to provide more information about their national situation.

The TIMSS database also provides comparable student, teacher, and school data files for Chinese Taipei, Singapore, the Czech Republic, Slovenia, and the United States in SAS and SPSS formats. The IEA International Database Analyzer software, described in the User Guide, allows for easy merging of the data across different sources (e.g., the school and student questionnaires). This feature was essential in helping to organize the data used for this analysis, which combines data from multiple questionnaires and across two levels (i.e., the school- and student-level). The database also provides software programs to properly calculate the standard errors and means for achievement data with plausible values.

The database provides background data almanacs, including weighted summary statistics for each item used in TIMSS 2007 by participating country. The almanacs present detailed descriptive statistics for each questionnaire item as well as its relationship to average achievement. For categorical data (e.g., teacher level of formal education), the almanacs provide the sample size, the number of valid cases, the weighted percentages of students in each response category, the weighted mean achievement values of students in each response category, and similar statistics for cases where there are no valid responses (e.g., not administered or omitted). Similar information is displayed for continuous variables (e.g., total school enrollment) except that instead of percentages of students in each response category, the weighted mean, mode, minimum, maximum, and the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles are shown.

## Developing School Effectiveness Factors

### *Student-Level Home Resources Control*

This dissertation research focused on identifying important school factors that affect student achievement in STEM areas above and beyond the student's home resources. Student-level factors, such as home educational resources and parental education are included in school effectiveness analyses as covariates so the added value of schooling can be separated from the impact of the home.

An examination of the *TIMSS 2007 Student Questionnaire* yielded twelve items that could be used to control for students' home resources. These items asked about: the number of books in the student's home; whether or not the student had any of the following possessions in his or her home—calculator, computer, study desk, dictionary, Internet connection, and four country-specific possessions; the education level of the student's mother; and the education level of the student's father.

### *School-Level Factors*

The set of overarching school effectiveness factors examined in this dissertation emerged through an iterative process of comparing factors in the conceptual model (see Figure 1.1) with items in the TIMSS 2007 questionnaires. The TIMSS 2007 questionnaires were thoroughly reviewed to identify individual items that matched the broader factors in the conceptual model. The review resulted in selecting 112 school-level mathematics items and 125 school-level science items for further exploratory analyses. The items were from the eighth grade versions of the *TIMSS 2007 Student*

*Questionnaire,* the *TIMSS 2007 Student Questionnaire for Separate Science Subjects*, the *TIMSS 2007 Mathematics Teacher Questionnaire*, the *TIMSS 2007 Science Teacher Questionnaire*, or the *TIMSS 2007 School Questionnaire* (which is not subject-specific).

Table 3.2 shows eight factors represented by the items selected that matched the three broad areas in the conceptual model—three factors related to school resources, three factors related to fidelity of curriculum implementation, and two factors related to school climate.

Table 3.2. School Effectiveness Factors Selected for Analysis

| School Resources |
|---|
| Teachers Preparation |
| General School Resources |
| Mathematics and Science Specific School Resources |
| **Fidelity of Curriculum Implementation** |
| Fidelity to Curriculum Structure |
| Fidelity to Curriculum Process |
| Participant Responsiveness |
| **School Climate** |
| Discipline and Attendance Problems |
| Support for Academic Success |

## *Exploratory Analyses*

Initially, the background data almanacs provided in the *TIMSS 2007 International Database* (Foy & Olson, 2009) were reviewed to determine whether the items of interest had a relationship with students' mathematics and science achievement in the five countries under study. Percentages of students in each response category were examined in conjunction with the average achievement of students in each country. Items positively related to achievement were identified for further review.

Before proceeding with further analyses, items were reviewed for national adaptations implemented by the countries under study. For example, the student questionnaire asks students if they have any of the following things in their home—calculator, computer, study desk, dictionary, or Internet connection. Each participating country also had the option of including up to four country-specific home possessions in this item. For example, the Czech Republic added "your own Discman or mp3 player", "digital camera (do not count a mobile phone with a camera)", "camcorder", and "dishwasher" to the list of possessions (Foy & Olson, 2009). Supplement 2 of the TIMSS 2007 database provides information on the items that countries adapted, did not administer, or modified to suit their national context.

## *Evaluating Items for Scaling*

Once a preliminary set of items having a relationship with mathematics and science achievement was identified, individual items were examined to determine if any could be combined into composite variables or scales. Combining sets of items into scales increases the level of measurement precision compared to a single item and can better reveal levels of an underlying construct or latent variable (DeVellis, 2003). Combined sets of items have an advantage in providing a more reliable measure of the construct, or underlying phenomenon that is not directly observable (Messick, 1989).

However, when creating scales measuring background characteristics, it is important to have evidence that the scale is a good representation of the underlying construct. Using a scale that does not measure what one assumes it measures can lead to incorrect conclusions about results. For valid research, it is important to ensure that scales have sound measurement properties so that results can provide valid information about the construct of interest (DeVellis, 2003). Therefore, sets of items that could be scaled for use in further multilevel analyses were first evaluated in terms of scale length, unidimensionality, reliability, and relationship with achievement.

First, in order to be able to assign meaning to the scores produced by a scale, there should be enough items to reliably represent the underlying construct. Scale length is positively related to the reliability of the scale (Crocker & Algina, 1986). It is preferable for scales to have at least 20 or so scale points (about five or six 4-point items) to be suitable for IRT scaling (Suen, 1990).

Second, factor analytic techniques were used to statistically confirm that the potential sets of items were measuring one latent variable. Specifically, principal components analysis (PCA) was used to establish the unidimensionality of each set of items. PCA is a data reduction technique used to determine the underlying factor structure of a group of items (Dunteman, 1989). PCA can reveal if the construct being measured by the instrument is unidimensional (drawing on one ability or trait) or multidimensional (drawing upon separate abilities or traits). Statistically, this is done by maximizing the variance of a linear combination of items using a vector that captures as much variance as possible between responses (Dunteman, 1989). Thus, PCA summarizes the variation among sets of items into smaller numbers of major, or principal, components. In this research, factor analyses were conducted using PCA implemented in IBM SPSS Statistics® 19.0.

In PCA, the proportion of shared variance represents the percent of variance in a set of items that can be explained by the latent variable (represented by the component) (Dunteman, 1989). It is preferable for a single factor to account for at least 50 percent of the scale variance with no second dominant factor. However, Reckase (1979) proposes a minimum criterion of 20 percent because scales composed of items whose first factor does not account for at least 20 percent of the variance can result in unstable item parameters.

In each analysis, the item-scale correlations, or factor loadings, were examined to determine which items most clearly belonged in each scale. A high factor loading indicates that an item has a high correlation with the respective factor. Factor loadings of 0.5 or above provide evidence of good item-scale correlation and of the item representing that particular factor. Items with factor loadings below 0.3 are considered unrelated to the scale due to the low correlation with the respective component (Crocker and Algina, 1986), and were removed from each scale in this analysis.

Third, the reliability was estimated for each potential scale. In this dissertation, Cronbach's alpha reliability statistic was used to estimate the degree of internal consistency among the items making up each scale. Following guidelines set out by Nunnally (1978), a Cronbach's alpha coefficient of 0.7 was used as a minimum reliability threshold. However, Nunnally (1978) suggests that the criteria used for reliability depends on the context and purpose for the research. Because of the exploratory nature of this dissertation research and the fact that no decisions will be made about individuals based on results, this threshold was not strictly met in all cases and some scales were created based on the theoretical importance of including particular constructs.

In addition to computing Cronbach's alpha based on all of the items for each scale, the Cronbach's alpha for the scale if an item were to be deleted from the scale was computed for each potential item in a set. If the overall Cronbach's alpha level of the scale increased if the item was deleted, this gave an indication that the item was not contributing to the reliability of the scale. Items that increased the Cronbach's alpha statistic if deleted were removed from the scale.

Finally, scales were evaluated in relation to STEM achievement. Scale scores that distinguish between high and low STEM achievement provide useful information to policymakers about differences in school effectiveness factors. The strength of the association between each scale and STEM achievement was measured by the Pearson product-moment correlation coefficient. In addition, scale categories were created to aid in scale interpretation by categorizing students with different levels of the construct being measured (e.g., low, medium, or high). Average achievement of students in each category was then examined. In general, scales where students in the high category had higher STEM achievement than students in the lower categories were preferred and were retained for further analyses.

### *Scaling the Student and School Context Items*

Once a proposed set of items was determined to be appropriate for scaling, the scale was created using 1-Parameter Item Response Theory (Rasch) Scaling. Previous research has indicated that Rasch scaling is appropriate for analyzing TIMSS 2007 questionnaire data (Preuschoff, 2010). Specifically, the Partial Credit Model was used because the TIMSS 2007 background questionnaire items have different numbers of response categories (between 2 and 5) and this model allows the number of response categories to vary across items on a scale (Masters, 1982).

The Rasch scaling analysis was conducted using the ConQuest 2.0 Generalized Item Response Modeling Software (Wu, Adams, Wilson, & Haldane, 2007). The scaling analysis was conducted in three steps. First, each component item was recoded into ordered categories so that the lowest category was 0. Second, item parameters were calibrated and the scale examined to see if it was functioning properly. Finally, individual student scores to be used in further multilevel analysis were estimated.

In calibrating the item parameters, the data from all five countries were combined so that each country's responses contributed equally to the parameter estimation and the resulting item parameter estimates were standard and comparable across countries (Masters & Wright, 1997). For each scale item, cumulative threshold parameters were estimated. Thresholds are points that separate the underlying continuum (the scale) into ordered categories (Andrich, 2010). The threshold parameter is the point on the scale at which respondents have a 50 percent probability of scoring in that category or higher, and a less than 50 percent probability of scoring below that category. Cumulative threshold parameters provide a common basis for including polytomous and dichotomous item types in a Rasch scale.

Typically, interpretation of Rasch scale results is facilitated by graphical displays, called Item or Wright Maps (Wilson, 2005). Item maps relate specific item responses to locations on a scale, with items representing high levels of a construct shown on the upper end and items representing low levels of a construct on the lower end. Students and items are displayed on the same scale so that item difficulty (threshold parameters) may be displayed in conjunction with the distribution of respondent scale scores.

The item map for each scale was examined to ensure that the item difficulty estimates (response thresholds) and person ability estimates (scale scores) were appropriately matched (Bond & Fox, 2007). Item thresholds should be located on that part of the scale covered by the scale score distribution, indicating that item difficulties are approximately equivalent to respondents' abilities (i.e., items are not all too easy or not all too difficult). Item thresholds measuring low levels of the construct will be located lower on the scale compared to item thresholds measuring high levels of the construct.

Each Rasch scale also was examined for item fit. Mean square item fit statistics are the mean of the squared residuals for that item and represent consistency of responses over the entire set of persons for each item (Bond & Fox, 2007). Mean square statistics may be weighted or unweighted. Weighted fit statistics (INFIT) give more weight to respondents whose response (i.e., ability) is closer to the item difficulty, providing more information about that item. Unweighted fit statistics (OUTFIT) are just that—unweighted, so they are more influenced by outlying scores. For this reason, INFIT statistics are usually preferred to OUTFIT statistics for the purposes of examining fit. Large INFIT values (values greater than 1.33) indicate that results are not consistent with model expectations and that a relatively large number of unexpected responses have occurred for that item (Adams and Khoo, 1996). This implies poor item fit for the scale. Items with INFIT values greater than 1.33 were examined for each scale.

Once item threshold parameters were determined to be functioning properly, scale scores were estimated for each student in each country using ConQuest's weighted maximum likelihood procedure. The Rasch student scale scores derived from the Rasch scaling are in logits (or a log odds unit) and are constrained to have a mean of zero and a standard deviation of 1 (Bond & Fox, 2007). Consistent with the approach used in IEA's CIVED study and adopted by TIMSS, the logit scores were converted into an international metric, with a mean of 10 and a standard deviation of 2. The student scores were then used as predictor variables in the subsequent multilevel modeling analyses.

When scaling sets of items was not appropriate, items were combined into composite variables with different categories. This was done in cases where the final combined variable did not have a clear theoretical continuum with one end representing high levels of the construct and one end representing low levels of the construct. For example, science teachers were asked nine separate questions about their major or main area of study in post-secondary education (biology, physics, chemistry, earth science, education-science, mathematics, education-mathematics, education-general, or other). However, since there was not one clear category that could be described as high on the continuum of teachers' major, these items were combined into one variable with two categories—1) major in science or science education and 2) no major in science or science education.

## Modeling School Effectiveness in Mathematics and Science

A series of multilevel models were created to investigate the relationships between school context factors and STEM achievement in each of the five countries studied, after controlling for student home resources. This investigation was designed to explore how school effectiveness factors contributing to student STEM achievement operate in the United States compared to Chinese Taipei, Singapore, the Czech Republic, and Slovenia. Results provide information about how school factors related to success in STEM achievement operate in some of the TIMSS 2007 top-performing countries.

Because the TIMSS 2007 students were clustered within schools and then in classes according to the sampling design, they often share the same learning contexts (e.g., they have the same teachers and classroom resources). Participants in TIMSS 2007 were sampled as part of intact classrooms and these classrooms were nested within schools. Mostly, for the participants in this dissertation research, two classrooms were chosen from each school to be representative of the student population. However, in some cases, only one classroom was sampled (see Table 3.1). Therefore, for this dissertation analysis, the classroom level could not be explicitly modeled, but rather was combined with the school level. That is, variance was not separated at the class level (between classrooms within schools), but only at the school level (between schools). The nesting of students in classrooms therefore, is synonymous with the nesting of students in schools for the purposes of this dissertation research.

To account for the multilevel structure of the nested data, this dissertation used the HLM 6 software program, which accommodates analyses with multilevel datasets

(Raudenbush, Bryk, & Congdon, 2004). This program is designed to carry out computations with five plausible values to effectively analyze achievement differences in the TIMSS 2007 dataset and can use the sampling weight variables of the dataset. Each statistic related to TIMSS achievement was estimated five times, for each of the five composite STEM achievement estimates (plausible values). Then the five estimates were averaged.

A series of multilevel models for each of the five countries examined how STEM achievement differs between schools after controlling for students' home resources. Following the overarching conceptual model of this research (presented in Figure 1.1), successive multilevel models were built for each country to examine the effects of school resources, fidelity of curriculum implementation, and school climate in relation to average STEM achievement. When the variance accounted for is low, school factors in the model do not contribute much to the explanation of students' STEM achievement differences (Rumberger & Palardy, 2004). In contrast, when the variance accounted for is high, school factors do contribute to the explanation of differences in students' STEM scores.

The difference between students' achievement scores (the "variance") accounted for by the factors in each model was compared across countries. Hanushek and Woessmann (2010) note that examining results across countries provides important information about whether any particular relationship found is specific to one country context or may be generalized to a broader context. The percent of variance accounted for was compared across countries to determine how important each school effectiveness factor was in different country contexts.

### Between School Variance in STEM Achievement

The multilevel approach to the analysis of school effectiveness allows the study of the impact of school factors while controlling for student factors. To determine the impact of school factors, the proportion of variance in student STEM achievement that is attributable to differences between schools must first be determined. This constitutes the upper limit on what can be explained by differences in school factors. When the difference between schools is small, there is less scope for effective school factors to explain such differences.

The proportion of the total variance in student achievement scores that is attributable to differences between schools can be determined by using what is known as the *unconditional model*. In this dissertation, the *unconditional model* had two levels—the school level (level-2) and the student-within-school level, also known as the student level (level-1). This model does not include any predictors at the student- or school-level, but simply allows average school STEM achievement (the “intercept”) to vary across schools.

Equation 3.1 presents the student-level and Equation 3.2 the school-level components of the *unconditional model*. The full “mixed model” used to represent both levels simultaneously is presented in Equation 3.3.

Student-level: $$STEM1 - STEM5 = \beta_0 + r_{ij} \quad (3.1)$$

School-level: $$\beta_0 = \gamma_{00} + u_{0j} \quad (3.2)$$

Mixed Model: $$STEM1 - STEM5 = \gamma_{00} + u_{0j} + r_{ij} \quad (3.3)$$

Where:

$STEM1 - STEM5$ = 5 STEM composite measures for student *i* in school *j*

$\beta_0$ = Mean STEM score for school *j*

$\gamma_{00}$ = Grand mean STEM score across all schools

$r_{ij}$ = Random error associated with student *i* in school *j*

$u_{0j}$ = Random error associated with school *j*

The *unconditional model* provides decomposition of the total achievement variance into two components: between-school and within-school. Between-school variance is the variation in student achievement that exists between schools, due to school-level characteristics and within-school variance is the variation that exists within schools, due to individual student characteristics. Variables measured at the school level, such as school resources cannot be used to explain variance directly at the student level because such variables are constant for all students within a school.

The *unconditional model* also provides information used to calculate the proportion of total variability that exists between schools without considering any explanatory variables (also known as the Intraclass Correlation Coefficient, or ICC). This is an estimate of the amount of nesting in the data and is calculated by dividing the variance that is between schools by the total variance. Statistically, this is represented as follows: $\hat{\rho} = \frac{\hat{\tau}_{00}}{(\hat{\tau}_{00}+\hat{\sigma}^2)}$, where $\tau_{00}$=between school variability and $\sigma^2$=within school variability (Raudenbush & Bryk, 2002). The proportion of variability that exists between schools from the *unconditional model* provides an indication of whether or not multilevel modeling is appropriate to use and represents the maximum amount of variance in STEM achievement between schools that can be explained by any predictor variables.

### *Modeling Student Achievement Within Schools*

When analyzing data using multilevel modeling, it is important to determine how the relationship within schools between student home background and achievement will be modeled. Typically, this will be represented by a straight line, i.e., a slope and intercept, for each school. If the slope is fixed, the relationship between student home background and achievement is assumed to be the same in all schools (i.e., has homogeneity of covariance). If the student-level slope is fixed in the model even though the slopes are actually heterogeneous (i.e., the relationship between student home background and STEM achievement actually varies across schools) this could bias the school-level estimates (Bickel, 2007). Raudenbush and Bryk (2002) recommend allowing the slope of a student-level coefficient to vary randomly unless the reliability is below 0.05. Therefore, in this dissertation research, given that the reliability of the slope was greater than 0.05, the slope for the student home resources control was allowed to vary randomly in the multilevel models for each country, using what is known as the random coefficients regression model.

Another analysis procedure used in multilevel models involves centering predictor variables (Raudenbush & Bryk, 2002). In mean centering, raw variables are transformed by subtracting them from the group mean or grand mean, which often improves the interpretability of coefficients and reduces multicollinearity and standard errors (Bickel, 2007). Grand mean centering is often recommended by researchers over group mean centering because it is more parsimonious (due to the smaller number of predictors required) and provides more precise mean estimates, especially when the size of the student-level sample (i.e., number of students in a school) is small (Paccagnella, 2006). Thus, all of the scales in this dissertation were entered into the multilevel analyses grand mean centered. Categorical variables (i.e., composite variables) were dummy coded so that the regression coefficients could be interpreted in terms of group comparisons.

### *Variance Explained by Multilevel Models*

As previously described, multilevel modeling allows for the decomposition of variance by level. In this dissertation research, residual variance is estimated between schools ($\tau_{00}$=between school variability) and within schools ($\sigma^2$=within school variability). These two variance components can be used to calculate 1) the proportion of student within school variance explained by the model, 2) the proportion of between school variance explained by the model, 3) the proportion of total variance in STEM achievement explained by the model, and 4) the proportion of total variance in STEM achievement due to differences between schools after considering the predictor variables in the model.

First, the amount of variance explained by student-level predictors in a multilevel model can provide information about the importance of the factors in the model for explaining differences in STEM achievement within schools. The proportion of student within school variance explained by the factors modeled may be calculated using the formula shown in Equation 3.4 (Raudenbush & Bryk, 2002). In this dissertation, only "Student Home Resources" is modeled at the student within school level. Thus, examining the proportion of student within school variance explained by the models in this dissertation indicates what proportion of the total variance in STEM achievement that is within schools is accounted for by student home resources.

$$\frac{\hat{\sigma}^2_{Uncond} - \hat{\sigma}^2_{Residual}}{\hat{\sigma}^2_{Uncond}} \tag{3.4}$$

Where:

$\hat{\sigma}^2_{Uncond}$ = The student-level residual variation from the unconditional model

$\hat{\sigma}^2_{Residual}$ = The student-level residual variation from the model with predictor variables

Second, for school-level factors, the proportion of between school variance explained by the factors modeled may be calculated using the formula shown in Equation 3.5 (Raudenbush & Bryk, 2002). Results indicate how much of the total variance in STEM achievement that is between schools can be explained by school-level factors.

$$\frac{\hat{\tau}_{Uncond} - \hat{\tau}_{Residual}}{\hat{\tau}_{Uncond}} \tag{3.5}$$

Where:

$\hat{\tau}_{Uncond}$ = The school-level residual variation from the unconditional model

$\hat{\tau}_{Residual}$ = The school-level residual variation from the model with predictor variables

Third, the residual variance components from both the student- and school-levels may be used to calculate the proportion of total variance in STEM achievement explained by all predictors in a multilevel model (see Equation 3.6) (Raudenbush & Bryk, 2002). The proportion of total variance explained gives information about how much of the total variance in STEM achievement (within and between schools) is explained by all predictors (student- and school-level). For this dissertation analysis, the predictors include school-level factors as well as the within-school effect of student home resources.

$$\frac{(\hat{\tau}_{00Uncond}+\hat{\sigma}^2_{Uncond})-(\hat{\tau}_{00Residual}+\hat{\tau}_{01Residual}+\hat{\sigma}^2_{Residual})}{(\hat{\tau}_{00Uncond}+\hat{\sigma}^2_{Uncond})} \quad (3.6)$$

Where:

$\hat{\tau}_{00Uncond}$ = The school-level residual variation from the unconditional model

$\hat{\sigma}^2_{Uncond}$ = The student-level residual variation from the unconditional model

$\hat{\tau}_{00Residual}$ = The school-level residual variation from the model with predictor variables

$\hat{\tau}_{01Residual}$ = The school-level residual variation associated with the student home resources slope from the model with predictor variables

$\hat{\sigma}^2_{Residual}$ = The student-level residual variation from the model with predictor variables

Finally, when predictors are added to the *unconditional model*, residual variance from the model provides information that may be used to calculate the proportion of total variance in STEM achievement due to differences between schools after considering the predictor variables in the model. For a random coefficients (i.e., varying slopes) regression model, variance in STEM achievement due to differences between schools is computed using Equation 3.7, which takes into account the variation due to the random slope associated with the student home resources control.

$$\hat{\rho}_{Cond} = \frac{\hat{\tau}_{00}}{(\hat{\tau}_{00}+\hat{\tau}_{01}+\hat{\sigma}^2)} \tag{3.7}$$

Where:

$\hat{\tau}_{00}$= The residual variation associated with the school-level (between school variability)

$\hat{\tau}_{01}$= The residual variation associated with the student home resources slope

$\hat{\sigma}^2$ = The residual variation associated with the student-level (within school variability)

The proportion of total variance in STEM achievement due to differences between schools after considering predictor variables in the model can be compared to the variation in STEM achievement due to differences between schools from the *unconditional model* to determine the change in proportion of variance due to the factors modeled. Ideally, the proportion of variance attributable to differences between schools would decrease to zero with the addition of the most explanatory factors (Bickel, 2007). In this dissertation, all of the models examining school-level factors use a model controlling for student home resources at the student level as their base (the *student home resources model*). Thus, the total variance between schools after controlling for student home resources in the *student home resources model* was used as a reference to see how much more variation in STEM achievement due to differences between schools after controlling for student home resources was accounted for by the addition of school-level factors.

### *Student Home Resources*

As previously discussed, this dissertation approaches school effectiveness research from the perspective that an effective school has the capacity to improve student achievement over and above what might be expected given the characteristics of the entering student body. That is, the characteristics students have when entering school have been shown to have a strong relationship with achievement and should be explicitly modeled in order to better isolate the effects of a school (Scheerens & Bosker, 1997).

In the context of schooling, student home resources can act to influence student achievement in two ways—1) through advantages accruing to the individual student as a result of an advantageous home background, and 2) through the aggregate effect of attending a school where many other students also have advantageous home backgrounds. Accordingly, student home resources were represented in two ways—1) individual student home resources, representing the characteristics of each student, and 2) school average home resources, representing the composition of the student body in a school. To best examine manipulable school factors in isolation, a model building strategy was adopted to account for the two aspects of student home resources at different points in the model building process.

The individual student home resources factor was controlled for at the beginning of the model building process so that school factors could be examined in isolation of student factors. The school average home resources aspect was considered at the end of the model building process so that school effectiveness factors could first be examined in relation to STEM achievement setting aside considerations of school differences related to the composition of the student body. After examining the effects of the school effectiveness factors, a measure of the school average home resources was added to the final model in each country so that how school effectiveness factors work in each country's particular context could be examined.

To begin with, a *student home resources model* was built to control statistically variance related to individual student differences. Specifically, to capture the contribution of individual student home resources to school achievement differences, a student–level model was built upon the *unconditional model* for each country. The student home resources control variables (intercept and slope) were added to each country's *unconditional model* to explore the extent to which controlling student home resources diminished differences in students' STEM achievement. The grand mean STEM achievement in each country was adjusted for differences in student home resources so that school-level factors could be examined in relation to the adjusted school STEM achievement. This analysis reveals the extent to which differences in STEM achievement are attributable to student home resources, before taking school factors into account. Results from the *student home resources model* also provides information about the residual variance components from the model, indicating how well the model explains variation in STEM achievement.

The statistical *student home resources models* are presented in Equation 3.8 (student-level), Equation 3.9 (school-level), and Equation 3.10 (mixed model), respectively.

Student-level: $STEM1 - STEM5 = \beta_0 + \beta_{1j}(HOME_{ij} - \overline{HOME}..) + r_{ij}$ (3.8)

School-level: $\beta_0 = \gamma_{00} + u_{0j}$ (3.9)

$\beta_1 = \gamma_{10} + u_{1j}$

Mixed Model: (3.10)

$$STEM1 - STEM5 = \gamma_{00} + \gamma_{10}\left(HOME_{ij} - \overline{HOME}..\right) + u_{0j} + u_{1j}\left(HOME_{ij} - \overline{HOME}..\right) + r_{ij}$$

Where:

$\gamma_{00}$= The average of the school means on STEM achievement across the population of schools

$y_{10}$= The average student home resources control slope across those schools (grand mean centered)

$u_{0j}$= The unique increment to the intercept associated with school *j*

$u_{1j}$= The unique increment to the slope associated with school *j*

### School Resources Models

Once the variance related to individual student home resources was controlled statistically through the *student home resources* model for each country, a series of models were built to examine the additive effects of school-level factors for explaining differences in STEM achievement.

To begin with, the answer to research question 1, "How do school resources (well-prepared teachers, general school resources, and resources specific to mathematics and science instruction) add to the explanation of STEM school effectiveness in each country?" was explored through a series of *school resources models*. Two separate models were built to examine 1) **Teacher Preparation** and 2) **General School Resources** as well as **Mathematics and Science Specific School Resources**. Both models were built upon the *student home resources model* for each country in order to better isolate the relationships between school factors and student STEM achievement.

First, a model was built to examine how **Teacher Preparation** adds to the explanation of school effectiveness. How well-prepared a school is at the beginning of an academic year to provide STEM education to students begins with how well-prepared its teachers are to teach the STEM topics. The literature reviewed in Chapter 2 suggests that there is a relationship between teachers having higher education levels and majors in the subject area they teach and student achievement. Therefore, the teacher level measures associated with the factor **Teacher Preparation** (identified in Table 3.2): "Mathematics Teachers' Level of Education and Major Area of Study" and "Science Teachers' Level of Education and Major Area of Study" were entered into a *teacher preparation model*.

Because the measures representing **Teacher Preparation** were categorical variables, they were dummy coded to be entered into the multilevel analyses. As previously mentioned, the classroom level was not explicitly modeled in this research, so classroom- and student-level data were aggregated to the school level for multilevel analyses. To maintain the interpretability of each category of the **Teacher Preparation** measures (e.g., the first dummy variable represented teachers with a postgraduate degree and major in mathematics or mathematics education), one teacher's data was randomly sampled from each school. While the teacher data could have been averaged across a school to give an indication of the proportion of well-qualified teachers in a school, this method was chosen to maintain interpretability of results. That is, maintaining the dummy coded variables allowed interpretation of the relationship with STEM achievement for each category of the variable (e.g., the effect of students having teachers with a postgraduate degree and major in mathematics compared to the effect of students having teachers without a major in mathematics).

Results from the *teacher preparation model* provide information about how teachers' education level and major area of study adds to the explanation of school effectiveness. The statistical form of the *teacher preparation model* includes the student home resources control at the student level and the teacher measures at the school level. Equation 3.11 presents the student-level model, which is equivalent to Equation 3.8. Equation 3.12 presents the school-level model, including the teacher preparation measures. The combined *teacher preparation model* that analyzes the data simultaneously at two levels is presented in Equation 3.13.

Student-level: $STEM1 - STEM5 = \beta_0 + \beta_{1j}(HOME_{ij} - \overline{HOME}_{\cdot\cdot}) + r_{ij}$ (3.11)

School-level: $\beta_0 = \gamma_{00} + \gamma_{01}(MathTeacher_j) + \gamma_{02}(ScienceTeacher_j) + u_{0j}$ (3.12)

$\beta_1 = \gamma_{10} + u_{1j}$

Mixed Model: (3.13)

$$STEM1 - STEM5 =$$

$$\gamma_{00} + \gamma_{01}(MathTeacher_j) + \gamma_{02}(ScienceTeacher_j) + \gamma_{10}(HOME_{ij} - \overline{HOME}_{\cdot\cdot}) +$$

$$u_{0j} + u_{1j}(HOME_{ij} - \overline{HOME}_{\cdot\cdot}) + r_{ij}$$

Where:

$y_{01}$ and $y_{02}$ = Regression coefficients associated with teacher preparation measures

A second school resources model was built to investigate how **General School Resources** as well as **Mathematics and Science Specific School Resources** add to the explanation of school effectiveness. When examining variance explained in multilevel models, the proportion of variance explained by any particular variable (or set of variables when entered simultaneously) is partially dependent on its relationship with other variables in the model (Pedhazur, 1997). Therefore, the *general and STEM resources model* was built separately from the *teacher preparation model* so that variation in student STEM scores due to **General School Resources** and **Mathematics and Science Specific School Resources** could be examined independently of variation related to **Teacher Preparation**.

The literature reviewed in Chapter 2 indicates that **General School Resources** as well as **Mathematics and Science Specific School Resources**, such as a science laboratory are associated with higher STEM achievement. Therefore, these two factors were examined together since the literature also indicates general resources are often a prerequisite for having mathematics and science resources. That is, **Mathematics and Science Specific Resources** provide information about the STEM resources a school has above and beyond **General School Resources**.

Two scales associated with the factors **General School Resources** and **Mathematics and Science Specific School Resources** were examined: “Instruction Affected by General Resource Shortages” and “Instruction Affected by Shortages in

Resources Specific to teaching Mathematics and Science." These two scales were entered into one model to determine how the combination of these resource measures adds to the explanation of school effectiveness.

The statistical form of the *general and STEM resources model* includes the student home resources control at the student level and the two resource shortage scales at the school level. The mixed model of the *general and STEM resources model* would, therefore, have an equation similar to Equation 3.13, but including measures representing **General School Resources** and **Mathematics and Science Specific School Resources** instead of measures representing **Teacher Preparation**.

### *Fidelity of Curriculum Implementation Models*

The answer to research question 2—"How does fidelity of curriculum implementation (fidelity to curriculum structure, fidelity to curriculum process, and participant responsiveness) add to the explanation of STEM school effectiveness in each country?" was explored through a series of *fidelity of curriculum implementation models*. Three successive *fidelity of curriculum implementation models* were built to examine the additive effects of **Fidelity to Curriculum Structure**, **Fidelity to Curriculum Process**, and **Participant Responsiveness** in each of the five countries.

Instruction begins with the content a teacher covers. The literature in Chapter 2 suggests that teachers who teach the intended curriculum with fidelity to structure (i.e., teach the topics in the curriculum) are associated with students who have higher STEM achievement. Therefore, the scale "Mathematics Topics Taught," representing the number of TIMSS mathematics topics a teacher has taught by eighth grade measuring

**Fidelity to Curriculum Structure** was entered into the first *fidelity of curriculum implementation model* (Model 1). This model provides information about how adherence to the curriculum adds to the explanation of school effectiveness.

Aside from content covered, the quality of the strategies teachers use to teach the curriculum to students in the class also is an important factor related to student achievement according to the literature reviewed in Chapter 2. Therefore, a second *fidelity of curriculum implementation model* (Model 2) was built adding scales representing the factor **Fidelity to Curriculum Process**: "Effective Instructional Strategies for Encouraging Mathematics Reasoning" and "Effective Instructional Strategies for Teaching Science Inquiry." The instructional strategies scales were included in addition to "Mathematics Topics Taught" because instructional strategies cannot be conducted in isolation of the content covered. Including the instructional strategies scales in a second model provides information about how much more variation in STEM achievement could be explained by the strategies teachers use in addition to curriculum content coverage.

**Participant Responsiveness** is another important factor relating to classroom instruction, represented in this research by student attitudes toward learning. Student attitudes toward learning aggregated to the school level provide information about instructional processes since they may be considered a report from the student's perspective on the effectiveness of a teacher's instruction. The literature reviewed in Chapter 2 suggests that a positive student attitude toward learning STEM subjects is associated with higher STEM achievement. Therefore, students may engage in instructional activities to a greater or lesser extent depending on their attitudes toward the subject taught. In order to examine the relationship between all fidelity of curriculum implementation factors together, scales representing **Participant Responsiveness**, "Students Like Learning Mathematics" and "Students Like Learning Science" were entered into a third *fidelity of curriculum implementation model* (Model 3) in addition to all of the other fidelity of curriculum implementation scales. Results provide information on how much more variation in STEM achievement can be explained by student attitudes toward learning in addition to curriculum content coverage and instructional strategies.

The statistical form of the complete *fidelity of curriculum implementation model* (Model 3) includes the student home resources control at the student level and the mathematics topics taught, effective instructional strategies, and student attitudes toward learning scales at the school level. The statistical form of the mixed model of the *fidelity of curriculum implementation model* is shown in Equation 3.14.

Mixed Model: (3.14)

$$STEM1 - STEM5 =$$

$$\gamma_{00} + \gamma_{01}(TopicsTaught_j - \overline{TopicsTaught_{..}}) + \gamma_{02}(InstructionalStrategies_j - \overline{InstructionalStrategies_{..}}) + \gamma_{03}(Attitudes_j - \overline{Attitudes_{..}}) + \gamma_{10}(HOME_j - \overline{HOME_{..}}) + u_{0j} + u_{1j}(HOME_j - \overline{HOME_{..}}) + r_{ij}$$

Where:

$y_{01} - y_{02}$= Regression coefficients associated with fidelity of curriculum implementation measures (grand mean centered)

### *School Climate Models*

The answer to research question 3—"How does school climate (discipline and attendance problems and support for academic success) add to the explanation of STEM school effectiveness in each country?" was explored through two *school climate models*. These two models were built to examine the effect of **Discipline and Attendance Problems** (i.e., negative aspects of school climate) and **School Support for Academic Success** (i.e., positive aspects of school climate). The first model examined the negative aspects of school climate in isolation through the scale "Schools Have Few Discipline and Attendance Problems." This model represents the problems a school encounters when providing instruction (e.g., bullying, student absenteeism, etc.). The second model included a scale representing the positive aspects of schools climate, "Schools Have a Climate Supportive of Academic Success" to see how positive support from school leaders and parents could mediate problems in a school.

The amount of variance explained in the final *school climate model* with the addition of the scale measuring the positive aspects of school climate was examined compared to the amount of variance explained by the first model (negative aspects only). The difference in variance explained gives an indication of how much more variation in STEM achievement positive school climate factors explain after discipline and attendance problems are accounted for. The statistical form of the mixed model of the final *school climate model* is shown in Equation 3.15.

Mixed Model: (3.15)

$$STEM1 - STEM5 = \gamma_{00} + \gamma_{01}\left(NegativeClimate_j - \overline{NegativeClimate_{..}}\right) + \gamma_{02}\left(PositiveClimate_j - \overline{PositiveClimate_{..}}\right) + \gamma_{10}\left(HOME_j - \overline{HOME_{..}}\right) + u_{0j} + u_{1j}\left(HOME_j - \overline{HOME_{..}}\right) + r_{ij}$$

Where:

$y_{01}$ and $y_{02}$= Regression coefficients associated with school climate factors (grand mean centered)

### *School Resources, Fidelity of Curriculum Implementation, and School Climate Models*

In actuality, the schooling process is a complex dynamic system in which all student background and school context factors act interdependently in influencing student outcomes. That is, these factors do not operate in isolation, but rather simultaneously in an education system. To answer research question 4, “How does the combination of school resources, fidelity of curriculum implementation, and school climate add to the explanation of STEM school effectiveness in each country?” three additive models were built to determine how factors in these three areas operate together to explain variation in STEM achievement after taking into account student home resources. That is, these models explored what level of resources a school had to deliver instruction, how the school used these resources to deliver instruction, and in what atmosphere or climate the school delivered instruction.

First, the *combined school resources model* was built to explore how teacher quality might be mediated by material resources in a school. For example, a teacher might be highly qualified to teach STEM subjects, but have no access to textbooks, calculators, or even adequate instructional space. In that case, material resource shortages could actually depress the association between teacher quality and student STEM achievement in a school. On the other hand, while **Teacher Preparation** may lead to higher STEM achievement without **General School Resources** and **Mathematics and Science Specific School Resources**, having material resources will not lead to STEM achievement without a qualified teacher to use them. These hypotheses were explored in the *combined school resources model*, which included both **Teacher Preparation** measures and **General School Resources** and **Mathematics and Science Specific School Resources** measures. Overall results from the *combined school resources model* show the extent to which school differences in STEM achievement are related to the complete set of school resources factors, after controlling for student home resources.

The *school resources combined model* was then used as a base upon which to build the remaining models. In the next model, the *school resources and curriculum implementation model*, all the school resources factors and all the factors representing fidelity of curriculum implementation were combined to provide information about how much variation in STEM achievement fidelity of curriculum implementation could explain above and beyond school resources. This gives an indication of how much teachers adhering to the curriculum content and using effective instructional strategies and students liking learning STEM subjects can add to a school's effectiveness beyond having human (teacher) and material resources.

A third model, *school resources, fidelity of curriculum implementation, and school climate* was then built, adding school climate factors to school resources and curriculum implementation factors to give an indication of how the academic atmosphere of a school adds to the explanation of STEM achievement.

### *Complete School Effectiveness Model*

Finally, all school effectiveness factors together with student home resources averaged at the school level were used to create the *complete school effectiveness model*. Controlling for student home resources at the individual student level within school is only one part of the complex relationship between student characteristics and student achievement. The other important piece of this relationship is the difference between schools in the social composition of the student body, which is reflected in the "School Average Student Home Resources" measure. For a complete picture, the magnitude of school effectiveness factors were examined once "School Average Student Home Resources" was taken into account, in addition to the individual effect of student home resources. Results from the *complete school effectiveness model* provide a picture of how the combined effectiveness factors add to the explanation of STEM school effectiveness, after accounting for both components of student home resources.

The mixed model for the *complete school effectiveness model* is presented in Equation 3.16.

Mixed Model: (3.16)

$$STEM1 - STEM5 = \gamma_{00} + \gamma_{01}(MathTeacher_j) + \gamma_{02}(ScienceTeacher_j) + \gamma_{03}(GeneralResources_j - \overline{GeneralResources_{..}}) + \gamma_{04}(STEMResources_j - \overline{STEMResources_{..}}) + \gamma_{05}(TopicsTaught_j - \overline{TopicsTaught_{..}}) + +\gamma_{06}(InstructionalStrategies_j - \overline{InstructionalStrategies_{..}}) + \gamma_{07}(Attitudes_j - \overline{Attitudes_{..}}) + \gamma_{08}(NegativeClimate_j - \overline{NegativeClimate_{..}}) + \gamma_{09}(PositiveClimate_j - \overline{PositiveClimate_{..}}) + \gamma_{010}(AvgHOME_j - \overline{AvgHOME_{..}}) + \gamma_{10}(HOME_j - \overline{HOME_{..}}) + +u_{0j} + u_{1j}(HOME_j - HOME) + r_{ij}$$

Where:

$y_{01}$– $y_{010}$= Regression coefficients associated with school effectiveness factors

Building each of the multilevel models separately for the five countries examined in this dissertation allowed for an investigation of the school effectiveness factors involved in within-country differences between schools. That is, for each country context, the variance in STEM achievement accounted for by each model may be compared. For example, the *school resources combined model* might account for 10 percent of the variation in student STEM achievement in Singapore, but the *school climate model* may account for 25 percent of the variation.

# Chapter 4

## *School Effectiveness Factors Related to STEM Achievement*

### Introduction

This chapter describes how school resources, fidelity of curriculum implementation, and school climate factors related to high mathematics and science achievement operated in five TIMSS 2007 countries—the United States, Chinese Taipei, Singapore, the Czech Republic, and Slovenia. First, measures of the school effectiveness factors were constructed using Rasch scaling, so the first section of the chapter presents the results of this activity, listing the scales created, their source descriptions, and their relationship to achievement. Second, and primarily, multilevel modeling was used in each of the five countries under study to investigate the relationship between the school effectiveness factors and STEM achievement, after controlling for student home resources. The majority of this chapter describes the results of the multilevel analyses, providing information about the school factors associated with effective schools in STEM areas.

### Measure of STEM Achievement

The STEM achievement outcome measure used in this dissertation research is a composite of students' TIMSS 2007 mathematics and science achievement scores. Considering the breadth of mathematics and science topics covered in the *TIMSS 2007*

*Assessment Frameworks*, this composite is a relatively comprehensive indicator of achievement in STEM areas. Table 4.1 presents the average STEM achievement and corresponding standard error for each country in this study. As expected based on countries' individual mathematics and science achievement scores in TIMSS 2007, Chinese Taipei and Singapore had the highest STEM achievement—approximately 1160, followed by the Czech Republic and Slovenia, with average scores of approximately 1040, and then the United States with an average STEM score of 1028.

Table 4.1. Descriptive Statistics for STEM Achievement

| Country | Average STEM Achievement | Standard Error |
|---|---|---|
| Chinese Taipei | 1159 | 8.1 |
| Czech Republic | 1043 | 4.1 |
| Singapore | 1160 | 8.2 |
| Slovenia | 1039 | 4.2 |
| United States | 1028 | 5.6 |

## Developing Measures of School Effectiveness Factors

Table 4.2 lists the Rasch scales developed for the multilevel analyses, including the student home resources control and the measures of the school factors. As described in Chapter 3, classifying the TIMSS 2007 eighth grade questionnaire items according to school resources, fidelity of curriculum implementation, and school climate resulted in the eight factors shown in Table 4.2. Prior to the multilevel analysis, the individual items representing each of the eight factors were combined into scales (or composite variables). Of the 11 school effectiveness measures, all except two are Rasch scales, which typically provide more reliable measures of constructs than individual items.

Table 4.2. Measures of School Effectiveness Included in the Multilevel Analyses*

| | |
|---|---|
| Student-Level Home Resources Control | |
| | Student Home Resources Scale |
| School Resources Factors | |
| **Teacher Preparation** | |
| | Mathematics Teachers' Level of Education and Major Area of Study |
| | Science Teachers' Level of Education and Major Area of Study |
| **General School Resources** | |
| | Instruction Affected by General Resource Shortages Scale |
| **Mathematics and Science Specific School Resources** | |
| | Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science Scale |
| Fidelity of Curriculum Implementation Factors | |
| **Fidelity to Curriculum Structure** | |
| | Mathematics Topics Taught Scale |
| **Fidelity to Curriculum Process** | |
| | Effective Instructional Strategies for Encouraging Mathematics Reasoning Scale |
| | Effective Instructional Strategies for Teaching Science Inquiry Scale |
| **Participant Responsiveness** | |
| | Students Like Learning Mathematics Scale |
| | Students Like Learning Science Scale |
| School Climate Factors | |
| **Discipline and Attendance Problems** | |
| | Schools Have Few Discipline and Attendance Problems Scale |
| **Support for Academic Success** | |
| | Schools Have a Climate Supportive of Academic Success Scale |

*Schools having science laboratories, yearly instructional hours spent on mathematics, and yearly instructional hours spent on science also were examined.

To determine suitability for inclusion in the multilevel modeling, the Rasch scales were evaluated according to a set of criteria as described in Chapter 3. For unidimensionality, it was preferred that a single factor accounted for at least 50 percent of the variance in the set of items with no second dominant factor, although 20 percent was a minimum (Reckase, 1979). Similarly, a Cronbach's alpha coefficient of 0.7 was used as a guideline for scale reliability (Nunnally, 1978), but several scales not meeting the criteria were created based on their theoretical importance. For item fit, items with INFIT values greater than 1.33 were examined for relatively large numbers of unexpected responses (Adams and Khoo, 1996). Also, the item map for each Rasch scale was examined to ensure that the item difficulty estimates (response thresholds) and person ability estimates (scale scores) were appropriately matched (Bond & Fox, 2007). Finally, scales having a positive correlation with STEM achievement or that provided greater distinctions between levels of STEM achievement were preferable to those that did not.

### *Student-Level Home Resources Control*

School effectiveness analyses include a student-level home resources control, which makes it possible to identify important school factors that affect student achievement above and beyond students' home resources; that is, to isolate the added value of schooling. Accordingly, the "Student Home Resources" scale was developed to control for student-level home resources in the multilevel analyses.

A principal components analysis of the TIMSS 2007 student home background items indicated that the number of books in the home, a composite of the number of home possessions, and the highest level of education of either parent constituted a single factor, accounting for 56 percent of the total variance. Cronbach's alpha coefficient of reliability for the "Student Home Resources" scale was relatively close to the reliability threshold of 0.7 ($\alpha$=0.60). As with all the Rasch scales developed for the multilevel analyses, the final "Student Home Resources" scale was transformed to have a mean of 10 and a standard deviation of 2.

In the Appendix, Table A.1 provides detailed information about the "Student Home Resources" scale, presenting each of the items together with its response categories. Data about home possessions included country-specific items added by individual countries to suit their national context, referred to as "national adaptations." Home possessions specific to each participating country's national context provided a more targeted measure of the level of resources in the home when added to possessions common across the international context. Table A.1 also includes information about the source of the data (i.e., data were collected from students).

Rasch scaling results were examined next. The INFIT statistics from the "Student Home Resources" scale ranged from 0.97–0.99, indicating good item fit. Figure A.1 presents the Rasch item map produced from the scaling analysis. Evaluating the item map showed that the "Student Home Resources" scale also had a good match between item threshold parameters and respondent scale scores since the distribution of respondent scores was on the part of the scale covered by the item thresholds. Moreover, item thresholds measuring low levels of the construct were located lower on the scale than thresholds measuring high levels of the construct, indicating the item threshold parameters were ordered as expected. For example, the item threshold representing students with 5 home possessions was lower on the scale than the item threshold representing students with 6 home possessions.

Table 4.3 presents the correlation between the "Student Home Resources" scale and average student STEM achievement in each country. The Pearson correlations for the five countries ranged from 0.35 to 0.51, indicating a positive relationship between the scale and STEM achievement. Table 4.3 also shows the percentages and corresponding average STEM achievement for three categories of eighth-grade students, those having "many resources," "some resources," and "few resources." Students assigned to the "many resources" category had a scale score equal to or higher than the scale point corresponding to students reporting they had more than 100 books in the home, at least 8 home possessions, and at least one parent who finished university or higher, on average. Students assigned to the "few resources" category had a scale score equal to or lower than the scale point corresponding to students reporting that on average they had 25 or fewer books in the home, no more than 6 home possessions, and neither parent going beyond upper-secondary education. All other students were assigned to the "some resources" category. For example, in Chinese Taipei, the 15 percent of students from homes with "many resources" averaged 237 points higher in STEM achievement than the 32 percent of students with "few resources."

Table 4.3. “Student Home Resources” Scale in Relation to STEM Achievement

| Country | Pearson Correlation | Many Resources | | Some Resources | | Few Resources | |
|---|---|---|---|---|---|---|---|
| | | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 0.50*** | 15 | 1274 (8.4) | 54 | 1200 (5.5) | 32 | 1037 (9.3) |
| Czech Republic | 0.39*** | 11 | 1143 (8.1) | 69 | 1048 (3.9) | 19 | 965 (4.7) |
| Singapore | 0.51*** | 11 | 1312 (7.9) | 55 | 1198 (7.0) | 34 | 1050 (11.6) |
| Slovenia | 0.35*** | 14 | 1116 (6.6) | 67 | 1046 (4.5) | 19 | 962 (7.6) |
| United States | 0.39*** | 27 | 1108 (5.0) | 57 | 1021 (5.0) | 17 | 929 (6.9) |

***p < .01
Note: Standard errors are shown in parentheses.

## *School Resources Factors*

The overarching area of school resources included three factors related to STEM achievement: **Teacher Preparation**, **General School Resources**, and **Mathematics and Science Specific School Resources**.

### Teacher Preparation

In TIMSS 2007 at the eighth grade, seven items were related to mathematics **Teacher Preparation** and ten items were related to science **Teacher Preparation**. These two sets of items were collapsed into two categorical variables with four categories each. For mathematics teachers, the categories were: 1) completed postgraduate degree and majored in mathematics or mathematics education, 2) completed bachelor's degree and majored in mathematics or mathematics education, 3) completed less than a bachelor's degree and majored in mathematics or mathematics education, and 4) did not major in mathematics or mathematics education. The science teacher preparation measure had the same categories, but specifying science or science education majors. Appendix Table A.2 provides further information about the teacher preparation measures.

Tables 4.4 and 4.5 show for mathematics and science teachers, respectively, the four categories of the teacher preparation measure together with the percentage of students that had teachers in each category. The percentages of students are consistent with teacher education policies in the five countries as explained in the *TIMSS 2007 Encyclopedia*.

Table 4.4. Mathematics Teachers' Level of Education and Major Area of Study in Relation to STEM Achievement

| Country | Completed Postgraduate Degree and Majored in Mathematics or Mathematics Education | | Completed Bachelor's Degree and Majored in Mathematics or Mathematics Education | | Completed Less than a Bachelor's Degree and Majored in Mathematics or Mathematics Education | | Did Not Major in Mathematics or Mathematics Education | |
|---|---|---|---|---|---|---|---|---|
| | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 19 | 1189 (17.0) | 64 | 1155 (8.3) | 2 | 1184 (33.0) | 16 | 1136 (20.7) |
| Czech Republic | 91 | 1047 (4.1) | 1 | ~ | 2 | 999 (31.9) | 7 | 1006 (15.5) |
| Singapore | 4 | 1163 (44.1) | 70 | 1165 (9.8) | 3 | 1129 (63.6) | 23 | 1143 (20.4) |
| Slovenia | 1 | ~ | 43 | 1038 (6.7) | 52 | 1038 (6.5) | 5 | 1043 (32.6) |
| United States | 41 | 1036 (8.4) | 28 | 1024 (13.4) | 0 | ~ | 31 | 1014 (10.9) |

Note: Standard errors are shown in parentheses.

Table 4.5. Science Teachers' Level of Education and Major Area of Study in Relation to STEM Achievement

| Country | Completed Postgraduate Degree and Majored in Science or Science Education | | Completed Bachelor's Degree and Majored in Science or Science Education | | Completed Less than a Bachelor's Degree and Majored in Science or Science Education | | Did Not Major in Science or Science Education | |
|---|---|---|---|---|---|---|---|---|
| | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 29 | 1174 (12.5) | 54 | 1166 (11.2) | 13 | 1151 (20.0) | 3 | 1028 (37.8) |
| Czech Republic | 88 | 1047 (4.3) | 1 | ~ | 2 | 1026 (13.4) | 10 | 1002 (9.9) |
| Singapore | 7 | 1245 (25.1) | 83 | 1179 (9.5) | 6 | 1144 (37.4) | 4 | 1187 (34.4) |
| Slovenia | 2 | 1078 (8.2) | 42 | 1040 (5.7) | 51 | 1036 (5.2) | 5 | 1042 (9.8) |
| United States | 47 | 1031 (2.7) | 27 | 1024 (2.5) | 0 | -- | 27 | 1022 (2.7) |

Note: Standard errors are shown in parentheses.

In Chinese Taipei, the majority of eighth grade students had teachers with bachelor's or master's degrees, but some students had science teachers with less than a bachelor's degree since, according to the *TIMSS 2007 Encyclopedia*, the policy before 1994 was to educate teachers in teacher colleges ("normal schools") (Jen et al., 2008). Nearly all students in the Czech Republic had teachers with postgraduate degrees, because most teachers receive a bachelor's degree and master's degree to become qualified teachers and are usually qualified in two specific subject areas (Basl & Tomasek, 2008). Although the majority of Singaporean teachers hold bachelor's degrees in their subject disciplines, it is understandable that 23 percent of the eighth grade students had mathematics teachers without majors in mathematics because some teachers hold diplomas in education or in subjects they are not currently teaching (Quek et al., 2008). In Slovenia, the policies regarding requirements for teacher education have been standardized since the Declaration of Bologna enacted in 2006, so that teachers acquire an educational mathematics or educational science degree at a university (Pavesic & Svetlik, Forthcoming). However, prior to 2006, some teachers were trained through 4-year study programs in education, which were not classified as bachelor's degree programs (Republic of Slovenia, Ministry of Education and Sport, 2011). In the United States, teachers are required to have a bachelor's degree to be qualified as teachers, but not necessarily to major in the subject area they plan to teach (Keene, 2008).

In Chinese Taipei and the Czech Republic, students whose teachers had university degrees and majored in mathematics or science areas had higher achievement than students whose teachers did not major in mathematics or science areas. In Singapore, Slovenia, and the United States there was little variation in average STEM achievement across teacher education categories.

## General School Resources

In TIMSS 2007 at the eighth grade, **General School Resources** was measured by a set of items asking school principals whether the quality of instruction was affected by any shortages in the following instructional materials: budget for supplies, school buildings and grounds, heating/cooling and lighting systems, instructional space, teachers, and computer support staff. The principal components analysis indicated the shortages could be summarized by one factor that accounted for 47 percent of the total variation in the items. Also, Cronbach's alpha reliability coefficient was $\alpha=0.81$. Given that these results provided evidence that the items were measuring one construct, the "Instruction Affected by General Resource Shortages" scale was developed and transformed to have a mean of 10 and a standard deviation of 2 like the other scales in this study.

Fit statistics for the "Instruction Affected by General Resource Shortages" scale indicated there was good item fit for all items (INFIT values ranged from 0.94–1.16). Appendix Table A.3 provides detailed information about the "Instruction Affected by General Resource Shortages" scale. Figure A.2 containing the Rasch item map shows the distribution of respondent scores was located somewhat higher on the scale than the item threshold parameters, indicating scale scores were somewhat positively skewed. Understandably, in these TIMSS 2007 high-performing countries, most school principals reported little if any impact from general resource shortages.

Table 4.6 presents the correlation between the "Instruction Affected by General Resource Shortages" scale and STEM achievement as well as the percentages of students affected by resource shortages to varying degrees (i.e., "none," "a little," "some or a lot"), together with their average STEM achievement. Because students in Chinese Taipei, the Czech Republic, Singapore, and Slovenia were mostly in equally well-resourced schools, the relationship with STEM achievement was weak (and slightly negative in Slovenia. However, the relationship between general resource shortages and STEM achievement was strongest in the United States where students with no instructional impact due to resource shortages had average achievement 58 points higher than students whose instruction was affected "some or a lot" by resources shortages.

Table 4.6. "Instruction Affected by General Resource Shortages" Scale in Relation to STEM Achievement

| Country | Pearson Correlation | None | | A Little | | Some or A Lot | |
|---|---|---|---|---|---|---|---|
| | | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 0.04*** | 12 | 1180 (17.4) | 60 | 1160 (11.2) | 29 | 1152 (13.1) |
| Czech Republic | 0.04*** | 12 | 1062 (21.8) | 74 | 1040 (5.2) | 15 | 1039 (14.9) |
| Singapore | 0.04*** | 36 | 1163 (14.7) | 61 | 1161 (9.8) | 3 | 1074 (54.7) |
| Slovenia | -0.03*** | 23 | 1034 (10.1) | 63 | 1039 (5.2) | 14 | 1048 (13.0) |
| United States | 0.11*** | 9 | 1060 (12.4) | 67 | 1035 (7.8) | 24 | 1002 (13.2) |

***p < .01
Note: Standard errors are shown in parentheses.

## Mathematics and Science Specific School Resources

The **Mathematics and Science Specific School Resources** factor was measured by two sets of items in TIMSS 2007, one about science laboratories and another about school resources specific to teaching both mathematics and science (e.g., computers).

Because the literature in Chapter 2 suggested that students in schools with laboratory equipment and materials have higher science achievement than students in less-equipped schools, two items relating to school laboratory resources were combined into one composite variable with three categories—Science Laboratory and Assistance Available, Science Laboratory Available (But No Assistance), and No Science Laboratory Available (see Appendix Table A.4 for item details). This measure worked well for the Czech Republic. However, as shown in Table 4.7, nearly all students in Chinese Taipei and Singapore attended schools with well-staffed science laboratories (little variation between schools) and there was no relationship with achievement in Slovenia and the United States. Because the measure developed from TIMSS 2007 did not help explain between school differences in STEM achievement in the majority of the five countries, it was not included in the multilevel analyses. Still, it is well worth noting that nearly all students in Chinese Taipei and Singapore (93%) attended schools with well-staffed science laboratories, compared to only 17 percent of students in the United States.

Table 4.7. Schools with Science Laboratories and Assistance Available When Students Conduct Experiments in Relation to STEM Achievement

| Country | Science Laboratory and Assistance Available | | Science Laboratory Available (But No Assistance) | | No Science Laboratory Available | |
|---|---|---|---|---|---|---|
| | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 93 | 1161 (8.0) | 7 | 1143 (39.2) | 0 | -- |
| Czech Republic | 0 | -- | 47 | 1055 (7.2) | 53 | 1031 (7.3) |
| Singapore | 93 | 1161 (8.5) | 7 | 1139 (39.8) | 0 | -- |
| Slovenia | 49 | 1035 (7.1) | 6 | 1058 (11.8) | 44 | 1042 (6.6) |
| United States | 17 | 1026 (15.7) | 63 | 1033 (7.7) | 20 | 1015 (12.6) |

Note: Standard errors are shown in parentheses.

As a second indicator of school resources pertinent to STEM achievement, principals were asked about any instructional impact due to resource shortages specific to mathematics and science instruction, including computers, computer software, calculators, library materials, audio-visual resources, and science laboratory equipment (see Appendix Table A.5 for items). The principal components analysis showed one factor explained 66 percent of the total variance in the items so, a Rasch scale was created (scale mean of 10 and standard deviation of 2), "Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science." The Cronbach's alpha reliability coefficient was $\alpha=0.95$.

Table 4.8 presents the results for the "Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science" scale in relation to STEM achievement. Similar to the "Instruction Affected by General Resource Shortages" scale, the relationship with STEM achievement was positive in the United States, but weak in Chinese Taipei, the Czech Republic, Singapore, and Slovenia. Again, these findings reflect the fact that the subset of TIMSS countries selected are top-performing TIMSS countries who do well in STEM subjects, so it is not surprising that they have mathematics and science resources to support instruction.

Appendix Figure A.3 contains the Rasch item map, indicating most respondents reported little impact due to resource shortages. Also, two misfitting items: instruction affected by a shortage of computers for math instruction (INFIT = 1.41) and instruction affected by a shortage of science laboratory equipment and materials (INFIT=1.79).

appear to further reflect the fact that many more respondents have laboratory equipment and materials than would be expected by the model.

Table 4.8. "Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science" Scale in Relation To STEM Achievement

| Country | Pearson Correlation | None | | A Little | | Some or A Lot | |
|---|---|---|---|---|---|---|---|
| | | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 0.04*** | 13 | 1190 (19.4) | 44 | 1154 (12.3) | 43 | 1157 (13.2) |
| Czech Republic | 0.02*** | 6 | 1058 (32.6) | 74 | 1042 (6.4) | 19 | 1040 (13.2) |
| Singapore | 0.07*** | 50 | 1169 (11.5) | 46 | 1154 (11.6) | 4 | 1089 (47.0) |
| Slovenia | -0.02 | 13 | 1035 (15.6) | 68 | 1040 (5.6) | 19 | 1043 (11.3) |
| United States | 0.06*** | 13 | 1070 (16.0) | 51 | 1020 (9.8) | 36 | 1027 (11.0) |

***p < .01

Note: Standard errors are shown in parentheses.

## *Fidelity of Curriculum Implementation Factors*

The three school factors related to the broad area of fidelity of curriculum implementation included: **Fidelity to Curriculum Structure**, **Fidelity to Curriculum Process**, and **Participant Responsiveness**.

### Fidelity to Curriculum Structure

As described in Chapter 2, **Fidelity to Curriculum Structure** includes adherence to the curriculum (i.e., extent of curriculum coverage) and duration of the implementation (i.e., instructional time). The "Mathematics Topics Taught" scale provides a measure of the extent of curriculum coverage in terms of the number of topics from the *TIMSS 2007 Mathematics Framework* that were taught by eighth grade, since the framework represents an agreed upon curriculum for participants in TIMSS. Appendix Table A.6 lists the 39 mathematics topics assessed by TIMSS 2007, according to content domain (e.g., algebra, geometry). The principal components analysis indicated that a single factor accounted for 25 percent of the total variance in the 39 topics taught items, which was low because the topics in the mathematics framework are diverse. However, the variance explained still met the minimum criteria of 20 percent and the reliability was high (Cronbach's alpha reliability coefficient $\alpha=0.92$), so the items were used to create a scale having a mean of 10 and standard deviation of 2.

The fit statistics for the "Mathematics Topics Taught" Rasch scale indicated good item fit (INFIT = 0.75-1.27) for all items but one: students having been taught translation, rotation, and reflection by eighth grade (INFIT = 1.37). Apparently, relatively small percentages of students were taught this geometry topic in Chinese Taipei (27%), the Czech Republic (35%), and Singapore (9%) compared to the other mathematics topics taught. More importantly, the Rasch item map in Figure A.4 indicates most students had been taught the topics in the number content domain (i.e., the four operations with whole numbers, fractions, and decimals) before eighth grade—not surprising given the countries selected for this analysis were high-achieving countries with rigorous curricula. Although the items in the number domain could have been excluded from the scale, they were retained so that the scale provided a more exact theoretical match to the topics included in the *TIMSS 2007 Mathematics Framework*.

Table 4.9 presents results for the "Mathematics Topics Taught" scale in relation to STEM achievement. The relationship between the scale and STEM achievement was positive in Singapore and the United States, as represented by the correlations 0.47 and 0.30, respectively. In these countries, students who had been taught all of the mathematics topics had higher achievement than students who had not been taught all of the topics.

Table 4.9. “Mathematics Topics Taught” Scale in Relation To STEM Achievement

| Country | Pearson Correlation | All Topics Taught By Eighth Grade and Some Number Topics Taught Before Eighth Grade | | Not All Topics Taught By Eighth Grade | |
|---|---|---|---|---|---|
| | | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 0.04*** | 51 | 1168 (10.9) | 49 | 1151 (10.5) |
| Czech Republic | 0.04*** | 39 | 1049 (9.5) | 61 | 1038 (6.7) |
| Singapore | 0.47*** | 62 | 1230 (8.4) | 39 | 1048 (17.2) |
| Slovenia | 0.04*** | 25 | 1047 (10.5) | 75 | 1036 (5.2) |
| United States | 0.30*** | 62 | 1063 (7.1) | 38 | 970 (8.3) |

***p < .01
Note: Standard errors are shown in parentheses.

TIMSS 2007 information about the number of instructional hours teachers spend on mathematics and science annually addresses the second aspect of **Fidelity to Curriculum Structure**, duration of the delivery (i.e., instructional time). Table 4.10 presents the results for the annual number of mathematics and science instructional hours (see Appendix Table A.7 for item details) in relation to STEM achievement. The number of hours teachers spend on mathematics and science instruction varies by country, especially when comparing the number of science instructional hours in the Czech Republic and Slovenia where science is taught as separate subjects. However, the number of instructional hours does not vary much from school to school within any one country, particularly, the Czech Republic, Singapore, and Slovenia, because number of instructional hours per curriculum area is determined by the Education Ministry as a matter of country-wide policy. Therefore, despite the importance of instructional time as a factor influencing higher student achievement, this "Yearly Instructional Hours" measure yields little information about instructional time differences between schools and was excluded from the multilevel analysis in this dissertation research.

Table 4.10. Yearly Instructional Hours Spent on Mathematics and Science in Relation to STEM Achievement

| Country | Mathematics Hours | | Science Hours | | STEM Achievement |
|---|---|---|---|---|---|
| | Hours | Standard Deviation | Hours | Standard Deviation | Country Average |
| Chinese Taipei | 159 | 37.8 | 145 | 30.7 | 1159 (8.1) |
| Czech Republic | 128 | 19.7 | 226 | 29.8 | 1043 (4.1) |
| Singapore | 124 | 15.6 | 109 | 18.8 | 1160 (8.2) |
| Slovenia | 113 | 3.7 | 165 | 38.2 | 1039 (4.2) |
| United States | 149 | 38.1 | 138 | 33.1 | 1028 (5.6) |

Note: Standard errors are shown in parentheses.

## Fidelity to Curriculum Process

**Fidelity to Curriculum Process** is addressed by two measures of effective instructional strategies—one for mathematics lessons and the other for science lessons. For mathematics, TIMSS 2007 focused on the degree to which instruction encouraged mathematical reasoning. Teachers reported on how often they asked students to explain their answers, relate the learning topics to their daily lives, decide on their own procedures for solving complex problems, and work on problems for which there is no immediately obvious solution method. A dominant factor summarizing the mathematics reasoning instructional items accounted for 54 percent of the total variance. Cronbach's alpha reliability coefficient was $\alpha=0.70$. The resulting Rasch scale (mean of 10 and standard deviation of 2) was entitled "Effective Instructional Strategies for Encouraging Mathematics Reasoning."

The “Effective Instructional Strategies for Teaching Science Inquiry” scale was based on how often teachers asked students to observe natural phenomena and describe what they see, watch the teacher demonstrate an experiment, design or plan experiments, conduct experiments, work together in small groups on experiments, use scientific formulae to solve problems, give explanations about something they are studying, and relate what they are learning to their daily lives. Summarized as a single component, these items accounted for 39 percent of the total variance. Cronbach’s alpha reliability coefficient was $\alpha=0.76$.

Detailed information about the items in the effective instructional strategies measures are provided in Table A.8 and Figures A.5–A.6 provide the Rasch item maps. The Rasch item map for the science instructional strategies scale showed the distribution of respondent scale scores was slightly negatively skewed in relation to the item threshold parameters. That is, teachers did not use many of the instructional strategies “every or almost every lesson.” However, INFIT values for both scales reflected good item fit (ranging from 0.95–1.10 for mathematics strategies and 0.86–1.06 for science strategies).

Table 4.11 shows the relationship between the “Effective Instructional Strategies for Encouraging Mathematics Reasoning” scale and STEM achievement. Except in Chinese Taipei and the United States, students whose teachers used the strategies more than half of the lessons had higher STEM achievement than students whose teachers used the strategies in some lessons or less.

Table 4.11. "Effective Instructional Strategies for Encouraging Mathematics Reasoning" Scale in Relation to STEM Achievement

| Country | Pearson Correlation | More Than Half the Lessons | | About Half the Lessons | | Some Lessons or Less | |
|---|---|---|---|---|---|---|---|
| | | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 0.01*** | 7 | 1140 (27.8) | 57 | 1166 (10.1) | 36 | 1153 (13.1) |
| Czech Republic | 0.04*** | 39 | 1047 (9.9) | 56 | 1042 (8.0) | 5 | 1017 (11.0) |
| Singapore | 0.08*** | 6 | 1254 (29.0) | 57 | 1153 (12.7) | 37 | 1151 (16.6) |
| Slovenia | 0.19*** | 28 | 1080 (7.7) | 59 | 1026 (6.0) | 13 | 1005 (15.6) |
| United States | 0.04*** | 25 | 1038 (11.4) | 60 | 1027 (6.0) | 15 | 1021 (13.8) |

***$p < .01$
Note: Standard errors are shown in parentheses.

Table 4.12 presents results for the "Effective Instructional Strategies for Teaching Science Inquiry" scale in relation to STEM achievement. The relationship between the scale and STEM achievement was weak, perhaps reflecting research discussed in Chapter 2, indicating some science instructional activities are positively associated with higher science achievement when used in moderation instead of frequently (what Braun et al., 2009 term the "Goldilocks" pattern). Nevertheless, the "Effective Instructional Strategies for Teaching Science Inquiry" scale was used in the multilevel analyses because the National Research Council and other researchers have identified the types of science inquiry instructional practices represented by the scale as important for STEM achievement (see National Research Council, 2011). The scale also provides parallel information to the Effective Instructional Strategies for Teaching Mathematics scale, giving a more complete picture of STEM instruction in each country.

Table 4.12. "Effective Instructional Strategies for Teaching Science Inquiry" Scale in Relation to STEM Achievement

| Country | Pearson Correlation | More Than Half the Lessons | | About Half the Lessons | | Some Lessons or Less | |
|---|---|---|---|---|---|---|---|
| | | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 0.01*** | 6 | 1174 (32.6) | 70 | 1162 (10.3) | 24 | 1155 (17.3) |
| Czech Republic | -0.03*** | 8 | 1020 (6.2) | 87 | 1046 (4.6) | 5 | 1029 (10.3) |
| Singapore | 0.05*** | 5 | 1292 (24.2) | 81 | 1176 (9.2) | 14 | 1186 (22.3) |
| Slovenia | 0.03*** | 7 | 1047 (11.1) | 89 | 1039 (4.2) | 5 | 1024 (13.3) |
| United States | -0.01*** | 18 | 1021 (14.6) | 77 | 1031 (6.1) | 5 | 1016 (28.6) |

***p < .01
Note: Standard errors are shown in parentheses.

## Participant Responsiveness

The factor, **Participant Responsiveness**, is measured in part, by the scale, "Students Like Learning Mathematics." Eighth grade students were asked how much they agreed they: 1) would like to take more mathematics in school, 2) enjoy learning mathematics, 3) think mathematics is boring (reverse coded), and 4) like mathematics (See Appendix Table A.9 for item details). The principal components analysis indicated that a single factor accounted for 71 percent of the total variance among these items. Cronbach's alpha reliability coefficient was α=0.86. Appendix Figure A.7 provides the Rasch item map, which demonstrated there was a good match between item threshold and respondent scale score distributions. Also, item fit was good for all items, with INFIT values ranging from 0.8 to 1.23.

For the corresponding "Students Like Learning Science Scale," eighth grade students in countries teaching science as an integrated subject (Chinese Taipei, Singapore, and the United States) responded to a set of items that paralleled those for mathematics (Appendix Table A.9). The principal components analysis indicated that summarizing the four science items with a single factor accounted for 74 percent of the total variance. Cronbach's alpha reliability coefficient was α=0.88. Also, there were no misfitting items (INFIT values ranged from 0.78-1.32) and the Rasch item map (see Figure A.8) demonstrated there was a good match between item threshold and respondent scale score distributions.

For the two countries teaching science as separate subjects (the Czech Republic and Slovenia), there were four times the number of items representing student attitudes toward science since the attitudes toward learning questions were asked for each of the different science content areas—biology, chemistry, physics, and earth science (Appendix Table A.9). The principal components analysis indicated that a single factor accounted for 31 percent of the total variance among the sixteen items. Cronbach's alpha reliability coefficient was α=0.85. (The item sets for the separate sciences also were explored as four separate scales, with the first principal component accounting for approximately the same amount of variance (71 to 74 percent). Also, four separate scales for only two countries would have unduly complicated the interpretation of student attitudes toward learning science in the multilevel modeling across countries.) The Rasch item scale including all sixteen items indicated that the scale functioned well in terms of item fit (INFIT values ranged from 0.88–1.17) and the location of the item threshold parameters matched the location of the respondent scale score distribution (see Appendix Figure A.9).

Table 4.13 presents the "Students Like Learning Mathematics" scale in relation to STEM achievement and Table 4.14 presents the "Students Like Learning Science" scale in relation to STEM achievement. In all five countries, students who like learning mathematics or science had higher achievement compared to students who do not like learning mathematics or science.

Table 4.13. "Students Like Learning Mathematics" Scale in Relation to STEM Achievement

| Country | Pearson Correlation | Like Learning Mathematics | | Somewhat Dislike Learning Mathematics | | Do Not Like Learning Mathematics | |
|---|---|---|---|---|---|---|---|
| | | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 0.42*** | 31 | 1255 (6.9) | 40 | 1171 (8.3) | 30 | 1050 (8.6) |
| Czech Republic | 0.19*** | 20 | 1094 (6.5) | 41 | 1042 (5.2) | 39 | 1018 (5.0) |
| Singapore | 0.24*** | 57 | 1197 (8.0) | 30 | 1128 (10.0) | 12 | 1069 (13.1) |
| Slovenia | 0.10*** | 15 | 1063 (11.1) | 40 | 1048 (4.6) | 45 | 1025 (4.3) |
| United States | 0.13*** | 37 | 1050 (6.6) | 38 | 1031 (6.2) | 25 | 996 (6.2) |

***p < .01
Note: Standard errors are shown in parentheses.

Table 4.14. "Students Like Learning Science" Scale in Relation to STEM Achievement

| Country | Pearson Correlation | Like Learning Science A Lot | | Like Learning Science A Little | | Somewhat Like or Do Not Like Learning Science | |
|---|---|---|---|---|---|---|---|
| | | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 0.27*** | 7 | 1294 (11.5) | 60 | 1178 (8.5) | 34 | 1099 (8.8) |
| Czech Republic | 0.06*** | 20 | 1058 (8.7) | 60 | 1041 (4.2) | 20 | 1034 (6.4) |
| Singapore | 0.22*** | 19 | 1228 (9.4) | 69 | 1156 (8.3) | 12 | 1081 (11.7) |
| Slovenia | 0.16*** | 18 | 1080 (8.2) | 58 | 1038 (4.8) | 24 | 1012 (5.5) |
| United States | 0.12*** | 15 | 1065 (7.2) | 61 | 1031 (5.9) | 24 | 1004 (6.5) |

***p < .01
Note: Standard errors are shown in parentheses.

### *School Climate Factors*

School Climate was measured in TIMSS 2007 by two sets of items—one asking principals about **Discipline and Attendance Problems** (negative aspects of school climate) and the other asking about **Support for Academic Success** (positive aspects). Appendix Table A.10 presents the questions in full. The set addressing negative aspects of school climate asked principals twelve questions about the severity of various problems in schools relating to discipline and attendance among eighth grade students. The principal components analysis indicated that a single factor explained 46 percent of the total variance among the 12 items. Also, Cronbach's alpha reliability coefficient was $\alpha$=0.90. Thus, the Rasch scale was created—"Schools Have Few Discipline and Attendance Problems." A second Rasch scale was created to measure the positive aspects of school climate—"Schools Have a Climate Supportive of Academic Success." This scale was based on five items asking principals to evaluate teachers' understanding of the school's curricular goals, teachers' degree of success in implementing the curriculum, teachers' expectations for student achievement, parental support for student achievement, and students' desire to do well in school (see Table A.11 for details). A single factor explained 60 percent of the total variance among these items, and Cronbach's alpha reliability coefficient was $\alpha$=0.83.

Appendix Figures A.10–A.11 contain the Rasch item maps. The item map for the "Schools Have a Climate Supportive of Academic Success" scale shows positively skewed scale scores, indicating principals did not often respond that there were very low levels of teacher and student support for academic success. However, there were no misfitting items in either of the Rasch scales (INFIT values ranged from 0.83–1.17).

Tables 4.15 and 4.16 present each of the two school climate scales, respectively, in relation to STEM achievement. In the Czech Republic, Singapore, and the United States, the scale measuring the negative aspects of school climate had a moderate relationship with student achievement. Students in schools with only minor or no discipline and attendance problems had higher achievement than students in schools where the principal reported serious behavior problems. Although all positive, the correlations between the scale and STEM achievement ranged from moderately strong in Singapore to very weak in Slovenia. Students attending schools in Chinese Taipei, Singapore, and the United States with very high or high support for academic success had higher achievement than those in schools that provided less support.

Table 4.15. "Schools Have Few Discipline and Attendance Problems" Scale in Relation to STEM Achievement

| Country | Pearson Correlation | Not a Problem | | Minor Problem | | Serious Problem | |
|---|---|---|---|---|---|---|---|
| | | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 0.01*** | 33 | 1153 (13.0) | 47 | 1167 (13.9) | 20 | 1141 (12.0) |
| Czech Republic | 0.16*** | 4 | 1114 (31.0) | 71 | 1050 (6.0) | 24 | 1009 (7.2) |
| Singapore | 0.27*** | 17 | 1269 (18.3) | 72 | 1146 (10.0) | 12 | 1083 (21.4) |
| Slovenia | 0.00 | 6 | 1050 (14.1) | 65 | 1038 (5.4) | 29 | 1041 (7.3) |
| United States | 0.17*** | 5 | 1047 (22.2) | 68 | 1046 (8.0) | 27 | 984 (14.0) |

***p < .01
Note: Standard errors are shown in parentheses.

Table 4.16. “Schools Have a Climate Supportive of Academic Success” Scale in Relation to STEM Achievement

| Country | Pearson Correlation | Very High or High Support | | Medium Support | | Low or Very Low Support | |
|---|---|---|---|---|---|---|---|
| | | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement | Percent of Students | STEM Achievement |
| Chinese Taipei | 0.15*** | 70 | 1177 (10.0) | 28 | 1122 (14.2) | 2 | 1077 (31.5) |
| Czech Republic | 0.09*** | 5 | 1059 (46.4) | 74 | 1048 (5.9) | 21 | 1018 (10.0) |
| Singapore | 0.37*** | 42 | 1241 (12.8) | 55 | 1103 (12.7) | 2 | 1006 (59.6) |
| Slovenia | 0.06*** | 15 | 1062 (10.7) | 81 | 1035 (5.0) | 4 | 1038 (13.2) |
| United States | 0.22*** | 39 | 1066 (8.4) | 56 | 1010 (7.2) | 5 | 944 (31.2) |

***$p < .01$
Note: Standard errors are shown in parentheses.

## *Adjusting the Rasch Scales for Missing Data*

After the final measures were selected for the multilevel analyses, the amount of missing data was analyzed. At the student level, the number of cases (students) missing a scale score for the “Student Home Resources” scale was minimal in each country, ranging from 0.04 percent in the Czech Republic to 0.9 percent in the United States. However, at the school level, the number of cases (schools) missing at least one score on the 11 school-level measures was greater. In Chinese Taipei, 4.6 percent of schools were missing a score on at least one scale; 1.4 percent were missing a score in the Czech Republic, 3.7 percent in Singapore, 4.7 percent in Slovenia, and 16.7 percent in the United States.

Since there was little missing data at the student level it was decided not to impute missing values for the “Student Home Resources” scale. However, as Raudenbush and Bryk (2002) point out, the sample size at the school level has more influence on estimates of power than the sample size at the student level, and so missing data at the school level were imputed. Specifically, conditional mean imputation was used to impute missing data for each school level measure with missing data in each country. Following the single non-stochastic regression method described by Little and Rubin (2002), a separate multiple regression equation was built to predict missing values for one measure conditional on the values of all of the other measures in the model. For example, if a country was missing scores for the “Effective Instructional Strategies for Encouraging Mathematics Reasoning” scale, all of the other school-level measures were entered into a regression equation used to predict the missing score. The missing values for each Rasch scale were then replaced with the predicted values from each model. Although there are more sophisticated methods to impute missing data, this straightforward approach was taken due to the minimal amount of missing data.

## Examining School Effectiveness Factors Related to STEM Achievement

How do school effectiveness factors related to teaching and learning STEM subjects operate together in the United States compared to top-achieving countries in TIMSS 2007? This research question was examined through the lens of a school effectiveness framework by conducting a series of multilevel analyses showing how these factors that operate in the United States, Chinese Taipei, Singapore, the Czech Republic, and Slovenia. This section presents the results of these analyses.

### *Between School Variance in STEM Achievement*

To begin, the percentage of variance in student STEM achievement that is attributable to differences between schools was determined for each of the five countries under study, using what is known as the *unconditional model*. This constitutes the upper limit on the variation that can be explained by differences in school factors. When variation in student achievement attributable to differences between schools is small, there is less scope for effective school factors to explain achievement differences (Bickel, 2007).

Table 4.17 presents the variance components produced from the *unconditional model* in each of the five countries in this study as well as the percentage of variability within and between schools that exists without considering any explanatory variables.

Table 4.17. Variance Components and Percentage of Total Variation in STEM Achievement Due to Differences Within and Between Schools

| Country | Variance Components | | Differences Within Schools | Differences Between Schools |
|---|---|---|---|---|
| | Student Level ($\sigma^2$) | School Level ($\tau_{00}$) | | |
| Chinese Taipei | 27964.9 | 7527.9 | 79% | 21% |
| Czech Republic | 13838.5 | 5372.6 | 72% | 28% |
| Singapore | 17256.7 | 19038.1 | 48% | 52% |
| Slovenia | 17014.1 | 1963.8 | 90% | 10% |
| United States | 14699.9 | 8428.7 | 64% | 36% |

There is a range in the amount of between school variance across the countries under study (Table 4.17), with Slovenia showing the smallest differences between schools (10% of total variability) and Singapore showing the largest differences between schools (52% of total variability). That is, in Singapore, more than half of the total variability in Singaporean student STEM achievement is attributable to differences between schools.

There are not any strict guidelines as to what qualifies as a minimum amount of variance between schools for further multilevel modeling of differences between schools (Scherbaum & Ferreter, 2009). Therefore, other literature examining international education systems through multilevel modeling was evaluated to provide a reference point for a reasonable inclusion criteria in this research. In PISA 2009, researchers examined differences in reading achievement that could be attributed to differences between schools in countries with percentages ranging from 8 percent in Finland to 98 percent in Argentina, with an overall average of 42 percent (Organisation for Economic Co-operation and Development, 2010b). However, school effectiveness research using TIMSS 1995 data at the eighth grade examined school effectiveness differences in countries with at least 10 percent variation, even though percentages ranged from 7 percent in Cyprus, Japan, Korea, Norway, and Slovenia to 51 percent in Romania for differences between school in science achievement; and 7 percent in Japan to 64 percent in the U.S. for differences between school in mathematics achievement (Martin, Mullis, Gregory, Hoyle, & Shen, 2000). Consistent with the previous TIMSS guideline, the data from the five countries in this study were analyzed using HLM analyses even though Slovenia was borderline with just 10 percent of the total variance between schools.

### *Student Home Resources*

The *student home resources model* (see Equation 3.10) was built to control for the variance in student achievement within schools related to student's home background, so that school effects could be better isolated. Table 4.18 summarizes the results from applying the *student home resources model* to the five countries studied in this dissertation.

Table 4.18. Variance Components of Student Home Resources Model

| Country | Percentage of Student Within School Variance Explained by the Student Home Resources Model | Percentage of Between School Variance Explained by the Student Home Resources Model | Percentage of Total Variance in STEM Achievement Explained by the Student Home Resources Model | Percentage of Total Variance in STEM Achievement due to Differences Between Schools, After Controlling for Student Home Resources Within School |
|---|---|---|---|---|
| Chinese Taipei | 17% | 55% | 25% | 13% |
| Czech Republic | 8% | 28% | 13% | 23% |
| Singapore | 14% | 23% | 18% | 49% |
| Slovenia | 11% | 31% | 13% | 8% |
| United States | 6% | 24% | 12% | 32% |

First, building on the percentage of total variation in STEM achievement due to differences within schools shown in Table 4.17, the *student home resources model* accounted for 17 percent of the student within school variance in STEM achievement in Chinese Taipei, 8 percent in the Czech republic, 14 percent in Singapore, 11 percent in Slovenia, and 6 percent in the United States. Also, although it was not modeled explicitly at the school level, the *student home resources model* accounted for part of the differences between schools (last column Table 4.17). This ranged from a low of 23 percent in Singapore to a high of 55 percent in Chinese Taipei.

Next, as shown in the third column of Table 4.18, the *student home resources model* explained from 12 percent to 25 percent of the total variance in student STEM achievement. For example, the 25 percent of total variance explained by the model in Chinese Taipei was composed of 17 percent of the 79 percent of within school differences (from Table 4.17) and 55 percent of the 21 percent of between school differences (from Table 4.17).

After fitting the *student home resources* model, the variation in STEM achievement due to differences between schools (last column of Table 4.17) was reduced somewhat. Controlling for student home resources within school resulted in a range of between school variance across the five countries (8% in Slovenia to 49% in Singapore) that can be explained by school factors.

Table 4.19 presents information about the straight line representing the relationship between mean student home resources and STEM achievement within schools. The slope of this line ranged from 16.2 (with an intercept of 1025.9) in the United States to 34.0 (with an intercept of 1164.1) in Chinese Taipei. That is, in the United States, for every one point increase in "Student Home Resources" over and above the grand mean, STEM achievement is predicted to increase by approximately 16 points. All of the scales in this dissertation research were transformed to have a mean of 10 and a standard deviation of 2. Thus, a one point increase represents approximately one half of a standard deviation increase in "Student Home Resources". Table 4.19 also indicates the slope for "Student Home Resources" varied significantly across schools in each country and was reliable ($\hat{\lambda}$=0.08 to 0.52). These results signify that in each country, the relationship between student home resources and STEM achievement within schools varies significantly across the population of schools.

Table 4.19. Student Home Resources Model

| Country | Intercept | | Slope | | Residual Variance for Student Home Resources Slope | |
|---|---|---|---|---|---|---|
| | Coefficient | S.E. | Coefficient | S.E. | Variance | Reliability |
| Chinese Taipei | 1164.1*** | (5.6) | 34.0*** | (1.7) | 150.6*** | 0.39 |
| Czech Republic | 1043.4*** | (5.7) | 23.0*** | (1.3) | 24.6*** | 0.13 |
| Singapore | 1165.7*** | (10.2) | 23.4*** | (1.6) | 201.3*** | 0.52 |
| Slovenia | 1040.3*** | (4.4) | 27.0*** | (1.4) | 16.3*** | 0.08 |
| United States | 1025.9*** | (5.5) | 16.2*** | (1.0) | 43.3*** | 0.25 |

## STEM School Effectiveness in the United States

### *School Resources Models*

In order to study how school resources operate in the United States, three models were built examining the factors **Teacher Preparation**, **General Resources**, and **Mathematics and Science Specific School Resources**. The first model estimated for the United States was the *teacher preparation model* (see Equation 3.13 for model), providing information about how teachers' education level and major area of study adds to the explanation of school effectiveness, after controlling for student home resources. Two measures representing differences in teacher education levels and majors for mathematics and science teachers (dummy coded into 2 variables each) were entered into the *teacher preparation model* at the school level in addition to the student home resources control at the student level.

Table 4.20 presents results from the *teacher preparation model* for the United States. **Teacher Preparation** did not explain any additional variance in STEM achievement that is between schools after controlling for student home resources, since the percentage remained the same as for the *student home resources model* (32%).

Table 4.20. School Resources in the United States

| | Student Home Resources | | Teacher Preparation | | General and STEM Resources | |
|---|---|---|---|---|---|---|
| Variance Explained | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 32% | | 32% | | 31% | |
| Percentage of Between School Variance Explained | 24% | | 23% | | 26% | |
| Percentage of Total Variance Explained | 12% | | 12% | | 13% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1025.9 *** | (5.5) | 1014.8 *** | (11.9) | 1025.6 *** | (5.4) |
| Student Home Resources Slope | 16.2 *** | (1.0) | 16.2 *** | (1.0) | 16.2 *** | (1.0) |
| Teacher Preparation | | | | | | |
| Postgraduate Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 15.5 | (13.4) | | |
| Bachelor's Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 20.8 | (13.6) | | |
| Postgraduate Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | -2.2 | (12.8) | | |
| Bachelor's Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | -0.2 | (14.1) | | |
| General School Resources | | | | | | |
| Instruction Affected by General Resource Shortages | | | | | 5.7 * | (3.2) |
| Mathematics and Science Specific School Resources | | | | | | |
| Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science | | | | | 3.6 | (3.3) |

***p < .01, **p < .05, *p<.10

Table 4.20 also presents results from the *general and STEM resources model* in the United States. “Instruction Affected by General Resource Shortages” and “Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science” yielded a one percentage point decrease in total variance between schools after controlling for home resources, compared to the *student home resources model*, primarily attributed to “Instruction Affected by General Resource Shortages,” which was a significant predictor of STEM achievement. After controlling for student home resources and holding everything else constant, for every one point increase in "Instruction Affected by General Resource Shortages" over and above the grand mean, STEM achievement is predicted to increase by approximately 6 points.

### *Fidelity of Curriculum Implementation Models*

Table 4.21 (left side) presents results from three successive *fidelity of curriculum implementation models* for the United States (see Equation 3.14 for final model). Examining the variance explained by the models shows that fidelity of curriculum implementation factors explain differences in STEM achievement even after controlling for student home resources. In particular, the "Mathematics Topics Taught" scale was a significant predictor in the U.S., decreasing the total variance between schools after controlling for home resources by three percentage points compared to the *student home resources model* for the U.S. (29% compared to 32%, respectively). After controlling for student home resources, for every one point increase in "Mathematics Topics Taught" over and above the grand mean, STEM achievement is predicted to increase by approximately 17 points.

There was little change in the effect of "Mathematics Topics Taught" across Models 1, 2, and 3, indicating that teaching the topics in the curriculum is important to STEM achievement in the United States, even after considering what strategies teachers use (Model 2) and student attitudes toward learning mathematics and science (Model 3).

Table 4.21. Fidelity of Curriculum Implementation and School Climate in the United States

| | Student Home Resources | | Fidelity of Curriculum Implementation | | | | | | School Climate | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| Variance Explained | | | | | | | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 32% | | 29% | | 28% | | 29% | | 30% | | 28% | |
| Percentage of Between School Variance Explained | 24% | | 34% | | 35% | | 34% | | 30% | | 37% | |
| Percentage of Total Variance Explained | 12% | | 16% | | 16% | | 16% | | 15% | | 17% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1025.9 *** | (5.5) | 1024.6 *** | (5.2) | 1024.7 *** | (5.3) | 1024.8 *** | (5.2) | 1023.5 *** | (5.4) | 1023.6 *** | (5.1) |
| Student Home Resources Slope | 16.2 *** | (1.0) | 16.0 *** | (1.0) | 16.0 *** | (1.0) | 16.0 *** | (1.0) | 16.0 *** | (1.0) | 16.0 *** | (1.0) |
| Fidelity to Curriculum Structure | | | | | | | | | | | | |
| Mathematics Topics Taught | | | 16.5 *** | (2.8) | 16.4 *** | (2.8) | 16.5 *** | (2.8) | | | | |
| Fidelity to Curriculum Process | | | | | | | | | | | | |
| Effective Instructional Strategies for Encouraging Mathematics Reasoning | | | | | -5.9 | (8.9) | -6.4 | (8.8) | | | | |
| Effective Instructional Strategies for Teaching Science Inquiry | | | | | 11.9 | (7.4) | 11.8 | (7.5) | | | | |
| Participant Responsiveness | | | | | | | | | | | | |
| Students Like Learning Mathematics | | | | | | | 0.7 | (3.1) | | | | |
| Students Like Learning Science | | | | | | | 0.7 | (3.1) | | | | |
| Discipline and Attendance Problems | | | | | | | | | | | | |
| Schools Have Few Discipline and Attendance Problems | | | | | | | | | 13.8 *** | (3.2) | 11.1 *** | (3.2) |
| Support for Academic Success | | | | | | | | | | | | |
| Schools Have a Climate Supportive of Academic Success | | | | | | | | | | | 11.2 *** | (2.2) |

***p < .01, **p < .05, *p<.10

## *School Climate Models*

Table 4.21 (right side) also shows the results for two *school climate models* for the United States (see Equation 3.15 for final model). **Discipline and Attendance Problems** (Model 4) in isolation is a significant predictor of STEM achievement and resulted in a two point reduction in the total variance between schools after controlling for student home resources compared to the *student home resources model* (30% compared to 32%, respectively). **School Support for Academic Success** (Model 5) is a significant predictor of STEM achievement also, even when considering the negative aspects of school climate. Including this factor in the model further reduced the amount of total variance between schools after controlling for student home resources by two percentage points. The results show that STEM achievement in the United States is higher in schools having few discipline and attendance problems and a climate supportive of academic success.

## *Complete School Effectiveness Model*

The next models explored how school resources, fidelity of curriculum implementation, and school climate operate together to explain differences in STEM achievement in the United States. Table 4.22 presents the results.

First, the *school resources combined model* (Model 2) included the teacher preparation measures and both resource shortages scales together with the student home resources control. For the United States, "Instruction Affected by General Resource Shortages" was a significant predictor of STEM achievement in the combined model. The magnitude of the teacher preparation measures for mathematics did

Table 4.22. Final Predictors of STEM School Effectiveness in the United States

| | Student Home Resources | | School Resources Combined | | School Resources and Fidelity of Curriculum Implementation | | School Resources, Fidelity of Curriculum Implementation, and School Climate | | Complete School Effectiveness Model | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| Variance Explained | | | | | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 32% | | 31% | | 28% | | 26% | | 19% | |
| Percentage of Between School Variance Explained | 24% | | 25% | | 35% | | 42% | | 62% | |
| Percentage of Total Variance Explained | 12% | | 13% | | 17% | | 19% | | 27% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1025.9 *** | (5.5) | 1020.8 *** | (12.0) | 1017.7 *** | (11.1) | 1017.7 *** | (10.8) | 1015.9 *** | (8.4) |
| Student Home Resources Slope | 16.2 *** | (1.0) | 16.2 *** | (1.0) | 16.0 *** | (1.0) | 15.9 *** | (1.0) | 14.9 *** | (1.0) |
| School Factors | | | | | | | | | | |
| School Average Student Home Resources | | | | | | | | | 50.8 *** | (5.1) |
| School Resources | | | | | | | | | | |
| Postgraduate Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 8.4 | (13.6) | 16.1 | (12.9) | 12.1 | (12.7) | 16.6 | (10.1) |
| Bachelor's Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 16.4 | (13.8) | 27.9 ** | (13.3) | 19.5 | (12.7) | 17.0 | (10.8) |
| Postgraduate Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | -1.8 | (12.6) | -12.4 | (11.8) | -7.9 | (11.4) | -6.6 | (9.6) |
| Bachelor's Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | -9.3 | (14.6) | -10.2 | (14.2) | -4.8 | (13.8) | -9.1 | (10.6) |

***$p < .01$, **$p < .05$, *$p<.10$

Table 4.22. Final Predictors of School Effectiveness in the United States (Continued)

| | Student Home Resources | | School Resources Combined | | School Resources and Fidelity of Curriculum Implementation | | School Resources, Fidelity of Curriculum Implementation, and School Climate | | Complete School Effectiveness Model | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| School Resources (Continued) | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Instruction Affected by General Resource Shortages | | | 5.5 * | (3.2) | 4.5 | (3.1) | 0.7 | (3.1) | -1.4 | (2.5) |
| Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science | | | 4.0 | (3.2) | 1.2 | (3.4) | 2.6 | (3.3) | 2.3 | (2.5) |
| Fidelity of Curriculum Implementation | | | | | | | | | | |
| Mathematics Topics Taught | | | | | 16.7 *** | (2.8) | 12.7 *** | (2.9) | 1.8 | (2.2) |
| Effective Instructional Strategies for Encouraging Mathematics Reasoning | | | | | -4.9 | (8.5) | -2.5 | (8.0) | -7.3 | (5.9) |
| Effective Instructional Strategies for Teaching Science Inquiry | | | | | 10.5 | (6.9) | 12.8 * | (6.5) | 10.3 ** | (5.1) |
| Students Like Learning Mathematics | | | | | 1.2 | (3.0) | 1.6 | (2.9) | 0.6 | (2.6) |
| Students Like Learning Science | | | | | 0.3 | (3.0) | 0.1 | (2.9) | 1.1 | (2.4) |
| School Climate | | | | | | | | | | |
| Schools Have Few Discipline and Attendance Problems | | | | | | | 5.7 * | (3.2) | 0.0 | (2.6) |
| Schools Have a Climate Supportive of Academic Success | | | | | | | 9.8 *** | (2.2) | 5.7 *** | (1.8) |

***p < .01, **p < .05, *p<.10

decrease with the addition of the resource shortage scales however, implying that well-resourced schools also have more highly-prepared teachers.

Second, school resources and fidelity of curriculum implementation were explored together controlling for student home resources in the *school resources and fidelity of curriculum implementation model* (Model 3). Compared to 32 percent total variance between schools after controlling for student home resources in the *student home resources model*, this model reduced the percentage of variance to 28 percent. In the United States (Model 3), teacher education level and mathematics major were significant predictors of STEM achievement when considered in conjunction with "Mathematics Topics Taught." In Model 3, adhering to the curriculum as measured by "Mathematics Topics Taught" was a significant predictor of higher STEM achievement.

Interestingly, **General School Resources** was no longer significant when considering fidelity of curriculum implementation factors, perhaps indicating that when teachers teach the topics and are qualified to do so, it can compensate for a lack of resources to some degree.

Next, Table 4.22 also presents results from the *school resources, fidelity of curriculum implementation, and school climate model* (Model 4), which reduced the total variance between schools after controlling for student home resources an additional two percentage points to 26 percent, a reduction of 6 percentage points total compared to the *student home resources* model. With the addition of school climate factors to the model, **Teacher Preparation** was no longer significant. However, in the United States even in the presence of school climate factors, "Mathematics Topics Taught" was still an important predictor of STEM achievement and “Effective Instructional Strategies for Teaching Science Inquiry” became a significant predictor of STEM achievement. Both school climate factors were significant predictors of STEM achievement beyond adhering to the curriculum and implementing it with effective instructional strategies.

Finally, all school effectiveness factors together with a measure of school average student home resources were used to create the *complete school effectiveness model* (Model 5). (See Equation 3.16 for model). Results from the complete model provide a picture of how the combined school effectiveness factors add to the explanation of STEM school effectiveness, after considering student home resources at both the individual and school levels. For each school, an average of the student home resources is used to model the effect of the socioeconomic composition of the school’s student body on STEM achievement.

“School Average Student Home Resources” is important in explaining between school differences in STEM achievement in the United States, reducing the total between school variance after controlling for student home resources to 19 percent. Because “Mathematics Topics Taught” and “Schools Have Few Discipline and Attendance Problems” no longer were significant predictors after controlling for the composition of the student body, it appears that schools with many students from well-resourced homes are likely to be those that teach the curriculum and have orderly school environments. However, even after controlling for school composition, “Effective Instructional Strategies for Teaching Science Inquiry” and “Schools Have a Climate Supportive of

Academic Success" were important for explaining STEM achievement differences between schools in the United States.

To summarize in the United States, teaching the mathematics curriculum, using instructional strategies to teach science inquiry, eliminating discipline and attendance problems, and providing a school climate supportive of academic success are school effectiveness factors that account for differences in STEM achievement between schools, even after controlling for differences between student's home resources. Although teaching the mathematics curriculum and maintaining an orderly school environment already are characteristics of schools attended by students from well-resourced homes.

## STEM School Effectiveness in Chinese Taipei

### *School Resources Models*

Following the model building process in the United States, the same series of models were built for Chinese Taipei, one of the top-performing countries in TIMSS 2007 at the eighth grade in both mathematics and science. School resources factors were explored first. Table 4.23 presents results from the two school resources models—the *teacher preparation model* and the *general and STEM resources model*. **Teacher Preparation** in science helped to explain some of the between school differences in STEM achievement after controlling for student home resources (the total variance between schools after controlling for home resources decreased by one percentage point compared to the *student home resources model*). As might be anticipated, teachers having a postgraduate degree and major in science or science education was a significant

Table 4.23. School Resources in Chinese Taipei

| | Student Home Resources | | Teacher Preparation | | General and STEM Resources | |
|---|---|---|---|---|---|---|
| Variance Explained | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 13% | | 12% | | 13% | |
| Percentage of Between School Variance Explained | 55% | | 57% | | 55% | |
| Percentage of Total Variance Explained | 25% | | 25% | | 25% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1164.1 *** | (5.6) | 1118.9 *** | (21.0) | 1164.0 *** | (5.6) |
| Student Home Resources Slope | 34.0 *** | (1.7) | 33.8 *** | (1.7) | 34.0 *** | (1.7) |
| Teacher Preparation | | | | | | |
| Postgraduate Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 34.6 | (21.7) | | |
| Bachelor's Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 28.4 | (20.4) | | |
| Less than a Bachelor's Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 6.8 | (25.5) | | |
| Postgraduate Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 41.7 ** | (17.6) | | |
| Bachelor's Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 16.7 | (13.4) | | |
| Less than a Bachelor's Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 45.8 ** | (17.3) | | |
| General School Resources | | | | | | |
| Instruction Affected by General Resource Shortages | | | | | -1.5 | (3.4) |
| Mathematics and Science Specific School Resources | | | | | | |
| Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science | | | | | 0.8 | (3.3) |

***p < .01, **p < .05, *p<.10

predictor of STEM achievement. However, perhaps because these are the more senior and experienced members of the teaching force, teachers with less than a bachelor's degree and major in science or science education also was significant. Differences in **General School Resources** and **Mathematics and Science Specific School Resources** did not help to explain differences in STEM achievement between schools and did not impact the relationship between teacher preparation and STEM achievement, likely because most schools are well resourced.

### *Fidelity of Curriculum Implementation Models*

Next, three successive models were built to examine the additive effects of **Fidelity to Curriculum Structure**, **Fidelity to Curriculum Process**, and **Participant Responsiveness**. Table 4.24 presents results from the three fidelity of curriculum implementation models for Chinese Taipei. “Mathematics Topics Taught” (Model 1) did not explain any additional variance in STEM achievement that is between schools after controlling for student home resources. Perhaps because all students in Chinese Taipei study a rigorous curriculum. Chinese Taipei has a centralized educational policy, characterized by national curriculum standards, which all teachers follow (Jen et al., 2008). However, both measures of **Fidelity to Curriculum Process**—“Effective Instructional Strategies for Encouraging Mathematics Reasoning” and “Effective Instructional Strategies for Teaching Science Inquiry” were significant predictors of differences in STEM achievement.

### *School Climate Models*

Table 4.24 also presents the results from the two *school climate models* for Chinese Taipei. These two factors did not explain any between school differences in STEM achievement after considering student home resources. It appears that there is little variation between schools in school climate in Chinese Taipei, with most students attending schools that are well disciplined and focused on academic success.

Table 4.24. Fidelity of Curriculum Implementation and School Climate in Chinese Taipei

| | Student Home Resources | | Fidelity of Curriculum Implementation | | | | | | School Climate | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| Variance Explained | | | | | | | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 13% | | 12% | | 7% | | 7% | | 13% | | 13% | |
| Percentage of Between School Variance Explained | 55% | | 56% | | 75% | | 75% | | 55% | | 55% | |
| Percentage of Total Variance Explained | 25% | | 25% | | 29% | | 29% | | 25% | | 25% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1164.1 *** | (5.6) | 1163.8 *** | (5.6) | 1163.2 *** | (4.7) | 1163.4 *** | (4.7) | 1164.2 *** | (5.6) | 1164.1 *** | (5.6) |
| Student Home Resources Slope | 34.0 *** | (1.7) | 33.8 *** | (1.7) | 33.3 *** | (1.7) | 33.3 *** | (1.7) | 34.0 *** | (1.7) | 34.0 *** | (1.7) |
| Fidelity to Curriculum Structure | | | | | | | | | | | | |
| Mathematics Topics Taught | | | 4.3 | (2.9) | 1.9 | (2.6) | 2.1 | (2.7) | | | | |
| Fidelity to Curriculum Process | | | | | | | | | | | | |
| Effective Instructional Strategies for Encouraging Mathematics Reasoning | | | | | 44.1 *** | (6.6) | 44.6 *** | (6.9) | | | | |
| Effective Instructional Strategies for Teaching Science Inquiry | | | | | 12.8 * | (7.4) | 12.0 * | (7.2) | | | | |
| Participant Responsiveness | | | | | | | | | | | | |
| Students Like Learning Mathematics | | | | | | | 1.2 | (1.7) | | | | |
| Students Like Learning Science | | | | | | | -3.5 | (2.6) | | | | |
| Discipline and Attendance Problems | | | | | | | | | | | | |
| Schools Have Few Discipline and Attendance Problems | | | | | | | | | -1.2 | (2.0) | -1.3 | (2.0) |
| Support for Academic Success | | | | | | | | | | | | |
| Schools Have a Climate Supportive of Academic Success | | | | | | | | | | | 1.9 | (2.4) |

***p < .01, **p < .05, *p<.10

### *Complete School Effectiveness Model*

Table 4.25 presents results from models exploring the combination of school effectiveness factors in Chinese Taipei. Examining school resources and fidelity of curriculum implementation (Model 3) decreased the percentage of total variance between schools after controlling for home resources by 5 percentage points compared to the model with only school resources (Model 2), primarily because “Effective Instructional Strategies for Encouraging Mathematics Reasoning” was a strong predictor. The *complete school effectiveness model* (Model 5), combining all school effectiveness factors together with a measure of school average home resources, decreased the variance between schools after controlling for home resources by only one percentage point compared to Model 4. However, **Teacher Preparation** in science and “Effective Instructional Strategies for Encouraging Mathematics Reasoning” remained significant in the model, indicating that these factors are unrelated to the socioeconomic composition of the student body in Chinese Taipei.

Table 4.25. Final Predictors of STEM School Effectiveness in Chinese Taipei

| | Student Home Resources | | School Resources Combined | | School Resources and Fidelity of Curriculum Implementation | | School Resources, Fidelity of Curriculum Implementation, and School Climate | | Complete School Effectiveness Model | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| Variance Explained | | | | | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 13% | | 12% | | 7% | | 7% | | 6% | |
| Percentage of Between School Variance Explained | 55% | | 56% | | 75% | | 75% | | 81% | |
| Percentage of Total Variance Explained | 25% | | 25% | | 29% | | 29% | | 30% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1164.1 *** | (5.6) | 1119.4 *** | (21.9) | 1120.4 *** | (21.4) | 1121.7 *** | (20.7) | 1124.8 *** | (20.8) |
| Student Home Resources Slope | 34.0 *** | (1.7) | 33.8 *** | (1.7) | 33.1 *** | (1.7) | 33.1 *** | (1.7) | 32.1 *** | (1.8) |
| School Factors | | | | | | | | | | |
| School Average Student Home Resources | | | | | | | | | 19.0 *** | (4.2) |
| School Resources | | | | | | | | | | |
| Postgraduate Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 33.1 | (23.9) | 31.9 | (22.9) | 29.4 | (22.4) | 27.1 | (20.7) |
| Bachelor's Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 27.6 | (23.3) | 30.5 | (21.9) | 28.4 | (21.3) | 32.5 | (19.8) |
| Less than a Bachelor's Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 5.8 | (28.8) | 32.0 | (26.6) | 29.7 | (26.3) | ment | (24.8) |

***p < .01, **p < .05, *p<.10

Table 4.25. Final Predictors of STEM School Effectiveness in Chinese Taipei (Continued)

| | Student Home Resources | | School Resources Combined | | School Resources and Fidelity of Curriculum Implementation | | School Resources, Fidelity of Curriculum Implementation, and School Climate | | Complete School Effectiveness Model | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| School Resources (Continued) | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Postgraduate Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 42.3 ** | (19.1) | 31.8 * | (16.9) | 33.3 * | (17.8) | 17.0 | (16.4) |
| Bachelor's Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 17.1 | (14.7) | 11.2 | (12.7) | 12.1 | (13.1) | 6.6 | (11.9) |
| Less than a Bachelor's Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 45.8 ** | (20.3) | 32.4 ** | (15.9) | 33.6 ** | (16.4) | 36.4 ** | (17.1) |
| Instruction Affected by General Resource Shortages | | | -0.9 | (3.4) | 1.8 | (3.0) | 2.3 | (3.1) | 2.1 | (2.8) |
| Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science | | | 0.1 | (3.6) | -0.2 | (3.0) | -0.2 | (3.0) | -0.6 | (2.9) |
| Fidelity of Curriculum Implementation | | | | | | | | | | |
| Mathematics Topics Taught | | | | | 2.3 | (2.7) | 2.5 | (2.6) | -1.1 | (2.6) |
| Effective Instructional Strategies for Encouraging Mathematics Reasoning | | | | | 45.9 *** | (6.5) | 45.8 *** | (6.6) | 40.5 *** | (6.6) |
| Effective Instructional Strategies for Teaching Science Inquiry | | | | | 8.8 | (6.9) | 8.1 | (7.0) | 1.2 | (7.0) |
| Students Like Learning Mathematics | | | | | 1.3 | (1.7) | 1.3 | (1.7) | 1.3 | (1.7) |
| Students Like Learning Science | | | | | -5.4 ** | (2.6) | -5.9 ** | (2.7) | -5.2 * | (2.7) |
| School Climate | | | | | | | | | | |
| Schools Have Few Discipline and Attendance Problems | | | | | | | -1.6 | (1.8) | -2.1 | (1.7) |
| Schools Have a Climate Supportive of Academic Success | | | | | | | 0.7 | (2.0) | -0.7 | (2.1) |

***p < .01, **p < .05, *p<.10

## STEM School Effectiveness in Singapore

### *School Resources Models*

The same progressive model building was completed next for Singapore, another high-performing Asian country in TIMSS 2007. Table 4.26 presents results from the two *school resources models* for Singapore. The percentage of total variance between schools after controlling for home resources did not change much compared to the *student home resources model* because none of the factors were significant predictors of STEM achievement. These results make sense based on the *TIMSS 2007 Encyclopedia* chapter on Singapore, which describes how the Ministry of Education sets high standards for teacher preparation and provides a wealth of resources to all schools for mathematics and science instruction (Quek et al., 2008).

Table 4.26. School Resources in Singapore

| | Student Home Resources | | Teacher Preparation | | General and STEM Resources | |
|---|---|---|---|---|---|---|
| **Variance Explained** | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 49% | | 50% | | 49% | |
| Percentage of Between School Variance Explained | 23% | | 22% | | 23% | |
| Percentage of Total Variance Explained | 18% | | 18% | | 18% | |
| **Student Home Resources Control** | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1165.7 *** | (10.2) | 1179.3 *** | (72.1) | 1165.4 *** | (10.2) |
| Student Home Resources Slope | 23.4 *** | (1.6) | 23.3 *** | (1.6) | 23.4 *** | (1.6) |
| **Teacher Preparation** | | | | | | |
| Postgraduate Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 2.2 | (77.5) | | |
| Bachelor's Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | -20.7 | (72.2) | | |
| Less than a Bachelor's Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | -14.7 | (76.6) | | |
| Postgraduate Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 60.9 | (53.1) | | |
| Bachelor's Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 5.0 | (18.0) | | |
| Less than a Bachelor's Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | -20.9 | (59.6) | | |
| **General School Resources** | | | | | | |
| Instruction Affected by General Resource Shortages | | | | | -0.1 | (5.5) |
| **Mathematics and Science Specific School Resources** | | | | | | |
| Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science | | | | | 4.3 | (7.0) |

***p < .01, **p < .05, *p<.10

### *Fidelity of Curriculum Implementation Models*

The next series of models built for Singapore explored fidelity of curriculum implementation factors. Table 4.27 presents results from the three fidelity of curriculum implementation models for Singapore. **“**Mathematics Topics Taught” was significantly related to STEM achievement as were “Effective Instructional Strategies for Encouraging Mathematics Reasoning” and “Effective Instructional Strategies for Teaching Science Inquiry”. Taken together (Model 2), **Fidelity to Curriculum Structure** and **Fidelity to Curriculum Process** reduced the total variance between schools after controlling for home resources 18 percentage points compared to the *student home resources model*.

### *School Climate Models*

Table 4.27 also shows that the two school climate factors, **Discipline and Attendance Problems** and **School Support for Academic Success** are significant predictors of STEM achievement in Singapore. Together, these two factors reduced the percentage of total variance between schools after controlling for home resources 12 percentage points.

Table 4.27. Fidelity of Curriculum Implementation and School Climate in Singapore

| | Student Home Resources | | Fidelity of Curriculum Implementation | | | | | | School Climate | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| Variance Explained | | | | | | | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 49% | | 41% | | 31% | | 31% | | 44% | | 37% | |
| Percentage of Between School Variance Explained | 23% | | 46% | | 65% | | 65% | | 37% | | 53% | |
| Percentage of Total Variance Explained | 18% | | 30% | | 40% | | 40% | | 25% | | 37% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1165.7 *** | (10.2) | 1161.3 *** | (8.5) | 1158.4 *** | (7.0) | 1158.3 *** | (7.0) | 1163.5 *** | (9.1) | 1161.5 *** | (7.8) |
| Student Home Resources Slope | 23.4 *** | (1.6) | 23.1 *** | (1.6) | 23.2 *** | (1.7) | 23.2 *** | (1.7) | 23.3 *** | (1.6) | 23.2 *** | (1.7) |
| Fidelity to Curriculum Structure | | | | | | | | | | | | |
| Mathematics Topics Taught | | | 31.1 *** | (3.6) | 22.4 *** | (3.4) | 22.1 *** | (3.5) | | | | |
| Fidelity to Curriculum Process | | | | | | | | | | | | |
| Effective Instructional Strategies for Encouraging Mathematics Reasoning | | | | | 49.8 *** | (12.0) | 49.5 *** | (12.3) | | | | |
| Effective Instructional Strategies for Teaching Science Inquiry | | | | | 61.9 *** | (11.8) | 62.8 *** | (11.9) | | | | |
| Participant Responsiveness | | | | | | | | | | | | |
| Students Like Learning Mathematics | | | | | | | -1.7 | (3.9) | | | | |
| Students Like Learning Science | | | | | | | 2.3 | (4.5) | | | | |
| Discipline and Attendance Problems | | | | | | | | | | | | |
| Schools Have Few Discipline and Attendance Problems | | | | | | | | | 22.3 *** | (3.8) | 18.2 *** | (3.5) |
| Support for Academic Success | | | | | | | | | | | | |
| Schools Have a Climate Supportive of Academic Success | | | | | | | | | | | 30.5 *** | (6.5) |

***p < .01, **p < .05, *p<.10

## *Complete School Effectiveness Model*

School resources, fidelity of curriculum implementation, and school climate factors were explored together in Singapore in the next series of models. Table 4.28 presents results from the *complete school effectiveness model* in Singapore, which includes a measure of school average student home resources in addition to all the other school effectiveness factors. Compared to the *student home resources model*, the *complete school effectiveness model* reduced the total variance between schools after controlling for home resources from 49 to 18 percent, demonstrating that the school effectiveness factors selected for analysis were important predictors that could help explain differences in STEM achievement between schools, even after controlling for student home resources.

Although “School Average Student Home Resources” was an important predictor for explaining differences in STEM achievement (the amount of total variance between schools after controlling for student home resources in Model 5 decreased 7 percentage points compared to Model 4), many other factors were still important for predicting differences in STEM achievement. More specifically, except for teachers with a postgraduate degree and major in science, all of the school factors that were significant in the previous models remained significant, even after accounting for the composition of the student body: **Fidelity to Curriculum Structure**, **Fidelity to Curriculum Process, Discipline and Attendance Problems**, and **School Support for Academic Success**.

Model 4 shows that Singapore is actually a model for how effective schools work. Even after controlling for student home resources within school, the more effective

Table 4.28. Final Predictors of STEM School Effectiveness in Singapore

| | Student Home Resources | | School Resources Combined | | School Resources and Fidelity of Curriculum Implementation | | School Resources, Fidelity of Curriculum Implementation, and School Climate | | Complete School Effectiveness Model | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| Variance Explained | | | | | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 49% | | 50% | | 31% | | 25% | | 18% | |
| Percentage of Between School Variance Explained | 23% | | 22% | | 64% | | 73% | | 83% | |
| Percentage of Total Variance Explained | 18% | | 18% | | 40% | | 44% | | 49% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1165.7 *** | (10.2) | 1177.7 *** | (71.1) | 1154.5 *** | (44.0) | 1156.3 *** | (39.2) | 1190.4 *** | (34.9) |
| Student Home Resources Slope | 23.4 *** | (1.6) | 23.3 *** | (1.6) | 23.1 *** | (1.7) | 23.0 *** | (1.7) | 22.2 *** | (1.7) |
| School Factors | | | | | | | | | | |
| School Average Student Home Resources | | | | | | | | | 50.6 *** | (6.9) |
| School Resources | | | | | | | | | | |
| Postgraduate Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 1.5 | (76.8) | -10.6 | (47.5) | -14.4 | (41.1) | -34.9 | (36.7) |
| Bachelor's Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | -21.2 | (71.7) | 3.4 | (43.9) | -1.4 | (38.3) | -30.4 | (34.5) |
| Less than a Bachelor's Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | -15.6 | (76.6) | 0.9 | (46.0) | 1.5 | (40.6) | -20.4 | (36.9) |

***p < .01, **p < .05, *p<.10

Table 4.28. Final Predictors of STEM School Effectiveness in Singapore (Continued)

| | Student Home Resources | | School Resources Combined | | School Resources and Fidelity of Curriculum Implementation | | School Resources, Fidelity of Curriculum Implementation, and School Climate | | Complete School Effectiveness Model | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| School Resources (Continued) | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Postgraduate Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 64.7 | (55.4) | 56.1 * | (30.8) | 48.4 ** | (19.5) | -9.6 | (17.4) |
| Bachelor's Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 7.2 | (18.6) | 1.3 | (13.5) | 1.9 | (12.4) | -9.4 | (11.0) |
| Less than a Bachelor's Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | -18.9 | (61.3) | -11.2 | (31.6) | 12.1 | (29.6) | 9.2 | (28.1) |
| Instruction Affected by General Resource Shortages | | | -1.6 | (5.6) | -0.6 | (4.3) | -0.4 | (4.0) | -0.3 | (3.4) |
| Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science | | | 5.4 | (7.1) | 0.7 | (5.3) | -0.8 | (4.5) | -1.7 | (3.8) |
| Fidelity of Curriculum Implementation | | | | | | | | | | |
| Mathematics Topics Taught | | | | | 22.0 *** | (3.5) | 15.5 *** | (3.7) | 7.0 * | (3.6) |
| Effective Instructional Strategies for Encouraging Mathematics Reasoning | | | | | 48.9 *** | (12.3) | 51.2 *** | (11.6) | 42.8 *** | (10.5) |
| Effective Instructional Strategies for Teaching Science Inquiry | | | | | 64.1 *** | (12.1) | 52.2 *** | (11.5) | 30.9 *** | (9.9) |
| Students Like Learning Mathematics | | | | | -2.4 | (3.6) | -4.7 | (3.4) | -1.9 | (3.3) |
| Students Like Learning Science | | | | | 2.7 | (4.5) | 0.0 | (4.1) | -4.3 | (3.6) |
| School Climate | | | | | | | | | | |
| Schools Have Few Discipline and Attendance Problems | | | | | | | 9.6 *** | (2.9) | 5.8 ** | (2.7) |
| Schools Have a Climate Supportive of Academic Success | | | | | | | 20.0 *** | (5.3) | 13.2 *** | (4.3) |

***p < .01, **p < .05, *p<.10

schools have better educated teachers (in science at least), teach the mathematics topics in the curriculum, apply effective instructional strategies in mathematics and science, and have supportive school climates with few problems. This remains true even after controlling for the social composition in the school (Model 5).

## STEM School Effectiveness in the Czech Republic

### *School Resources Models*

The next set of multilevel models was created for the Czech Republic, where at the eighth grade, science is taught as separate subjects. In TIMSS 2007, the Czech Republic had science achievement similar to Chinese Taipei and Singapore (although achievement was somewhat lower in mathematics).

The two *school resources models* were built first for the Czech Republic. Table 4.29 shows that the **Teacher Preparation** factors, specifically, the measures representing teachers with a postgraduate degree and major in mathematics or mathematics education, teachers with a postgraduate degree and major in science or science education, and teachers with less than a bachelor's degree and major in science or science education were significant predictors of STEM achievement after controlling for student home resources. The Czech Republic policy requiring teachers to receive a bachelor's degree and master's degree in two specific subject areas appears to be related to high STEM achievement.

Table 4.29. School Resources in the Czech Republic

| | Student Home Resources | | Teacher Preparation | | General and STEM Resources | |
|---|---|---|---|---|---|---|
| Variance Explained | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 23% | | 23% | | 23% | |
| Percentage of Between School Variance Explained | 28% | | 30% | | 27% | |
| Percentage of Total Variance Explained | 13% | | 14% | | 13% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1043.4 *** | (5.7) | 979.5 *** | (15.4) | 1043.2 *** | (5.7) |
| Student Home Resources Slope | 23.0 *** | (1.3) | 22.9 *** | (1.3) | 23.0 *** | (1.3) |
| Teacher Preparation | | | | | | |
| Postgraduate Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 37.3 *** | (11.0) | | |
| Less than a Bachelor's Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 2.6 | (16.0) | | |
| Postgraduate Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 31.2 * | (17.1) | | |
| Less than a Bachelor's Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 47.1 ** | (22.6) | | |
| General School Resources | | | | | | |
| Instruction Affected by General Resource Shortages | | | | | 3.5 | (3.6) |
| Mathematics and Science Specific School Resources | | | | | | |
| Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science | | | | | -1.5 | (5.0) |

***p < .01, **p < .05, *p<.10

### *Fidelity of Curriculum Implementation Models*

Next, the three fidelity of curriculum implementation models were built for the Czech Republic. Table 4.30 presents results from the three models and indicates there was no change in the variance explained from the *student home resources model*. That is, fidelity of curriculum implementation factors do not help to explain differences in STEM achievement after considering student home resources.

### *School Climate Models*

Table 4.30 also presents results from explorations into the positive and negative aspects of school climate in the Czech Republic through two *school climate models*. The second *school climate model* (Model 5) parallels the pattern of results in the United States and Singapore, indicating that **School Support for Academic Success** relates to achievement in addition to the absence of problems in a school, reducing the total variance between schools after controlling for home resources one additional percentage point beyond **Discipline and Attendance Problems** in isolation.

Table 4.30. Fidelity of Curriculum Implementation and School Climate in the Czech Republic

| | Student Home Resources | | Fidelity of Curriculum Implementation | | | | | | School Climate | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| Variance Explained | | | | | | | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 23% | | 23% | | 23% | | 23% | | 21% | | 20% | |
| Percentage of Between School Variance Explained | 28% | | 28% | | 28% | | 27% | | 39% | | 41% | |
| Percentage of Total Variance Explained | 13% | | 14% | | 13% | | 13% | | 16% | | 17% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1043.4 *** | (5.7) | 1043.2 *** | (5.6) | 1043.3 *** | (5.6) | 1043.6 *** | (5.6) | 1042.5 *** | (5.3) | 1042.5 *** | (5.2) |
| Student Home Resources Slope | 23.0 *** | (1.3) | 23.0 *** | (1.3) | 23.0 *** | (1.3) | 22.9 *** | (1.3) | 22.9 *** | (1.3) | 22.9 *** | (1.3) |
| Fidelity to Curriculum Structure | | | | | | | | | | | | |
| Mathematics Topics Taught | | | 6.8 | (5.5) | 6.4 | (5.4) | 6.7 | (5.4) | | | | |
| Fidelity to Curriculum Process | | | | | | | | | | | | |
| Effective Instructional Strategies for Encouraging Mathematics Reasoning | | | | | 4.0 | (3.7) | 3.8 | (3.6) | | | | |
| Effective Instructional Strategies for Teaching Science Inquiry | | | | | -0.4 | (5.0) | -0.5 | (4.9) | | | | |
| Participant Responsiveness | | | | | | | | | | | | |
| Students Like Learning Mathematics | | | | | | | 2.3 | (9.0) | | | | |
| Students Like Learning Science | | | | | | | -6.2 | (9.0) | | | | |
| Discipline and Attendance Problems | | | | | | | | | | | | |
| Schools Have Few Discipline and Attendance Problems | | | | | | | | | 12.0 *** | (2.8) | 11.1 *** | (2.6) |
| Support for Academic Success | | | | | | | | | | | | |
| Schools Have a Climate Supportive of Academic Success | | | | | | | | | | | 8.7 * | (4.5) |

***p < .01, **p < .05, *p<.10

### *Complete School Effectiveness Model*

School resources, fidelity of curriculum implementation, and school climate factors were explored together in the Czech Republic in the next series of models. Table 4.31 shows the *school resources, fidelity of curriculum implementation, and school climate model* (Model 4) reduced the total variance between schools after controlling for home resources by 3 percentage points from the *student home resources model*. Two **Teacher Preparation** measures were still significant after considering school climate factors, although the relationship between teacher quality and STEM achievement was somewhat mediated by school climate. **Discipline and Attendance Problems** also was a significant predictor of STEM achievement in this model.

The addition of "School Average Student Home Resources" over and above all the other school effectiveness factors was explored in the *complete school effectiveness model* for the Czech Republic. Compared to Model 4, the *complete school effectiveness model* (Model 5) reduced the amount of total variance between schools after controlling for home resources another 9 percentage points. In the Czech Republic, "School Average Student Home Resources" was a powerful predictor for explaining differences in STEM achievement between schools. Apparently, students from well-resourced homes are more likely to be in schools with well-prepared teachers, since the effect of the teacher preparation variables was reduced. Nevertheless, several factors were still important for predicting differences in STEM achievement, including **Teacher Preparation** (specifically, teachers with a postgraduate degree and major in mathematics), **Discipline and Attendance Problems**, and **School Support for Academic Success**. Also,

Table 4.31. Final Predictors of STEM School Effectiveness in the Czech Republic

| | Student Home Resources | | School Resources Combined | | School Resources and Fidelity of Curriculum Implementation | | School Resources, Fidelity of Curriculum Implementation, and School Climate | | Complete School Effectiveness Model | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| Variance Explained | | | | | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 23% | | 23% | | 23% | | 20% | | 11% | |
| Percentage of Between School Variance Explained | 28% | | 29% | | 30% | | 42% | | 71% | |
| Percentage of Total Variance Explained | 13% | | 14% | | 14% | | 17% | | 25% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1043.4 *** | (5.7) | 980.3 *** | (15.6) | 974.6 *** | (17.6) | 985.8 *** | (17.6) | 1008.1 *** | (13.3) |
| Student Home Resources Slope | 23.0 *** | (1.3) | 22.9 *** | (1.3) | 22.8 *** | (1.3) | 22.8 *** | (1.3) | 21.5 *** | (1.4) |
| School Factors | | | | | | | | | | |
| School Average Student Home Resources | | | | | | | | | 57.5 *** | (6.7) |
| School Resources | | | | | | | | | | |
| Postgraduate Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 35.6 *** | (11.3) | 37.7 *** | (12.8) | 33.7 *** | (12.7) | 28.7 *** | (10.3) |
| Less than a Bachelor's Degree and Major in Mathematics or Mathematics Education Compared to No Mathematics or Mathematics Education Major | | | 3.0 | (17.4) | 12.2 | (24.3) | 24.5 | (24.6) | 30.4 | (26.2) |
| Postgraduate Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 32.0 * | (17.1) | 36.2 * | (19.0) | 27.1 | (17.7) | 10.4 | (13.7) |
| Less than a Bachelor's Degree and Major in Science or Science Education Compared to No Science or Science Education Major | | | 45.3 ** | (21.6) | 49.8 ** | (21.2) | 37.3 * | (21.5) | 13.1 | (26.7) |

***p < .01, **p < .05, *p<.10

Table 4.31. Final Predictors of STEM School Effectiveness in the Czech Republic (Continued)

| | Student Home Resources | | School Resources Combined | | School Resources and Fidelity of Curriculum Implementation | | School Resources, Fidelity of Curriculum Implementation, and School Climate | | Complete School Effectiveness Model | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| School Resources (Continued) | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Instruction Affected by General Resource Shortages | | | 2.8 | (3.6) | 2.9 | (3.8) | -1.8 | (4.3) | -0.8 | (2.8) |
| Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science | | | -0.9 | (5.0) | -0.9 | (5.2) | -0.3 | (4.9) | 0.5 | (3.6) |
| Fidelity of Curriculum Implementation | | | | | | | | | | |
| Mathematics Topics Taught | | | | | 6.7 | (5.4) | 6.0 | (4.3) | 3.7 | (3.4) |
| Effective Instructional Strategies for Encouraging Mathematics Reasoning | | | | | 3.9 | (3.4) | 2.8 | (3.3) | 0.8 | (2.2) |
| Effective Instructional Strategies for Teaching Science Inquiry | | | | | -1.4 | (5.4) | 0.5 | (4.8) | -0.9 | (4.0) |
| Students Like Learning Mathematics | | | | | 4.9 | (9.3) | 12.3 | (8.6) | 17.3 ** | (6.8) |
| Students Like Learning Science | | | | | -9.4 | (8.5) | -7.9 | (7.7) | 4.7 | (5.8) |
| School Climate | | | | | | | | | | |
| Schools Have Few Discipline and Attendance Problems | | | | | | | 11.3 *** | (2.6) | 6.5 *** | (1.7) |
| Schools Have a Climate Supportive of Academic Success | | | | | | | 7.8 | (4.9) | 7.0 * | (4.1) |

***p < .01, **p < .05, *p<.10

"Students Like Learning Mathematics" became a significant predictor of STEM achievement.

## STEM School Effectiveness in Slovenia

### *School Resources Models*

Slovenia was selected for this research to represent a country that performed especially well in science in TIMSS 2007, most likely because science is taught as separate subjects at the eighth grade. However, Slovenia is a small, uniform country with little differences in STEM achievement between schools (only 10% of the total variance). **Teacher Preparation**, **General School Resources**, and **Mathematics and Science Specific School Resources** had little relationship with STEM achievement after controlling for student home resources (see Tables 4.4–4.6 and 4.8). Thus, the series of *school resources models* were not included in the final multilevel analyses for Slovenia.

### *Fidelity of Curriculum Implementation Models*

Fidelity of curriculum implementation factors seemed to relate to STEM achievement in Slovenia according to the explorations conducted in the first section of this chapter. Therefore, a series of models were built for Slovenia to examine the additive effects of **Fidelity to Curriculum Structure**, **Fidelity to Curriculum Process**, and **Participant Responsiveness** in relation to STEM achievement. Table 4.32 presents the results from the three models in Slovenia.

Examining the variance explained by the final *fidelity of curriculum implementation model* for Slovenia (Model 3) shows that fidelity of curriculum

Table 4.32. Fidelity of Curriculum Implementation and School Climate in Slovenia

| | Student Home Resources | | Fidelity of Curriculum Implementation | | | | | | School Climate | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | | Model 1 | | Model 2 | | Model 3 | | Model 4 | | Model 5 | |
| Variance Explained | | | | | | | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 8% | | 8% | | 8% | | 7% | | 8% | | 8% | |
| Percentage of Between School Variance Explained | 31% | | 34% | | 32% | | 39% | | 31% | | 30% | |
| Percentage of Total Variance Explained | 13% | | 13% | | 13% | | 14% | | 13% | | 13% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1040.3 *** | (4.4) | 1040.4 *** | (4.3) | 1040.2 *** | (4.3) | 1040.6 *** | (4.0) | 1040.1 *** | (4.4) | 1040.1 *** | (4.4) |
| Student Home Resources Slope | 27.0 *** | (1.4) | 26.9 *** | (1.4) | 26.9 *** | (1.4) | 27.0 *** | (1.4) | 27.1 *** | (1.4) | 27.1 *** | (1.4) |
| Fidelity to Curriculum Structure | | | | | | | | | | | | |
| Mathematics Topics Taught | | | 6.9 ** | (3.3) | 6.7 * | (3.5) | 5.6 * | (3.2) | | | | |
| Fidelity to Curriculum Process | | | | | | | | | | | | |
| Effective Instructional Strategies for Encouraging Mathematics Reasoning | | | | | 4.0 | (6.3) | 0.9 | (5.7) | | | | |
| Effective Instructional Strategies for Teaching Science Inquiry | | | | | 3.9 | (9.5) | 5.8 | (8.4) | | | | |
| Participant Responsiveness | | | | | | | | | | | | |
| Students Like Learning Mathematics | | | | | | | 9.4 ** | (3.9) | | | | |
| Students Like Learning Science | | | | | | | 7.1 *** | (2.4) | | | | |
| Discipline and Attendance Problems | | | | | | | | | | | | |
| Schools Have Few Discipline and Attendance Problems | | | | | | | | | 1.9 | (2.0) | 1.8 | (1.9) |
| Support for Academic Success | | | | | | | | | | | | |
| Schools Have a Climate Supportive of Academic Success | | | | | | | | | | | -2.2 | (6.2) |

***p < .01, **p < .05, *p<.10

implementation factors reduced the amount of total variance between schools after controlling for home resources by one percentage point compared to the *student home resources model*. **Fidelity to Curriculum Structure** as measured by "Mathematics Topics Taught" and **Participant Responsiveness** were statistically significant predictors of STEM achievement, including both "Students Like Mathematics" and "Students Like Science".

### School Climate Models

Table 4.32 also presents the results from the two *school climate models* for Slovenia. These factors did not explain any between school differences in STEM achievement after considering student home resources.

### Complete School Effectiveness Model

Table 4.33 shows the combined models for Slovenia (keeping in mind the *school resources models* were not used). Model 2 was built to explore the additive effects of fidelity of curriculum implementation factors and school climate factors in relation to STEM achievement. "Students Like Learning Mathematics" and "Students Like Learning Science" were significant predictors of STEM achievement in the model. (When examined together with school climate factors, "Mathematics Topics Taught" was no longer significant as shown in the final *fidelity of curriculum implementation model* in Table 4.32).

Finally, the *complete school effectiveness model* with the "School Average Student Home Resources" measure was examined. Unlike the other countries studied, "School Average Student Home Resources" was not significantly related to STEM achievement in Slovenia after considering individual student home resources and the other school effectiveness factors. This supports the other results found in this study demonstrating the relative similarities between schools in Slovenia.

Table 4.33. Final Predictors of STEM School Effectiveness in Slovenia

| | Student Home Resources | | Fidelity of Curriculum Implementation and School Climate | | Complete School Effectiveness Model | |
|---|---|---|---|---|---|---|
| | Model 1 | | Model 2 | | Model 3 | |
| Variance Explained | | | | | | |
| Percentage of Total Variance Between Schools After Controlling for Home Resources | 8% | | 7% | | 7% | |
| Percentage of Between School Variance Explained | 31% | | 38% | | 39% | |
| Percentage of Total Variance Explained | 13% | | 14% | | 14% | |
| Student Home Resources Control | Coefficient | S.E. | Coefficient | S.E. | Coefficient | S.E. |
| Intercept | 1040.3 *** | (4.4) | 1040.5 *** | (4.0) | 1040.2 *** | (4.0) |
| Student Home Resources Slope | 27.0 *** | (1.4) | 27.0 *** | (1.4) | 27.0 *** | (1.4) |
| School Factors | | | | | | |
| School Average Student Home Resources | | | | | -10.4 | (8.2) |
| Fidelity of Curriculum Implementation | | | | | | |
| Mathematics Topics Taught | | | 5.4 | (3.3) | 5.3 | (3.2) |
| Effective Instructional Strategies for Encouraging Mathematics Reasoning | | | 0.8 | (5.9) | 0.0 | (5.7) |
| Effective Instructional Strategies for Teaching Science Inquiry | | | 5.6 | (8.3) | 7.3 | (8.0) |
| Students Like Learning Mathematics | | | 9.1 ** | (3.8) | 8.9 ** | (3.7) |
| Students Like Learning Science | | | 7.2 *** | (2.4) | 7.1 *** | (2.3) |
| School Climate | | | | | | |
| Schools Have Few Discipline and Attendance Problems | | | 0.6 | (1.8) | 0.3 | (1.7) |
| Schools Have a Climate Supportive of Academic Success | | | -2.2 | (6.1) | -2.5 | (6.0) |

***p < .01, **p < .05, *p<.10

## Summary of STEM School Effectiveness Factors Across Countries

**Teacher Preparation** measures were significant in Chinese Taipei, Singapore, and the Czech Republic after considering all of the other school effectiveness factors as well as individual student home resources.

Reflecting the fact that most schools are well-resourced in the five countries studied, neither **General Resource Shortages** nor **Mathematics and Science Specific Resource Shortages** were significant predictors of STEM achievement in this research.

Many fidelity of curriculum implementation factors still persisted in their importance to STEM achievement even after considering school resources and school climate, and controlling for student home resources. **Fidelity to Curriculum Structure** as measured by "Mathematics Topics Taught" was a significant predictor of STEM achievement in the United States and Singapore

Measures representing **Fidelity to Curriculum Process** were significant predictors for three of the five countries studied. Both "Effective Instructional Strategies for Encouraging Mathematics Reasoning" and "Effective Instructional Strategies for Teaching Science Inquiry" were significant predictors of STEM achievement in Singapore's complete model, even after considering all other factors. In Chinese Taipei, "Effective Instructional Strategies for Encouraging Mathematics Reasoning" was significantly related to STEM achievement and "Effective Instructional Strategies for Teaching Science Inquiry" actually became a significant predictor of STEM achievement in the United States in the final model after considering school resources and school climate factors.

In the *complete school effectiveness* models, **Participant Responsiveness** was a significant predictor of STEM achievement in the Czech Republic ("Students Like Learning Mathematics") and Slovenia ("Students Like Learning Mathematics" and "Students Like Learning Science").

School climate factors remained important predictors of STEM achievement in the United States, Singapore, and the Czech Republic, even after considering all other school effectiveness factors and controlling for student home resources. Results indicated that both **School Support for Academic Success** and absence of **Discipline and Attendance Problems** are associated with higher STEM achievement.

# Chapter 5

## *Discussion*

### Overview of the Study

As previously noted, results from TIMSS 2007 showed a gap in mathematics and science achievement between students in the United States and those in the top–performing countries, particularly in more complicated content and more complex reasoning. Thus, to provide a perspective for these achievement differences, TIMSS 2007 data were used to investigate how school effectiveness factors known to be strongly associated with higher STEM achievement operated in the United States compared to Chinese Taipei, the Czech Republic, Singapore, and Slovenia. Multilevel modeling was used to control for the within school effect of student home resources and examine STEM achievement in relation to school resources, fidelity of curriculum implementation, and school climate factors in the five countries. The policy implications of these findings are discussed, focusing on how the U.S. can improve STEM education.

A secondary purpose of this dissertation research was to help prepare for multilevel modeling planned for the TIMSS and PIRLS 2011 data. This research involving five countries provided a good preliminary experience in conducting multilevel modeling in an international context. For TIMSS and PIRLS 2011 data, two different analyses at the fourth grade involving nearly 40 countries will be conducted, one comparing school effectiveness factors across mathematics, science, and reading, and the other comparing the relative efforts schools devote to reading, mathematics, and science. Thus, the methodological implications for analyzing TIMSS and PIRLS 2011 are presented.

## School Effectiveness Factors

The effective schools literature and findings from TIMSS show the following are important school factors for increasing STEM achievement at the eighth grade.

A rigorous well-articulated curriculum is fundamental to achievement in STEM areas. Countries with established standards and clearly defined curricular goals provide straightforward expectations for the content to be delivered and student benchmarks to be reached.

### *School Resources*

Teacher preparation also is important for STEM achievement since teachers must have the capacity to teach the mathematics and science topics in the curriculum. Research has shown teachers' degree of success in delivering the curriculum, as measured by student performance, is related to their preparedness to teach, including their level of formal education completed and whether or not they hold a subject–specific academic degree. That is, level of formal education together with content focus is an important factor related to student STEM achievement.

Also, the resources available in a school can facilitate teachers' delivery of the curriculum. Of course general school resources, such as instructional materials, budget for supplies, heating and lighting systems, and instructional space are required for students to achieve in all subjects. In order to successfully implement a STEM curriculum, teachers must also have access to the technological and other resource components required to deliver such a curriculum (e.g., computer hardware, networking capabilities, etc.). Thus, STEM achievement has been associated with students' access to resources specific to mathematics and science learning, such as specialized instructional materials as well as laboratory equipment.

### *Fidelity of Curriculum Implementation*

Fidelity of curriculum implementation factors, adherence to the curriculum, instructional strategies used, and student attitudes toward learning also are important for STEM achievement. Higher–performing schools tend to actually teach the intended mathematics and science curriculum and use effective instructional strategies to do so. Student positive attitudes toward learning mathematics and science can foster higher STEM achievement.

### *School Climate*

Finally, school climate factors are key for STEM achievement. School climates that support academic success are associated with higher achievement, including those having few discipline and attendance problems, or having a more positive environment for student learning. Schools with high levels of student behavioral problems do not constitute a climate conducive to learning since disruptions, bullying, and absenteeism impede students' opportunity to learn. On the other hand, school climates that promote an atmosphere of success through support from principals, teachers, and parents can lead to increased STEM achievement by creating a positive professional community in which teachers can successfully deliver the curriculum and students can learn.

## Major Findings

For the most part, schools in the United States as well as Chinese Taipei, Singapore, the Czech Republic, and Slovenia have the characteristics of effective schools to some degree. All five of the countries can be considered to have well-functioning economies and well-developed education systems. However, in each country, the school effectiveness characteristics were not spread entirely evenly across schools, with some schools being more effective than others.

Table 5.1 presents the percentage of total variation in STEM achievement due to differences within and between schools without considering any explanatory variables (Columns 1 & 2). There is a range in the amount of between school variance across the countries under study, with Slovenia showing the smallest differences between schools (10% of total variability) and Singapore showing the largest differences between schools (52% of total variability).

As customary in school effectiveness research, the first step in examining school effectiveness is to control statistically for the variance in student achievement that is due to student home background. Therefore, a series of multilevel models were built to examine school effectiveness factors, beginning with modeling the relationship between student home resources and STEM achievement. Table 5.1 presents the percentage of student within school variance and percentage of between school variance explained after controlling for student home resources (Columns 3 &4). Across the five countries studied, controlling for student home resources accounted for 6 percent (in the United States) to 17 percent (in Chinese Taipei) of the total variation in STEM achievement due to differences within schools. Also, controlling for student home resources accounted for between 23 percent in Singapore and 55 percent in Chinese Taipei of the total variation in STEM achievement due to differences between schools.

Table 5.1. Percentage of Total Variation in STEM Achievement Due to Differences Within and Between Schools and Percentage of Within and Between School Variance Explained by Student Home Resources

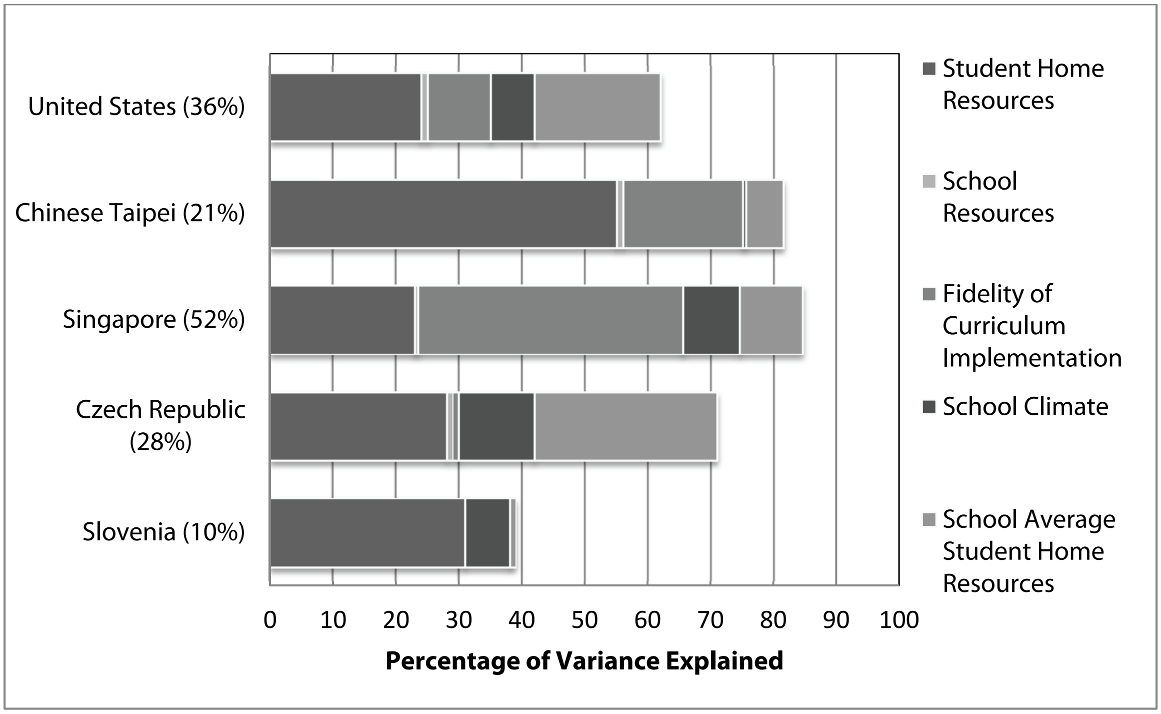
| | Percentage of Total Variance Within and Between Schools | | Percentages of Within and Between School Variance Accounted for by Student Home Resources | |
|---|---|---|---|---|
| Country | Percentage of Total Variation in STEM Achievement Due to Differences Within Schools | Percentage of Total Variation in STEM Achievement Due to Differences Between Schools | Percentage of Student Within School Variance Explained by Student Home Resources | Percentage of Between School Variance Explained by Student Home Resources |
| Chinese Taipei | 79% | 21% | 17% | 55% |
| Czech Republic | 72% | 28% | 8% | 28% |
| Singapore | 48% | 52% | 14% | 23% |
| Slovenia | 90% | 10% | 11% | 31% |
| United States | 64% | 36% | 6% | 24% |

Examining the percentage of between school variance in STEM achievement explained by successive school effectiveness models after controlling for student home resources provided an indication of the degree to which school effectiveness factors were associated with STEM achievement. First, school resource factors were explored in relation to STEM achievement, controlling for student home resources. Second, fidelity of curriculum implementation factors were explored after considering school resource factors and controlling for student home resources. Third, school climate factors were examined, after considering school resources and fidelity of curriculum implementation factors and controlling for student home resources. Finally, the social composition of the student body was examined after considering all other school effectiveness factors and controlling for individual student home resources.

Figure 5.1 presents the percentage of between school variance in STEM achievement explained by the additive school effectiveness models applied to each of the five countries studied. Specifically, each segment of a country's bar shows the between school variance in STEM achievement explained by a set of factors once the influence of the other factors in the model was accounted for. For example, in the United States, after considering school resources and controlling for individual student home resources, fidelity of curriculum implementation factors accounted for an additional 10 percent of the between school variance in STEM achievement compared to when school resources and individual student home resources were examined alone. Taken together, the school effectiveness models explained between 39 percent (in Slovenia) and 84 percent (in Singapore) of the between school variance in STEM achievement.

More specifically, the multilevel modeling for the United States indicated that more effective schools in the U.S. taught the TIMSS mathematics topics, used effective instructional strategies, and had supportive environments.  Even controlling for student home background, schools teaching more mathematics topics and more frequently using effective instructional strategies for teaching science inquiry had higher STEM achievement. The different school climates for learning also were important for STEM achievement. Everything else being equal, students had higher STEM achievement if they attended schools with few student discipline and attendance problems, and where there was support from principals, teachers, and parents for academic achievement.

Figure 5.1. Percentage of Between School Variance in STEM Achievement Explained by Successive School Effectiveness Models




( ) Percentage of total variation in STEM achievement due to differences between schools. See Table 5.1

**Student Home Resources**
Student Home Resources

**School Resources**
Teacher Preparation
Instruction Affected by General Resource Shortages
Instruction Affected by Mathematics and Science Specific Resource Shortages

**Fidelity of Curriculum Implementation**
Mathematics Topics Taught
Effective Instructional Strategies for Encouraging Mathematics Reasoning
Effective Instructional Strategies for Teaching Science Inquiry
Students Like Learning Mathematics
Students Like Learning Science

**School Climate**
Schools Have Few Discipline and Attendance Problems
Schools Have a Climate Supportive of Academic Success

**School Average Student Home Resources**
School Average Student Home Resources

In comparison, results from the multilevel analyses for Chinese Taipei indicated that students uniformly are taught the same rigorous curriculum. Also, they all are in schools with mostly well-behaved students focused on academic success. However, differences in teacher preparation and instructional strategies explained differences in STEM achievement above and beyond student home resources. Specifically, teacher preparation in science and teachers using effective instructional strategies for encouraging mathematics reasoning were associated with higher STEM achievement in Chinese Taipei.

The results from Singapore were relatively similar to those for the United States in that higher STEM achievement was associated with schools where more mathematics topics from the curriculum were being taught and where more effective instructional strategies used to encourage mathematics reasoning or to teach science inquiry. Also, schools with orderly environments for learning and greater support for academic success had higher STEM achievement than schools without.

Multilevel analysis results revealed that in the Czech Republic, teacher preparation in mathematics, student attitudes toward learning, and school climate factors were important for predicting differences in STEM achievement between schools. Similar to the United States, differences in students' learning environments were important for explaining between school differences in STEM achievement. In particular, schools with fewer problems and with greater support for learning were associated with higher STEM achievement.

Slovenia is a small country with a high degree of uniformity across schools (Pavesic, & Svetlik, Forthcoming). Despite the similarity of schools though, schools where students had positive attitudes toward learning mathematics and science had higher STEM achievement.

## Education Policy Implications for the United States

Results from the multilevel analysis indicated that schools teaching the curriculum and using instructional strategies involving reasoning and inquiry were positively related to STEM achievement. Currently, most U.S. states are in the process of adopting the Common Core State Standards (CCSS) that set high expectations for student achievement (Common Core State Standards Initiative, 2010). While setting higher curricular standards is a step in the right direction, having a rigorous curriculum is not enough. The U.S. needs well-prepared teachers to teach these standards, qualified in STEM areas. In Chinese Taipei and Singapore, education is a valued discipline, reflected in teacher education and recruitment policies. In Singapore, for example, the Ministry of Education recruits teachers from the top third of each university cohort (Quek et al., 2008). Teacher education in the Czech Republic reflects a focus on higher education levels and specializations in the subject area teachers plan to teach through policies requiring that teachers receive a master's degree and specialize in two subject areas. To make significant improvements in STEM education in the United States, the teaching profession needs to become "desirable" enough to attract individuals highly competent in mathematics and science.

In addition, teacher education programs will need to adapt to providing high level training experiences in pedagogical effectiveness. Teacher education programs should prepare teachers to use instructional strategies that have been found to be effective in preparing students for success in STEM subjects. In mathematics, these strategies include asking students to explain their answers or decide on their own procedures for solving complex problems and in science, asking students to design or plan an experiment or give explanations about something they are studying.

Teachers in the U.S. also must have the support and time to teach all of the topics in the intended curriculum. Results from this dissertation research indicated that teaching the mathematics topics in the curriculum was related to higher STEM achievement. When new curricula are introduced in the U.S., policymakers should implement support programs for teachers to discuss the content, goals, and approaches to teaching the curriculum.

Also, this research showed that school environment was strongly associated with high STEM achievement. Both absence of discipline and attendance problems as well as a school climate supportive of academic success were important predictors of student achievement in three of the countries, including the United States. There is room for improvement in many schools in the United States in terms of reducing the negative school climate factors related to bullying, absenteeism, and general disruptive behavior. The positive factors related to school climate also could be improved by increasing teachers' understanding of the school's curricular goals, supporting teachers in implementing the school curriculum, and increasing teacher expectations for student achievement, parental support for student achievement, and students' desire to do well.

School effectiveness methodology focuses on examining differences between schools, but sometimes important explanatory factors do not differ between schools, making it difficult to demonstrate the effectiveness of the factors. This can be a positive situation when most schools are providing equitable educational experiences at a high level of quality. For example, almost all of the students in Singapore and Chinese Taipei attended schools with well-staffed science laboratories, so this variable was not included in the study. However, it is clear that science laboratories are fundamental to improving STEM achievement and because only 17 percent of students in the United States attend schools with science laboratories and assistance available when students conduct experiments, the United States could consider investing in this area.

It also should be recognized that the other four participants in this study are relatively small countries with highly centralized education systems. Thus, they can mandate and regulate criteria for teacher preparation, curriculum development and implementation, and school environments. There was less scope for finding differences in the effectiveness factors in these countries than in the United States. For example, having an orderly school environment was not an explanatory factor in Chinese Taipei and Slovenia, but this was because their schools had few discipline and attendance problems. In general, however, an orderly school environment is undoubtedly a prerequisite for high STEM achievement.

Another challenge in interpreting results from school effectiveness studies is that differences between schools can arise because of organizational, structural, or cultural reasons. For example, in Singapore and the Czech Republic, students are streamed or tracked into different courses of study that occur in different schools. In the United States, education is funded locally and students attend schools in their neighborhoods. Thus, some schools in the United States are much better resourced than others. All of the factors studied are important to STEM achievement, although a few of them are not apparent in the results because they go hand in hand with well-resourced schools in high socioeconomic neighborhoods.

## Methodological Implications for Analyzing TIMSS and PIRLS 2011

This dissertation research showed the potential of using TIMSS data as a basis for conducting school effectiveness analyses to examine factors related to achievement in different country contexts. However, the multilevel models used to examine the context of schooling become complicated very quickly (e.g., modeling student achievement within schools as well as examining factors between schools). It is therefore recommended that the models used be as parsimonious as possible while still providing information about differences between educational contexts. For example, teacher preparation needs to be modeled more effectively.

Even when studying only five countries, interpretation of the results was complicated by organizational, structural, and cultural considerations within each of the countries. This dissertation research has shown the importance of having a conceptual framework to guide model-building decisions. For school effectiveness analyses using all TIMSS and PIRLS countries, there needs to be a well-articulated framework of school effectiveness factors and increased understanding of how these factors are likely to operate country to country.

There will be much more variation in how school effectiveness factors operate across all of the TIMSS and PIRLS countries in 2011. The applicability of the school effectiveness factors studied across these diverse contexts should be considered. For example, teacher preparation has a widely different context in Honduras where there are many challenges to prepare teachers adequately compared to Qatar, which relies on teachers educated in different countries to meet the needs of their education system. More research will need to be conducted to provide information about how the different school effectiveness factors operate across the diverse set of countries in TIMSS and PIRLS. Smaller investigations using groups of similar countries could be conducted to see how the selected factors operate in different contexts.

Presentation of results also is important to consider, particularly when using more complex statistical models. This research has shown that presenting comprehensive results that also are comparable across countries is challenging. More work will need to be done to explore graphical presentations that could be used for showing results from different school effectiveness models for nearly 40 countries.

## Limitations of the Study

A number of limitations of educational research generally have been addressed by using the TIMSS 2007 database. More specifically, each country's data was collected from a large, nationally representative sample of schools and students. Also, the STEM achievement measure was extremely robust, and the school effectiveness factors were measured by reliable scales representing the constructs.

Accordingly, the data were an excellent basis for a descriptive cross national study. It is the case that descriptive studies, including those using multilevel modeling, cannot provide cause and effect results. The findings reported herein are based on correlational data and are suggestive rather than definitive.

# References

Abadzi, H. (2007). *Absenteeism and beyond: Instructional time loss and consequences.* (World Bank Policy Research Working Paper Number 4376). Washington, D.C.: The World Bank Independent Evaluation Group.

Adams, R. J. and Khoo, S. T (1996). *Quest - The interactive test analysis system*. Camberwell, Australia: Australian Council for Educational Research.

Akiba, M., LeTendre, G. K., & Scribner, J. P. (2007). Teacher quality, opportunity gap, and national achievement in 46 countries. *Educational Researcher, 36*(7), 369-387.

Aladejana, F., & Aderibigbe, O. (2007). Science laboratory environment and academic performance. *Journal of Science Education and Technology, 16*(6), 500-506.

Andrich, D. (2010). Understanding the response structure and process in the polytomous Rasch model. In M. L. Nering, & R. Ostini (Eds.), *Handbook of polytomous item response theory models* (pp. 123-152). New York, NY: Taylor & Francis Group.

Archibald, S. (2006). Narrowing in on educational resources that do affect student achievement. *Peabody Journal of Education, 81*(4), 23-42.

Austin, G., & Bailey, J. (2008). *What teachers and other staff tell us about California schools: Statewide results of the 2004-06 California school climate survey.* (Report prepared for the California Department of Education Safe and Healthy Kids Program Office). Sacramento, CA: WestEd.

Barber, M., & Mourshed, M. (2007). *How the world's best-performing school systems come out on top.* London: McKinsey & Company.

Barton, P. E. (2003). *Parsing the achievement gap: Baselines for tracking progress.* (Policy Information Report). Princeton, NJ: Educational Testing Service.

Basl, J., & Tomasek, V. (2008). The Czech Republic. In I. V. S. Mullis, M. O. Martin, J. F. Olson, D. R. Berger, D. Milne, & G. M. Stanco. (Eds.). (2008). *TIMSS 2007 encyclopedia: A guide to mathematics and science education around the world*. (pp. 621-634). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Bickel, R. (2007). *Multilevel analysis for applied research: It's just regression!*. New York, NY: The Guilford Press.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Braun, H., Coley, R., Jia, Y., & Trapani, C. (2009). *Exploring what works in science instruction: A look at the eighth-grade science classroom.* (Policy Information Report). Princeton, NJ: Educational Testing Service.

Buchmann, C. (2002). Measuring family background in international studies of education: Conceptual issues and methodological challenges. In A. C. Porter, & A.

Gamoran (Eds.), *Methodological advances in cross-national surveys of educational achievement* (pp. 150–197). Washington, DC: National Academy Press.

Bureau of Labor Statistics. (2007). *"STEM" occupations and job growth.* Retrieved from http://www.bls.gov/opub/ted/2007/jun/wk4/art04.htm

Byrnes, J. P., & Miller, D. C. (2007). The relative importance of predictors of math and science achievement: An opportunity-propensity analysis. *Contemporary Educational Psychology, 32*(4), 599-629.

Chevalier, A., & Lanot, G. (2002). The relative effect of family characteristics and financial situation on educational achievement. *Education Economics, 10*(2), 165-181.

Committee on Highly Successful Schools or Programs in K-12 STEM Education, & National Research Council. (2011). *Successful K-12 STEM education: Identifying effective approaches in science, technology, engineering, and mathematics*. Washington, D.C.: The National Academies Press.

Common Core State Standards Initiative. (2010). *About the standards.* Retrieved from http://www.corestandards.org/about-the-standards

Creemers, B. P. M., & Kyriakides, L. (2006). Critical analysis of the current approaches to modeling educational effectiveness: The importance of establishing a dynamic model. *School Effectiveness and School Improvement, 17*(3), 347-366.

Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. New York, NY: Taylor & Francis.

Creemers, B. P. M., Kyriakides, L., & Sammons, P. (2010). *Methodological advances in educational effectiveness research*. New York, NY: Routledge.

Crocker, L., & Algina, J. (1986). *Introduction to classical & modern test theory*. NY: Holt, Rinehart & Winston.

Crone, L. J., Lang, M. H., Teddlie, C., & Franklin, B. J. (1995). Achievement measures of school effectiveness: Comparison of model stability across years. *Applied Measurement in Education, 8*(4), 353-365.

Darling-Hammond, L. (2000). Teacher quality and student achievement: A review of state policy evidence. *Education Policy Analysis Archives, 8*(1) Retrieved from http://epaa.asu.edu/epaa/v10n12/

DeVellis, R. F. (2003). *Scale development: Theory and applications* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.

Dunteman, G. H. (1989). *Principal components analysis*. Newbury Park, CA: Sage Publications, Inc.

Erberber, E., Arora, A., & Preuschoff, C. (2008). Developing the TIMSS 2007 background questionnaires. In J. F. Olson, M. O. Martin & I. V. S. Mullis (Eds.), *TIMSS 2007 technical report* (pp. 45-62). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Fogleman, J., McNeill, K. L., & Krajcik, J. (2011). Examining the effect of teachers' adaptations of a middle school science inquiry-oriented curriculum unit on student learning. *Journal of Research in Science Teaching, 48*(2), 149-169.

Foy, P., Galia, J., & Li, I. (2008). Scaling the data from the TIMSS 2007 mathematics and science assessments. In J. F. Olson, M. O. Martin & I. V. S. Mullis (Eds.), *TIMSS 2007 technical report* (pp. 225-279). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Foy, P., & Olson, J. F. (Eds.). (2009). *TIMSS 2007 international database and user guide*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Fulton, K., & Britton, T. (2011). *STEM teachers in professional learning communities: From good teachers to great teaching.* Washington, D.C.: National Commission on Teaching and America's Future.

Goldhaber, D. D., & Brewer, D. J. (2000). Does teacher certification matter? High school teacher certification status and student achievement. *Educational Evaluation and Policy Analysis, 22*(2), 129-145.

Gonzales, P., Williams, T., Jocelyn, L., Roey, S., Kastberg, D., & Brenwald, S. (2008). *Highlights from TIMSS 2007: Mathematics and science achievement of U.S. fourth and eighth-grade students in an international context.* Washington, D.C.: National Center for Education Statistics, Institute of Education Sciences, U.S. Department of Education.

Grossman, P., Schoenfeld, A., & Lee, C. (2005). Teaching subject matter. In L. Darling-Hammond, & J. Bransford (Eds.), *Preparing teachers for a changing world: What teachers should learn and be able to do* (pp. 201-231). San Francisco, CA: John Wiley & Sons, Inc.

Hanushek, E., Jamison, D. T., Jamison, E. A., & Woessmann, L. (2008). Education and economic growth. *Education Next, 8*(2), 62-70.

Hanushek, E. A., & Woessmann, L. (2010). *The economics of international differences in educational achievement.* (NBER Working Paper 15949). Cambridge, MA: National Bureau of Economic Research.

Hattie, J. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. New York, NY: Taylor & Francis.

Hill, H. C., & Lubienski, S. T. (2007). Teachers' mathematics knowledge for teaching and school context. *Educational Policy, 21*(5), 747-768.

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

House, J. D. (2005). Classroom instruction and science achievement in Japan, Hong Kong, and Chinese Taipei: Results from the TIMSS 1999 assessment. *International Journal of Instructional Media, 32*(3), 295-311.

Hoy, W. K., Tarter, C. J., & Hoy, A. W. (2006). Academic optimism of schools: A force for student achievement. *American Educational Research Journal, 43*(3), 425-446.

Husen, T. (1973). Foreword. In L. C. Comber, & J. P. Keeves (Eds.), *Science education in nineteen countries: An empirical study* (pp. 9-16). Stockholm: Almqvist & Wiksell.

Husen, T. (1996). A brief history of the international association for the evaluation of educational achievement. *Assessment in Education: Principles, Policy & Practice, 3*(2), 129-141.

Jen, T., Lee, C., Lo, P., Chang, C., & Chen, K. (2008). Chinese Taipei. In I. V. S. Mullis, M. O. Martin, J. F. Olson, D. R. Berger, D. Milne, & G. M. Stanco. (Eds.). (2008). *TIMSS 2007 encyclopedia: A guide to mathematics and science education around the world*. (pp. 621-634). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Johnson, B., & Stevens, J. J. (2006). Student achievement and elementary teachers' perceptions of school climate. *Learning Environments Research, 9*(2), 111-122.

Joncas, M. (2008). TIMSS 2007 sampling weights and participation rates. In J. F. Olson, M. O. Martin, & I. V. S. Mullis (Eds.), *TIMSS 2007 technical report* (pp. 153-192). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Keene, L. (2008). United States. In I. V. S. Mullis, M. O. Martin, J. F. Olson, D. R. Berger, D. Milne, & G. M. Stanco. (Eds.). (2008). *TIMSS 2007 encyclopedia: A guide to mathematics and science education around the world*. (pp. 621-634). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Kellaghan, T. (1996). IEA studies and educational policy. *Assessment in Education: Principles, Policy & Practice, 3*(2), 143-160.

Koth, C. W., Bradshaw, C. P., & Leaf, P. J. (2008). A multilevel study of predictors of student perceptions of school climate: The effect of classroom-level factors. *Journal of Educational Psychology, 100*(1), 96-104.

Lavy, V. (2010). *Do differences in School's instruction time explain international achievement gaps in math, science, and reading? Evidence from developed and developing countries.* (Working Paper 16227). Cambridge, MA: National Bureau of Economic Research.

Lee, J. W., & Barro, R. J. (2001). Schooling quality in a cross–section of countries. *Economica, 68*(272), 465-488.

Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Hoboken, NJ: John Wiley & Sons, Inc.

Lubienski, S. T. (2006). Examining instruction, achievement, and equity with NAEP mathematics data. *Education Policy Analysis Archives, 14*(14), 1-33.

Lubienski, S. T., Lubienski, C., & Crane, C. (2008). Achievement differences and school type: The role of school climate, teacher certification, and instruction. *American Journal of Education, 115*(1), 97-138.

Luyten, H., Visscher, A., & Witziers, B. (2005). School effectiveness research: From a review of the criticism to recommendations for further development. *School Effectiveness and School Improvement, 16*(3), 249-279.

Ma, X., & Xu, J. (2004). Determining the causal ordering between attitude toward mathematics and achievement in mathematics. *American Journal of Education, 110*(3), 256-280.

Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: The role of home and school factors. *Educational Research and Evaluation, 12*(2), 105-128.

Martin, M. O., Mullis, I. V. S., & Foy, P. (2008). *TIMSS 2007 international science report: Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Martin, M. O., Mullis, I. V. S., Gregory, K. D., Hoyle, C., & Shen, C. (2000). *Effective schools in science and mathematics.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Martins, L., & Veiga, P. (2010). Do inequalities in parents' education play an important role in PISA students' mathematics achievement test score disparities? *Economics of Education Review, 29*(6), 1016-1033.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika, 47*(2), 149-174.

Masters, G. N., & Wright, B. D. (1997). The partial credit model. In W. J. van der Linden, & R. K. Hambleton (Eds.), *Handbook of modern item response theory* (pp. 101-121). New York, NY: Springer Verlag.

Mayer, D. P., Mullens, J. E., & Moore, M. T. (2000). *Monitoring school quality an indicators report* (NCES Statistical Analysis Report No. 2001-03). Washington, D.C.: U.S. Department of Education.

McLaughlin, M., McGrath, D. J., Burian-Fitzgerald, M., Lanahan, L., Scotchmer, M., Enyeart, C., & Salganik, L. (2005). *Student content engagement as a construct for the measurement of effective classroom instruction and teacher knowledge.* Washington, D.C.: American Institutes for Research.

McNeill, K. L., Pimentel, D. S., & Strauss, E. G. (2011). The Impact of High School Science Teachers' Beliefs, Curricular Enactments and Experience on Student Learning During an Inquiry-based Urban Ecology Curriculum. *International Journal of Science Education*, 1–37.

Merchlinsky, S., & Hansen-Grafton, B. (2007). *Considering the role of fidelity of implementation (FOI) in science education research: Evaluation and science specialists' role in collecting FOI data in a large school district*. Paper Presented at the Annual Meeting of the American Educational Research, Chicago, IL.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (3rd ed.). New York, NY: The American Council on Education.

Minner, D. D., Levy, A. J., & Century, J. (2009). Inquiry-based science instruction—What is it and does it matter? Results from a research synthesis years 1984 to 2002. *Journal of Research in Science Teaching, 47*(4), 474–496.

Mourshed, M., Chijioke, C., & Barber, M. (2010). *How the world's most improved school systems keep getting better.* London: McKinsey & Company.

Mowbray, C. T., Holter, M. C., Teague, G. B., & Bybee, D. (2003). Fidelity criteria: Development, measurement, and validation. *American Journal of Evaluation, 24*(3), 315-340.

Mullis, I. V. S., & Martin, M. O. (2007). TIMSS in perspective: Lessons learned from IEA's four decades of international mathematics assessments. In T. Loveless (Ed.), *Lessons learned* (pp. 9–36). Washington, DC: Brookings Institution Press.

Mullis, I. V. S., Martin, M. O., & Foy, P. (2008). *TIMSS 2007 international mathematics report: Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & & Chrostowski, S. J. (2004). *TIMSS 2003 international mathematics report: Findings from IEA's trends in international mathematics and science study at the fourth and eighth grades.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Gonzalez, E. J., & Kennedy, A. M. (2003). *PIRLS 2001 international report: IEA's study of reading literacy achievement in primary schools.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Olson, J. F., Berger, D. R., Milne, D., & Stanco, G. M. (Eds.). (2008). *TIMSS 2007 encyclopedia: A guide to mathematics and science* *education around the world*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., Arora, A., & Erberber, E. (2005). *TIMSS 2007 assessment frameworks.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Ruddock, G. J., O'Sullivan, C. Y., & Preuschoff, C. (2009). *TIMSS 2011 assessment frameworks.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Murnane, R. J., & Raizen, S. A. (Eds.). (1988). *Improving indicators of the quality of science and mathematics education in grades K-12*. Washington, D.C.: National Academy Press.

Murphy, P. (1996). The IEA assessment of science achievement. *Assessment in Education: Principles, Policy & Practice, 3*(2), 213-232.

National Center for Education Statistics. (2009). *The nation's report card: Mathematics 2009.* Washington, D.C.: Institute of Education Sciences, U.S. Department of Education.

National Center for Education Statistics. (2011). *The nation's report card: Science 2009.* Washington, D.C.: Institute of Education Sciences, U.S. Department of Education.

National Council of Teachers of Mathematics. (2011). *Principles and standards for school mathematics.* Retrieved from http://www.nctm.org/uploadedFiles/Math_Standards/12752_exec_pssm.pdf

National Research Council. (2004). *On evaluating curricular effectiveness: Judging the quality of K-12 mathematics evaluations.* (Committee for a Review of the Evaluation Data on the Effectiveness of NSF-Supported and Commercially Generated Mathematics Curriculum Materials, Mathematical Sciences Education Board, Center for Education, Division of Behavioral and Social Sciences and Education). Washington, DC: National Academies Press.

National Research Council. (2006). *America's lab report: Investigations in high school science.* (Committee on High School Science Laboratories: Singer, S.R.; Hilton, M.L.; Schweingruber, H.A., Editors. Board on Science Education, Center for Education, Division of Behavioral and Social Sciences and Education). Washington, D.C.: The National Academies Press.

National Research Council. (2011). *A framework for K-12 science education: Practices, crosscutting concepts, and core ideas.* (Committee on a Conceptual Framework for New K-12 Science Education Standards. Board on Science Education, Division of Behavioral and Social Sciences and Education). Washington, DC: The National Academies Press.

Nunnally, J. (1978). *Psychometric theory*. New York, NY: McGraw-Hill.

O'Donnell, C. L. (2008). Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in K–12 curriculum intervention research. *Review of Educational Research, 78*(1), 33-84.

Olson, J. F., Martin, M. O., & Mullis, I. V. S. (Eds.). (2008). *TIMSS 2007 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Organisation for Economic Co-operation and Development. (2005). *PISA 2003 technical report.* Paris, France: OECD Publishing.

Organisation for Economic Co-operation and Development. (2010a). *The high cost of low educational performance: The long-run economic impact of improving PISA outcomes.* Paris, France: Author.

Organisation for Economic Co-operation and Development. (2010b). *PISA 2009 results: Overcoming social background - Equity in learning opportunities and outcomes (Volume II).* Paris, France: Author.

Organisation for Economic Co-operation and Development. (2010c). *Strong performers and successful reformers in education: Lessons from PISA for the United States.* Paris, France: Author.

Organisation of Economic Co-operation and Development. (2011). *What PISA is.* Retrieved from http://www.pisa.oecd.org/pages/0,3417,en_32252351_32235907_1_1_1_1_1,00.html

Paccagnella, O. (2006). Centering or not centering in multilevel models? The role of the group mean and the assessment of group effects. *Evaluation Review, 60*, 66-85.

Papanastasiou, C. (2008). A residual analysis of effective schools and effective teaching in mathematics. *Studies in Educational Evaluation, 34*(1), 24-30.

Pavesic, B. J., & Svetlik, K. (Forthcoming). Slovenia. In I. V. S. Mullis, & M. O. Martin, (Eds.). *TIMSS 2011 encyclopedia: Education policy and curriculum in mathematics and science*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Pedhazur, E. J. (1997). *Multiple regression in behavioral research: Explanation and prediction* (3rd ed.). Florence, KY: Wadsworth.

Preuschoff, C. (2010). *Using TIMSS and PIRLS to construct global indicators of effective environments for learning.* (Doctoral dissertation, Boston College, http://hdl.handle.net/2345/1947)

Quek, G., Goh, K., Yen, Y. P., Liu, Y. M., Tan, Q. L., Chin, T. Y., & Lau, C. Y. (2008). Singapore. In I. V. S. Mullis, M. O. Martin, J. F. Olson, D. R. Berger, D. Milne, & G. M. Stanco. (Eds.). (2008). *TIMSS 2007 encyclopedia: A guide to mathematics* *and science education around the world*. (pp. 621-634). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

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

Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2004). *Hierarchical linear and nonlinear modeling: HLM for windows (version 6.00)*

Reckase, M. D. (1979). Unifactor latent trait models applied to multifactor tests: Results and implications. *Journal of Educational Statistics, 4*(3), 207-230.

Remillard, J. T. (Summer 2005). Examining key concepts in research on teachers' use of mathematics curricula. *Review of Educational Research, 75*(2), 211-246.

Republic of Slovenia, Ministry of Education and Sport. (2011). *Compulsory basic education in Slovenia.* Retrieved from http://www.mss.gov.si/en/areas_of_work/compulsory_basic_education_in_slovenia/

Reynolds, D., & Teddlie, C. (2000). The processes of school effectiveness. In C. Teddlie, & D. Reynolds (Eds.), *The international handbook of school effectiveness research* (pp. 134-159). New York, NY: Falmer Press.

Rice, J. K. (2003). *Teacher quality: Understanding the effectiveness of teacher attributes.* Washington, D.C.: Economic Policy Institute.

Rothon, C., Head, J., Klineberg, E., & Stansfeld, S. (2010). Can social support protect bullied adolescents from adverse outcomes? A prospective study on the effects of bullying on the educational achievement and mental health of adolescents at secondary schools in East London. *Journal of Adolescence, 34*(3), 579-588.

Ruddock, G. J., O'Sullivan, C. Y., Arora, A., & Erberber, E. (2008). Developing the TIMSS 2007 mathematics and science assessments and scoring guides. In J. F. Olson, M. O. Martin & I. V. S. Mullis (Eds.), *TIMSS 2007 technical report* (pp. 13-44). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Rumberger, R. W., & Palardy, G. J. (2004). Multilevel models for school effectiveness research. In D. Kaplan (Ed.), *The sage handbook of quantitative methodology for the social sciences* (pp. 235-258). Thousand Oaks, CA: Sage Publications, Inc.

Sammons, P. (2007). *School effectiveness and equity: Making connections, a review of school effectiveness and improvement research and its implications for practitioners and policy makers.* Reading, UK: CfBT Education trust.

Scheerens, J., & Bosker, R. J. (1997). *The foundations of educational effectiveness*. Oxford, England: Pergamon.

Scherbaum, C. A., & Ferreter, J. M. (2009). Estimating statistical power and required sample sizes for organizational research using multilevel modeling. *Organizational Research Methods, 12*(2), 347-367.

Schmidt, W. H., McKnight, C. C., Valverde, G. A., Houang, R. T., & Wiley, D. E. (1997). *Many visions, many aims: A cross-national investigation of curricular intentions in school mathematics*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Schmidt, W. H., Raizen, S. A., Britton, E. D., Bianchi, L. J., & Wolfe, R. G. (1997). *Many visions, many aims: A cross-national investigation of curricular intentions in school science*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Schreiber, J. B. (2002). Institutional and student factors and their influence on advanced mathematics achievement. *The Journal of Educational Research, 95*(5), 274-286.

Schütz, G., Ursprung, H. W., & Wößmann, L. (2008). Education policy and equality of opportunity. *Kyklos: International Review for Social Sciences, 61*(2), 279-308.

Shelley, M. C., & Su, W. (2011). Effects of student-level and school-level characteristics on the quality and equity of mathematics achievement in the United States: Using factor analysis and hierarchical linear models to inform education policy. *Mapping Equity and Quality in Mathematics Education, 2*, 145-161.

Singh, K., Granville, M., & Dika, S. (2002). Mathematics and science achievement: Effects of motivation, interest, and academic engagement. *The Journal of Educational Research, 95*(6), 323-332.

Sirin, S. R. (2005). Socioeconomic status and academic achievement: A meta-analytic review of research. *Review of Educational Research, 75*(3), 417-453.

Straková, J. (2007). The impact of the structure of the education system on the development of educational inequalities in the Czech Republic. *Sociologický Časopis [Czech Sociological Review], 43*(3), 589-609.

Suen, H. K. (1990). *Principles of test theories.* Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Taylor, J. A., Van Scotter, P., & Coulson, D. (2007). Bridging research on learning and student achievement: The role of instructional materials. *Science Educator, 16*(2), 44-50.

Teddlie, C. (2010). The legacy of the school effectiveness research tradition. In A. Hargreaves, E. Lieberman, M. Fullan & D. Hopkins (Eds.), *Second international handbook of educational Change* (pp. 523-554). New York, NY: Springer.

Teddlie, C., & Reynolds, D. (2000). *The international handbook of school effectiveness research*. New York, NY: Falmer Press.

Townsend, T. (Ed.). (2007). *International handbook of school effectiveness and improvement*. Dordrecht, The Netherlands: Springer.

Tsui, M. (2005). Family income, home environment, parenting, and mathematics achievement of children in China and the United States. *Education and Urban Society, 37*(3), 336-355.

U.S. Department of Education. (2007). *Report of the academic competitiveness council.* Washington, D.C.: Author.

U.S. Department of Education. (2009). *Race to the top program executive summary.* (Report No. 20202). Washington, D.C.: Author. Retrieved from http://www2.ed.gov/programs/racetothetop/executive-summary.pdf

U.S. Department of Education. (2011). *President's FY 2012 budget request for the U.S. department of education: Improving science, technology, engineering, and mathematics (STEM) education.* Retrieved from http://www2.ed.gov/about/overview/budget/budget12/index.html

Uline, C., & Tschannen-Moran, M. (2008). The walls speak: The interplay of quality facilities, school climate, and student achievement. *Journal of Educational Administration, 46*(1), 55-73.

United States Government Accountability Office. (2005). *Higher education: Federal science, technology, engineering, and mathematics programs and related trends.* (Report to the Chairman, Committee on Rules, House of Representatives). Washington, D. C.

Wang, A. H., Coleman, A. B., Coley, R. J., & Phelps, R. P. (2003). *Preparing teachers around the world.* (Policy Information Report). Princeton, NJ: Educational Testing Service.

Wang, D. B. (2004). Family background factors and mathematics success: A comparison of Chinese and US students. *International Journal of Educational Research, 41*(1), 40-54.

Wayne, A. J., & Youngs, P. (2003). Teacher characteristics and student achievement gains: A review. *Review of Educational Research, 73*(1), 89-122.

WestEd. (2011). *Technology and engineering literacy framework for the 2014 National Assessment of Educational Progress (pre-publication edition).* Washington, D.C.: National Assessment Governing Board.

Willms, J. D. (2006). *Learning divides: Ten policy questions about the performance and equity of schools and schooling systems.* Montreal, Canada: UNESCO Institute for Statistics.

Wilson, M. (2005). *Constructing measures: An item response modeling approach*. New York, NY: Taylor & Francis Group, LLC.

Woessmann, L. (2004). *How equal are educational opportunities? Family background and student achievement in Europe and the US.* (CESifo Working Paper No. 1162). Munich, Germany.

Wu, M. L., Adams, R. J., Wilson, M., & Haldane, S. A. (2007). *ConQuest—Generalized item response modeling software*. Camberwell, Australia: Australian Council for Educational Research.

Ysseldyke, J., Spicuzza, R., Kosciolek, S., Teelucksingh, E., Boys, C., & Lemkuil, A. (2003). Using a curriculum-based instructional management system to enhance math achievement in urban schools. *Journal of Education for Students Placed at Risk, 8*(2), 247-265.

# Appendix

## *Information About the Measures of School Effectiveness Factors*

Table A.1. Items in the "Student Home Resources" Scale

| **Number of Books in the Home** | |
|---|---|
| Source (Includes National Adaptations) | Based on student responses to the following question in the Student Questionnaire – Eighth Grade:<br><br>About how many books are there in your home? (Do not count magazines, newspapers, or your school books.) |
| Response Categories | 1= None or very few (0-10 books)<br>2= Enough to fill one shelf (11-25 books)<br>3= Enough to fill one bookcase (26-100 books)<br>4= Enough to fill two bookcases (101-200 books)<br>5= Enough to fill three or more bookcases (more than 200 books) |
| Recoded Response Categories | 0= None or very few (0-10 books) (Original category 1)<br>1= Enough to fill one shelf (11-25 books) (Original category 2)<br>2= Enough to fill one bookcase (26-100 books) (Original category 3)<br>3= Enough to fill two bookcases (101-200 books) (Original category 4)<br>4= Enough to fill three or more bookcases (more than 200 books) (Original category 5) |
| **Number of Home Possessions** | |
| Source (Includes National Adaptations for items 6-9) | Based on student responses to the following questions in the Student Questionnaire – Eighth Grade:<br><br>International Version<br>1. Do you have a calculator in your home?<br>2. Do you have a computer in your home? (do not include PlayStation®, GameCube®, XBox®, or other TV/video game computers)<br>3. Do you have a study desk/table for your use in your home?<br>4. Do you have a dictionary in your home?<br>5. Do you have an Internet connection in your home? |

Table A.1. Items in the "Student Home Resources" Scale (Continued)

| | |
|---|---|
| | Chinese Taipei<br>6. Do you have a compact disc, software, or videotape for learning math in your home?<br>7. Do you have a reference book for learning math in your home?<br>8. Do you have a compact disc, software, or videotape for learning science in your home?<br>9. Do you have a reference book for learning science in your home?<br><br>Czech Republic<br>6. Do you have your own Discman or mp3 player?<br>7. Do you have a digital camera (do not count a mobile phone with a camera)?<br>8. Do you have a camcorder?<br>9. Do you have a dishwasher in your home?<br><br>Singapore<br>6. Do you have a car?<br>7. Do you have domestic help (e.g., a maid)?<br>8. Do you have a mobile phone?<br>9. Do you have a piano/organ/violin?<br><br>Slovenia<br>6. Do you have your own room in your home?<br>7. Do you have your own computer in your home?<br>8. Do you have more than one bathroom in your home?<br>9. Do you go to outside school activities that your parents pay for?<br><br>United States<br>6. Do you have an encyclopedia (book or CD-ROM) in your home?<br>7. Do you have a PlayStation, GameCube, Xbox, or other TV/Video game system in your home?<br>8. Do you have a VHS or DVD player in your home?<br>9. Do you have three or more cars, small trucks, or sport utility vehicles? |
| Response Categories | 1=yes 2=no |
| Recoded Response Categories | 0= 0–4 home possessions<br>1= 5 home possessions<br>2= 6 home possessions<br>3= 7 home possessions<br>4= 8 home possessions<br>5= 9 home possessions |

Table A.1. Items in the “Student Home Resources” Scale (Continued)

| **Highest Education Level of Either Parent** | |
|---|---|
| Source | Based on student responses to the following questions in the Student Questionnaire – Eighth Grade:<br><br>1. What is the highest level of education completed by your mother (or stepmother or female guardian)?<br>2. What is the highest level of education completed by your father (or stepfather or male guardian)? |
| Response Categories (Includes National Adaptations) | International Version*<br>1= Some <ISCED Level 1 or 2> or did not go to school [Some primary or lower secondary education or did not go to school]<br>2 = <ISCED 2> [Lower secondary education]<br>3= <ISCED 3> [Upper secondary education]<br>4 = <ISCED 5B> [Post-secondary non-tertiary education]<br>5 = <ISCED 5A, first degree> [First stage of tertiary education]<br>6= Beyond <ISCED 5A, first degree> [Second stage of tertiary education]<br>7= I don’t know<br><br>*TIMSS 2007 uses International Standard Classification of Education (ISCED) levels in questionnaire items relating to education so that responses will be internationally comparable. The names for each of the ISCED levels used in the student questionnaire are provided in brackets.<br><br>Chinese Taipei<br>1 = High school, vocational school<br>2 = Option not administered or data not available<br>3 = Junior teacher college, junior college (2-year or 5-year system) / College of Technology (2-year system)<br>4 = University or college (including 4-year Institute of Technology)<br>5 = Master's degree or Ph.D.<br>6 = I don't know<br><br>Czech Republic<br>1 = Secondary vocational education without Maturate / Full secondary vocational education with Maturate / Full secondary technical education with Maturate / Full secondary general education with Maturate<br>2 = Option not administered or data not available<br>3 = Higher professional school or conservatory<br>4 = Bachelor's degree<br>5 = Master's degree or higher<br>6 = I don't know |

Table A.1. Items in the "Student Home Resources" Scale (Continued)

| | |
|---|---|
| | Singapore<br>1 = Upper secondary<br>2 = Postsecondary nontertiary (e.g., JC/ Pre-U/ ITE)<br>3 = Diploma (e.g., Polytechnic)<br>4 = University (first degree)<br>5 = Beyond university (first degree) (e.g., master's degree or Ph.D.)<br>6 = I don't know<br><br>Slovenia<br>1 = Secondary school<br>2 = Professional education after secondary school<br>3 = 3 years of university studies (called higher studies)<br>4 = 4 to 5 years of university studies<br>5 = Master's degree or Ph.D.<br>6 = I haven't decided yet / I can't answer, because I don't know the levels of education<br><br>United States<br>1 = High school<br>2 = Vocational/technical education after high school<br>3 = Community or junior college<br>4 = Bachelor's degree at a college or university<br>5 = Beyond a bachelor's degree<br>6 = I don't know |
| Recoded Response Categories | 0= ISCED 1 or 2 or Did Not Go to School (Original categories 1 and 2)<br>1= ISCED 3 (Original category 3)<br>2= ISCED 4 and 5B (Original categories 4 and 5)<br>3= ISCED 5A, first degree and beyond (Original category 6) |

Figure A.1. Item Map from "Student Home Resources" Scale

```
================================================================================
MAP OF LATENT DISTRIBUTIONS AND THRESHOLDS
--------------------------------------------------------------------------------
                |
  2             |
                |
                |
                |
                |
               X|
               X|
              XX|
              XX|
               X|1.4
             XXX|
  1           XX|
             XXX|2.5
           XXXXX|
            XXXX|3.3
         XXXXXXX|
           XXXXX|1.3
          XXXXXX|
          XXXXXX|
        XXXXXXXX|
        XXXXXXXX|3.2
      XXXXXXXXXX|2.4
       XXXXXXXXX|
  0     XXXXXXXX|
        XXXXXXXX|
        XXXXXXXX|
        XXXXXXXX|
      XXXXXXXXXX|1.2
          XXXXXX|
         XXXXXXX|2.3
           XXXXX|
            XXXX|
           XXXXX|3.1
             XXX|
 -1          XXX|2.2
              XX|
              XX|
               X|
               X|
               X|1.1
               X|
                |
               X|2.1
                |
                |
                |
 -2             |
                |
================================================================================
Each 'X' represents 148.0 cases. The labels for thresholds show the levels of
item, and step, respectively
================================================================================
```

<u>Books in the Home</u>
1.1= 11-25 books
1.2= 26-100 books
1.3= 101-200 books
1.4= More than 200 books

<u>Home Possessions</u>
2.1= 5 home possessions
2.2= 6 home possessions
2.3= 7 home possessions
2.4= 8 home possessions
2.5= 9 home possessions

<u>Parent's Education Level (Highest of Mother or Father)</u>
3.1= ISCED 3
3.2= ISCED 4 and 5B
3.3= ISCED 5A, first degree and beyond

Table A.2. Items in the “Teacher Preparation” Composite Variables

| **Highest Level of Formal Education** | |
|---|---|
| Source | Based on teacher responses to the following question in the Mathematics or Science Teacher Questionnaire – Eighth Grade:<br><br>What is the highest level of formal education you have completed? |
| Response Categories (Includes National Adaptations) | International Version*<br>1= Some <ISCED Level 1 or 2> or did not go to school [Some primary or lower secondary education or did not go to school]<br>2 = <ISCED 2> [Lower secondary education]<br>3= <ISCED 3> [Upper secondary education]<br>4 = <ISCED 5B> [Post-secondary non-tertiary education]<br>5 = <ISCED 5A, first degree> [First stage of tertiary education]<br>6= Beyond <ISCED 5A, first degree> [Second stage of tertiary education]<br><br>*TIMSS 2007 uses International Standard Classification of Education (ISCED) levels in questionnaire items relating to education so that responses will be internationally comparable. The names for each of the ISCED levels used in the Teacher Questionnaire are provided in brackets.<br><br>Chinese Taipei<br>1 = Did not finish senior high school, vocational school, or equivalent level<br>2 = Teacher school, senior high school, vocational school<br>3 = Junior teacher college, junior college (2-year or 5-year system)<br>4 = College of technology (2-year system)<br>5 = University or college (including 4-year Institute of Technology)<br>6 = Master’s degree or Ph.D.<br><br>Czech Republic<br>1 = Did not complete secondary school<br>2 = Secondary school<br>3 = Extension study (postsecondary, nontertiary level)<br>4 = Higher professional school or conservatory<br>5 = Bachelor’s degree<br>6 = Master’s or higher degree<br><br>Singapore<br>1 = Did not complete upper secondary education<br>2 = Upper secondary education<br>3 = Postsecondary nontertiary (e.g., JC/ Pre-U/ ITE)<br>4 = Diploma (e.g., Polytechnic)<br>5 = University (first degree)<br>6 = University (master’s degree) or higher |

Table A.2. Items in the “Teacher Preparation” Composite Variables (Continued)

| | |
|---|---|
| | Slovenia<br>1 = Did not complete secondary school<br>2 = Secondary school<br>3 = Vocational courses after secondary school<br>4 = 2 to 3 years of university studies (called higher studies)<br>5 = 4 years of university studies<br>6 = Master’s degree or Ph.D.<br><br>United States<br>1 = Did not complete high school<br>2 = High school<br>3 = Vocational/technical certificate after high school<br>4 = Associate’s degree (A.A.) in a vocational/technical program<br>5 = Academic associate’s or bachelor’s degree<br>6 = Academic master’s degree, postgraduate certificate program (e.g., teaching), or first professional degree (e.g., law, medicine, dentistry) / Doctorate (Ph.D. or Ed.D.) |
| Recoded Response Categories | 0 = Less than bachelor’s degree (Original categories 1 to 4)<br>1 = Bachelor’s degree (Original category 5)<br>2 = Postgraduate degree (Original category 6) |
| **Mathematics Teachers’ Major or Main Area of Study** | |
| Source | Based on teacher responses to the following questions in the Mathematics Teacher Questionnaire – Eighth Grade:<br><br>1. During your <post-secondary> education, was your major or main area(s) of study mathematics?<br>2. During your <post-secondary> education, was your major or main area(s) of study education-mathematics?<br>3. During your <post-secondary> education, was your major or main area(s) of study science?<br>4. During your <post-secondary> education, was your major or main area(s) of study education-science?<br>5. During your <post-secondary> education, was your major or main area(s) of study education-general?<br>6. During your <post-secondary> education, was your major or main area(s) of study other? |
| Response Categories | 1=yes 2=no |

Table A.2. Items in the “Teacher Preparation” Composite Variables (Continued)

| **Science Teachers’ Major or Main Area of Study** | |
|---|---|
| Source | Based on teacher responses to the following questions in the Science Teacher Questionnaire – Eighth Grade:<br><br>1. During your <post-secondary> education, was your major or main area(s) of study biology?<br>2. During your <post-secondary> education, was your major or main area(s) of study physics?<br>3. During your <post-secondary> education, was your major or main area(s) of study chemistry?<br>4. During your <post-secondary> education, was your major or main area(s) of study earth science?<br>5. During your <post-secondary> education, was your major or main area(s) of study education-science?<br>6. During your <post-secondary> education, was your major or main area(s) of study mathematics?<br>7. During your <post-secondary> education, was your major or main area(s) of study education-mathematics?<br>8. During your <post-secondary> education, was your major or main area(s) of study education-general? |
| Response Categories | 1=yes 2=no |

Table A.3. Items in the "Instruction Affected by General Resource Shortages" Scale

| Instruction Affected by General Resource Shortages | |
|---|---|
| Source | Based on principal responses to the following question in the School Questionnaire:<br><br>1. Is your school's capacity to provide instruction affected by a shortage or inadequacy of instructional materials (e.g., textbook)?<br>2. Is your school's capacity to provide instruction affected by a shortage or inadequacy of budget for supplies (e.g., paper, pencils)?<br>3. Is your school's capacity to provide instruction affected by a shortage or inadequacy of school buildings and grounds?<br>4. Is your school's capacity to provide instruction affected by a shortage or inadequacy of heating/cooling and lighting systems?<br>5. Is your school's capacity to provide instruction affected by a shortage or inadequacy of instructional space (e.g., classrooms)?<br>6. Is your school's capacity to provide instruction affected by a shortage or inadequacy of teachers?<br>7. Is your school's capacity to provide instruction affected by a shortage or inadequacy of computer support staff? |
| Response Categories | 1= None<br>2= A little<br>3= Some<br>4= A lot |
| Recoded Response Categories | 0= A lot (Original category 4)<br>1= Some (Original category 3)<br>2= A little (Original category 2)<br>3= None (Original category 1) |

Figure A.2. Item Map from "Instruction Affected by General Resource Shortages" Scale

```
================================================================================
MAP OF LATENT DISTRIBUTIONS AND THRESHOLDS
================================================================================
                |
  3             |
                |
                |
               X|
                |
               X|
              XX|
  2           XX|
              XX|
             XXX|
             XXX|
            XXXX|
           XXXXX|
           XXXXX|
         XXXXXXX|
  1      XXXXXXX|
         XXXXXXX|
          XXXXXX|
       XXXXXXXXX|
         XXXXXXX|
      XXXXXXXXXX|7.3
      XXXXXXXXXX|5.3
  0      XXXXXXX|
        XXXXXXXX|
        XXXXXXXX|2.3 3.3 6.3
        XXXXXXXX|
       XXXXXXXXX|4.3
          XXXXXX|1.3
          XXXXXX|7.2
 -1        XXXXX|
             XXX|3.2 5.2
            XXXX|
              XX|4.2 6.2
              XX|2.2
              XX|
               X|1.2
               X|3.1 5.1
 -2             |7.1
               X|6.1
                |
                |
                |
                |4.1
                |
 -3             |2.1
                |1.1
                |
                |
================================================================================
Each 'X' represents 149.3 cases
The labels for thresholds show the levels of item, and step, respectively

================================================================================
```

Note: The item labels correspond to the question numbers in Table A.3 and their recoded response categories (the zero category is not shown). For example:

<u>Shortage or inadequacy of instructional materials</u>
1.1= Some
1.2= A Little
1.3= None

Table A.4. Items in the "Science Laboratories Available" Composite Variable

| **Schools with Science Laboratories and Assistance Available When Students Conduct Experiments** | |
|---|---|
| Source | Based on teacher responses to the following questions in the School Questionnaire – Eighth Grade:<br><br>a. Does your school have a science laboratory?<br>b. Do teachers usually have assistance available when students are conducting science experiments? |
| Response Categories | 1=yes 2=no |
| Recoded Categories | 2= Yes to part a (School has science laboratory) and Yes to part b (teachers have assistance available)<br>1= Yes to part a (School has science laboratory) and No to part b (teachers have assistance available)<br>0= No to part a (School has science laboratory) and No to part b (teachers have assistance available) |

Table A.5. Items in the "Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science" Scale

| Instruction Affected by Mathematics and Science Resource Shortages | |
|---|---|
| Source | Based on principal responses to the following question in the School Questionnaire:<br><br>1. Is your school's capacity to provide instruction affected by a shortage or inadequacy of computers for mathematics instruction?<br>2. Is your school's capacity to provide instruction affected by a shortage or inadequacy of computer software for mathematics instruction?<br>3. Is your school's capacity to provide instruction affected by a shortage or inadequacy of calculators for mathematics instruction?<br>4. Is your school's capacity to provide instruction affected by a shortage or inadequacy of library materials relevant to mathematics instruction?<br>5. Is your school's capacity to provide instruction affected by a shortage or inadequacy of audio-visual resources for mathematics instruction?<br>6. Is your school's capacity to provide instruction affected by a shortage or inadequacy of science laboratory equipment and materials?<br>7. Is your school's capacity to provide instruction affected by a shortage or inadequacy of computers for science instruction?<br>8. Is your school's capacity to provide instruction affected by a shortage or inadequacy of computer software for science instruction?<br>9. Is your school's capacity to provide instruction affected by a shortage or inadequacy of calculators for science instruction?<br>10. Is your school's capacity to provide instruction affected by a shortage or inadequacy of library materials relevant to science instruction?<br>11. Is your school's capacity to provide instruction affected by a shortage or inadequacy of audio-visual resources for science instruction? |
| Response Categories | 1= None<br>2= A little<br>3= Some<br>4= A lot |
| Recoded Response Categories | 0= A lot (Original category 4)<br>1= Some (Original category 3)<br>2= A little (Original category 2)<br>3= None (Original category 1) |

Figure A.3. Item Map from the "Instruction Affected by Shortages in Resources Specific to Teaching Mathematics and Science" Scale

```
================================================================================
MAP OF LATENT DISTRIBUTIONS AND THRESHOLDS
================================================================================
  6             |
                |
                |
               X|
  5           XX|
              XX|
              XX|
  4          XXX|
             XXX|
             XXX|
  3         XXXX|
            XXXX|
          XXXXXX|
  2        XXXXX|
           XXXXX|
          XXXXXX|
        XXXXXXXX|8.3
  1       XXXXXX|6.3
      XXXXXXXXXX|2.3 7.3
      XXXXXXXXXX|1.3 10.3 11.3
  0     XXXXXXXX|4.3 5.3
       XXXXXXXXX|
      XXXXXXXXXX|9.3
 -1     XXXXXXXX|
       XXXXXXXXX|1.2 2.2 3.3 6.2 7.2 8.2
           XXXXX|
           XXXXX|5.2
 -2         XXXX|4.2 9.2 10.2 11.2
             XXX|3.2
             XXX|8.1
 -3           XX|2.1 7.1
              XX|1.1 6.1
               X|
 -4            X|10.1
                |5.1 11.1
                |4.1
 -5             |3.1 9.1
                |
                |
                |
 -6             |
================================================================================
Each 'X' represents 162.7 cases
The labels for thresholds show the levels ofitem, and step, respectively
================================================================================
```

Note: The item labels correspond to the question numbers in Table A.5 and their recoded response categories (the zero category is not shown). For example:

<u>Shortage or inadequacy of computers for mathematics instruction</u>
1.1= Some
1.2= A Little
1.3= None

Table A.6. Items in the “Mathematics Topics Taught” Scale

| Mathematics Topics Taught | |
|---|---|
| Source | Based on student responses to the following question in the Mathematics Teacher Questionnaire:<br><br>The following list includes the main topics addressed by the TIMSS mathematics test. Choose the response that best describes when students in the TIMSS class hae been taught each topic. If a topic was taught half this year but not yet completed, please choose “mostly taught this year.” If a topic is not in the curriculum, please choose “Not yet taught or just introduced.”<br><br>Number Topics<br><br>1. Whole numbers including place value, factorization, and the four operations<br>2. Computations, estimations, or approximations involving whole numbers<br>3. Common fractions including equivalent fractions and ordering of fractions<br>4. Decimal including place value, ordering, and converting to common fractions (and vice versa)<br>5. Representing decimals and fractions using words, numbers, or models (including number lines)<br>6. Computations with fractions<br>7. Computations with decimals<br>8. Representing, comparing, ordering, and computing with integers<br>9. Ratios (equivalence, division of a quantity by a given ratio)<br>10. Conversion of percents to fractions or decimals and vice versa<br><br>Algebra Topics<br><br>11. Numeric, algebraic, and geometric patterns or sequences (extension, missing terms, generalization of patterns)<br>12. Sums, products, and powers of expressions containing variables<br>13. Evaluating expressions for given numeric value<br>14. Simplifying or comparing algebraic expressions<br>15. Modeling situations using expressions<br>16. Evaluating functions/formulas for given values of the variables<br>17. Simple linear equations and inequalities, and simultaneous (two variables) equations<br>18. Equivalent representations of functions as ordered pairs, tables, graphs, words, or equations |

Table A.6. Items in the "Mathematics Topics Taught" Scale (Continued)

| | |
|---|---|
| | Geometry Topics<br><br>19. Angles - acute, right, straight, obtuse, reflex<br>20. Relationships for angles at a point, angles on a line, vertically opposite angles, angles associated with a transversal cutting parallel lines, and perpendicularity<br>21. Properties of geometric shapes: triangles, quadrilaterals, and other common polygons<br>22. Construct or draw triangles and rectangles of given dimensions<br>23. Congruent figures (triangles, quadrilaterals) and their corresponding measures<br>24. Similar triangles and recall their properties<br>25. Relationships between two-dimensional and three-dimensional shapes<br>26. Pythagorean Theorem (not proof ) to find length of a side<br>27. Measurement, drawing, and estimation of the size of angles, the lengths of lines, areas, and volumes<br>28. Measurement formulas for perimeters, circumferences, areas of circles, surface areas, and volumes<br>29. Measures of irregular or compound areas (e.g., by covering with grids or dissecting and rearranging pieces)<br>30. Cartesian plane - ordered pairs, equations, intercepts, intersections, and gradient<br>31. Line and rotational symmetry for two-dimensional shapes<br>32. Translation, reflection, and rotation<br><br>Data and Chance Topics<br><br>33. Reading data from tables, pictographs, bar graphs, pie charts, and line graphs<br>34. Organizing and displaying data using tables, pictographs, bar graphs, pie charts, and line graphs<br>35. Characteristics of data sets including mean, median, range, and shape of distribution (in general terms)<br>36. Interpreting data sets (e.g., draw conclusions, make predictions, and estimate values between and beyond given data points)<br>37. Data displays that could lead to misinterpretation (e.g., inappropriate grouping and misleading or distorted scales)<br>38. Using data from experiments to predict chances of future outcomes<br>39. Taught using the chances of a particular outcome to solve problems |
| Response Categories | 1= Mostly taught before this year<br>2= Mostly taught this year<br>3= Not yet taught or just introduced |
| Recoded Response Categories | 0= Not yet taught or just introduced (Original category 3)<br>1= Mostly taught this year (Original category 2)<br>2= Mostly taught before this year (Original category 1) |

Figure A.4. Item Map from “Mathematics Topics Taught” Scale

```
================================================================================
MAP OF LATENT DISTRIBUTIONS AND THRESHOLDS
================================================================================
                  |
  2              X|
                  |
                  |39.2
                  |38.2
                  |
                 X|
                 X|
                 X|36.2 37.2
                  |18.2 26.2 35.2
                 X|12.2 14.2 16.2 17.2 30.2
                XX|24.2 28.2
                 X|15.2 25.2
                 X|13.2 38.1 39.1
  1             XX|
                XX|37.1
                XX|
               XXX|11.2
               XXX|
             XXXXX|36.1
              XXXX|29.2 32.2
             XXXXX|
             XXXXX|
           XXXXXXX|31.2 34.2 35.1
            XXXXXX|24.1
         XXXXXXXXX|32.1 33.2
          XXXXXXXX|23.2
  0        XXXXXXX|25.1
          XXXXXXXX|27.2 30.1
          XXXXXXXX|
          XXXXXXXX|29.1
        XXXXXXXXXX|34.1
            XXXXXX|21.2 22.2 33.1
           XXXXXXX|
              XXXX|9.2 20.2 31.1
               XXX|17.1 18.1
              XXXX|
                XX|
                XX|
                XX|19.2
 -1              X|
                 X|16.1 26.1
                 X|
                  |9.1
                 X|11.1
                  |
                  |
                  |27.1 28.1
                  |10.2
                  |1.2 2.2
                  |8.2 22.1
                  |20.1 23.1
                  |
 -2               |
                 X|
                  |5.2 14.1 15.1 19.1 21.1
                  |1.1 2.1 3.1 3.2 4.1 4.2 5.1 6.1 6.2 7.1 7.2 8.1 10.1 12.1 13.1
================================================================================
Each 'X' represents 173.5 cases
The labels for thresholds show the levels ofitem, and step, respectively

================================================================================
```

Note: The item labels correspond to the question numbers in Table A.6 and their recoded response categories (the zero category is not shown). For example:

Whole numbers
1.1= Mostly taught this year
1.2= Mostly taught before this year

Table A.7. Items in the "Yearly Instructional Hours" Composite Variables

| **Yearly Instructional Hours Spent on Mathematics** | |
|---|---|
| Source | Based on teacher responses to the following question in the Mathematics Teacher Questionnaire:<br><br>1. How many minutes per week do you teach mathematics to the TIMSS class?<br>(Write in the number of minutes per week)<br><br>Based on principal responses to the following questions in the School Questionnaire:<br><br>1. How many days per year is your school open for instruction?<br>(Write in the number of days)<br>2. In one calendar week, how many days is the school open for instruction? |
| Response Categories | 1= 6 days<br>2= 5 ½ days<br>3= 5 days<br>4= 4 ½ days<br>5= 4 days<br>6= Other<br>7= Please specify |
| Calculated Response Method | Instructional Hours per Year = (Teacher Reports of Weekly Mathematics Instructional Hours / Principal Reports of School Days per Week) x Principal Reports of School Days per Year |
| **Yearly Instructional Hours Spent on Science** | |
| Source | Based on teacher responses to the following question in the Science Teacher Questionnaire:<br>1. How many minutes per week do you teach science to the TIMSS class?<br>(Write in the number of minutes per week)<br><br>Based on principal responses to the following questions in the School Questionnaire:<br>1. How many days per year is your school open for instruction?<br>(Write in the number of days)<br>2. In one calendar week, how many days is the school open for instruction? |

Table A.7. Items in the "Yearly Instructional Hours" Composite Variables (Continued)

| | |
|---|---|
| Response Categories | 1= 6 days<br>2= 5 ½ days<br>3= 5 days<br>4= 4 ½ days<br>5= 4 days<br>6= Other<br>7= Please specify |
| Calculated Response Method | $$\text{Instructional Hours per Year} = \frac{\text{Teacher Reports of Weekly Science Instructional Hours}}{\text{Principal Reports of School Days per Week}} \times \text{Principal Reports of School Days per Year}$$ |

Table A.8. Items in the "Effective Instructional Strategies" Scales

| Effective Instructional Strategies for Encouraging Mathematics Reasoning | |
|---|---|
| Source | Based on teacher responses to the following question in the Mathematics Teacher Questionnaire:<br><br>1. In teaching mathematics to the students in the TIMSS class, how often do you usually ask them to explain their answers?<br>2. In teaching mathematics to the students in the TIMSS class, how often do you usually ask them to relate what they are learning in mathematics to their daily lives?<br>3. In teaching mathematics to the students in the TIMSS class, how often do you usually ask them to decide on their own procedures for solving complex problems?<br>4. In teaching mathematics to the students in the TIMSS class, how often do you usually ask them to work on problems for which there is no immediately obvious method of solution? |
| Response Categories | 1= Every or almost every lesson<br>2= About half the lessons<br>3= Some lessons<br>4= Never |
| Recoded Response Categories | 0= Never (Original category 4)<br>1= Some lessons (Original category 3)<br>2= About half the lessons (Original category 2)<br>3= Every or almost every lesson (Original category 1) |
| **Effective Instructional Strategies for Teaching Science Inquiry** | |
| Source | Based on teacher responses to the following question in the Science Teacher Questionnaire:<br><br>1. In teaching science to the students in the TIMSS class, how often do you usually ask them to observe natural phenomena and describe what they see?<br>2. In teaching science to the students in the TIMSS class, how often do you usually ask them to watch you demonstrate an experiment or investigation?<br>3. In teaching science to the students in the TIMSS class, how often do you usually ask them to design or plan experiments or investigations?<br>4. In teaching science to the students in the TIMSS class, how often do you usually ask them to conduct experiments or investigations?<br>5. In teaching science to the students in the TIMSS class, how often do you usually ask them to work together in small groups on experiments or investigations? |

Table A.8. Items in the "Effective Instructional Strategies" Scales (Continued)

| | |
|---|---|
| | 6. In teaching science to the students in the TIMSS class, how often do you usually ask them to use scientific formulae and laws to solve routine problems?<br>7. In teaching science to the students in the TIMSS class, how often do you usually ask them to give explanations about something they are studying?<br>8. In teaching science to the students in the TIMSS class, how often do you usually ask them to relate what they are learning in science to their daily lives? |
| Response Categories | 1= Every or almost every lesson<br>2= About half the lessons<br>3= Some lessons<br>4= Never |
| Recoded Response Categories | 0= Never (Original category 4)<br>1= Some lessons (Original category 3)<br>2= About half the lessons (Original category 2)<br>3= Every or almost every lesson (Original category 1) |

Figure A.5. Item Map from “Effective Instructional Strategies for Encouraging Mathematics Reasoning” Scale

```
==========================================================================
MAP OF LATENT DISTRIBUTIONS AND THRESHOLDS
==========================================================================
   4                |
                    |
                    |
                   X|4.2
                   X|
   3               X|
                   X|
                  XX|
                 XXX|
                 XXX|
   2            XXXX|3.2 4.1
                XXXX|
             XXXXXXX|
               XXXXX|2.2
              XXXXXX|
   1          XXXXXX|
            XXXXXXXX|3.1
            XXXXXXXX|
          XXXXXXXXXX|
            XXXXXXXX|1.2
   0        XXXXXXXX|
            XXXXXXXX|2.1
            XXXXXXXX|
            XXXXXXXX|
          XXXXXXXXXX|1.1
  -1          XXXXXX|
             XXXXXXX|
                XXXX|
                XXXX|
               XXXXX|
  -2             XXX|
                 XXX|
                  XX|
                  XX|
                   X|
  -3               X|
                   X|
                   X|
                   X|
                   X|
  -4                |
                    |
                    |
                    |
==========================================================================
Each 'X' represents 155.8 cases
The labels for thresholds show the levels ofitem, and step, respectively

==========================================================================
```

Note: The item labels correspond to the question numbers in Table A.8 and their recoded response categories (the zero category is not shown). For example:

<u>Ask students to explain their answers</u>
1.1= Some lessons
1.2= About half the lessons
1.3= Every or almost every lesson

Figure A.6. Item Map from “Effective Instructional Strategies for Teaching Science Inquiry” Scale

```
==============================================================================
MAP OF LATENT DISTRIBUTIONS AND THRESHOLDS
==============================================================================
                |
  3             |
                |3.2
                |
                |4.2
                |2.2
                |5.2
               X|
  2            X|3.1
               X|
               X|6.2
              XX|
              XX|1.2
             XXX|
            XXXX|4.1 5.1
           XXXXX|2.1
  1       XXXXXX|
          XXXXXX|1.1
          XXXXXX|
        XXXXXXXX|6.1 7.2
         XXXXXXX|
      XXXXXXXXXX|
       XXXXXXXXX|8.2
  0     XXXXXXXX|
        XXXXXXXX|
        XXXXXXXX|
        XXXXXXXX|
        XXXXXXXX|
          XXXXXX|7.1
          XXXXXX|
 -1          XXX|
            XXXX|8.1
             XXX|
             XXX|
              XX|
              XX|
               X|
 -2            X|
               X|
               X|
                |
               X|
                |
                |
                |
 -3             |
==============================================================================
Each 'X' represents 333.5 cases
The labels for thresholds show the levels of item, and step, respectively

==============================================================================
```

Note: The item labels correspond to the question numbers in Table A.8 and their recoded response categories (the zero category is not shown). For example:

Observe natural phenomena and describe what they see
1.1= Some lessons
1.2= About half the lessons
1.3= Every or almost every lesson

Table A.9. Items in "Students Like Learning" Scales

| | |
|---|---|
| **Students Like Learning Mathematics** | |
| Source | Based on student responses to the following questions in the Student Questionnaire:<br><br>1. How much do you agree with this statement about learning mathematics? - I would like to do more mathematics in school<br>2. How much do you agree with this statement about learning mathematics? - I enjoy learning mathematics<br>3. How much do you agree with this statement about learning mathematics? - Mathematics is boring<br>4. How much do you agree with this statement about learning mathematics? - I like mathematics |
| Response Categories | 1= Agree a lot<br>2= Agree a little<br>3= Disagree a little<br>4= Disagree a lot |
| Recoded Response Categories | 0= Disagree a lot (Original category 4)<br>1= Disagree a little (Original category 3)<br>2= Agree a little (Original category 2)<br>3= Agree a lot (Original category 1) |
| **Students Like Learning Science (General)** | |
| Source | Based on student responses to the following questions in the Student Questionnaire:<br><br>1. How much do you agree with this statement about learning science? - I would like to do more science in school<br>2. How much do you agree with this statement about learning science? - I enjoy learning science<br>3. How much do you agree with this statement about learning science? - Science is boring<br>4. How much do you agree with this statement about learning science? - I like science |
| Response Categories | 1= Agree a lot<br>2= Agree a little<br>3= Disagree a little<br>4= Disagree a lot |
| Recoded Response Categories | 0= Disagree a lot (Original category 4)<br>1= Disagree a little (Original category 3)<br>2= Agree a little (Original category 2)<br>3= Agree a lot (Original category 1) |

Table A.9. Items in "Students Like Learning" Scales (Continued)

| Students Like Learning Science (Separate Science Subjects) | |
|---|---|
| Source | Based on student responses to the following questions in the Student Questionnaire- Separate Science Subjects<br><br>*Biology*<br>How much do you agree with each statement about learning biology?<br>1. I would like to do more biology in school<br>2. I enjoy learning biology<br>3. Biology is boring<br>4. I like biology<br><br>*Earth Science*<br>How much do you agree with each statement about learning earth science?<br>5. I would like to do more earth science in school<br>6. I enjoy learning earth science<br>7. Earth science is boring<br>8. I like earth science<br><br>*Chemistry*<br>How much do you agree with this statement about learning chemistry?<br>9. I would like to do more chemistry in school<br>10. I enjoy learning chemistry<br>11. Chemistry is boring<br>12. I like chemistry<br><br>*Physics*<br>How much do you agree with each statement about learning physics?<br>13. I would like to do more physics in school<br>14. I enjoy learning physics<br>15. Physics is boring<br>16. I like physics |
| Response Categories | 1= Agree a lot<br>2= Agree a little<br>3= Disagree a little<br>4= Disagree a lot |
| Recoded Response Categories | 0= Disagree a lot (Original category 4)<br>1= Disagree a little (Original category 3)<br>2= Agree a little (Original category 2)<br>3= Agree a lot (Original category 1) |

Figure A.7. Item Map from “Students Like Learning Mathematics” Scale

```
==============================================================================
MAP OF LATENT DISTRIBUTIONS AND THRESHOLDS
==============================================================================

  6                |
                   |
                  X|
  5               X|
                  X|
                 XX|
                  X|
  4              XX|
                 XX|
                XXX|
  3            XXXX|1.3
               XXXX|2.3
              XXXXX|4.3
  2            XXXX|3.3
              XXXXX|
             XXXXXX|
  1        XXXXXXXX|
           XXXXXXXX|1.2
         XXXXXXXXXX|
           XXXXXXXX|
  0        XXXXXXXX|2.2 3.2 4.2
           XXXXXXXX|
           XXXXXXXX|
 -1        XXXXXXXX|
          XXXXXXXXX|1.1
             XXXXXX|
 -2          XXXXXX|2.1 3.1 4.1
               XXXX|
                XXX|
 -3            XXXX|
                 XX|
                 XX|
                 XX|
 -4              XX|
                  X|
                  X|
 -5               X|
                  X|
                  X|
 -6               X|
                   |
                   |
 -7                |
==============================================================================
Each 'X' represents 160.5 cases
The labels for thresholds show the levels of item, and step, respectively

==============================================================================
```

Note: The item labels correspond to the question numbers in Table A.9 and their recoded response categories (the zero category is not shown). For example:

Would like to do more mathematics in school
1.1= Disagree a little
1.2= Agree a little
1.3= Agree a lot

Figure A.8. Item Map from “Students Like Learning Science” Scale (General Science)

```
================================================================================
MAP OF LATENT DISTRIBUTIONS AND THRESHOLDS
================================================================================

   7                |
                    |
                    |
   6               X|
                   X|
                  XX|
   5              XX|
                  XX|
                  XX|
   4             XXX|
                 XXX|
                XXXX|
   3            XXXX|
              XXXXXX|
                XXXX|
   2           XXXXX|1.3
               XXXXX|2.3 3.3 4.3
             XXXXXXX|
   1        XXXXXXXX|
          XXXXXXXXXX|
           XXXXXXXXX|
   0       XXXXXXXXX|
           XXXXXXXXX|1.2
           XXXXXXXXX|
  -1       XXXXXXXXX|2.2 3.2 4.2
           XXXXXXXXX|
  -2          XXXXXX|
             XXXXXXX|
                XXXX|
  -3            XXXX|
                XXXX|1.1 3.1 4.1
                  XX|2.1
  -4              XX|
                  XX|
                   X|
  -5               X|
                   X|
                   X|
  -6               X|
                   X|
                   X|
  -7                |
                    |
                    |
 ===============================================================================
Each 'X' represents   98.9 cases
The labels for thresholds show the levels of item, and step, respectively

================================================================================
```

Note: The item labels correspond to the question numbers in Table A.9 and their recoded response categories (the zero category is not shown). For example:

Would like to do more science in school
1.1= Disagree a little
1.2= Agree a little
1.3= Agree a lot

Figure A.9. Item Map from "Students Like Learning Science" Scale (Separate Science)

```
================================================================================
MAP OF LATENT DISTRIBUTIONS AND THRESHOLDS
================================================================================

                  |
                  |
                  |
                  |
  2               |
                  |
                  |13.3
                 X|14.3
                  |16.3
                 X|
                 X|9.3
                 X|1.3 5.3
                XX|2.3 6.3 10.3 15.3
                XX|13.2
  1            XXX|8.3 12.3
               XXX|4.3
              XXXX|3.3 7.3 11.3
             XXXXX|
             XXXXX|9.2 14.2
             XXXXX|
           XXXXXXX|1.2 5.2 16.2
            XXXXXX|
          XXXXXXXX|
        XXXXXXXXXX|10.2
  0        XXXXXXX|15.2
          XXXXXXXX|6.2 12.2
           XXXXXXX|2.2 13.1
          XXXXXXXX|8.2
          XXXXXXXX|4.2 11.2
            XXXXXX|7.2
            XXXXXX|3.2 16.1
             XXXXX|9.1 14.1
               XXX|5.1
               XXX|15.1
 -1            XXX|1.1 12.1
                XX|10.1
                XX|7.1 8.1 11.1
                 X|2.1 4.1 6.1
                 X|3.1
                 X|
                 X|
                 X|
                 X|
                  |
                  |
 -2               |
                  |
                  |
================================================================================
Each 'X' represents   63.3 cases
The labels for thresholds show the levels of item, and step, respectively
================================================================================
```

Note: The item labels correspond to the question numbers in Table A.9 and their recoded response categories (the zero category is not shown). For example:

<u>Would like to do more biology in school</u>
1.1= Disagree a little
1.2= Agree a little
1.3= Agree a lot

Table A.10. Items in the “Schools Have Few Discipline and Attendance Problems” Scale

| **Schools Have Few Discipline and Attendance Problems Scale** | |
|---|---|
| Source | Based on principal responses to the following question in the School Questionnaire:<br><br>1. If the behavior occurs, how severe is the problem? - Arriving late at school<br>2. If the behavior occurs, how severe is the problem? - Absenteeism (i.e., unjustified absences)<br>3. If the behavior occurs, how severe is the problem? - Skipping class <hours/periods><br>4. If the behavior occurs, how severe is the problem? - Violating dress code<br>5. If the behavior occurs, how severe is the problem? - Classroom disturbance<br>6. If the behavior occurs, how severe is the problem? - Cheating<br>7. If the behavior occurs, how severe is the problem? - Profanity<br>8. If the behavior occurs, how severe is the problem? - Vandalism<br>9. If the behavior occurs, how severe is the problem? - Theft<br>10. If the behavior occurs, how severe is the problem? - Intimidation or verbal abuse of other students<br>11. If the behavior occurs, how severe is the problem? - Physical injury to other students<br>12. If the behavior occurs, how severe is the problem? - Intimidation or verbal abuse of teachers or staff |
| Response Categories | 1= Not a problem<br>2= Minor problem<br>3= Serious problem |
| Recoded Response Categories | 0= Serious problem (Original category 3)<br>1= Minor problem (Original category 2)<br>2= Not a problem (Original category 1) |

Figure A.10. Item Map from "Schools Have Few Discipline and Attendance Problems" Scale

```
================================================================================
MAP OF LATENT DISTRIBUTIONS AND THRESHOLDS
================================================================================
                  |
   6              |
                  |
                 X|
                 X|
   5              |
                XX|
                 X|
               XXX|
   4           XXX|
                XX|
               XXX|
   3           XXX|
              XXXX|
             XXXXX|
             XXXXX|
   2       XXXXXXX|10.2
            XXXXXX|1.2 5.2
           XXXXXXX|
           XXXXXXX|2.2 7.2
   1     XXXXXXXXX|4.2
          XXXXXXXX|
        XXXXXXXXXX|
           XXXXXXX|8.2 9.2
   0      XXXXXXXX|
          XXXXXXXX|6.2 11.2
          XXXXXXXX|
  -1     XXXXXXXXX|3.2
        XXXXXXXXXX|
          XXXXXXXX|12.2
          XXXXXXXX|
  -2          XXXX|10.1
              XXXX|
              XXXX|1.1 2.1 5.1
               XXX|7.1
  -3            XX|
                XX|
                XX|11.1
                 X|
  -4             X|4.1
                 X|3.1 12.1
                 X|8.1
  -5              |9.1
                 X|
                  |
                  |6.1
  -6              |
================================================================================
Each 'X' represents 136.4 cases
The labels for thresholds show the levels of item, and step, respectively

================================================================================
```

Note: The item labels correspond to the question numbers in Table A.10 and their recoded response categories (the zero category is not shown). For example:

<u>Arriving late at school</u>
1.1= Minor problem
1.2= Not a problem

Table A.11. Items in “Schools Have a Climate Supportive of Academic Success” Scale

| Schools Have a Climate Supportive of Academic Success Scale | |
|---|---|
| Source | Based on principal responses to the following question in the School Questionnaire:<br><br>1. How would you characterize teachers' understanding of the school's curricular goals within your school?<br>2. How would you characterize teachers' degree of success in implementing the school's curriculum?<br>3. How would you characterize teachers' expectations for student achievement?<br>4. How would you characterize parental support for student achievement?<br>5. How would you characterize students' desire to do well? |
| Response Categories | 1= Very high<br>2= High<br>3= Medium<br>4= Low<br>5= Very low |
| Recoded Response Categories | 0= Low or Very low (Original categories 4 and 5)<br>1= Medium (Original category 3)<br>2= High (Original category 2)<br>3= Very high (Original category 1) |

Figure A.11. Item Map from “Schools Have a Climate Supportive of Academic Success” Scale

```
================================================================================
MAP OF LATENT DISTRIBUTIONS AND THRESHOLDS
================================================================================
                      |
   4                 X|
                     X|5.3
                    XX|
                    XX|
   3                XX|4.3
                    XX|
                   XXX|2.3
                   XXX|
   2             XXXXX|3.3
                 XXXXX|1.3
                XXXXXX|
   1            XXXXXX|
               XXXXXXX|
             XXXXXXXX|4.2
          XXXXXXXXXX|5.2
   0        XXXXXXXX|
            XXXXXXXX|
           XXXXXXXXX|
           XXXXXXXXX|
  -1      XXXXXXXXXX|2.2
            XXXXXXXX|
             XXXXXXX|3.2
                XXXXX|
  -2           XXXXXX|1.2 4.1
                  XXXX|
                  XXXX|
                   XXX|
  -3                XX|
                     X|
                     X|
  -4                 X|
                      |
                      |3.1 5.1
                      |1.1
  -5                  |2.1
                      |
================================================================================
Each 'X' represents 166.2 cases
The labels for thresholds show the levels ofitem, and step, respectively
================================================================================
```

Note: The item labels correspond to the question numbers in Table A.11 and their recoded response categories (the zero category is not shown). For example:

<u>Teachers' understanding of the school's curricular goals</u>
1.1= Medium
1.2= High
1.3= Very high