FACTORS RELATED TO STUDENT ACHIEVEMENT IN MATHEMATICS AND COMPARISON OF THE U.S. WITH OTHER COUNTRIES: A STUDY BASED ON TIMSS 2007 REPORT

by

Venkata Subbaiah Patnam
A Dissertation
Submitted to the
Graduate Faculty
of
George Mason University
in Partial Fulfillment of
The Requirements for the Degree
of
Doctor of Philosophy
Education

Committee:

____ Chair

____

____

____ Program Director

____ Dean, College of Education and Human Development

Date: ______________________________________ Summer Semester 2013
George Mason University
Fairfax, VA

Factors Related to Student Achievement in Mathematics and Comparison of the U.S. with Other Countries: A Study Based on TIMSS 2007 Report

A Dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy at George Mason University

by

Venkata Subbaiah Patnam
Bachelor of Education
Mysore University, 1976
Master of Science
Patna University, 1961
Bachelor of Science
Andhra University, 1959

Director: Margret A. Hjalmarson, Associate Professor
Graduate School of Education

Summer Semester 2013
George Mason University
Fairfax, VA

Dedication

To Sri Sri Thakur Anukulchandra, my Master and Guru,
Who was an embodiment of Love and Wisdom,
Who led multitudes of seekers on enlightened paths of *Being* and *Becoming*,
Who was a friend, philosopher, and guide to many,
And with whose blessings, inspiring encouragement, and at whose behest,
I have spent long years in enlivening labor of love for *Learning*.

His messages on education include:

1. 'Education is to know existence in consonant contrast to environment by doing and discerning', and
2. 'To bring out and materialize a congruity from among varieties and unify them meaningfully discovering their relation to existence— is the essence of education'. (*The Message*, Vol. VIII, pp.11-12, Satsang Publishing House, Deoghar, 1995.)

Acknowledgements

I am grateful to my dissertation committee, Dr. Margret Hjalmarson, Dr. Dimiter Dimitrov, and Dr. Jennifer Suh for their support and guidance. Throughout the development of my dissertation, Dr. Hjalmarson helped me with her meticulous scrutiny and suggestions for improvement. Dr. Dimitrov provided me with very useful cues for my methods and Dr. Suh provided me with effective feedback and tips for improvement.

I had the pleasure of working as Research Assistant to Dr. Robert Smith for 4 years on *Teaching for Meaning* project in Educational Leadership. This experience considerably augmented my efforts in my main research efforts and also gave me an opportunity to co-present a paper at the annual meeting of American Educational Research Association in April 2012 at Vancouver, Canada. Thanks a lot to you, Dr. Smith.

Further, I also owe thanks to several of my other Professors and Administrators in the Education department of George Mason University for their help and encouragement. In particular, I would like to highlight the munificent contribution of Dr. Gary Galluzzo, who was instrumental in fine-tuning my vision for dissertation research. Ever since I came into contact with him in my Ph. D. program, I always found in him a ready resource for moral support and holistic guidance. I am beholden to you, Dr. Galluzzo.

My pursuit of this research work has been possible only with the care and help of my family. Apart from providing me with affectionate personal attention and unsparing service, my wife Vijaya helped me with occasional proof-reading and data entry tasks. My sons, Ravi and Krishna helped me to update my computer techniques and have been constant troubleshooters. They stood by me from the beginning and proved to be pillars of strength to keep my perseverance aloft. I express my gratitude, love, and unbounded appreciation to them.

My thanks and appreciation also go to my dearest friend and brother-in-spirit, Arun Ganguly, my loving niece Saroja, and my uncles, Ramana Rao (a former Judge, who always permeates peace and amity around) and Kesava Chandran (an altruistic former bank manager, who always seeks to play the role of a Good Samaritan) and their lovely families. I also thank my old students and family-friends, Dr. Kartik Viswanathan, M.D., Dr. Suresh Prabhu, Ph.D., and Dr. Vijaya Subrahmaniam, Ph.D. for their active encouragement to me and my family. Finally, I express my eternal gratitude and veneration to my late mother Ratnamma, who lived a simple life of virtue, piety, and self-denial and, in spite of great bereavements in her life, tenaciously nurtured and nourished my body and soul with her unbounded love.

Table of Contents

List of Tables

List of Figures

Abstract

FACTORS RELATED TO STUDENT ACHIEVEMENT IN MATHEMATICS AND COMPARISON OF THE U.S. WITH OTHER COUNTRIES: A STUDY BASED ON TIMSS 2007 REPORT

Venkata Subbaiah Patnam, Ph.D.

George Mason University, 2013

Dissertation Director: Dr. Margret A. Hjalmarson

A purpose of this study was to explore the inter-relations among eighth grade mathematics student achievements in the content domains of *Number, Algebra, Geometry, and Data & Chance* and the cognitive domains of *Knowing, Applying, and Reasoning*, in the context of the Trends in International Mathematics and Science Study (TIMSS), 2007. A second purpose was to find the associations between student achievements in the cognitive domains with student-related, teacher-related, school-related, and home-related variables. The variables were selected mainly on the basis of Carroll's model of school learning. Further, the math achievement of the students of the United States in each of the domains was compared with *Bulgaria, Chinese Taipei, Republic of Korea, Singapore,* and *Thailand*. After analysis, interpretations were made on the findings of the relationships among student achievements in each of the content and cognitive domains and also on the associations between variables and math achievement

in the four content domains. Further, comparative performance of the U.S. with each of the five selected countries in each of the seven domains was analyzed.

For finding the relationships among student performance scores across all the countries in eighth grade mathematics in different domains, Pearson's product moment coefficient of correlation was used as the measure. For finding the associations of the scores with the variables, multiple regression method was employed. Finally, for comparing student performances of the U.S. with other states, independent samples *t*-test was used.

The findings of this study indicated that student performance in all the content and cognitive domains were highly correlated. The second finding was that the variable '*School Resources*' was significantly related to student performance in each of the content domains, whereas '*Overall Home Involvement*' had moderately significant relation with student achievement in *Number* and *Data & Chance*. The findings revealed no significant relation between the other variables considered and student achievement in the four content domains. The third finding showed that *Chinese Taipei*, *Republic of Korea*, and *Singapore* outperformed the U.S. in all the domains, whereas the U.S. performed better than *Bulgaria* and *Thailand* in each of the domains. These results provide some clues for improving classroom instruction and streamlining administrative priorities in order to improve student performance in the United States. Furthermore, the results indicate paths for further research in order to augment the educational practice not only in the U.S., but in all countries, for securing better student achievement in mathematics.

Chapter One: Introduction

The Background

> *Improving mathematics and science education in the United States is a vitally important task, one that will only be accomplished if state policymakers and interested stakeholders work together to meet the needs of both teachers and students – (Zinth & Dounay,2006, p 8).*

Moses (2001) sounded like a sentinel with his observation, ‘In today's world, economic access and full citizenship depend crucially on math and science literacy” (p.5). Echoing these remarks, Zinth and Dounay (2006) further clarified that in recent times, the international scenario in respect of global economy lent itself to sophisticated procedures that use high technology, which proved to be more advantageous to nations that have several individuals with good quality mathematics and science education in terms of global competition. Consequently, public interest and national priorities in mathematics and science education have assumed high importance. In the United States, with growing public awareness of the need to improve the performance of schools in the wake of the release of *A Nation at Risk* (National Commission for Excellence in Education, 1983), and *A Nation Prepared: Teachers for the 21st Century* (Carnegie Forum on Education and the Economy, 1986), a climate of educational reform took shape. Much before the launching of the Trends in International Mathematics and Science Study (TIMSS) in 1995, Stevenson, Lee, and Stigler (1986), who observed that student performance of the

United States was poor as compared to some other countries, raised a concern among leaders that the graduating students of the United States might not be adequately equipped to meet the demands of global economy. Subsequently, the following developments took place in the educational scenario of the United States, as mentioned by Wang, Haertel, and Walberg (1993b):

> Reports of students' poor performance also led to America 2000, the first federally initiated reform of U.S. schools (U.S. Department of Education, 1991). America 2000: An Education Strategy Sourcebook was built around six national goals agreed on by the 50 governors to increase the achievement of U.S. students. America 2000 has four tracks designed to promote educational excellence by improved accountability, new technology, lifelong learning, and greater parental and community involvement. The America 2000 program was adopted by all 50 states and implemented in thousands of communities across the United States. (p.251)

Further, international comparative studies such as the TIMSS and the Program for International Student Assessment (PISA) have enhanced the interest in student achievement in mathematics and science (Sfard, 2005). In this regard, the policy statement of TIMSS calls for attention (TIMSS, 2007, p.1).

> There is almost universal recognition that the effectiveness of a country's educational System is a key element in establishing competitive advantage in what is an increasingly global economy. Education is fundamentally implicated

> not only in a country's economic and social development, but also in the personal development of its citizens. It is considered one of the primary means whereby inequities, social and economic can be reduced. Attendant on this growing recognition of the importance and centrality of education has been the recognition, worldwide, of the importance of regular monitoring of educational performance and its antecedents.

Some probing studies have been made to examine the mutual influence of large-scale mathematics assessments and standards efforts in the United States. The National Assessment of Education Progress (NAEP) and TIMSS appear to have influenced the shaping of the *Principles and Standards for School Mathematics* (2000) from the original *Curriculum and Evaluation Standards for School Mathematics* (1989), which was developed by the National Council of Teachers of Mathematics (NCTM). On the other hand, it is speculated that the *Principles and Standards for School Mathematics* would help reshape the future frameworks of NAEP and the TIMSS (Lindquist, 2001). The NCTM explained its outlook on the *Standards* as follows:

> "What it means to be mathematically literate in a world that relies on calculators and computers to carry out mathematical procedures and in a world where mathematics is rapidly growing and is extensively being applied in diverse fields" (NCTM, 1989, p. 1).

In this context, it may be pertinent to note that Schmidt, Wang, and McKnight (2005), who made a comparative study of the coherence and standards in mathematics among the highest performing countries in TIMSS with those of some states and school districts in the United States, made a pointed suggestion to the policy makers in the US that curriculum standards ought to have a national scope, in order that coherence and rigor might be ensured.

Statement of the Problem

Fifty nine countries participated in TIMSS (2007). This study gave a wide data base of student achievement of all the participating countries in science and mathematics at fourth grade and eighth grade levels. A study of student achievement corresponding to particular variables will show the comparative performances of various nations. Inferences could be drawn on the connections between the various factors in learning situations and student achievement, so that they could be applied with advantage in other countries for possible augmentation in student performance. However, some scholars like Ravitch (1986), Bennett (1987), Finn, (1992), Berliner and Biddle (1995), and Bracey (1997) were cynical regarding the use of the inferences from large-scale comparative assessment studies across culturally diverse nations in order to shape the educational policy in the United States or in any other country. They contended that global cultural aspects influence the national patterns of schooling in any country, according to anthropological and sociological studies. So, utilizing the data from TIMSS would be tantamount to comparing results that accrued from contrasting situations, especially with regard to the non-instructional duties of teachers and the personal living styles of students. LeTendre, Baker, and Akiba (2001) suggested that the emphasis ought to be on the identification of the effective features of teacher work, followed by efforts to explore ways to improve the instructional conditions in the native cultural background. They argued that education policy ought not to try to change the indigenous value system, simply because some other nations with different value systems produced better student achievement. Regarding this issue, observing that the standards of attainment in school mathematics and the belief systems regarding the values and purpose are closely connected, Macnab (2000) commented that the TIMSS reports did not reflect a correct vision for mathematics education that could be effectively implemented in the classrooms.

Again, as pointed out by Wang (2001), the validity of the reports based on TIMSS data in varying international contexts in projecting a particular country's rankings came to be looked upon with some suspicion by some scholars. They argued that instrument construction, curricular inequalities, and statistical outliers are a matter of grave concern for using TIMSS results as benchmarks for school reforms and for interpreting TIMSS results. In the same vein, Holliday (1999) felt that it would be more pertinent and valuable to compare the results of culturally and economically similar nations and advised science teachers to ignore the studies based on TIMSS data on student achievement and encourage students in problem solving, developing concepts, and in knowing science facts.

However, these points about the validity and appropriateness of common test questions for all countries were comprehensively addressed in the TIMSS documents as follows: (TIMSS, 2007, PP. 27 -28).

… To ensure the reliability, validity, and comparability of the data through careful planning and documentation, cooperation among participating countries, standardized procedures, and rigorous attention to quality control throughout. The data are collected according to rigorous scientific standards detailed in manuals, and countries receive training every step of the way…

The student sampling for TIMSS 2007 was conducted with careful attention to quality and comparability. The sampling was designed to ensure that the data provided accurate and economical estimates of the student population. ... If procedures did not satisfy the TIMSS standards, the data are annotated in the report (or not reported at all).

Adherence to the test administration procedures was monitored through the use of international quality control observers arranged by the IEA Secretariat, and within-country quality control procedures. … Reliability data were collected for within-country scoring and across assessment cycles using special procedures developed by the IEA Data Processing and Research Center… The IEA Data Processing and Research Center checked each country's data files for internal consistency and accuracy, and interacted with countries to resolve data issues.

…In order to measure trends in mathematics achievement across assessments, the TIMSS achievement scales for mathematics were designed to provide reliable measures on a common metric established originally with the 1995 assessment, and now spanning the 1995, 1999, 2003, and 2007 assessments.

It is obvious that progress in academic achievement can only be known by comparing one's current achievement with that in the previous years and with the achievement of others. To this purpose, TIMSS serves as a veritable repository of useful data, and as Ferrini-Mundy and Schmidt (2005) pointed out, the assessments such as TIMSS and PISA serve as rich resources for researchers in mathematics education by providing data not only about student performance, but also on several contextual variables. Otherwise, without having baseline information about curriculum and the nature of instruction, it would not be possible to understand the shifts in student performance. Moreover, according to them, the framework and the tools used in these surveys, such as questionnaires on students' home backgrounds, resources for learning etc. could be utilized in several research studies.

A close analysis of the TIMSS (2007) data on student achievements in mathematics in different countries will reveal the nature of the various learning situations that promote better student achievement in varying cultural contexts in the content and the cognitive domains. How far these insights could be used to more effectively harness the learning situations in the United States or any other country could be examined later, depending upon the prevailing policy and fiscal situations.

Purpose of the Study

The purpose of this study is to explore the connections between student achievement in the content and cognitive domains of mathematics at the 8th grade level in different learning situations across several nations, by analyzing the data of TIMSS (2007). I have chosen to study student achievement in mathematics because of the importance of the subject. Also, I have chosen the eighth grade level, since it is a transitional point for high school stage, where student learning is in a formative stage in terms of approach and attitude , the ‘formal operational stage’ according to Piaget’s ‘cognitive development theory’. I intend to analyze the data from the TIMMS (2007) report and other resources related to eighth grade mathematics achievement in Singapore, Chinese Taipei, Korea, United States, Bulgaria, and Thailand. I have chosen the first three countries, because their overall performances in mathematics proficiency have been the best, according to the TIMSS (2007) report and also according to the TIMSS (1999) report. In addition, they were among the top four countries as per the TIMSS (2003) report. I wanted to compare them with the student achievement in the United States, because the United States is the most progressive country in terms of education at higher levels, apart from being the world leader in science and technology. The current downward trend in the giant economy of the United States, coupled with the claimed relationship between the economy and student achievement in math and science, has added weight to my decision to have comparisons of student achievement in other countries with that in the U.S. I have included Bulgaria since its average performance, which was slightly better than the United States as per the TIMSS (1999) report, has fallen to less than the international average, as per the TIMSS (2003) and TIMSS (2007) reports. Lastly, Thailand, which is an Asian country like the top three ones, has been scoring consistently midway among the countries below the international average, as per all the TIMSS reports.

I propose to analyze the data of the average scores of student achievement in the areas of content and cognitive domains pertaining to various strands of variables, such as student and family characteristics, student computer use, student activities outside of school, student engagement in mathematics, teacher background, teacher preparation and collaboration, classroom characteristics, classroom instruction, role of homework, school characteristics, school resources, home involvement, and school climate and safety. After analysis, interpretation will be made to indicate plausible linkages between student achievement in the content and cognitive domains and the learning situations, as depicted by the variables.

What is TIMSS?

Since 1959, the International Association for the Evaluation of Educational Achievement (IEA) has been conducting studies of student achievement across several nations. TIMSS is one of its projects, which is conducted assessment regularly every four years in 1995, 1999, 2003, and 2007 at the fourth grade and eighth grade levels. A cohort of students at the fourth grade level would be at the eighth grade level in the next cycle of TIMSS assessment, so that progress across grades could be known. It also provides "comparative perspectives on trends in achievement in the context of different educational systems, school organizational approaches, and instructional practices, TIMSS collect a rich array of background information" (TIMSS 2007, p. 14). The most recent assessment, TIMSS 2007 involved approximately 425,000 students from 59 countries around the world.

Eighth grade mathematics assessment in TIMSS (2007) was organized around two dimensions, one in the content domain and the other in the cognitive domain. The content domain covers four subject areas – number, algebra, geometry, and data & chance. The cognitive domain encompasses three areas of mathematical thinking, viz., *knowing*, *applying*, and *reasoning*. The 'knowing' area measures students' knowledge of mathematical facts, procedures, and concepts, whereas the 'applying' area focuses on the ability to apply knowledge and conceptual understanding to solve problems. Assessment in the 'reasoning' area is meant to gauge the students' capacity to tackle unfamiliar situations, complex contexts, and multi-step problems (Mullis, et al., 2005). The variables are classified broadly into nine categories: (i) student and family characteristics; (ii) student activities outside of school; (iii) student engagement in mathematics; (iv) teacher preparation and collaboration; (v) classroom characteristics; (vi) classroom instruction; (vii) role of homework; (viii) school resources; and (ix) school climate and safety (TIMSS, 2007).

Research Questions

The study will be aimed to address the following set of research questions:

1. Are the TIMSS 2007 eighth-grade mathematics scores in the content domains of Number, Algebra, Geometry, and Data & Chance and the scores in the cognitive domains of Knowing, Applying, and Reasoning across the countries associated with one another?

2. Are the TIMSS 2007 eighth-grade mathematics scores in Number, Algebra, Geometry, and Data & Chance across the countries associated with the variables related to factors in student background (amount of time spent on math homework), classroom instruction (frequency of applying facts, concepts and procedures), home involvement (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), school resources (for math instruction), and mathematics teacher professional development (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy /instruction, content, and assessment)?

3. How does the performance of the United States in the content domains Number, Algebra, Geometry, and Data & Chance and the cognitive domains of Knowing, Applying, and Reasoning compare with the performances of Bulgaria, Chinese Taipei, Republic of Korea, Singapore, and Thailand?

Significance of the Study

There has been a continual improvement in the United States student performance in school mathematics. 'Eighth-graders scored higher on average in 2009 than in all previous assessments' (NAEP, 2009). The average grade 8 scores of 2009 increased by

20 points over those of 1999 (Aud & Hannes, 2009a). In the international arena, as reflected by TIMSS scores, the performance of the students of U.S. in mathematics at the eighth grade level has been on a steady rise. Starting from below international average in 1995, the United States has come to occupy the 9th rank in overall score in TIMSS (2007).

After analyzing the TIMSS (2007) results, Gonzales (2009, pp. iii - iv) reported the following important derived facts on the relative performance of the United States in the international comparative study in mathematics (see Appendix 1a and Appendix 1b).

- In 2007, the average mathematics scores of both U.S. fourth-graders (529) and eighth-graders (508) were higher than the TIMSS scale average (500 at both grades).* At eighth grade, the average U.S. mathematics score was higher than those of students in 37 of the 47 other countries, lower than those in 5 countries (all of them located in Asia), and not measurably different from those in the other 5 countries.

- Compared to 1995, the average mathematics scores for both U.S. fourth- and eighth-grade students were higher in 2007. At eighth grade, the U.S. average mathematics score in 2007 was 508, 16 points higher than the 1995 average of 492.

- In 2007, 6 percent of U.S. eighth-graders scored at or above the advanced international benchmark in mathematics. At grade eight, seven countries had higher percentages of students performing at or above the advanced mathematics benchmark than the United States: Chinese Taipei, Korea, Singapore, Hong Kong SAR, Japan, Hungary, and the Russian Federation. These seven countries include the five countries that had higher average overall mathematics scores than the United States, as well as Hungary and the Russian Federation.

- In 2007, the average science scores of eighth-graders (520) were higher than the TIMSS scale average (500 at both grades). At eighth grade, the average U.S. science score was higher than the average scores of students in 35 of the 47 other countries, lower than those in 9 countries (all located in Asia or Europe), and not measurably different from those in the other 3 countries.

(* TIMSS provides two overall scales—mathematics and science—as well as several content and cognitive domain subscales for each of the overall scales. The scores are reported on a scale from 0 to 1,000, with the TIMSS scale average set at 500 and standard deviation set at 100.)

However, the connections between student achievement in mathematics in the content and cognitive domains, as well as their associations with different variables connected to students, teachers, instruction, school, and home are not known. These, along with the knowledge of the performance position of the United States in comparison to other countries in both the content and cognitive domains might help fill the gaps in the curricular priorities and in the learning situations in the United States. This dissertation study aims at comparing the student achievement in the United States with that in the other countries studied, as affected by the variables related to student background, teacher, instruction, school, and home associated with student achievement. The findings are expected to provide some clues for augmenting the educational practice in the United States.

Theoretical Framework

Learning is conceived more "as an active cognitive process rather than as a product of environmental conditions. Because of their long developmental history, cognitive developmental processes influence child learning and help to determine the effects of instruction upon learning and transfer" (Wittrock, 1987, p. 137). Bloom (1981) considered three variables, viz., entry level student characteristics, quality of instruction, and learning outcomes to form the model of his theory of school learning. According to him, education should aim at equality of learning outcomes rather than equality of opportunity.

Earlier, Dewey (1910) had stated that reflective inquiry had five stages -- suggestion, problem, hypothesis, reasoning, and testing. Later, Bruner (1977) stressed the factors that are basic to learning as: importance of structure; readiness for learning; intuitive and analytical thinking; motives for learning; and aids to teaching. Again, in 1963, Carroll (1963) developed a model of school learning, based on five variables – aptitude, opportunity to learn, perseverance, quality of instruction, and ability to understand instruction. Aptitude was defined as the amount of time needed to learn a given task, implying that the more the time needed, the less is the aptitude. Opportunity to learn was defined as the amount of time given for learning. Perseverance was taken as the time a student is willing to spend for learning. Quality of instruction and ability to understand relate to achievement. Carroll’s model, however, did not specifically indicate the characteristics of the quality of instruction, apart from stating that learners must be told about what they are to learn, given contact with learning materials, and planning the steps for learning. As regards the ability to understand, Carroll indicated that it is “the learners' ability to figure out for themselves what the learning task is and how to go about learning it” (p.26). Further, he suggested that “when the variables of quality of instruction and opportunity to learn are properly managed, the variable of student perseverance—willingness to learn—will take care of itself” (p. 30).

All said and done, what are the influences of the educational, psychological, and social factors on learning? An answer to this question was sought by Wang, Haertel, and Walberg (1993a), who made a study by gathering evidence from 61 research experts, 91 meta-analyses, and 179 handbook chapters and narrative reviews. Their analyses were three-pronged: content analyses, expert ratings, and results from meta-analyses. As per the findings of their study, the knowledge base for school learning depends mainly on proximal variables in the psychological, instructional and home environments, and not so much on variables on the distal variables of demography, policy, and organization. They also made a candid observation that “findings on what affects learning might aid practitioners and policymakers in the construction of more tightly focused, coherent, and cost-conscious agenda for reforming U.S. schools” (Wang, et al., 1993b, p. 367). During the past decade, there have been mainly two approaches to link student achievement to teaching effectiveness. In the first approach, for e.g., the study of Borko (2004), researchers concentrated more on the global aspects of teaching and analyzed the teaching patterns by using large-scale statistical surveys. The techniques involved controlling of the external variables by employing sophisticated statistical models, as discussed by Rowan, Correnti, and Miller (2002) and Raudenbush, (2004). On the other hand, in the second type of approach, researchers like Snow et al. (1980) concentrated more on the specific domains of knowledge that characterized multidimensional , i.e., cognitive, affective, and metacognitive aspects of learning.

The evaluation of student achievement is the touchstone of instructional effectiveness. Seidel and Shavelson (2007) made revealing observations with regard to this aspect (p.462):

> Evaluation of learning characterizes teaching acts that aim to assess student progress toward learning goals. And the regulation and monitoring component includes teaching acts such as feedback and support or teaching students strategies of self-regulation and self-monitoring.

Therefore, an in-depth analysis of the reports of TIMSS, which is the biggest and most comprehensive database of student achievement on a global scale, appears to be the right means to know the comparative student achievements and the finer nuances of the influences of a variety of independent variables. This dissertation study draws data from the TIMSS (2007) on student achievement with respect to variables related to student background, teacher, instruction, school, and home. Since the variables considered in this study more or less tally with the variables of Carroll’s model, this study conforms to a great extent to Carroll’s model of school learning.

Limitations

Using the assessment of achievement of a sample of the population to estimate the achievement of the total population entails the possibility of errors that are made in the collection and processing of data. Apart from these non-sampling errors, some sampling errors, in the form of standard errors, are also likely to occur when there are discrepancies between sample estimates and population characteristics. When the standard errors are significant, the interpretations and conclusions will not be reliable (Aud & Hannes, 2009b).

There was scope for non-response bias in the U.S. TIMSS samples. There were some indications for potential bias from regional and community-type differences. Moreover, some schools having large numbers of minority students were less likely to participate, although, when substitute schools were included, the measurable differences were small (TIMSS, 2007). Moreover, school entrance requirements and tracking practices in different countries can also impact the population tested. There’s also a limit to what can be learned from a survey and a multiple choice test, by design.

Definitions

Eighth-grade students: In the TIMSS study, eighth-grade students are defined as all students enrolled in the upper of the two adjacent grades that contained the largest proportion of 13-year old students at the time of testing.

Scores: The scores, which are reported on a scale from 0 to 1,000, with the TIMSS scale average set at 500 and standard deviation set at 100.

Chapter Two: Literature Review

This chapter is organized in seven major sections: (i) history of international assessments of mathematics achievement, (ii) theoretical framework, (iii) student-related factors and student achievement, (iv) institution and home-related factors and student achievement, (v) instruction and teacher-related factors and student achievement, (vi) factors related to achievement in algebra and geometry, and (vii) student achievement in the United States. At the end of each section, a summary and comments pertaining to the section are given. Lastly, a comprehensive chapter summary of the literature review is given, followed by final comments and hints at connected future directions.

History of International Assessments of Mathematics Achievement

The following section traces out the salient stages in the development of international large scale assessments of student achievement from the beginning to the current situation.

Historical development. A nation's security and potential for economic development depends to a considerable extent on the quality of education in mathematics and science of vast numbers of its citizens (Moses, 2001; Zinth & Dounay, 2006; Akiba, LeTendre, & Scribner, 2007). This issue was highlighted in the 5th annual report of the U.S. Department of Education.

> In order to strengthen our nation's competitiveness in the global marketplace, as well as our security at home, we must be certain that teacher proficiency in mathematics, science, technology, and foreign languages is sufficient to enable America's students to achieve at grade level and above in these subjects. (Akiba et al., 2007, citing U.S. Department of Education, 2006, p. iii)

By knowing the progress of education and student achievement of other countries, every country stands to gain, because some improvements can be incorporated after comparisons with student achievement in various countries and finding their influencing factors.

In 1959, the United Nations Educational, Scientific and Cultural Organization (UNESCO) initiated the *Pilot Twelve-Country Study*, the first large-scale international achievement assessment. The study involved Belgium, England, Finland, France, Federal Republic of Germany, Israel, Poland, Scotland, Sweden, Switzerland, United States and Yugoslavia (EER, 1960). The UNESCO envisaged:

> …educational systems cannot be transferred from one country to another, but ideas, practices, and devices developed under one set of conditions can always prove suggestive for improvement even where the conditions are somewhat different" (EER, 1960, p. 621).

According to Forshay et al. (1962), this study was made in three languages--French, English, and German and later translated by the participating countries into 8 languages. However, the test was actually administered in 1961 to samples of 13-year old students across the 12 countries, with two objectives: (1) to investigate whether some indications of the intellectual functioning could be deduced from the patterns of student responses across countries; and (2) to discover the possibilities and the difficulties attending a large-scale international study. Over a period of nearly half a century, 29 assessments of international student achievement were conducted by the International Association for the Evaluation of Educational Achievement (IEA), the International Assessment of Education Progress (IAEP), and the Organization for Economic Co-operation and Development (OECD). Gradually, the coverage of subject areas in these assessments increased to Math, Science, Reading, English, Literature Education, English as a Foreign Language, French as a Foreign Language Education, Writing, and Civic Education. In terms of the student populations, the assessments expanded to fourth, eighth, and twelfth grades. It is interesting to note that out of the 29 international achievement assessments conducted by IEA, 13 were mathematics assessments (IEA, 2007).

According to the Institute of Education Sciences (IES, 2007), with the coordination of OECD, the Program for International Student Assessment (PISA) started in 2000 and is administered every 3 years. Its focus is on the capabilities of 15 year old students in reading literacy, mathematics literacy, and science literacy, while emphasizing on the functional skills of and incorporating measures of competencies like problem solving. The Progress in International Reading Literacy Study (PIRLS), which began in 2001 and is administered every 5 years, is an international comparative study of reading literacy of the 4th grade students. The Trends in International Mathematics and

Science Study (TIMSS) commenced its first study in 1995 and repeats every 4 years to provide reliable comparative data on the mathematics and science achievement of the 4th and 8th grade students of the U.S. with that of the students in other countries.

TIMSS. TIMSS has a grade-specific structure with emphasis on curriculum unlike PISA, which has an age-specific structure with emphasis on the functional aspect of literacy, according to Lagrange (2007). Moreover, matching of curricular areas with the items of TIMSS was more than 90% for most of the countries that participated in TIMSS, while the U.S. and Hungary matched totally. Some details of the process of data collection, as mentioned by TIMSS (2007, p. 189) are:

> …Participants provided information about various educational policies and the curriculum topics covered in their respective curriculum guidelines (intended curriculum). Inclusion in the country's curriculum, however, does not guarantee students' opportunity to learn. Just as important is what their teachers choose to teach them. The lessons provided by the teachers ultimately determine the mathematics students are taught (implemented curriculum).

Apart from providing quantitative measures of student achievement in various strands, TIMSS has another feature in the form of video survey. However, even as this feature has projected some insights into a variety of learning situations, it has also exposed some of the challenges of studying classrooms across cultures and also some of the typical ways of instruction. There are some explicit benefits and also some inherent limitations of video survey approach (Stigler, Gallimore, & Hiebert, 2000). The advantages are:

> Video records of classroom lessons provide opportunities to discover unanticipated ideas and alternative analytic categories…The concrete nature of video means that it is not as theory bound as other methods of data collection. ..Video is not only interesting to researchers from different perspectives; it also has a longer shelf life than other kinds of data. ..Video provides concrete referents for the words and concepts used to describe instructional processes. ..Perhaps the greatest advantage of video is that it allows us to integrate qualitative and quantitative methods of analysis. (p. 90)

The limitations of video survey approach are that each classroom can be videotaped only a few times. In fact, in the TIMSS and TIMSS-R video studies, only one lesson was videotaped in each classroom, and therefore could not portray the contexts fully. Again, since video cannot give a complete picture of the real situation, some omissions of the real happenings are likely to occur, and camera effect also poses a potential problem. Taking all the above factors into consideration, the authors concluded that video surveys provide a good vista on classroom practice.

Assessment of student achievement, reforms, and standards. The gains of some nations from the TIMSS in the assessment of their own students' achievement in mathematics and science have been particularly rewarding when they did not have their own national assessment programs. In particular, the Federal Republic of Germany used the TIMSS for developing its own testing models for school improvement (Baumert & Köller, 2000). In the United States also, some standards-based reform efforts such as 'SciMath', the Minnesota's standards-based systemic reform in math and science education, were accentuated by the reports on TIMSS (Peterman & Linder-Scholer, 1999). In general, there has been a wave of introspection about student achievement, particularly in mathematics, in the wake of comparative studies using data from TIMSS. In the United States, it is now being realized that ‘U.S. students need to be challenged more in science and math, and this challenging work should be embedded in a supportive, humane environment that encourages students' interest in the subjects’ (Moore, 1998, p. 655). Standards-based curricula help students to do better when testing involves thinking, reasoning, and problem solving, but standards need to provide for equity also according to Stein (2007). In this connection, it is worth noting that the NCTM Standards (1989) and the Principles and Standards (2000) made explicit provisions for equity.

International comparative studies help policy makers in making proper decisions, with the knowledge of the correlates of achievement in a variety of educational systems by providing data on inputs, processes, and outputs simultaneously, as pointed out by Kellaghan (1996). But, according to Stedman (1997b), the achievements of different countries were not mutually comparable due to selection bias of samples, mismatch between the various curricula, bias due to nature of tests, and bias due to cultural difference. However, intervening in this debate, Zuzovsky (2002) argued against these reservations, stating that the variation among different countries gives information to diagnose the relative strengths and weaknesses of influencing factors, identify effective practices, and thus help in modifying standards and streamline accountability norms. Accordingly, all these factors would greatly help policy makers in tuning their systems of education.

However, curiously, the issue about the performance of the U.S. in mathematics, as reflected by the various reports that are based on TIMSS data has become a point of controversy among some scholars. Bracey (1997) stated that the performance of U.S. students in mathematics was much higher than the reports on TIMSS actually portrayed. Stedman (1997a) contradicted this view and suggested that it was wrong to think that the attainment of a majority of the students in the U.S. is world-class. He pointed out the deficiencies in mathematics achievement among U.S. students and even suggested the exclusion of Asian countries from international comparisons. Nevertheless, he clarified that the performance of U.S. students in mathematics has not been consistently poor, but varied according to the subject and grade level (Stedman, 1997b). Further, asserting that the mass of U.S. students was not unduly compared to the small academic elite of other countries, he pointed out that the poor performance of U.S. students was not due to a cultural conception of the learning systems and called for a fundamental school reform. Baker (1997) also supported this view and remarked that Bracey, who had alleged that the officials of the Education Department and the press had unduly raised serious concerns about the poor performance of U.S. students, had overstated about the performance of U.S. students.

Going a further step in understanding about mathematics achievement in the U.S. schools as compared to their international peers, some studies have probed into the curricular materials and teaching practices in different countries, based on data collected from the TIMSS (1999, 2003, and1995) reports. In its *Summary of Findings*, the U.S. National Research Center reported that the U.S. curricula in mathematics cover more topics than many others (U.S. TIMSS, 1996). Up to the ninth grade, the number of topics covered in the U.S. mathematics curricula was above the 75th percentile internationally. Many topics in U.S. curricula have been retained over several grades, unlike in several countries, ostensibly to develop mastery of the subject over the years. However, they do not focus on key goals, linking with other areas in the content, and do not set high demands on students. In the U.S., basics in mathematics are arithmetic, fractions, and a little of algebra, whereas in Germany, Japan, and several other countries, they mean algebra and geometry. However, the reform textbooks represent a significant shift for teachers and the materials are somewhat limited in terms of assisting teachers with implementing problem-based teaching which is often a departure for them. In the United States, since the curriculum is broader than in other countries, it appears that there is more demand on U.S. teachers to teach more topics in mathematics than in many other countries, including Germany and Japan. It was also pointed out that, as compared to others, the U.S. teachers teach more problem solving activities in mathematics. According to the report, the load of actual classroom teaching work on U.S. teachers is 30 hours per week, whereas it is 20 hours per week in Germany and Japan. The report recommended some drastic measures in terms of teacher quality, working conditions, curricula and textbook content, and student expectations. It may be pertinent to note in this context that recently, the U.S. government extended the time frame for teachers to reach up to the expectations of No Child Left Behind (NCLB) provisions. Recently, in December 2011, the Institute of Educational Sciences (IES) has taken a new step to understand the efforts of eighth grade students through an online study, *Access to Algebra 1: The Effects of Online Mathematics for Grade 8 Students*.

Summary. The economic development of a nation depends much on the quality of education in mathematics and science (Moses, 2001; Zinth & Dounay, 2006; Akiba, LeTendre, & Scribner, 2007). With this premise, in order to stand firm in global competitiveness, countries, including the U.S. to take steps to strengthen their education systems to ensure qualitative improvement in mathematics, science, technology, and foreign languages (Akiba et al., 2007, citing U.S. Department of Education, 2006). The UNESCO initiated the Pilot Twelve-Country Study, the first large-scale international achievement assessment in 1959 with a view to help all nations to gain in terms of ideas, practices, and devices to improve their own systems (EER, 1960). Three large-scale international studies were started with the coordination of OECD; TIMSS in 1995, PISA in 2000, and PIRLS in 2001.

TIMSS, with a grade-specific structure and emphasis on curriculum, provides quantitative measures of student achievement in various strands. It provides video surveys also for qualitative analysis. Based on TIMSS reports, some standards-based reform efforts such as ‘SciMath’, the Minnesota's standards-based systemic reform in math and science education took place in the U.S. (Peterman & Linder-Scholer, 1999). In spite of some controversies concerning the performance of the United States in the TIMSS, the role of international comparative studies in helping policy makers in taking proper decisions is viewed positively (Kellaghan, 1996).

Theoretical Framework

The basic theoretical framework for the present study is Carroll's model of school learning, which involves five measurable aspects of student achievement. However, along with further details of this model, some developments of significant thoughts about learning are also briefly discussed in this section.

According to Dewey (1910), the process of understanding begins with "some perplexity, confusion, or doubt (p. 12)". This idea was exemplified in Polya's (1957) remarks:

> "A great discovery solves a great problem but there is a grain of discovery in the solution of any problem. Your problem may be modest; but if it challenges your curiosity and brings into play your inventive faculties, and if you solve it by your own means, you may experience the tension and enjoy the triumph of discovery (p. v)."

Vygotsky (1978) seemed to further clarify this idea with his concept of 'proximal development' as the space in which the student's struggle could be productive. The knowledge of the subject that accrues as a result of the struggle was described as "the fruit of the undertakings that transform a problematic situation into a resolved one" by Dewey (1929, p. 242-243).

Carroll (1963) developed a model to explain the features of student success. He mentioned five variables: aptitude, opportunity to learn, perseverance, quality of instruction, and ability to understand instruction. All these five variables are expressible in terms of time and hence, measurable quantities. Aptitude was defined as the amount of time needed to learn a given task, implying that the more the time needed, the less is the aptitude. Opportunity to learn was defined as the amount of time given for learning, while perseverance was indicated by the time a student is willing to spend for learning. Quality of instruction and ability to understand are connected to achievement. Carroll's model of school learning suggested that the degree of learning was a function of the ratio of the amount of time spent in learning to the amount of time actually needed to for mastery learning. Speaking about Carroll's model for school learning, Hiebert and Grouws (2007) commented that it links teaching and learning with "considerations of students' entry knowledge, the nature and purpose of the tasks and activities, the likelihood of engagement" (p. 369). By taking time as well as topic coverage as an important variable, it had a preeminent influence on researchers. For the purposes of educational research, a correlational approach "is useful to create initial maps of the terrain when few hypotheses exist to suggest which features of teaching might facilitate which kinds of learning" (p. 397).

Further, the National Research Council (2004), while defining 'opportunity to learn' as "circumstances that allow students to engage in and spend time on academic tasks. . ." (p.333), pronounced that it "is widely considered the single most important predictor of student achievement" (p. 334). However, it is important to note here that 'opportunity to learn' is not indicted by only the amount of time spent for learning as in Carroll's definition, but takes into account the creation of the environment needed to support academic tasks.

Summary. While incorporating a correlational approach, this study involves the variables in TIMSS 2007 so chosen as to correspond with Carroll's five indicators to a large extent for finding correlations with student achievement. These chosen variables are related to factors in student background (amount of time spent on math homework), classroom instruction (frequency of applying facts, concepts and procedures), home involvement (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), school resources (for math instruction), and teacher professional development (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy /instruction, content, and assessment).

Student-Related Factors and Student Achievement

Gender and student beliefs are the two factors that were considered predominantly in the studies on the relationship between student-related factors and student achievement, using TIMSS reports.

Gender. The traditional notion that mathematics learning was predominantly a male stronghold seemed to be the subjected to scrutiny in some research studies that used the TIMSS student achievement data in respect of middle grade mathematics. Neuschmidt (2008) studied the TIMSS (1995-2003) reports to investigate the effect of gender differences among eighth grade students. Although in all these three surveys, boys performed better than girls in the 16 countries (Flemish Belgium, Cyprus, England, Hong Kong, Hungary, Iran, Japan, Korea, Latvia, Lithuania, New Zealand, Romania, Russian

Federation, Singapore, Slovak Republic, and the U.S.) which were in focus for this study, no significant changes in terms of gender differences were found in overall mathematics achievement. However in 1995, whereas girls outperformed boys only in Lithuania, boys outperformed girls in Iran, Japan, Korea, and the Slovak Republic, while in the remaining countries there was no significant difference in the performance of boys and girls. In 2003, girls outperformed boys in Cyprus and Singapore, while in Flemish Belgium, Hungary, and the United States boys outperformed girls. The study adopted the regression approach, as explained in Gonzalez and Miles (2001) and used the jackknife replication procedure for sampling errors, as explained in Martin and Kelly (1997). The findings of the study further revealed that from 1995 to 2003, the gap between the performances of girls and boys changed in favor of girls. In Korea, the significantly better performance of boys in 1995 decreased to just marginally higher performance than girls in 2003. Again, in Iran, boys scored significantly better than girls in 1995; but by 2003, the girls overtook the boys. Moreover, apart from these findings, some specific trends of achievement in the three content areas of algebra, measurement, and geometry were noticed over the three years. In the content area of algebra, boys performed better than girls in all the three years of the study, whereas girls performed better consistently in the area of measurement. But, in geometry, the results were not consistent; two countries showed better performances of girls in 1995 and 2003, four in 1995, one in 1999, whereas three countries showed that boys performed better than girls in 2003.

The above studies lead one to ponder whether there are any particular types of test items that are more amenable for drawing better performances from girls, and similarly, whether there are certain types of items at which boys perform better than at other items. Bielinski and Davison (2001) sought to find answers to these questions. They studied the interaction between sex difference and item difficulty, using multiple-choice questions from math test data from the 1992 National Assessment of Educational Progress (NAEP), the National Educational Longitudinal Study (NELS) of 1988, and the TIMSS (1995). The data in respect of NAEP participants involved 9,414 fourth graders (50.8% male), 10,291eighth graders (51.5% male), and 9,499 12th graders (49.6% male) and that in respect of NEL participants, 23,648 in eighth grade (50.8% male), 17,793 in 10th grade (51.5% male), and 14,236 in 12th grade (49.6% male). In respect of TIMSS (1995) participants from the U.S., the data covered 11,115 fourth graders (50.6% male), and 10,973 eighth graders (49.5% male). Item difficulty was estimated for each item by using BILOG-MG program, which, according to the authors, "constrains the item discrimination and the pseudo-guessing parameter of each item to be equal across groups …and constrains the ability parameter estimates for the reference group to have a mean equal to 0.0 and a standard deviation equal to 1.0" (p. 59). In the study, by considering the male sample as the reference group, the interaction between sex difference and item difficulty was assessed by correlating the item difficulty difference with item difficulty, and by computing on the combined male/female sample. The findings of this study, which used 8 data sets, indicated that in 5 data sets, there was significant interaction of sex difference and item difficulty. Out of the remaining, two were the smallest samples, and hence, according to the authors, might not have yielded reliable results. In conclusion, the study indicated the existence of interaction between sex difference and item difficulty in multiple-choice items on several mathematics tests to varying degrees. A significant finding of this study was that easy items tended to be easier for females than males, whereas difficult items tended to be harder for females than males.

Penner (2003) also examined the correlation between item difficulty and gender, using the data from TIMSS (1995) in respect of middle grade mathematics and science, pertaining to 10 countries (the United States, Canada, Australia, New Zealand, Lithuania, the Czech Republic, Sweden, Austria, South Africa, and Cyprus). The percentage of students who answered an item incorrectly was considered as item difficulty, which was calculated for boys and separately, and also overall for all of the students -- in respect of each country separately, and for all the ten countries put together. Then, the interaction between item difficulty and gender was computed by means of regression of the logit of gender-specific item difficulty, weighted by the variance of gender-specific item difficulty. In consonance with the results of Bielinski and Davison (2001), the findings of this study also indicated that in general, as the level of item difficulty increased, so also did the magnitude of boys' better performance over the girls.

In yet another study on the same lines, Innabi and Dodeen (2006) studied TIMSS (1999) data in respect of eighth grade mathematics in Jordan, to analyze gender-related differential item functioning (DIF). The study involved 5299 students and 124 multiple-choice test items and adopted the Mantel-Haenszel (MH) DIF procedure, which is based on the principle that 'differential item functioning occurs when two groups with the same ability level do not have the same probability of getting an item correct'. The indicator of DIF, MH-Alpha was then calculated as the quotient of the ratios of the average odds of the right response of the two groups. The findings indicated that although there was no significant difference between boys and girls on the total score, most of the DIF items in algebra and data analysis were in favor of females, whereas the items on measurement were favor of males. As regards the DIF items in number sense and geometry, there was no pattern. However, most of the DIF items that favored males involved judgment or expectation, whereas, DIF items involving fixed answers and familiar items found favor with females.

Do gender differences matter in the acquisition of spatial-mechanical skills? Casey, Nuttall, and Pezaris (2001) probed into this aspect by conducting a study that involved 187 students of the eighth grade belonging to middle income group (91 boys and 96 girls) in a suburban middle school in a Northeastern State of the United States. The purpose of the study was to compare the spatial-mechanical skills with self-confidence as a mediator of gender differences in mathematics. The study used gender-based items from TIMSS (1995), selecting 15 items that showed the largest gender difference favoring males (TIMSS Male-subtest) and 15 items that showed the largest gender difference favoring females (TIMSS Female-subtest). These items were administered to a selected cohort of 187 students. For measuring the spatial-mechanical skills, the study employed Vandenberg Mental Rotation Test (Vandenberg & Kuse, 1978); the Mechanical Reasoning subtest of the Differential Aptitude Tests (DAT; Bennett, Seashore, and Wesman (1990); and a Piagetian test of horizontality, the Water Levels Test (WLT, based on Sholl, 1989). To measure self-confidence, the study used six items on mathematics self-confidence and six items on English self-confidence. The items were taken from a questionnaire, cited by Parsons et al. (1980). The findings indicated that there was significant correlation with the TIMSS male-subtest score, whereas there was no significant correlation with TIMSS female-subtest score. The findings also showed significant correlation between mathematics and English self-confidence and both the TIMSS subsets. However, mathematics self-confidence and English self-confidence did not show any mutual correlation. Again, spatial-mechanical skills had significant correlation with mathematics self-confidence, but not with English self-confidence. The implication of this study, according to the authors, is that when boys perform well in solving mathematics problems, they are more likely to employ their visualization skills than girls who perform well. As such, the authors called upon teachers to include greater use of spatial strategies in mathematics curriculum.

The impact of gender differences on the learning situations was probed by Meelissen and Luyten (2008). Their study examined the gender gap in mathematics, by analyzing the influence of differences in learning contexts between schools and classes on mathematics achievement and attitudes of Dutch grade 4 students. The study used the mathematics data from TIMSS (2003) pertaining to grade 4 students of the Netherlands. Using concentric circles to represent multilevel models and employing MLwiN software (Rasbash et al., 2000), the study treated student characteristics as the 1st model, school-related variables as the 2nd model, the contextual and cultural school characteristics as the 3rd model, and finally, the contextual and cultural teacher characteristics as the 4th model. Analysis of the results showed that a gender difference due to self-confidence in mathematics was a major issue. It was found that boys had higher levels of self-confidence than girls. But, ‘girls from higher socioeconomic background have more confidence in their mathematics abilities than do girls from lower SES background, and that SES appears to have little influence for boys’ (p. 91). The study also indicated that differences between schools had significant relationship with the stereotyped views of the students and student achievement. It was also found that, only in the case of girls, the feeling safety at school was an important factor for performance. In another similar study, Louis and Mistele (2011) investigated the connections between student’s achievement scores in mathematics, student gender, and self-efficacy. The researchers used TIMSS (2007) report and employed the statistical measures ANOVA and MANCOVA to examine whether gender impacted the achievement scores for Algebra, Geometry, Data, Number, and the overall achievement scores in mathematics. The results showed that males exhibited higher levels of self-efficacy than females in mathematics. However, females scored higher than males in algebra, while there was no significant difference in the scores in other mathematics subjects, Data, Number, Geometry, and the overall mathematics scores. These findings run counter to the results of the study of Neuschmidt (2008). However, the added dimension of self-efficacy’s impact on performance in this study has led to this finding.

Beliefs. The psychological basis for any achievement seems to lie in belief in oneself. House (2000a) investigated the relationship between self-beliefs and achievements of the students of Ireland, using TIMSS (1995) data in science, covering 5881 students of the middle grades. He employed a two-stage cluster design, using the schools and classrooms in the first stage, and then took student characteristics, instructional activities, out-of-school activities, family characteristics, learning resources, and mathematics and science achievement in the second stage. The study involved jackknife variance procedures that provided unbiased variance estimates for applying statistical tests. The findings showed that student self-beliefs were significantly related to science achievement. Those students who expressed that they enjoyed learning science or felt that science was important had better test scores, whereas the students who indicated that either good luck or natural talent were necessary for success in science at school tended to obtain lower test scores. House (2003) repeated this study to find the relationship between self-beliefs and achievements of the students of Hong Kong, using TIMSS data in mathematics, covering 6476 students of the middle grades. He adopted similar procedures with two-stage cluster design and obtained similar results that showed significant relationship between student beliefs and achievement. The students who indicated that they enjoyed learning mathematics and felt that mathematics was important in life tended to achieve higher scores than others. Also, those who felt that mathematics was boring corresponded to lower test scores. Carrying this type of investigation further, House (2005) examined the TIMSS (1999) mathematics data in respect of 4668 eighth grade students of Japan. Using again the same procedures of two-stage cluster design, he confirmed his earlier results that showed significant relationship between the beliefs and achievements of students. Those who enjoyed learning mathematics and felt mathematics was easy showed better achievement in tests, whereas those who attributed success in mathematics to external factors, such as good luck, showed lower achievement.

Again, in an extended replication of these studies, House (2006a) examined the TIMSS (1999) data using 4615 students of Japan. However, this time, along with student beliefs, he included instructional activities also as independent variables and considered achievement in algebra as the dependent variable. Retaining, as in earlier cases, the two-stage cluster design and jackknife variance estimation procedures, he obtained results that showed that even after considering the effects of instructional practices, student mathematics beliefs were significantly related to algebra achievement. Like in his earlier studies, he found that the students who enjoyed learning mathematics showed higher scores, whereas the students with negative attitudes scored less. Additionally, the study showed that frequent homework assignment, beginning homework in class and discussing homework in the class, as well as trying to solve mathematics problems with examples from everyday life related to higher algebra scores. These results were in agreement with the results of the study of House (2002b), involving students of Chinese Taipei. In a further replication of his studies to find the relationship between student mathematics beliefs and math achievement, House (2009) examined the U.S. mathematics data of TIMSS (2003) in respect of all the participating 130 eighth grade Native American students. Applying the stratified two-stage cluster design and jackknife variance estimation procedures as in his earlier studies, he obtained similar results that showed significant relationship between student beliefs and student achievement. Students who indicated enjoying learning mathematics and felt that they do well in mathematics tended to show better results, whereas those students who had negative beliefs and lacked in self-efficacy about their mathematics learning abilities showed lesser achievement.

Summary. There is ample literature on the relation between gender and student performance. In some of the studies, topics and item types were examined from TIMSS data, to relate them to gender. Neuschmidt (2008) reported from his study that boys were better in algebra, whereas girls were better in measurement, with geometry being an area in which boys and girls performed equally. But, among the students with high self-efficacy, girls scored higher than boys in algebra, according to the study of Louis and Mistele (2011). Regarding item difficulty levels, girls tended to experience more difficulty in answering the difficult items and those that involved visualization skills as compared to boys (Bielinski & Davison, 2001; Penner, 2003; Innabi & Dodeen, 2006). Regarding self-confidence in mathematics, boys generally appeared to be slightly better than girls. Among the girls, those from higher SES backgrounds were found better than those with lower SES background.

Students' beliefs appear to have a significant effect on their achievement. Several studies pointed out that those who enjoyed learning mathematics and felt mathematics was easy tended to get higher scores, whereas those who attributed success in mathematics to good luck or natural talent got poor scores (House, 2003a; House, 2005; House, 2006a; House, 2009). Similar result was found for Science, according to House (2000a). Some studies used data from TIMSS to find out the influence of institution-related, home-related, and teacher-related factors on student achievement. School resources, especially the availability of effective teachers, and parents' educational background, mother's perceptions about the importance of mathematics learning were found to be important factors on student achievement. Interestingly, although several studies pointed out student's attitude also as an important factor, Schreiber (2002) found that it was not a significant factor where teacher's influence was strong enough to offset the attitudinal deficiencies of the student.

Institution and Home-Related Factors and Student Achievement

Among the important variables that influence student achievement are those that are connected to institution and home. There are several studies that dealt with this issue. Some representative samples of such studies are discussed below.

Institution-related factors. The influence of institutional factors on student achievement in mathematics was examined by Schreiber (2002). The study involved the data related to 162 schools and 1,839 students of the high school senior students of the United States from TIMSS (1995) and adopted a multilevel (hierarchical) linear regression model. The findings indicated that student related factors did not have any influence on achievement, with the exception of parent education and attitudes. The study revealed that the students, whose parents had higher levels of formal education, scored higher. However, institutional factors, especially, school resources and the availability of teachers with relevant skills to teach advanced topics had significant relation with the mean school achievement. Another interesting finding in this study was that student attitude was not significantly related to achievement in some schools where teacher factors were strong, and according to the author, seemingly offset deficiencies in student attitudes. While considering the influence of institutional factors on student achievement, it is apt that the following questions arise in the mind. Do schools function as professional communities, and to what extent? And, how far this issue is an indicator of student achievement? Probing about this interesting phenomenon of the professional functioning of schools was the focus of the study of Lomos, Hofman, and Bosker (2011). They analyzed the mathematics data from TIMSS (2003) in respect of the Netherlands to explore the extent to which Dutch schools were functioning like professional communities. The study further aimed at finding the relationship between the aspect of functioning like professional communities and student achievement in Dutch schools. A three-level hierarchical linear modeling (HLM) method was used, taking student-teacher, teacher-classroom, and classroom-school as the three nesting levels and using the HLM statistical package. The sample comprised 2706 students, 117 schools and 117 teachers and departments. Items pertaining to the professional community concept, such as interaction with other teachers and shared values, were chosen from the TIMSS data set. The results indicated that the departments acted like professional communities as compared with a few decades ago, and that schools with high professional community scores had the best scores of student achievement. In yet another study to probe the a impact of a complexity of influences on student achievement, Patterson, Perry, and Decker (2003) studied the relation between the independent variables of gender, parents' opinions, and amount of time spent on studying and the dependent variables student achievement in mathematics literacy and advanced mathematics performance. The researchers used the data from TIMSS (1995) concerning the final year high school students of 211 United States schools. The study involved computation of statistical measures like Chi-square tests, analysis of variance (ANOVA), multivariate regression, and multivariate analysis of covariance (MANCOVA). The findings indicated that for both male and female students, math attitude was higher in cases where the parents thought that doing well in mathematics was important. The study also indicated that math attitude and time spent on studying mathematics were related to achievement in mathematics literacy and performance. However, the results also showed that males had higher mean math attitude than girls.

In a connected study, Zuzovsky (2008) explored the data concerning Israel from TIMSS (2003) to find out the factors that accentuate the achievement gap between two student sub-populations of a country. The study examined the 108 Hebrew-speaking schools with a student population of 3163, and 38 Arabic-speaking schools of Israel, having1092 students. The data on teachers were obtained from the TIMSS teacher questionnaires, involving 315 mathematics teachers and 284 science teachers, whereas the data on schools were obtained from the participating Principals; 97 Hebrew-speaking, and 36 Arabic-speaking. The study adopted the pair-wise deletion method of Hierarchical Linear Modeling procedure. The analysis considered the variables at four levels: student level, school level, subject level, and teacher level. The student level variables were: aspiration to finish university education and having one of the parents with academic qualifications, having high-level confidence, and the number of books at home. The school level variables were: ethnic affiliation of the school, and the class aggregate of academic aspiration and of the academic home background of the students. The subject variables in mathematics were: practicing four basic operations, and relating what is learnt to real life situations, whereas in science, listening to the teacher was the variable. The teacher level variables were the percentage of students who participated in teacher's professional development activities, and the teacher's involvement in such activities. The analyses showed that percentage of students aspiring to study at university and having at least one parent who had academic qualifications was nearly twice in Hebrew-speaking student population than in their Arabic-speaking counterparts. According to the author, this factor pointed towards better achievement in Hebrew-speaking schools in science. On the other hand, the study found that the Arabic-speaking schools' performance in mathematics was better, even though traditional methods of teaching were practiced in both Arabic-speaking and Hebrew-speaking schools. This was explained by the author thus: 'in mathematics, some traditional instructional modes, such as drill and practice, arithmetic without calculators, reviewing homework, relating what is learnt to daily life, and listening to teacher lectures, occur more frequently and interact negatively with the ethnic affiliation of schools, thus favoring students in Arabic-speaking classes' (p.116).

Howie (2003) examined the impact of schooling conditions on mathematics achievement by analyzing the eighth grade mathematics data of TIMSS-R (1998) in respect of South Africa. The study, involving 200 schools and 8,146 students, used not only the achievement tests in mathematics and science and the national option test for English language proficiency, but also the principal's questionnaire on enrollment, demographics, subject selection, as well as administrative, curricular, budgetary and social issues. The method of analysis involved developing constructs with frequencies in all possible school, class, and student level factors and forming a correlational matrix.

Analysis of principal components and reliability analysis were done on sets of items that referred to one factor or construct. The findings indicated that the factors that influenced mathematics achievement were: the location of the school, the home language of the student, and the teachers' unions.

Wang and O'Dwyer (2011) examined the TIMSS reports of 2003 and 2007 to find connections between teacher directed student use of technology and mathematics achievement at the eighth grade. Their analysis, which involved the use of Hierarchical Linear Modeling method with student, teacher, and school as covariates revealed that technology use accounted only for a very small percent of the total variance in mathematics test scores. Expanding the scope of topic, House and Telese (2012) analyzed the TIMSS (2007) report on eighth grade mathematics achievement using multiple regression method to examine the effects of lesson activities and computer use on algebra achievement. However, their study was confined only to the students of the United States and Japan. Data involved students' reports on the use of the computer at home, school, and elsewhere (public library, friend's home, and internet café). Further, the reports indicated the frequency of their use of computer and also the amount of time they spent outside of school playing computer games or using internet. The study revealed that high levels of algebra achievement corresponded with higher use of computers at home and at school in both countries. On the other hand, the more the students of either country spent time outside of school playing computer games, the lower were their algebra test scores. In addition, interestingly, the more frequently the students used computers for schoolwork in mathematics, the lower they scored in algebra tests.

Another study involving the students of the U.S. and Japan and using the TIMSS (2007) report was also completed by Yashino (2012),which investigated the relationship between math self-concept of the students and their math achievement. The investigation considered the potential role of culture by comparing math self-concept of the students with mothers' and fathers' education levels and the quantity of books held in the students' household. Data on self-concepts of students was provided in TIMSS (2007, p. 300) report through students' ratings on "(1) I enjoy learning math; (2) Math is boring; (3) I like math; (4) I usually do well in math; (5) I would like to take more math classes; (6) Math is more difficult for me; (7) Math is not one of my strengths; and (8) I learn things quickly in math". The method of analysis involved *t*-test, the analysis of variance (ANOVA), and multiple regression models. The findings of the study showed that there was a positive association between students' math self-concept and their math achievement in the samples of both the countries. However, Japanese students' self-concept was found to be significantly lower than that of the American students at the same level of math achievement. Additionally, the independent factors of mothers' education and fathers' education were positively associated with the students' math achievement.

In a similar study, Choi, Choi, and McAninch (2012) examined the background questionnaires of the TIMSS (2007) report to investigate the psychological conditions of high achievers in mathematics at the eighth grade level. Three psychological indices, viz., Positive Affect toward Mathematics (PATM), Students' Valuing Mathematics (SVM), and Self-Confidence in Learning Mathematics (SCM), drawn from TIMSS background questionnaires were used as data for comparison using *t*-test, *z*-test, and Chi-Square tests for analysis. Ten countries were chosen for comparison and placed in groups A and B, the former with higher percent of students who reached advanced level and the latter with lesser percent of such student. Group A consisted of the five countries, Chinese Taipei, Republic of Korea, Singapore, Hong Kong SAR, and Japan, whereas, Group B comprised Hungary, England, Russian Federation, Lithuania, and the United States. The two research questions were "(1) whether high achieving students (the top 5 %) in the TIMSS 2007 mathematics assessment possess three psychological conditions in comparison to their peers (the remaining 95 %); and (2) how high achieving students exemplify trends in the three elements of psychological conditions differently by comparison of (a) Group A and Group B and (b) Korea versus the other nine countries." (p. 197)

The study showed that high achieving students possessed significantly more of all the three psychological conditions (PATM, SVM, and SCM) in comparison to their peers. Whereas there were no significant proportional differences between Korea and Group A, Korea and Group B, and Korea and each of the nine countries for the PATM index, there was a significant difference in group-level country-level comparisons in respect of the other two conditions. Interestingly, the results about the index SVM indicated that high achieving students in Group B were more likely to consider mathematics and its learning more valuable for their everyday living and future success than their peers in Group A. As regards the index SCM also, another apparent anomaly was noticed in the findings inasmuch as the high achieving Korean students were found to be less likely to show confidence in their mathematical ability than the others in Group A and all in Group B.

The impact of class size on student achievement continues to be the topic of an ongoing debate. The effect of class size on student achievement was the theme in the study of Pong and Pallas (2001), who explored the eighth grade math data from TIMSS (1995), pertaining to Australia, Canada, France, Germany, Hong Kong, Korea, Iceland, Singapore, and the United States to find connections between class size and achievement. Taking gender and family background as control variables, the study adopted hierarchical linear modeling method to estimate the student-level and classroom-level effects simultaneously. The findings showed that Korea had the highest average class size of 50, followed by Hong Kong and Singapore, each having average number of students between 35 and 40. Iceland had the lowest class size with less than 20 students on the average; whereas all the rest of the countries had average class size 25 or 26. Findings also indicated that the total enrollment in a school was strongly associated with class size. Larger classes were situated in the urban areas in the United States, Germany, Iceland, Singapore, and Korea. In general, the class size had positive correlation with average class socio economic status and negative correlation with low student achievement. However, in contrast, in the centralized Asian systems of education where the average class size was higher, student achievement was also higher. The study revealed that after controlling for school and classroom factors, except in the case of the United States, the class size was not linearly associated with math achievement. Singularly, in the U.S., after controlling for the effects of teacher, school, and classroom variables, there was negative association between class size and math achievement. In a similar study, Wobmann and West (2006) also explored the effects of class-size on school systems, using data from TIMSS (1993) in respect of 11 countries. The method adopted was the least-squares regression analysis of test scores on class size, by controlling for a set of family-background characteristics. The results suggested that the effect of class size depends upon the school system.

Through an analytical study, Akiba, LeTendre, and Baker (2002) examined the issue of the impact of student victimization and school violence on student achievement, using data from TIMSS (1995). The student questionnaire that referred to self and peer victimization was taken as a measure. The analysis employed the multivariate model method. The study used TIMSS data to find how prevalent school victimization was across the nations and explored the national-level correlates of school violence from the victimization reports of the U.N., and INTERPOL. Then, the effect of school-level variables on achievement was assessed, controlling for those national-level variables. The findings indicated that school violence was prevalent in all the 37 countries that participated in TIMSS. The rates of school of violence were related to school-system variables and to some social indicators such as deprivation and age distribution, but not to income inequality, social integration, or general crime rates in these countries. Also, countries with greater achievement differences between high-achieving and low-achieving students tended to record more violence in their schools.

Home-related factors. Dealing with the domain of attitudinal and motivational aspects, Hammouri (2004) probed the relation between student-related attitudinal and motivational variables to achievement in mathematics. His study explored the TIMSS (1999) data pertaining to 3736 eighth grade students of Jordan, using a structural equation modeling procedure. In this study, the independent variables were: the perceptions of student's mother, friends, and self about the importance of mathematics learning and the student's attribution of success—hard work, or luck. The findings showed that the independent variables with significant positive relation to achievement were: mother's perception of the importance of mathematics learning, student's attitude and attribution of success to hard work, self-perception about the importance of mathematics learning, and friends' perception, in that order. It was also found that the student's attribution of success to luck and friends' perception of mathematics importance on mathematics achievement had a retrograde effect on student achievement.

In a study on similar lines, using the data in respect of middle grade mathematics from TIMSS (2007) for his doctoral dissertation study, Liou (2010) investigated the relationships between student math achievement and student motivational attitudes for learning mathematics, the pattern of the cultural factors, the relationships of student- and school-level factors associated with student mathematics achievement within the country. The researcher adopted factor analysis and HLM methods to arrive at the results, which indicated that self-efficacy was consistently positively related to math achievement. Also, parental involvement and school resources had significant positive relation to student achievement.

Post and Pong (2000) investigated the effect of student employment during middle school years on achievement. The study used the data from National Educational

Longitudinal Study (NELS, 1988) in respect of about 26,000 eighth grade students of the United States, and TIMSS (1995) data in respect of the U.S. and 22 other nations. With a view to find the impact of paid employment, the authors adopted a method of descriptive and exploratory analysis, regressions on the eighth grade on student achievement in mathematics and science were made after controlling for prior achievement, to find the impact of paid employment. The findings indicated significant negative effects of adolescent employment on math and science achievement for boys as well as girls, although they were more in the case of boys.

Summary. Student achievement is associated with the educational qualifications of parents, school resources, and the availability of qualified teachers, according to some studies. (Zuzovsky, 2008; Schreiber, 2002). The study of Yashino (2012) also revealed that mothers' education and fathers' education were positively associated with the students' math achievement. Liou (2012) also found that parental involvement and school resources were significantly related to student achievement. Moreover, according to Lomos, Hofman, and Bosker (2011), the schools that acted like professional communities produced better student scores. Some studies revealed that class size in itself was not a significant factor for student achievement. However, it mattered in conjunction with the school system. In the U.S., it was found to be inversely related to student achievement (Akiba, LeTendre, & Baker, 2002). Student adolescent employment was found to have a negative effect on achievement (Post & Pong, 2000). Apart from students' own perceptions about success, such as attribution of success to hard work, friends' perceptions also are associated with achievement, as per the findings in some studies

(Hammouri, 2004). The study of Choi, Choi, and McAninch (2012) revealed that high achieving students possessed significantly more positive attitude and self-confidence than their peers. In the same way, the dissertation study of Liou (2010) also revealed that self-efficacy was consistently positively related to math achievement.

Comment. The above studies considered several school situations that influence student achievement, including availability of qualified teachers, class size, functioning like professional bodies etc. However, in the available literature, the impact of the availability of resources for mathematics instruction on student achievement has not been probed. The present study considers this aspect in respect of the research question 2.

Instruction and Teacher-Related Factors and Student Achievement

The following section deals with the studies that are based on finding the impact of factors belonging to the most important aspect of instruction and teacher background.

Instruction-related factors. In order to probe the effect of spending additional time at school, Luyten (2006) adopted a multilevel modeling approach in study, drawing data from TIMSS (2003) and covering 45 countries. He took grade level and date of birth as independent variables and student achievement in mathematics and science as dependent variables to assess the effect of one extra year of schooling on student achievement. The findings showed that the effect of an extra year of schooling was lower in cases where there was a high level of achievement and stronger in schools that had modest levels of achievement in the lower grade. However, the effect of age varied from country to country, with England showing the highest effect on student achievement and Singapore, the lowest.

House and Telese (2011) investigated the relationship between computer activities and classroom lesson strategies on the one hand and motivation for mathematics learning on the other, using the TIMSS (2007) data for the population of the 8th grade students of the United States and Korea. They adopted variance estimation procedures for sampling designs to identify the specific computer activities and classroom lesson strategies that were significantly related to motivation for learning mathematics at the eighth grade level. Students' computer usage for schoolwork and using the internet outside of school were found to be the computer activities that were significantly related to motivation for learning mathematics. House (2002a) also examined the use of computer in the classrooms of Japan. He used the TIMSS Videotape classroom study material and examined 50 transcripts in his qualitative study and inferred that in all the cases several of Gagne's nine events of instruction (gaining attention, describing the goal, stimulating recall of prior knowledge, presenting the material to be learned, providing guidance for learning, eliciting practice of performance, providing informative feedback, assessing performance test, and enhancing retention and transfer) were incorporated in the lessons. The impact of the use of calculators on student achievement was the subject of the study of Tarr, Mittag, and Uekawa (2000), who used the TIMSS data in respect of the United States, Japan, and Portugal. The study, using HLM model, found that whereas there was rampant use of the calculator in the United States and Portugal, in Japan the use of calculator was very scarce. Moreover, interestingly, in Japan, the use of calculators was negatively correlated to student performance, unlike in the other two countries.

In order to find the correlates of mathematics achievement from an integral point of view for her dissertation study, Phan (2008) explored the TIMSS (2003) data in respect of the eighth grade mathematics of students from Canada, Egypt, South Africa, and the United States. She adopted a two-level hierarchical linear model, using five models incorporating: student background, home resources, instructional practices, teacher background, school background related factors, and a model, which was built by including all the statistically significant predictors in the earlier four models to predict math achievement. Findings from this study indicated that instructional practices model suited best for the United States. This model involved opportunity to learn math topics, activities in math lessons, amount of homework assignment, and average math instructional hours per year. However, for Egypt the teacher background model worked best, whereas for the other two countries, the last model that included all the significant factors of the first four models proved to be the most effective one.

Teacher-related factors. The beliefs of the providers of instruction, i.e., the teachers, also appear to hold sway on student learning. In this direction, Kupari (2003) conducted a study to probe teachers' practices and beliefs, using the data concerning Finland in TIMSS (1999). The author considered lesson structure, teacher and student interaction, and homework as the concepts of teachers' practices and analyzed teachers' beliefs from their opinions on the nature of mathematics learning and their responses to questions on what they thought about the practices and skills, in order for students to be good in mathematics. The study adopted a 'hierarchical cluster analysis' method of distribution of the instruction variables and the belief items. This method involved merging two closest clusters into a new cluster. The clustering would proceed hierarchically and go on until there was only one cluster left. Results revealed that, in a vast majority of the cases, teachers' practices involved students working individually with the assistance of the teacher. Whole-class instruction that included exercises was also found common. Regular homework was also found to be a feature of the teachers' practice. Regarding beliefs, the study found that most of the teachers believed mathematics to be a 'practical and structured guide for addressing real situations' (p. 251). Very few believed that mathematics should be learned as a set of rules and algorithmic procedures. Also, a large majority of the teachers believed that representation should be used in teaching the topics in mathematics and that teachers' liking and understanding of students was essential for teaching mathematics. The study also found that on one side, those teachers who had constructive beliefs used several representational forms while trying to understand the students, whereas on the other side, some teachers who believed that mathematics was an abstract subject employed procedural methods, using rules and algorithms. The imperatives of this study, according to the author, were that in the preservice and in-service programs of professional development of teachers should take into account the beliefs as well as the practices of the teachers.

The study of House and Telese (2011) revealed that the classroom lesson strategies that were significantly related to motivation to learning mathematics at the eighth grade were: writing of equations and functions to represent relationships, relating what is learnt in mathematics to daily lives, deciding on own procedures for solving complex problems, listening to the teacher in a lecture-style presentation, and working problems on their own. Again, student achievement was found to be better in classes where the teachers explained rules and definitions, solved examples related to new topics, gave frequent homework, or solved problems related to real life experiences, as per the findings of the study of House (2001). The study used the data related to about 10,000 students of eighth grade of Japan from TIMSS (1995) for a study to examine the effect of instructional activities on student achievement in mathematics and science. This study considered three specific instructional activities as independent variables– teaching activities for new topics, homework activities, and classroom instructional activities to find their effect on student achievement. The method of analysis adopted a cluster design, with schools in the first stage of sampling and classrooms in the second stage. The findings also indicated that very frequent collaborative student learning activities, as well as frequent student checking of each other's homework during mathematics lessons was related to lower student achievement. Again, House (2002b) repeated this study using the data related to the students of eighth grade of Chinese Taipei from TIMSS (1999). The results were in consonance with the earlier results of House (2001). In addition, it was found that there was significant negative correlation between computer use and student achievement, whereas student working on mathematics projects and using things from everyday life for solving mathematics problems were positively related to achievement. Further again, House (2004) repeated this study with data from TIMSS (1999) in respect of students from Japan. However, in this study, he considered only two independent variables; homework activities and teaching strategies for new topics in mathematics. The results were similar to those in the earlier two studies for the given variables.

In another interesting exploratory study to find the connections between teacher quality and student achievement and cognitive skills, Xin, Xu, and Tatsuoka (2004) used the mathematics data in respect of students of eighth grade students of the Netherlands, the United States, Korea, and Japan from TIMSS (1999). The study employed the 'Rule-Space methodology', to identify each student's knowledge and skill attributes that explain the performance on the test items. The findings indicated no "positive correlations between teacher quality effect as measured by teacher's degree level, major field, certification status and experience and student achievement" (p.218). Is the class size a factor on the impact of the teacher on student achievement? According to Wobmann and West (2006), who explored the effects of class-size, only if the average capability of the teachers is low, smaller classes have an advantage in terms of student achievement.

Summary. Computer activities were found to be significantly related to motivation for learning mathematics according as per the findings of a study by House and Telese (2011). But, as House and Telese (2012) found that high levels of algebra achievement corresponded with higher use of computers at home and at school in the U.S. and Japan. Further, the more the students spent time outside of school playing computer games or used computers for schoolwork in mathematics, the lower were their algebra test scores. However, the actual use of computers was found to have a negative impact on student achievement, although student working on mathematics projects and using things from everyday life for solving mathematics problems were positively related to achievement (House, 2002b). Technology use, according to the study of Wang and

O'Dwyer (2011), accounted only for a very small percent of the total variance in mathematics test scores. In the U.S. and Portugal, the use of calculators was found to be rampant, whereas in Japan, the use of calculator was very scarce. And, interestingly, in Japan, the use of calculators was negatively correlated to student performance, unlike in the other two countries, according to a study by Tarr, Mittag, and Uekawa (2000).

Teachers' instructional methods appeared to depend on their beliefs. Those who had constructive beliefs tried more to understand the students and adopted representational forms, whereas those who thought mathematics was an abstract subject employed procedural methods, using rules and algorithms (Kupari, 2003). Further, the studies by House and Telese (2011) revealed that the classroom lesson strategies representations, linkages with daily lives, helping students to develop their own procedures for solving complex problems were positively associated with student achievement. Also, when teachers explained rules and definitions, solved examples related to new topics, gave frequent homework, student achievement was better (House (2001). However, teacher's degree level, major field, certification status and experience and student achievement were found to have no positive correlation with student achievement, according to the study of Xin, Xu, and Tatsuoka (2004). Again, the impact of class size on student achievement was found to be significant only when the teacher's capability was low (Wobmann & West, 2006).

Comment. There are several studies concerned with instructional factors such as concerning student's use of computers, calculators, working on projects etc. Also, some studies of correlations between student achievement and several teacher attributes, such as beliefs, qualifications, and methodological issues were also discussed. However, it appears imperative that if an analytical study of the TIMSS reports is expected to be helpful in finding clues for improvement in the quality of educational processes in general, then it may be necessary for every teacher to have deep insights into the process of assessment itself, so as to modify her or his own teaching approach for a better effectiveness. With this view, the present study considers the impact of variable conditions of teacher professional development in the area of assessment on student achievement.

Factors Related to Achievement in Algebra and Geometry

Student achievement in specific subject areas was probed for a close view by some researchers. The learning of mathematics ought to connect with the life experiences of the students, as recommended by the Diversity in Mathematics Education Center (2007). Algebra, especially, is the area in which students learn the symbolism which is essential for developing technological knowledge. In the area of K-12 mathematics, it is desirable to have the initial focus in research and development on algebra, because algebra is the gateway to most areas of mathematics, science, and engineering, as per RAND Mathematics Study Panel (2003). The Algebra Project for middle grade students, incorporating activities that are supported by culture and inquiry, was an effort in this direction. However, it is important to note that in the development of algebraic thought, technology is an inadequate tool without teacher's effective role, according to Kieran (2007), who after examining the TIMSS (2003) report for eighth grade mathematics, found that in algebra, the average score of the 48 participating countries was 25 out of a possible 53, whereas that of the U.S. was 29. Interestingly, Presmeg (2007) made strong connections between technology and culture, with her observation that "technology, by entering all avenues of life, influences not only the mathematics of the curriculum and the ontology of mathematics itself, but also the culture of the future, through its children" (p. 451).

Strauss (2003) made an in depth analysis of the eighth grade mathematics data of TIMSS (1999) in respect of Slovenia, to investigate whether there were any topics in algebra in which the scores were not as high as the overall achievement scores. The purpose of the study was to locate areas in Slovene curriculum that required strengthening, by comparing the Slovene student achievement in algebra with 14 other European countries, both overall and at different ability levels. The framework for the study was centered on finding the gaps between the intended, the implemented, and the attained curricular levels. The study adopted a methodology of pooled analysis of the IRT algebra scores of European countries in the database of TIMSS (1999) at four ability levels—the 25th, the 50th, the 75th, and the 90th percentile. Slovene students, whose scores differed from the European benchmarks by not more than 15 points in each of these levels, were grouped in the corresponding groups. Then, for all the items, the percentages of correct answers were computed and compared across the ability groups in European countries. Jackknife procedure was used for computing standard errors to identify significant differences between Slovenia and the other European countries. The results pointed out that on the average and at 25th and 50th percentile ability levels, Slovene eighth grade student achievement was significantly better than 7 or 8 European countries, and no European country was significantly better than Slovenia. At the 75th and 90th percentile levels, only two countries were better than Slovenia. The study identified 18 items at which Slovene students at the higher ability levels of 75th and 90th percentiles performed significantly lower than European average. However, at the lower ability levels of 25th and 50th percentiles, Slovene achievement was not significantly different from the European average.

In a similar study, Battista (1999) analyzed the geometry scores in respect of both elementary and middle grades from the data of TIMSS (1999). The item-wise analysis showed that the U.S. students in the grades 3 and 4 scored less than the international average only in the items that involved estimation problems that involved metric units. At the middle grades level, the U.S. students scored higher than or equal to the international average in all items, except those that involved visualization, solving geometry problems, and three-dimensional geometry. On the basis of these findings, the author suggested that teachers should be made aware of the research findings about geometric reasoning and helped with tasks that would enable them to understand their students' geometric reasoning. A question arises as to whether attitudes matter in the learning of geometry. House (2006b) touched upon the psychological undercurrents in the learning of geometry, by investigating the relationship between self-beliefs and mathematics achievement, as reflected by the scores in geometry in the data from TIMSS (1999), in respect of the middle grade students in Japan. The method included finding the correlation coefficients between each mathematics belief variable and geometry test scores. Further, to assess the relative contribution of each mathematics belief variable toward the explanation of geometry achievement test scores, multiple regression procedures were employed. These analyses were made separately for male and female students and also for the whole sample. The findings, which were similar for boys and girls, indicated that positive attitudes toward mathematics corresponded with higher geometry achievement. Students who indicated that they enjoyed learning mathematics, or felt that mathematics was an easy subject, tended to score higher. Those students who attributed success in mathematics to memorizing, natural talent, or good luck showed lower scores.

Summary. Student achievement in specific subject areas was probed by very few researchers. Kieran (2007) compared student performance of 48 countries in algebra, using the TIMSS (2003) report for eighth grade mathematics. The results showed that whereas the average score of the all participating countries was 25 out of a possible 53, the score of U.S. was 29. Using the data of TIMSS (1999) in respect of Slovenia and 14 other European countries, Strauss (2003) investigated whether there were any topics in algebra in which the scores were not as high as the overall achievement scores, in order to locate areas in Slovene curriculum that required strengthening.

A study in geometry by Battista (1999) analyzed the geometry scores in respect of both elementary and middle grades from the data of TIMSS (1999). Findings indicated that the U.S. students in the grades 3 and 4 scored less than the international average only in the items that involved estimation problems that involved metric units. At the middle grades level, the U.S. students scored higher than or equal to the international average in all items, except those that involved visualization, solving geometry problems, and three-dimensional geometry. Using the scores in geometry concerning geometry in the data from TIMS (1999), House investigated the relationship between self-beliefs and mathematics achievement among students of Japan. The results showed that the students who indicated that they enjoyed learning mathematics, or felt that mathematics was an easy subject, tended to score higher. Those students who attributed success in mathematics to memorizing, natural talent, or good luck showed lower scores.

Student Achievement in the United States

Schmidt, Wang, and McKnight (2005) studied the coherence and rigor in the mathematics standards of the highest-achieving countries in the Third International Mathematics and Science Study (TIMSS). They tallied them with U.S. national standards in science and mathematics and also with the standards developed by some states and school districts in the U.S. The study found that the organizing principle of the top-performing countries was different in quality as well as degree from the national, state, and district standards in the U.S. According to the authors, coherence, which is an important factor in the framing high-quality standards, is reflected when topics and their sequencing are clearly spelt out at each grade and across the grades, while specifying the depth in which they are to be studied. The study revealed that whereas in the highest-achieving countries this structure was evident, ‘US structure [was] so diffuse and seemingly arbitrary’ (p. 555). The researchers made a pointed suggestion to the policy makers in the US that curriculum standards ought to have a national scope, in order that coherence and rigor might be ensured.

Again, within the United States, student achievement in mathematics was found to be significantly at variance across the districts due to variation in content coverage. This suggested that the *opportunity to learn* (OTL) was not provided equally in these districts. Using TIMSS cross-national curriculum data, Schmidt et al. (2011) developed an ‘international grade placement index’ (IGP), which indicated the difficulty level of each topic, according to its instructional focus. This was further weighted, taking into account three aspects of OTL -- the mathematics content itself, instruction time for each topic, and rigor or content difficulty (as estimated from international curriculum data). They applied this nomenclature in their multilevel analysis of student achievement in mathematics based on TIMSS-R (2009) results. In their analysis, they employed a three level Hierarchical Linear Modeling (HLM) technique, with student model as level 1, teacher/classroom model as level 2, and school/district as model 3. They found that none of the 13 districts in the U.S., which were considered, met the international standards of course coverage in eighth grade. Based on these findings, they argued that even though TIMSS incorporated Carroll’s notion of OTL as a measurable quantity that is expressed in terms of the time allowed for learning, socioeconomic status (SES) and content coverage are factors that modify the effect of OTL on student achievement. Thus comprehensively, a wider concept of OTL is that which embeds considerations of socioeconomic status and content coverage, and student achievement is a consequence of this OTL. The researchers made a far-reaching observation that unequal OTL was inherent in the U.S. education system. Further, they remarked that the results pointed at not only at the threat of some children being left behind, but on a larger scale, indicated that the entire student population in some districts and states in the United States might be left behind due to the decisions made in those districts and states about the content.

Are the mathematics achievements of U.S. students related to teaching methods – reformed and traditional? This has been the subject of the doctoral dissertation study of Cheng (2012). Utilizing the data from TIMSS (2007) and adopting the hierarchical linear modeling method for analysis, the researcher attempted to find out whether the teaching methods had any relation to the problem-solving skills, reasoning, and overall performance of different ethnic groups among the eighth grade mathematics students of the United States. The finding of the research showed significant positive relation with all the three measures in the case of Caucasian, African-American, and Hispanic students, but not in the case of Asian-American students with respect to any of the three performance considered measures. Again, the traditional teaching method also had a positive significant relation in respect of all the three measures in the case of Caucasian and Hispanic students, but not related to African-American and Asian-American students. These results lead to a rather strange implication that the performance of Asian-Americans is impervious to the teaching method, inasmuch as neither of the methods had a relation to their performance. So also, another implication is that the performance of Caucasian and Hispanic students was also not influenced by the teaching method, since both methods had a significant positive relation to their performance. Lastly, the study showed that only African-American students appeared to gain in performance from reformed teaching methods. The researcher tried to clarify that these results were at variance from other studies, since this study used teachers' self-reports and had other inherent limitations such as “treatment effect of reformed or traditional type of teaching approaches cannot be identified in this study as there were no control or treatment groups; also, causal relationships between the two types of teaching and students’ mathematics achievement cannot be identified in this study” (p.117). Finally, not using the socio-economic status (SES) of the students could have been a lacuna, as the researcher explains. I feel that the study would have been more comprehensive if SES and other school and home related variables were considered.

In another study using the data from the Third International Mathematics and Science Study (TIMSS-R), Birenbaum, Tatsuoka, and Yamada (2004) analyzed the eighth grade mathematics performance of the students of Japan, the U.S., and Israel. The study involved 4,411 students from the U.S., 2,371 from Japan, and 1,684 Jews and 408 Arabs from Israel. The researchers employing a Rule-Space (RS) probabilistic model for cognitive diagnosis by classifying students’ responses on the basis of the strengths and weakness of the test items according to the attributes of procedure, skill, or content knowledge which is needed for completing the task. The results revealed that not only in mathematics knowledge, but also in process, and skill, the Japanese students were significantly better than the students of U.S. and Israel. In the matter of mathematical thinking, which has three facets -- logical thinking, inductive thinking, and also divergent thinking, the students from the U.S. and Israel were found to be deficient. The findings pointed out that in Japan, the curriculum was found coherent and challenging, unlike in the U.S. and Israel, where the number of topics covered was more, but lacking in depth. This observation is in consonance with the finding of Schmidt, Wang, and McKnight

(2005). The study also mentioned cultural contexts and different styles of instruction as a contributing factor. In this respect, they observed that ‘mathematics education in Japan is a composite of three elements, upbringing, regular classroom experience, and supplementary schools’ (p. 165). The authors also made a pertinent remark that the salaries of teachers in Japan were higher than those in the U.S. and Israel, implying that teacher motivation and involvement were higher in Japan. The study recommended the creation of communities of learning and adoption of the technique of *Lesson Study*, a Japanese method of school-level teacher preparation of instructional programs for improving student performance in the U.S. and Israel. This point was reinforced by the findings of Baker, and Jones (2005), who explored TIMSS and PISA data on student achievement in Science in New Zealand. According to them, since ‘standardized goals and prescriptive curricula do not result in higher achievement’ (p. 154), it is worthwhile redeveloping curricula and involve teachers in curricular goals. In this context, they quoted the cases of Japan and Germany, where such practices have been in vogue.

On the same lines, the studies of Schümer (1999), which examined the Japanese survey of Supplementary School system and also TIMSS, found that students spend more time in Japan on mathematics and have more supervised study opportunities than those in the U.S. or Germany. Aubrecht (1999) also, who used case studies after examining the performance of U.S. students from TIMSS data, inferred that U.S. students lacked depth of knowledge.

Cogan, Schmidt, and Wiley (2001) analyzed TIMSS data to find out the lacunae in the instructional practices of mathematics at the middle grades in the U.S. They probed into the instructional practices with structured questions about opportunities provided by the school to eighth-grade students, the various types of math courses offered, the mathematics studied by students, the text books that influenced the students, teachers' emphasis in their instruction, the type of most challenging mathematics studied by the students, and whether these challenges related to achievement. The study found that for about a third of the eighth-grade student population, there was no consonance among the mathematics course title and the textbook used in the course. Moreover, there were notable differences in learning opportunities on the basis of location, size, and minority enrollment. In this regard, the authors made a recommendation to consider students' classroom mathematics learning experiences, as a matter of policy, while framing curricular opportunities. The authors reflected that the inclusion of algebra and geometry in the curriculum for all eighth grade students might be a worthwhile goal, but efforts to provide for adequate learning opportunities ought to be ensured at the same time; lest disparities in student achievement should persist. In a study with a narrow focus on finding areas of difficulty, Thompson, and Preston, (2004) examined the analyzed the results of TIMSS and NAEP and found that the topic 'measurement' was an area of major weakness among U.S. students, particularly the minority students. The authors attributed this to the fact that the students needed to know two different measurement systems and it also depended upon the actual way in which they did measurement tasks at school. Rodriguez (2004) also probed the U.S. portion of TIMSS to investigate the deficiencies of U.S. students in mathematics, using an unconditional HLM model, that is, without using explanatory notes. He found that classes in which frequent and moderate levels of homework was given, student performance was better. On the basis of his researches, he suggested that teacher's assessment practice was related to student characteristics and performance.

Valverde, and Schmidt (2000) explored TIMSS curriculum data and examined the issue of improving student achievement in the U.S., with the following pointed questions (p. 653):

- What are the expectations held by the educational systems of countries performing among the best in the world? and
- How do expectations in the USA compare to the academic standards of this select group of countries?

Even as later echoed by Schmidt, Wang, and McKnight (2005), this study revealed that in the U.S., there was considerable lack of coherence and focus compared to the high-achieving countries in the TIMSS study. Whereas in the U.S., topics lingered longer in the curriculum, in the highest performing countries, unlike in the U.S., they had well defined patterns of cognitive expectations. However, the researchers hastened to caution against unreflective emulation of the highest performing countries in educational practice. Rather, they commented that benchmarking was the right approach to reconsider the educational practices in the U.S. in the light of the prevailing practices in the highest performing countries. One significant point was suggested by the researchers in this regard; that some topics need to be introduced in the curriculum at a later stage. The researchers called upon the policy makers in the U.S. to reconsider the most common practices and ponder over the convictions regarding standards and curricula. Discussing this issue in a more concerted way, Menon (2000) questioned whether the U.S. should emulate the Singapore system of education. In his short, but interesting discourse, the author contended that standardized tests laid emphasis on algorithmic responses, and therefore countries which preferred procedural knowledge excelled in TIMSS. Moreover, in the Singapore system of education, students are geared to take examinations. Private tutoring is the norm there, so much so that "some teachers have resigned their jobs to go full time into this lucrative tutoring profession!" (p. 346). In addition, the public ranking of schools attracted some parents to go all out to get their children admitted in high ranking schools, and even buy houses near them for facilitating their children's studies. Even though there are parents who utilize tutoring significantly and purchase homes depending on school quality in the U.S. also, the proportion of such cases cannot possibly be to the extent prevailing in a tiny country like Singapore. According to the author, the U.S. students do not memorize much, but are led to develop more of conceptual understanding. It is therefore no wonder that their performance in the TIMSS has not been very superlative. By citing a personal anecdote, the author revealed that even though the students of Singapore excelled in TIMSS and other major public examinations, they lacked the skills needed to succeed at the workplace. The U.S. students, on the other hand, had better problem solving skills and conceptual understanding of topics, even in comparison to those who did much better in TIMSS. Summing up his discussion, the author recommended a systemic change in education in the U.S. in terms of teacher preparation and culture of instruction, rather than attempting piecemeal solutions to improve the scores in the TIMSS.

Summary. Schmidt, Wang, and McKnight (2005) studied the coherence and rigor in the mathematics standards of the highest-achieving countries in TIMSS and the U.S. The study revealed that whereas in the highest-achieving countries the structure showed that the topics and their sequencing are clearly spelt out at each grade and across the grades and clearly specified the depth in which they are to be studied, in the U.S., it seemed arbitrary. Again, according to Schmidt et al. (2011) none of the 13 districts in the U.S., which were considered in their study, met the international standards of course coverage in eighth grade. In consonance with these findings, the study of Birenbaum, Tatsuoka, and Yamada (2004), which analyzed the eighth grade mathematics performance of the students of Japan, the U.S., and Israel pointed out that in Japan, the curriculum was found coherent and challenging, unlike in the U.S. and Israel, where the number of topics covered was more, but lacked in depth.

Some studies probed the instructional practices. The study of Cogan, Schmidt, and Wiley (2001) indicated that for about a third of the eighth-grade student population in the U.S., there was no consonance among the mathematics course title and the textbook used in the course. Moreover, there were notable differences in learning opportunities on the basis of location, size, and minority enrollment. However, the study of Rodriguez (2004) indicated that in the U.S., in classes where frequent and moderate levels of homework was given, student performance was better. In this context, it may be pertinent to note that students spend more time in Japan on mathematics and have more supervised study opportunities than those in the U.S. or Germany, as per the study of Schümer (1999). Reformed teaching methods were found to be significantly positively related to achievement, as compared to traditional teaching methods, especially in the case of African-American students, according to the dissertation study of Cheng (2010).

Should the U.S. emulate the systems of countries that performed much higher in the TIMSS? This question was discussed by Menon (2000) with particular reference to Singapore. However, according to the article, since the U.S. students do not memorize much, but are led to develop more of conceptual understanding, it is not a surprise that their performance in the TIMSS has not been very superlative. The author recommended a systemic change in education in the U.S. in terms of teacher preparation and culture of instruction, rather than attempting piecemeal solutions to improve the scores in the TIMSS.

In the present study, comparisons of student achievement in the U.S. and the three highest performing countries of TIMSS (2007) and two other countries in each of the content and cognitive domains will be made, as mentioned in the research question 3. In the light of the foregoing discussions, particularly the observations of Menon (2000), a matter of particular interest will be the comparison in the cognitive domains of Knowing, Applying, and Reasoning. The findings of this part of the present study, in conjunction with the study related to research question 2, which concerns the impact of some variables in student-related, school-related, instruction-related, teacher-related, and home-related backgrounds, might provide a few insights into the possible potential factors that could improve the effectiveness of the educational system in the U.S.

Chapter Summary

On a global scale, economic development has been found to be positively associated with the quality of education in mathematics (Moses, 2001; Zinth & Dounay, 2006; Akiba, LeTendre, & Scribner, 2007). With the initiative of the UNESCO, large scale assessment of mathematics achievement began in 1959. Later, three large-scale international studies were started with the coordination of OECD; TIMSS in 1995, PISA in 2000, and PIRLS in 2001. TIMSS, the large scale international program for assessing student achievement for 4th grade and 8th grade students in mathematics repeats every 4 years. In spite of some controversies concerning the performance of the United States in the TIMSS, the role of international comparative studies in helping policy makers in taking proper decisions is viewed positively (Kellaghan, 1996; Zuzovsky (2002).

The present study has Carroll's model of school learning as its basic framework. The five indicators of Carroll's model are aptitude, opportunity to learn, perseverance, quality of instruction, and ability to understand instruction. To correspond with them and also with a view to fill according to felt importance, the gaps in the available literature, variables from TIMSS (2007) are chosen. These variables are related to factors in student background (amount of time spent on math homework), classroom instruction (frequency of applying facts, concepts and procedures), home involvement (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), school resources (for math instruction), and teacher professional development (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy /instruction, content, and assessment).

Gender and beliefs are the main student-related factors in several studies that analyzed the TIMSS student achievement scores. The studies pointed out that in algebra, boys performed better than girls, whereas girls scored better in measurement. Also, as the difficulty level in the test items grew harder, or included more of spatial skills, or involved more of judgment or expectation, the average performance of boys tended to be better than that of the girls. (Neuschmidt, 2008; Gonzalez & Miles, 2001; Bielinski & Davison, 2001; Innabi & Dodeen, 2006; Nuttall & Pezaris (2001). But, girls with high self-efficacy tended to score higher than boys in algebra, according to the study of Louis and Mistele (2011).

Several studies involving student beliefs indicated that the students who indicated that they enjoyed learning mathematics and felt that they do well in mathematics tended to show better results, whereas those students who had negative beliefs and lacked in self-efficacy about their mathematics learning abilities showed lesser achievement. Similarly, the students who felt that mathematics was important in life tended to achieve higher scores than others, whereas those who felt that mathematics was boring corresponded to lower test scores. However, according to some studies, in general, boys had more of self-confidence than girls in solving mathematics problems; and among the girls, those who are from better socio-economic backgrounds, scored higher than those from lower socio-economic backgrounds (House, 2003, 2005, 2006a, 2009). It is observed that several important factors concerning student-related factors, other than gender and beliefs, are not dealt with in the available research studies. It may be noted in this context that the present study considers one important factor i.e. the amount of time spent on math homework from among those related to student background.

The association between student achievement and school-related variable factors, such as the availability of qualified teachers, functioning like professional bodies, and general school resources were probed and found to be positive, according to the studies of Zuzovsky (2008) and Schreiber (2002). The qualifications and the involvement of parents were positively related to student achievement according to the studies of Yashino (2012) and Liou (2012). Class size in the U.S., and also in other countries where teachers are of low capabilities had a negative association with student achievement (Akiba, LeTendre, & Baker, 2002). Adolescent employment was also another factor negatively associated with student achievement. However, there were no studies which probed into the connections between school resources for mathematics instruction and achievement in mathematics. The present study embeds this consideration.

Coming to instructional and teacher background concerns, several studies examined their impact on student achievement. Although computer activities were found to be significantly related to motivation in learning mathematics (House & Telese (2011), their use was found to exert negative influence on student achievement, unless used for working out projects (House, 2002b). Another study linking the use of calculator found that in Japan it is scarce, and interestingly, was negatively associated with student achievement there. In general, the use of technology was found to account for a very small percent of the total variance in mathematics achievement, according to the study of

Wang and O’Dwyer (2011). Student achievement was also found to be better when teachers explained rules and definitions, solved examples related to new topics, gave frequent homework, student achievement was better (House, 2001). As the impact of teacher professional development in the area of assessment on student achievement, it is included in the present study.

Very few studies looked into student achievement in specific subject areas. A few studies related to algebra and geometry. In the present study, it is proposed to consider the TIMSS (2007) scores in all the seven domains of mathematics at the eighth grade level for finding mutual associations among them, as well as for comparing the U.S. with five other countries. Moreover, the impact of five variables on each of the content domains, i.e., Number, Algebra, Geometry, and Data & Chance will be probed.

Comparative studies using TIMSS data concerning mathematics and involving the U.S. pointed out that the curricular structure in the U.S. was rather arbitrary and lacked coherence as compared to high achieving countries (Schmidt, Wang, and McKnight (2005). Moreover, in several districts in the U.S. did not meet the international standards of course coverage in eighth grade (Schmidt et al., 2011). Reformed teaching methods were found to be more effective than traditional teaching methods, especially in the case of African-American students, according to the dissertation study of Cheng (2010). Menon (2000) however, argued that due to the emphasis on conceptual understanding, the scores of the U.S. in TIMSS have not been very impressive. This contention will be probed in the present study, as according the research question 3, the study involves comparison of the performance of the U.S. with the three top performing countries and two others in the cognitive domains of Knowing, Applying, and Reasoning, along with scores in the content domains.

Comments and Future Directions

The existing literature did not show any study to find associations between student achievement in one area in the content domain or cognitive domain with another. Again, several important factors concerned with student-related, school-related, instruction-related, teacher-related, home-related backgrounds are not probed for association with student achievement. Some such ones that correspond to the indicators of Carroll's model of school learning are included in the present study. These variables are: the amount of time spent on math homework (student background), frequency of applying facts, concepts and procedures (classroom instruction), ensuring that student completes homework (home involvement), availability of resources for math instruction (school resources), and teacher professional development in mathematics assessment (teacher background). Also, analytical studies that compare the achievements of U.S. students with the top performing countries separately in all the areas of the content domains and in the cognitive domain are not available. The present study is focused on the above issues.

It is hoped that the answers to the research question 3 of this study on the comparisons of student achievement in the U.S. and the three highest performing countries of TIMSS (2007) and two other countries in each of the content and cognitive domains, in conjunction with the study related to research question 2, which concerns the impact of some variables might provide a few insights into the possible potential factors that could improve the effectiveness of the educational system in the U.S. With the additional inputs in the answer to the research question 1 about linkages between the scores in all the domains, some tangible clues pertaining to the needed modifications in curricular priorities and instructional practice in mathematics at the middle grade level in the U.S. might emerge, so as to provide some modest help in improving the student achievement in future.

As an ongoing extension of this study, it is felt that some studies could be taken up with a focus on replication of the results of this study in some modified form, like taking some other variables, countries for countries, and considering other TIMSS and also other international large-scale assessment studies.

Chapter Three: Method

This chapter is organized into the following sections: the purpose, research questions, research design, and data analysis.

Purpose of the Study

The purpose of this study is to explore the connections between eighth grade mathematics student achievements in the content domains and in the cognitive domains, in the context of TIMSS 2007. A second purpose is to find the connections between student achievements in the cognitive domains with the variables of factors in student background, classroom instruction, home involvement, school resources, and teacher professional development. Further, the math achievement of the students of the United States in all the domains will be compared to those of five selected countries. After analysis, interpretation will be made to indicate relationships between student achievement in the content and cognitive domains and also between the achievement in the content domains and the variables. The following research questions are related to the purpose of this study.

Research Questions

The study will be aimed to address the following set of research questions:

1. Are the TIMSS 2007 eighth-grade mathematics scores in the content domains of Number, Algebra, Geometry, and Data & Chance and the

scores in the cognitive domains of Knowing, Applying, and Reasoning across the countries associated with one another?

2. Are the TIMSS 2007 eighth-grade mathematics scores in Number, Algebra, Geometry, and Data & Chance across the countries associated with the variables related to factors in student background (amount of time spent on math homework), classroom instruction (frequency of applying facts, concepts and procedures), home involvement (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), school resources (for math instruction), and mathematics teacher professional development (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy /instruction, content, and assessment)?

3. How does the performance of the United States in the content domains Number, Algebra, Geometry, and Data & Chance and the cognitive domains of Knowing, Applying, and Reasoning compare with the performances of Bulgaria, Chinese Taipei, Republic of Korea, Singapore, and Thailand?

Research Design

This section deals with data source, sampling procedures, data collection methods, and country profiles.

Data source. In this study, data concerning the eighth grade mathematics achievement from the Trends in International Mathematics and Science Study (TIMSS) 2007, which was conducted by the International Association for the Evaluation of Educational Achievement (IEA), will be used. As an independent international cooperative of research institutions and government agencies, IEA has been conduction cross-national studies of student achievement in a wide range of subjects since 1959.

TIMSS has assessed mathematics and science in 1995, 1999, 2003, and 2007, and the work for the 2011 assessment is in progress. Apart from monitoring trends in achievement at the fourth and eighth grades, TIMSS collects a vast array of background information on trends in achievement in the context of varying educational systems, organizational approaches of schools, and instructional practices. TIMSS 2007, which covered 59 countries around the world, involving about 425,000 students, has provided information about educational achievement across countries with a view to serve as a resource for improving teaching and learning of mathematics and science (TIMSS, 2007).

Sampling procedures. TIMSS 2007 employed a two-stage stratified cluster design for selecting representative samples of students from each participating country. In the first stage, a sampling of schools was taken, followed in the second stage by a sampling of intact classrooms from the target grade in the sampled schools. The selection of schools was made with probability proportional to size, whereas classrooms had equal probabilities. About 150 schools were sampled in most countries, with one or two intact classrooms from each school. This approach ensured a representation of a sample size of at least 4,500 students in each country. With a view to ensure efficient sampling design and implementation, each country had national coordinators, who were trained in selecting school and student samples by using '*WinW$_3$S*' software provided by the IEA Data Processing and Research Center' (TIMSS 2007, p. 382).

Data collection. TIMSS 2007 was conducted in many languages, involving a huge work of translating all the assessment instruments. Representatives from the participating countries were involved in developing the TIMSS tests for 2007. Item writing workshops were conducted by the TIMSS & PIRLS International study Center. Participants were given guidelines for translating the instruments into their respective language, making efforts to adapt them to their appropriate contexts. Quality control of the items was supervised by observers within each country and also by international quality control observers arranged by the IEA Secretariat. All the items generated in each country were reviewed by subject-matter experts.

Data collection was made by contacting the schools and sampling classes to prepare materials for data collection. Assessment was then administered, followed by taking quality control measures, scoring, and creation of the data files. For constructed responses, scoring rubrics were applied across all the countries and scoring was done independently by two scorers. TIMSS calculated the reliability of the measurement of student achievement median 'KR-20' reliability across the test booklets. 'Reliability was generally high -- 0.8 to 0.9 in most countries' (TIMSS 2007, p. 401). For reporting the

TIMSS 2007 achievement data, item response theory (IRT) scaling methods were adopted. An overview of its framework of eighth grade mathematics tests and a distribution of the test items are shown in Table 1 and Table 2 respectively.

Table 1

Overview of TIMSS 2007 Eighth Grade Mathematics Framework

Content		Cognitive	
Domain	Percentage	Domain	Percentage
Number	30%	Knowing	35%
Algebra	30%	Applying	40%
Geometry	20%	Reasoning	25%
Data & Chance	20%		

Table 2

Distribution of Mathematics Items

Type	Domain	Number of Multiple Choice Items	Number of Constructed Response Items	Total Number of Items	Total Number of Score Points	Percentage of Score Points
Content	Number	35	28	63	72	30
	Algebra	34	30	64	69	29
	Geometry	31	16	47	50	21
	Data & Chance	17	24	41	47	20
	Total	117	98	215	238	100
Cognitive	Knowing	54	27	81	83	35
	Applying	48	40	88	98	41
	Reasoning	15	31	46	57	24
	Total	117	98	215	238	100

For the present study, mathematics achievement at the eighth grade level has been chosen because of its influence on education in the United States and other countries (Rodriguez, 2004; TIMSS, 1995, 1999, 2003, & 2007). Data of average scores across all the countries that participated was considered for the purpose of the first two research questions. For the purpose the third research question however, data concerning only six countries sought to be compared is considered. These countries are: the United States, the three top scoring countries, viz., Chinese Taipei, Korea, and Singapore, along with Bulgaria, which had scored better than the United States in 1995, but has slightly fallen behind the U.S. in the last two TIMSS assessments, and Thailand, which has consistently been scoring midway among the countries that scored below the international average. Details of the number of students, number of schools, and class sizes in the countries compared are provided in Table 3.

Table 3

Summary of the Sample Included for Comparison in the Study

Country	No. of students	No. of schools	Class size
Bulgaria	4019	163	22
Chinese Taipei	4046	150	35
Korea, Rep. of	4240	150	37
Singapore	4599	164	38
Thailand	5412	150	38
United States	7377	239	24

Country Profiles

The countries selected for a comparative study in Research Question 3 are: Bulgaria, Chinese Taipei, Republic of Korea, Singapore, Thailand, and United States. Brief profiles of these countries with regard to location, population, composition of ethnic groups, languages spoken, and expenditure on education are provided in Table 4, followed by further details about the education systems in these countries. This information will help understand the contexts in which educational efforts are being carried out in these countries, so that comparisons of student achievement in these countries can be understood with more relevance to their backgrounds.

Table 4

Comparative Statement of Country Profiles (Source: CIA)

Country	Location	Population	Ethnic Groups	Languages	Education Expenditures
Bulgaria	Southeastern Europe, bordering the Black Sea, between Romania and Turkey	7,093,635 (July 2011 est.)	Bulgarian 83.9%, Turk 9.4%, Roma 4.7%, other 2% (including Macedonian, Armenian, Tatar, Circassian) (2001 census)	Bulgarian (official) 84.5%, Turkish 9.6%, Roma 4.1%, other and unspecified 1.8% (2001 census)	4.1% of GDP (2007)
Chinese Taipei	Eastern Asia, islands bordering the East China Sea, Philippine Sea, South China Sea, and Taiwan Strait, north of the Philippines, off the southeastern coast of China	23,071,779 (July 2011 est.)	Taiwanese (including Hakka) 84%, mainland Chinese 14%, indigenous 2%	Mandarin Chinese (official), Taiwanese (Min), Hakka dialects	NA
Korea, South	Eastern Asia, southern half of the Korean Peninsula bordering the Sea of Japan and the Yellow Sea	48,754,657 (July 2011 est.)	homogeneous (except for about 20,000 Chinese)	Korean, English (widely taught in junior high and high school)	4.2% of GDP (2007)
Singapore	Southeastern Asia, islands between Malaysia and Indonesia	4,740,737 (July 2011 est.)	Chinese 76.8%, Malay 13.9%, Indian 7.9%, other 1.4% (2000 census)	Mandarin (official) 35%, English (official) 23%, Malay (official) 14.1%, Hokkien 11.4%, Cantonese 5.7%, Teochew 4.9%, Tamil (official) 3.2%, other Chinese dialects 1.8%, other 0.9% (2000 census)	3% of GDP (2009)
Thailand	Southeastern Asia, bordering the Andaman Sea and the Gulf of Thailand, southeast of Burma	66,720,153 (July 2011 est.)	Thai 75%, Chinese 14%, other 11%	Thai, English (secondary language of the elite), ethnic and regional dialects	4.1% of GDP (2009)
United States	North America, bordering both the North Atlantic Ocean and the North Pacific Ocean, between Canada and Mexico	313,232,044 (July 2011 est.)	white 79.96%, black 12.85%, Asian 4.43%, Amerindian and Alaska native 0.97%, native Hawaiian and other Pacific islander 0.18%, two or more races 1.61% (July 2007 estimate)	English 82.1%, Spanish 10.7%, other Indo-European 3.8%, Asian and Pacific island 2.7%, other 0.7% (2000 census)	5.5% of GDP (2007)

An overview of the educational systems in the countries considered is given below:

Bulgaria. Bulgaria has public schools and private educational institutions at all levels. Public schools do not charge fees. Secondary schools are of two types: general secondary schools and professional schools. In some professional secondary schools, the medium of teaching can be chosen from English, German, French, Russian, Italian, Spanish, apart from Bulgarian language. Education in Bulgaria is compulsory between the ages of 7 and 16. Although, the official language of instruction is Bulgarian, students may study their mother tongue in municipal schools under the protection and control of the state. There are three components in the curriculum: compulsory, elective, and optional. However, different types of schools have different correlations between these components. Children with special needs are well taken care of, by either integrating them into mainstream schools or by placing them in special needs schools, the latter often being boarding schools. An interesting feature in parental cooperation in Bulgaria is that there are some forums, such as *The International Women's Club of Sofia* and 'bg-mamma.com' which network parents to discuss the issues concerned with their children's educational and other problems. (Source: 'Angloinfo')

The grading system is uniform at all levels of schooling and is based on oral testing, homework, and in-class participation. Grades are indicated by numerals; where 6 is excellent, 5 is very good, 4 is good, 3 is satisfactory, 2 is poor, and 1 is very poor. Students who get less than 3 in three subjects need to repeat the year. However, a supplementary examination is allowed to improve the grade. Homework plays an important role in Bulgarian education. Students from grades 1 through 12 normally spend half a day in school, while the other half is spent doing homework and independent study at home.

Education in Bulgaria has significant foreign influences, although it has a fundamental national character. Schools in Bulgaria are well equipped with traditional audiovisual aids, but lack in contemporary computer hardware and software. Educators, however, are generally quite open to new instructional methods. They usually get their professional improvement through online exchange of information and methods. Bulgaria's educational reform is aimed to bring standards in line with the European context and to harmonize the educational process with that of Western Europe. UNESCO, the Council of Europe, the European Union, and the World Bank sponsored several educational initiatives through programs like PHARE, TEMPUS, and COPERNICUS to improve the structure and content of education in Bulgaria. (Source: Bulgaria-Educational System)

The source for the profiles of the remaining five countries described below with respect to their educational systems is TIMSS (2007) Encyclopedia, except where otherwise indicated.

Chinese Taipei. The administration of education in Chinese Taipei is under the Ministry of Education in the central government and the Bureau of Education in the local government. Education is compulsory up to grade 9. Five mathematics strands are administered according to the grade level: number and quantity, geometry, algebra, statistics and probability, and mathematical connection. The language of instruction in

Chinese Taipei is Mandarin Chinese. Schools work 5 days a week for 200 days. Mathematics in middle grades is taught in periods of 45 minutes for approximately 125 hours. Use of technology is encouraged in the subject area learning of elementary and secondary education. Basic Competency Test for junior high school students, conducted twice a year, is used as the main criteria for high school studies. There are three kinds of institutions for preparing teachers: normal universities, education-related departments, and teacher education centers in universities. To qualify as teacher, one has to pass the preservice teacher education program, followed by a half-year practicum. Later, teachers need to undergo periodical professional development programs, conducted by the Bureaus of Education of local governments. According to Ministry of Education, Republic of China (Taiwan), in Taiwan's teacher education system, graduates of normal universities are qualified to teach at junior and senior high schools, whereas graduates of universities of education are qualified to teach at kindergartens and elementary schools.

Republic of Korea. Korean education policy professes to assist people in perfecting their individual character to develop an ability to achieve an independent life and acquire the qualifications of democratic citizens, and to be able to participate in building a democratic state and promoting the prosperity of all humankind. Educational autonomy at the local level was promoted with the implementation of new modes of operation. In the national common basic curriculum, Mathematics is organized as a differentiated curriculum. Each of the ten grade levels of mathematics courses are organized in two sublevels that are operated on semester-basis. Students who fail at a level must take special supplementary classes. Each level has six content domains having in-depth processes, which are taken by students who have higher capabilities. Differentiation of the curriculum is made at the school level or district level. Students who need extra help with the content can take supplementary classes.

Mathematics instructional time in middle schools in Korea is 136 hours yearly, covered in instructional periods of 45 minutes each. With a single-track 6-3-3-4 system, 6 years primary, 3 years of both middle and high school, and 4 years of a university education, the Korean educational system is streamlined to ensure that there is no discrimination. School enrolment in primary and middle schools is compulsory and free of cost, with Korean as the medium of instruction. Middle school students are evaluated at the end of each academic semester. Students who complete a grade move onto the next grade automatically. However, in mathematics and English, students need to reach a level to move onto the next grade. Otherwise, they need to take special supplementary classes. On the other hand, in order to provide specialized education to gifted students, separate classes are arranged for them. About 0.3 percent of the entire student population has been identified as gifted. The public education system ran 451 classes for 8,200 gifted students in 2004. In addition, the Education Center for the Gifted ran 862 classes for 16,500 students. The use of a calculator is recommended to perform complicated calculations and to help students understand mathematical concepts and to solve mathematical problems, but not for teaching and learning about developing calculation skills.

With a view to improving the quality and professionalism of teachers, Korea is provides training programs through education offices and universities. Distance training programs have also been launched. Through these programs, which are provided at educational training centers under the city or provincial education offices, university-affiliated education training centers, and private training centers through in class, distance training, and learning at home modes, teachers acquire higher qualifications. Parents in Korea usually give much attention to their children's education, according to the 'The Economist' (2012), which stated that on November 10, 2011, the sole concern of the country was on the University Entrance Test for students, with parents being solely concerned with their children's academic ambitions. Also, it is noteworthy that, speaking about South Korean education, 'Asia Society' mentions that the South Korean society deems that for social and economic improvement, lifelong learning needs to be taken as a precept.

Singapore. Typically, there are four official languages in Singapore: Malay, Chinese (Mandarin), Tamil, and English. Although, Malay is the national language, English is the language of administration. Singapore's education system has a centralized curriculum where guidelines for syllabi for various subjects and assessment are provided by the ministry. The education system recognizes the differences in abilities, learning styles, and interests of students, and is geared to give the students flexibility to progress along the best suited pathways in order to realize their fullest potential. The Ministry of Education, Singapore conducts seminars on primary and secondary education to support parents and students in making informed decisions about education pathways (MOE Seminars for Parents). At the end of grade 6, students appear for the Primary School Leaving Examination (PSLE), "which assesses their abilities for placement in a secondary school course that suits their learning pace and aptitude" (TIMSS 2007

Encyclopedia, p. 540). Flexibility and choice are achieved through innovations at the school level, involving school leaders and teachers for taking the initiative to develop their own school-based programs to meet the learning needs of their students to encourage schools to embark on such innovations with the support of the education ministry.

A common curriculum framework is used at all levels; however, emphases differ at every level. There are five components in the curriculum: concepts, skills, processes, metacognition, and attitudes, problem-solving abilities. The framework sets out directions for the teaching, learning, and assessment of mathematics. Beginning teachers need to undergo preservice pedagogical and instructional training, followed by a practicum in school.

Thailand. Nine years of basic education is compulsory in Thailand. Basic education is provided by the state free for 12 years. In order to overcome the shortage of teachers, the Ministry of Education provides training in subject content and pedagogy to teachers who do not have relevant training. Under the aegis of the 'Institute for the Promotion of Teaching Science and Technology', several professional development programs for teachers are being undertaken in various ways, such as face-to-face, by correspondence, via satellite in collaboration with the Distance Learning Foundation and educational television, via the Internet, and at symposia organized by different agencies.

The mathematics curriculum has set standards in the following six subject content areas: numbers and operations, measurement, geometry, algebra, data analysis and probability, and mathematical skills and processes. Computer literacy, which is mandatory in secondary education, is an integral part of the core curriculum. The instructional time in secondary schools is 150 hours per semester in Thailand. Usually the periods of instruction are of 60 minutes duration. During the last decade, the Ministry of Education has implemented projects for developing software, media, and curriculum improvement, along with technologies for education under the ‘School Net Project’. If a student does not pass a common test, the school is made responsible for providing remedial teaching and retesting the student.

United States. In the United States, nearly 88 percent of elementary and secondary school students attend public schools. The U.S. Department of Education distributes and monitors funding for specific programs created by Congress. It organizes collection of data and dissemination of research, ensuring key educational issues such as equal access to education. However, the school districts decide the curricula. States develop curriculum frameworks and implement accountability to curriculum standards. In U.S. middle schools students move from one classroom another for different subjects with different teachers, and change their peer groups for different classes. Students are offered some choice in the selection of elective courses. Typically, there is no official language in the United States. But, since about 92 per cent of the population speaks English, the medium of instruction is English. However, there is demand for teachers of English as second language. Although all middle and secondary teachers (grades 6–12) are required to be specialists within their fields, some states allow teachers with elementary-level certification to work in middle schools. Technology use is high in U.S. middle schools, about 89 per cent in 2005, and more now. The amount of time students at a given grade level are expected to spend on mathematics is decided by the policies established by local school districts. Student assessments other than long-term and multistep projects are usually completed in the classroom itself.

The No Child Left Behind Act (NCLB) of 2001 stipulates that states should test students annually to assess whether schools are making adequate yearly progress toward proficiency benchmarks. All states require standardized tests to be administered to students from elementary, middle, and high school. All public school teachers must have a bachelor's degree and fulfill their state's requirements for certification. Beginning middle and high school teachers must pass a specified academic subject test. Curricular emphasis is on mastering basic skills, understanding concepts, and applying mathematics in real-life contexts. Individual school districts need to manage resources to provide schools with the instructional tools. Even though there are no national policies that prescribe the content and methods of professional development programs, several workshops and summer institutes are organized. More than 90% of teachers receive some professional training support.

Instruments and Variables

The source of the instruments and variables, considered in this dissertation study was TIMSS 2007 technical report.

Instruments. The TIMSS 2007 assessment contained 215 items in mathematics at the eighth grade. It involved assembling the items into 14 blocks of items, and then assembling the blocks into 14 booklets. Each student was administered a single booklet. These were developed over a period of 2 years, from January 2005 to November 2006.

Updating the 2007 assessment frameworks. The content domains were presented separately for the fourth and eighth grades. To increase the potential for analyzing and reporting the results according to cognitive domains, the cognitive domains were updated with support from the U.S. National Center for Education Statistics. Moreover, the allocation of blocks to booklets was modified to include fewer blocks in a booklet and to have the design fully balanced. Each booklet in TIMSS 2007 included 2 mathematics blocks and 2 science blocks, with half the booklets having the mathematics blocks first and half having the science blocks first. The time given to students to complete a block was increased from 15 to 2.5 minutes at eighth grade.

At eighth grade, the 2003 measurement domain was eliminated and the topics covered were redistributed to geometry (length, area, volume, angle, perimeter, and circumference) or number (time, speed, mass/weight, and temperature). The cognitive domains for mathematics were reduced from four to three: knowing, applying, and reasoning. The previous two cognitive domains, using concepts and solving routine problems, were split across the three new domains.

Blueprints for item development were prepared by considering number of items needed in the assessment based on the score points in each content domain and by distributing this number of items across the mathematics main topic areas according to their breadth of content, and by scaling up the number of items to be developed to allow for attrition during the item selection and field-testing process. Scoring guides were written at the same time items were drafted.

Trend items. The trend items from 2003 were mapped into the content and cognitive categories described in the TIMSS 2007 frameworks. Also, more new constructed-response items were added. To increase efficiency, the field test blocks were organized to represent the desired assessment as much as possible. Where the field test was successful, materials did not have to be reformatted. A total of 415 items were included in the field test, 214 in mathematics and corresponded to a total of 283 score points in mathematics. The field-test item pool was divided into two sets, "preferred" and "alternate".

Item selection for the TIMSS 2007 data collection. Based on the item analysis of the international results of the field test, selection of items for the TIMSS 2007 for data collection was made. Data almanacs containing basic item statistics for each country and internationally were developed by taking into consideration the difficulty levels for each item, capacity to discriminate between high- and low-performing students, effectiveness of distractors in multiple-choice items, the frequency of occurrence of diagnostic codes used in the scoring guides, and scoring reliability for constructed-response items. Appendix 2 shows the distribution of new and trend items in the TIMSS 2007 mathematics and science assessments by item format for eighth grade, and the scoring guidelines are shown in Appendix 3.

Assessment booklet design. The TIMSS design for 2007 divided the 353 items at fourth grade and 429 items at eighth grade into 28 item blocks at each grade, 14 mathematics blocks labeled M01 through M14, and 14 science blocks labeled S01 through S14. This general block design is shown in Appendix 4 and is the same for each grade level. However, the assessment time was 22.5 minutes for eighth grade blocks. In the TIMSS 2007 design, the 28 blocks of items were distributed across 14 student booklets, as shown in Appendix 5. Each booklet consisted of four blocks of items. To enable linking between booklets, each block appears in two booklets. The assessment time for individual students was 90 minutes at eighth grade, which is comparable to that in the 1995, 1999, and 2003 assessments. The booklets were organized into 2 two-block sessions (Parts I and II), with a break in between each part.

Assembling item blocks (see Appendix 6). The assessment blocks were assembled to create a balance across blocks and booklets with respect to content domain, cognitive domain, and item format. Depending on the exact number of multiple-choice and constructed-response items in each block, the total number of mathematics items in a block ranged from 11 –18 at eighth grade.

Updating TIMSS 2007 background questionnaires (see Appendices 7 and 8). TIMSS 2007 included four types of background questionnaires to collect information regarding the contexts in which students learn mathematics and science. The *Curriculum Questionnaire* collected information from the participating countries about the organization of the curriculum and the topics intended to be covered. The *School Questionnaire* asked the students' school principals to provide information about the school contexts and the resources available for instruction. The *Teacher Questionnaire* collected information from the students' teachers about the teachers' backgrounds, preparation, and professional development. It also asked about instructional activities and detailed information about the topics in the subject matter taught to students, and the

Student Questionnaire addressed students' home and school lives and their experiences in learning.

Documenting scoring reliability. In order to establish the reliability of the scoring within each country, two different scorers scored independently a random sample of 200 responses for each constructed-response item. The random samples were selected from 14 booklets, using WinW3S software.

Creating the TIMSS 2007 data files. A Windows-based program, called WinDEM to accommodate data entry and data verification. This software also offered data and file management capabilities, a convenient checking and editing mechanism, interactive error detection, and reporting and quality-control procedures.

Field test. In order to avoid any possible problems during data collection, the TIMSS 2007 field test involved approximately 1,400 students per grade tested in each participating country. About half of the achievement items were released into the public domain. Items that replaced the released ones were tried out in the field test in order to investigate the psychometric characteristics of the achievement items and make well-informed decisions about the best replacements. The field test in mathematics involved 7 newly developed item blocks. The field test was conducted from March–April 2006.

Scaling methodology. The IRT scaling approach used by TIMSS was developed originally by the Educational Testing Service for use in the U.S. National Assessment of Educational Progress (NAEP). It is based on psychometric models that were first used in the field of educational measurement in the 1950s and were extensively used since the 1970s for use in large-scale surveys, test construction, and computer adaptive testing.

Omitted and not-reached responses. An item was considered not reached when—within part 1 or part 2 of the booklet—the item itself and the item immediately preceding it were not answered, and there were no other items completed in the remainder of that part of the booklet. In estimating the values of the item parameters, items in the TIMSS 2007 assessment booklets that were considered not to have been reached by students were treated as if they had not been administered. However, not-reached items were always considered as incorrect responses when student proficiency scores were generated.

Variables. The selection of variables for this study has been mainly prompted by the conceptual model (Carroll, 1963), the review of existing literature on contextual factors related to student achievement in mathematics at the eighth grade (see Chapter Two), and also the practical implications of the variables to applicability in educational practice in the United States.

The variables of the study vary for each research question. For the first research question, the correlated variables are the IRT-based scores across all the participating countries in all the seven domains -- four in the content domains of *Number*, *Algebra*, *Geometry*, and *Data & Chance* and three cognitive domains of *Knowing*, *Applying*, and *Reasoning*. For the second research question, the dependent variables are the scores across all participating countries in the four content domains, whereas the independent variables are factors in student background (amount of time spent on math homework), classroom instruction (frequency of applying facts, concepts and procedures), home involvement (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), school resources (for math instruction), and teacher professional development (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy /instruction, content, and assessment). And for the third research question, the comparison variables are again the scores in all the seven domains, but only across six countries -- Bulgaria, Chinese Taipei, Republic of Korea, Singapore, Thailand, and the United States.

Table 5 shows a mapping between the variables selected to answer the second research question in this study and the variables identified in Carroll's (1963) conceptual model of school learning. It may be noted that the choice of variables for this part of the study was made from among the available 346 variables in the secondary data of TIMSS with a view to conform to a considerable extent to the norms of Carroll's conceptual model, giving priority to fill the gaps in the light of the available literature, reviewed in Chapter Two: Literature Review.

Table 5

Mapping of Variables in Carroll's Model with Variables in the Study

Variables in Carroll's Model	*Variables in the Study*
1) Aptitude – the amount of time needed to learn under optimal instructional conditions	1) Student background (amount of time spent on math homework)
2) Ability to understand instruction	2) Classroom instruction (how often ask/apply facts and concepts)
3) Perseverance – the amount of time the learner is willing to engage actively in learning	3) Home involvement (ask parents\ensure complete homework, ask parents\raise funds for school, ask parents\serve on school committees, ask parents\attend special events, ask parents\volunteer for school projects)
4) Opportunity to learn – time allowed for learning	4) School resources available for math learning
5) Quality of instruction – the extent to which instruction is presented so that no additional time is required for mastery beyond that required in regard to aptitude	5) Math teacher professional development (participation in developing I.T. into math participation in developing critical thinking, participation in developing math curriculum, participation in developing math pedagogy, participationin developing math content, participation in developing math assessment)

Descriptions of the variables considered for the study are given in Table 6.

Table 6

Descriptions of the Variables Considered in the Study

Variables in the Study	*Description of Variable*
1) Student background (amount of time spent on math homework)	**Index** of time spent on math homework (TMH) **ID**: BSDMTMH This index is based on two questions given to students about the frequency with which they are assigned homework and the amount of time they spend doing it. Responses to the frequency of homework assignment are coded on a 5-point scale for each item as follows. **Values**: *Every day* = 1*; 3 or 4 times a week* = 2*; 1 or 2 times a week* = 3*; Less than once a week* = 4*; Never = 5*. Responses to the amount of time spend on homework are coded for each item as follows: *Zero minutes* = 1; *1–15 minutes* = 2; *16–30 minutes* = 3; *31–60 minutes* = 4; *61–90 minutes* = 5; *More than 90 minutes* = 6. Responses are categorized for the index variable such that *High* = students report that they receive homework at least 3 or 4 times a week and spend more than 30 minutes on each assignment; *Low* = students report that they receive homework no more than twice a week and spend 30 minutes or less on each assignment; *Medium* = all other combinations of answers.
2) Classroom instruction (how often ask/apply facts and concepts)	**Full Title**: In teaching mathematics to the students in the TIMSS class, how often do you usually ask them to apply facts, concepts and procedures to solve routine problems? **ID**: BT4MASAC, **Values**: Every or almost every lesson, About half the lessons, Some lessons, Never
3) Home involvement (ask parents/ensure complete homework, ask parents\raise funds for school, ask parents\serve on school committees, ask parents\attend special events, ask parents\volunteer for school projects?	**Full Title**: Does your school ask parents to ensure that their child completes his/her homework? **ID**: BC4GAPCH; **Values**: Yes, No **Full Title**: Does your school ask parents to raise funds for the school? **ID**: BC4GAPRF; **Values**: Yes, No **Full Title**: Does your school ask parents to serve on school committees? **ID**: BC4GAPSC; **Values**: Yes, No **Full Title**: Does your school ask parents to attend special events? **ID**: BC4GAPSE; **Values**: Yes, No **Full Title**: Does your school ask parents to volunteer for school projects, programs, and trips? **ID**: BC4GAPVO; **Values**: Yes, No
4) School resources available for math learning	**Index** of available school resources for math instruction (ASRMI); **ID**: BCDSRMI This index is based on principals' responses about the availability of general school resources: instructional materials (e.g., textbook), budget for supplies (e.g., paper, pencils), school buildings and grounds, heating/cooling and

	lighting systems, instructional space (e.g., classrooms); and instructional materials for mathematics/science: computers for mathematics/science instruction, computer software for mathematics/science instruction, calculators for mathematics/science instruction, library materials relevant to mathematics/science instruction, and audio-visual resources for mathematics/science instruction. Responses were coded on a 4-point scale for each item as follows. **Values**: *None* = 1; *A little* = 2; *Some* = 3; *A lot* = 4. Responses were categorized for the index variable such that *High* = the average of general materials is less than 2, and the average of mathematics/science-specific materials is than 2; *Low* = average of general materials is greater than or equal to 3, and the average of mathematics/science-specific materials is greater than or equal to 3; *Medium* = all other response combinations.
5) Math teacher professional development (participation in developing I.T. into math participation in developing critical thinking, participation in developing math curriculum, participation in developing math pedagogy, participationin developing math content, participation in developing math assessment)	**Full Title**: In past two years, have you participated in professional development in integrating information technology into mathematics? **ID**: BT4MPDIT; **Values**: Yes, No **Full Title**: In the past two years, have you participated in professional development in improving students' critical thinking or problem solving skills? **ID**: BT4GPDCT; **Values**: Yes, No **Full Title**: In the past two years, have you participated in professional development in mathematics curriculum? **ID**: BT4MPDMC; Values: Yes, No **Full Title**: In the past two years, have you participated in professional development in mathematics pedagogy/instruction? **ID**: BT4MPDMP; **Values**: Yes, No **Full Title**: In the past two years, have you participated in professional development in mathematics content? **ID**: BT4MPDMT; **Values**: Yes, No **Full Title**: In the past two years, have you participated in professional development in mathematics assessment? **ID**: BT4MPDMA; **Values**: Yes, No

Data Analysis

The data used in this study are derived from a secondary source, TIMSS 2007 report, pertaining to international student achievement in mathematics at the eighth grade.

Secondary data analysis. It has become a common practice for researchers to use secondary data for researches in social and behavioral sciences. However, their advantages and disadvantages need to be considered while choosing to use them.

Advantages. An important advantage of using secondary data for research is the availability of ready samples for research, for which, collection of data and preparation of statistical packages have already been done. For effective generalizations, research studies need large-scale data. When such data can be obtained from a secondary data source, not only there is an enormous conservation of time and expense (Rosenberg et al., 2006). Secondary data become more important when they are collected by public agencies and established governmental organizations, such as the National Science Foundation (NSF), the National Center for Education Statistics (NCES), and the Association for Institutional Research (AIR). Again, from large-scale secondary data sources, several researchers can conduct different studies to answer different research questions. Testing different theoretical frameworks and hypotheses can be done from the same large-scale secondary data source (Kiecolt & Nathan, 1985).

Disadvantages. There are also some disadvantages associated with using secondary data. The purposes of using secondary data for analysis and the purpose of collecting the secondary data may be different. Moreover, the availability and accessibility of advanced statistical methodologies and specialized statistical software that must be used for analysis of complex, large-scale secondary data may also pose a problem (Gonzales, 2001). Again, data quality, including missing data, how latent constructs are defined, and the quality of supporting documentation for the data source are also limitations for the use of secondary data (Rosenberg, et al., 2006).

Measures used in the study. Three different measures are used in this study. For the first research question, I used Pearson's product moment coefficient of correlation, which determines the presence (or absence) of a linear relationship between two variables X and Y, its direction (positive or negative), and its strength. For the second research question, I employed the method of multiple linear regression, which explains or predicts a dependent variable from two or more independent variables. However, causality is not implied. Lastly, for the third research question, Independent samples *t*-test, which tests the (null) hypothesis that a relation targeted in the research question does not exist.

Analyses. The data analysis for Research Question 1 will consist of 42 correlations among the average scores of the three content domains (Number, Algebra, Geometry, and Data & Chance) and the three cognitive domains (Knowing, Applying, and Reasoning). Using the data related to the achievement scores of all the participating countries in each of these domains, a correlational matrix 7 x 7 will be generated, using SPSS software for this purpose. Leaving out the self-correlations, the remaining 42 will be analyzed for finding associations.

Pearson's product-moment correlation co-efficient (r) will be used as the measure for this analysis. According to Dimitrov (2008), "In general, a positive linear relationship (positive correlation) between two variables, X and Y, occurs when high values on X tend to be associated with high values on *Y* and, conversely, low values on *X* tend to be associated with low values on *Y*" (p.133). In the same way, "negative linear relationship (negative correlation) between two variables, *X* and *Y*, occurs when high values on *X* tend to be associated with low values on *Y* and, conversely, low values on *X* tend to be associated with high values on *Y*" (p.135). The values of *r* range from -1.0 to 1.0, the former indicating perfect negative correlation and the latter indicating perfect positive correlation.

As regards the Research Question 2, the data will comprise the scores (dependent variables) of all the participating countries in each of the four content domains, while the five chosen independent variables operate one at one time. The independent variables are: factors in student background (amount of time spent on math homework), classroom instruction (frequency of applying facts, concepts and procedures), home involvement (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), school resources (for math instruction), and teacher professional development (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy /instruction, content, and assessment). Here, the multiple regression method of analysis will be employed to find the amount of shared variance and strength of relationship between the chosen variables. For this purpose, the overall scores will be computed for of each attribute of the variables by giving proportionate weightages to each of the alternative values in respect of different variables. Since five attributes are grouped under the variable Home Involvement and six attributes for Math Teacher professional development, the respective averages will be calculated to arrive at the overall scores for those variables.

Lastly, for the Research Question 3, the data will comprise the average student achievement scores of the United States and the five other comparison countries in each of the seven subject domains. In this case, since comparison of student achievement is to be made with each of the other five countries, independent samples *t*-test will be used as the measure.

Rationale. For finding association (RQ1), linear relationship among the scores is to be found, which is indicated by Pearson's product moment coefficient of correlation 'r'. For RQ2, how the dependent variable student achievement is predicted or explained by each of the independent variables is investigated. For this purpose, the multiple regression method is amply suitable. Using SPSS, we can obtain the value of the F-statistic that provides indication of evidence if the variance in the dependent variables is accounted for by the predictors does not equal to zero for the population. Again, the coefficient of determination R^2, will give the percentages of differences in student achievement that are accounted for by each variable. Further, the regression coefficients will indicate the contribution of each variable to R^2. Thus, we will be able to interpret if and to what extent each independent variable predicts and explains the differences in student achievement.

For RQ3, the scores of the U.S., taking one domain at a time (P_0), and the scores of one of the 5 countries, Bulgaria, Chinese Taipei, Korea, Singapore, and Thailand, one by one (P_1) are considered. As these samples are independent of each other, testing of the null hypothesis H_0: $P_0 = P_1$ is done against the alternative hypothesis H_a: $P_0 \neq P_1$, using *t*-test for independent samples.

Limitations. In case Pearson's product moment coefficient of correlation (r) = 0, it shows that there is no linear relationship between two variables. However, there could be some non-linear relationship. Also, there could be a positive linear relationship between some subgroups of the variables, while other subgroups might have a negative linear relationship. However, this situation can be overcome by examining separate subgroups.

For multiple regression, like for simple regression, there are four assumptions -- linearity, independence, normality, and homoscedasticity i.e., the condition, where the variance of the dependent variable is the same for all the data. However, all these are supposed to be embedded in TIMSS data. Again, multicollinearity, when it occurs, can vitiate the predictive inferences of the variables. But, in this study, as the independent variables do not contain much of the same information, they will not have high inter-correlations and so, multicollinearity is not likely to occur.

Chapter Four: Results and Data Analysis

Introduction

The researcher explored the connections between eighth grade mathematics student achievements in the content domains and in the cognitive domains, in the context of TIMSS 2007 and the connections between student achievements in the cognitive domains with the variables of factors in student background, classroom instruction, home involvement, school resources, and teacher professional development. Also, the researcher made a study to compare the math achievement of the students of the United States in all the domains with those of five selected countries.

The data analysis and results of the study are presented in this chapter, which is organized in five sections: (a) introduction, (b) data collection, (c) data analysis, (d) findings, and (e) summary. After the introduction, the researcher presents a discussion of the data collected, descriptive statistics, and instrument reliability for the collection of data. A summary of the number of samples of the participating schools and students is also given. The third section contains the quantitative data analysis. In the fourth section, the results of the research questions are discussed. Finally, the chapter concludes with a summary of the important points.

Data Collection

This section describes the sampling procedures adopted in the collection of data, the scale anchoring analysis, the methods used for estimating standard errors, the background questionnaires, preparation of data files, and the methods used for the verification of international translations. It also shows the descriptive statistics of the data and discusses how the issues of reliability and validity are dealt with.

Sampling procedures. The sampling method adopted in TIMSS (2007) was by means of a two-stage cluster design. The first stage was for sampling of schools, and the second one for classrooms from the sampled schools, which were selected with probability proportional to size, and classrooms with equal probabilities.

Scale anchoring analysis. "For the scale anchoring analysis, the students' achievement results from all the participating countries were pooled, so that the benchmark descriptions refer to all students achieving at that level. Thus, in determining performance in relation to the benchmarks, it does not matter what country a student is from, only how he or she performed on the test. Considering students' mathematics achievement scores, criteria were applied to identify the sets of items that students reaching each international benchmark were likely to answer correctly and that those at the next lower benchmark were unlikely to answer correctly." (TIMSS 2007, p. 407)

Estimating standard errors. "Because the statistics presented in this report are estimates of national performance based on samples of students—rather than on the values that could be calculated if every student in every country had answered every question—it is important to have measures for the degree of uncertainty of the estimates"

(TIMSS 2007, p. 409). Jackknife procedure was used to estimate the error due to sampling and also the error due to variation between the plausible values generated for each student. Based on these errors, 95% confidence intervals were generated to make inferences about population means and proportions.

Background questionnaires. “TIMSS 2007 included four types of background questionnaires to collect information regarding the contexts in which students learn mathematics and science” (TIMSS 2007 Technical Report, p.47). These questionnaires were prepared in English by the TIMSS International Study Center, with contributions from the National Research Coordinators (NRCs) of participating countries. The NRCs nominated an International Control Monitor (QCM) for their country in order to ensure the quality of TIMSS data. The IEA hired the QCMs, who received training by the TIMSS International Study Center.

Whereas information about the organization of the mathematics curriculum and the topics intended to be covered was gathered from the participating countries through the Curriculum Questionnaire, the School Questionnaire drew information about the school contexts and the resources available for mathematics instruction. Again, the Teacher Questionnaire gave information about teachers’ backgrounds, preparation, and professional development, apart from instructional activities. Lastly, the Student Questionnaire dealt with students’ home and school lives and their learning experiences.

Data files. Each participating country prepared data files after administering a field test and sent them to IEA Data Processing and Research Center, where the data were transformed into the international format after due processing, to be sent to the TIMSS International Study Center for analysis and review.

International verification of the translations. After verification at the respective national centers, the translations were subjected to further verification by an international translation company, followed by a check by International quality Control Monitors. Finally, TIMSS International Study Center reviewed each country's assembled test instruments. A summary of the samples of the 49 participating schools and students is given in Table 7.

Table 7

Summary of the Samples Included in the Study

Country	No. of Students	No. of Schools	Country	No. of Students	No. of Schools
Algeria	5447	149	Kuwait	4091	158
Armenia	4689	148	Lebanon	3786	136
Australia	4069	228	Lithuania	3991	142
Bahrain	4230	74	Malaysia	4466	150
Bosnia	4220	150	Malta	4670	59
Botswana	4208	150	Morocco	3060	131
Bulgaria	4019	163	Norway	4627	139
Chinese Taipei	4046	150	Oman	4752	146
Columbia	4873	148	Palestine	4378	148
Cyprus	4399	67	Qatar	7184	66
Czech Republic	4855	147	Romania	4198	149
Egypt	6582	233	Russian Fed	4472	210
El Salvador	4063	145	Saudi Arabia	4243	165
England	4025	137	Scotland	4070	129
Georgia	4178	135	Serbia	4045	147
Ghana	5294	163	Singapore	4599	164
Hong Kong	3470	120	Slovenia	4043	148
Hungary	4111	144	Sweden	5215	159
Indonesia	4203	149	Syria	4650	150
Iran	3981	208	Thailand	5412	150
Israel	3294	146	Tunisia	4080	150
Italy	4408	170	Turkey	4498	146
Japan	4312	146	Ukraine	4424	146
Jordan	5251	200	United States	7377	239
Korea, Rep. of	4240	150			

Descriptive statistics. "One reason for organizing data and using statistics is so that we can draw meaningful conclusions" (Jackson, 2005, p. 90). The means of average scores, average standard deviations, and the standard errors of Mean in respect of all the scales in the content and cognitive domains are presented in Table 8. The average mean ranged from 451.938 to 470.122, with standard deviations ranging from 53.98 to 74.26. The standard errors of Mean ranged from 3.18 to 3.54.

Table 8

Mean Values and Standard Deviations of All Variables

Variable Name	N*	Mean	Std Deviation	Std Error Mean
Number Score	48	452.042	71.85	3.258
Algebra Score	48	452.500	68.32	3.431
Geometry Score	48	451.938	71.56	3.538
Data & Chance Score	48	452.438	71.23	3.183
Knowing Score	48	452.500	74.26	3.310
Applying Score	48	452.021	72.78	3.325
Reasoning Score	41	470.122	53.98	3.376

Note. *The scores in respect of Morocco are not considered due to low participation rates. Average scores of 7 other countries in *Reasoning* could not be estimated.

The data pertaining to the Research Question 3 consists of the average scores and the standard errors of the countries compared in each of the Content and Cognitive domains. They are shown in Table 9 and Table 10.

Table 9

Average Scores and Standard Errors in Content Domains

CONTENT DOMAIN *Country*	Stu Size	*NUM* Av Score	*NUM* Std Error	*ALG* Av Score	*ALG* Std Error	*GEO* Av Score	*GEO* Std Error	*D &C* Av Score	*D &C* Std Error
Bulgaria	4019	458	4.7	476	5.1	468	5	440	4.7
Chinese Taipei	4046	577	4.2	617	5.4	592	4.6	566	3.6
Korea	4240	583	2.4	596	3.0	587	2.3	580	2.0
Singapore	4599	597	3.5	579	3.7	578	3.4	574	3.9
Thailand	5412	444	4.8	433	5.0	442	5.3	453	4.1
USA	7377	510	2.7	501	2.7	480	2.5	531	2.8

Table 10

Average Scores and Standard Errors in Cognitive Domains

COGNITIVE DOMAIN *Country*	Stu Size	Knowing *Av Score*	Knowing *Std Error*	Applying *Av Score*	Applying *Std Error*	Reasoning *Av Score*	Reasoning *Std Error*
Bulgaria	4019	477	4.7	458	4.8	455	4.7
Chinese Taipei	4046	594	4.5	592	4.2	591	4.1
Korea	4240	596	2.5	595	2.8	579	2.3
Singapore	4599	581	3.4	593	3.6	579	4.1
Thailand	5412	436	4.8	446	4.7	456	4.4
USA	7377	514	2.6	503	2.9	505	2.4

Reliability and validity.

Steps were taken by the TIMSS to ensure reliability of scoring and identification of variables. Reliability coefficient was also computed for the data pertaining to each country.

Scoring reliability. TIMSS adopted procedures to ensure that the tests were reliable, that is, that they were constructed with adequate number of items to provide reliable measurement and were supported by effective administrative and scoring procedures. In order to document the scoring reliability from one TIMSS cycle to the next, Trend Scoring and Reliability Scoring Software and Manual (TSRS) (IEA, 2006b) was developed. Reliability of the scoring within each country was ascertained by having two independent scoring on random samples of 200 responses. NRCs analyzed the results of the agreement between the two scorers, as well as between each of their TIMSS 2007 scorers and the scores that were awarded in 2003. If agreement on any comparison was below 85 percent, the scorers were retrained.

The identification variable (ID) cleaning. Each record in a data file has a unique identification number. Student Activity area questionnaire file, which contained the ID variables needed to be cleaned with the cooperation of the NRCs, as these included details pertaining to the student participation and exclusion status, as well as the dates of birth and dates of testing.

Test reliability coefficient. This is the median of Cronbach's Alpha reliability across the 14 test booklets. The median reliabilities in eighth grade mathematics were generally high, with the international median as 0.88. However, a few countries had median reliabilities below 0.70 at one or both grades in mathematics' they were Algeria, Botswana, El Salvador, Ghana, Kuwait, Qatar, and Saudi Arabia. Details of the values of Cronbach's Alpha in respect of all the countries in Mathematics Test in TIMSS (2007) are given in Table 11.

Validity. The issue of comparative validity was addressed by ensuring that inferences made about achievement differences between countries can be substantiated through quality assurance programs.

Reliability tests. The purpose of a reliability analysis is to determine how consistently the selected variables measure some construct. Cronbach's alpha (α) is a coefficient of reliability (or consistency) that can be written as a function of the number of test items and the average inter-correlation among the items. To establish the reliability (or consistency) of a measure, a strong correlation coefficient, usually in the .80 or .90 between the two variables is desired (Jackson, 2005). The formula that determines alpha (α makes use of the number of items in the scale (k) and the average correlation between pairs of items (r) is: $\alpha = k\, r\, /\, [1 + (k\text{-}1)\, r]$. As the number of items in the scale (k) increases, the value of α becomes larger, and if the inter-correlation between the items is large, the corresponding α will also be large (George & Mallery, 2008). A summary of country-wise values of Cronbach's alpha for eighth grade mathematics test of TIMSS (2007) are provided in Table 11.

Table 11

Cronbach's Alpha Reliability Coefficient – TIMSS 2007 Mathematics Test

Country	Reliability Coefficient	Country	Reliability Coefficient
Algeria	0.66	Kuwait	0.69
Armenia	0.88	Lebanon	0.84
Australia	0.89	Lithuania	0.89
Bahrain	0.80	Malaysia	0.88
Bosnia	0.84	Malta	0.89
Botswana	0.69	Morocco	0.76
Bulgaria	0.90	Norway	0.84
Chinese Taipei	0.93	Oman	0.80
Colombia	0.77	Palestine	0.83
Cyprus	0.88	Qatar	0.64
Czech Republic	0.88	Romania	0.90
Egypt	0.84	Russian Fed	0.90
El Salvador	0.63	Saudi Arabia	0.62
England	0.90	Scotland	0.89
Georgia	0.84	Serbia	0.89
Ghana	0.68	Singapore	0.92
Hong Kong	0.92	Slovenia	0.88
Hungary	0.90	Sweden	0.87
Indonesia	0.83	Syria	0.79
Iran	0.84	Thailand	0.88
Israel	0.90	Tunisia	0.78
Italy	0.87	Turkey	0.91
Japan	0.91	Ukraine	0.88
Jordan	0.88	United States	0.89
Korea	0.92		

Data Analysis

For the first Research Question 1, Pearson's product-moment correlation co-efficient (r) was used as the measure for finding correlations among the average scores in the three content domains (Number, Algebra, Geometry, and Data and chance) and the three cognitive domains (Knowing, Applying, and Reasoning). For Research Question 2, the data comprised the scores (dependent variables) of all the participating countries in each of the four content domains, while the five chosen independent variables were factors in Student Activity area(amount of time spent on math homework), Classroom Instruction (frequency of applying facts, concepts and procedures), Overall Home Involvement (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), School Resources (for math instruction), and teacher's Overall Professional Development (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy /instruction, content, and assessment). Analysis for this was made using multiple regression method. Finally, for Research Question 3, the data comprised the average student achievement scores of Bulgaria, Chinese Taipei, Republic of Korea, Singapore, and Thailand in each of the seven subject domains. In this case, in order to compare the United States with each of the other five countries, independent samples *t*-test was used as the measure.

Correlation analysis. Although 49 countries participated in TIMSS (2007) assessment in mathematics for grade 8, Morocco has not been considered, since, with overall participation of only 55%, it "did not satisfy guidelines for sample participation rates" (TIMSS 2007, p. 406). Moreover, in respect of seven countries out of the remaining 48 countries, "the average achievement could not be accurately estimated" (TIMSS 2007, p. 102) in *Reasoning* (vide Table 12). Therefore, correlations among the average eighth grade mathematics scores of the 48 countries in all the four content domains i.e., *Number*, *Algebra*, *Geometry*, and *Data & Chance* and also two of the three cognitive domains, viz., *Knowing* and *Applying* are calculated, whereas in the remaining domain of *Reasoning*, correlations among the average scores of only 41 countries are calculated. However, in order to develop a more comprehensive picture of the correlations of the scores in all the seven domains among all the 48 countries, the scores at five percentile levels of 10%, 25%, 50%, 75%, and 90% are also compiled separately and their corresponding correlations are also presented separately.

Table 12

Average Scale Scores of All Countries in All the Domains

Country	*Number*	*Algebra*	*Geometry*	*Data & Chance*	*Knowing*	*Applying*	*Reasoning*
Algeria	403	349	432	371	371	412	**
Armenia	492	532	493	427	507	493	489
Australia	503	471	487	525	487	500	502
Bahrain	388	403	412	418	395	403	413
Bosnia and Herzegovina	451	475	451	437	478	440	452
Botswana	366	394	325	384	376	351	**
Bulgaria	458	476	468	440	477	458	455
Chinese Taipei	577	617	592	566	594	592	591
Colombia	369	390	371	405	364	384	416
Cyprus	464	468	458	464	468	465	461
Czech Republic	511	484	498	512	502	504	500
Egypt	393	409	406	384	392	393	396
El Salvador	355	331	318	362	336	347	**
England	510	492	510	547	503	514	518
Georgia	421	421	409	373	427	401	389
Ghana	310	358	275	321	313	297	**
Hong Kong SAR	567	565	570	549	574	569	557
Hungary	517	503	508	524	518	513	513
Indonesia	399	405	395	402	397	398	405

Iran, Islamic Rep. of	395	408	423	415	403	402	427
Israel	469	470	436	465	473	456	462
Italy	478	460	490	491	476	483	483
Japan	551	559	573	573	560	565	568
Jordan	416	448	436	425	432	422	440
Korea, Rep. of	583	596	587	580	596	595	579
Kuwait	347	354	385	366	347	361	**
Lebanon	454	465	462	407	464	448	429
Lithuania	506	483	507	523	508	511	486
Malaysia	491	454	477	469	477	478	468
Malta	496	473	495	487	490	492	475
Norway	488	425	459	505	458	477	475
Oman	363	391	387	389	372	368	397
Palestinian Nat'l Auth.	366	382	388	371	365	371	381
Qatar	334	312	301	305	307	305	**
Romania	457	478	466	429	470	462	449
Russian Federation	507	518	510	487	521	510	497
Saudi Arabia	309	344	359	348	308	335	**
Scotland	489	467	485	517	481	489	495
Serbia	478	500	486	458	500	478	474
Singapore	597	579	578	574	581	593	579
Slovenia	502	488	499	511	500	503	496
Sweden	507	456	472	526	478	497	490
Syrian Arab Republic	393	406	417	387	393	401	396
Thailand	444	433	442	453	436	446	456
Tunisia	425	423	437	411	421	423	425
Turkey	429	440	411	445	439	425	441
Ukraine	460	464	467	458	471	464	445
United States	510	501	480	531	514	503	505

(** = the average achievement could not be accurately estimated (TIMSS, 2007 Report, p.121).

Correlation method describes the relationship between two or more measured variables (Jackson, 2005). Pearson's product-moment correlation coefficient, usually referred to as Pearson's *r*, was calculated to measure variables and assessing relationships between them. For this study, an alpha level of .05 was used for all statistical analyses.

When the coefficient *r* is large and positive, close to +1, the two variables are highly correlated in a positive way (Aczel & Sounderpandian, 2006). The data were analyzed by using Statistical Program for the Social Sciences, Version 19.0 (SPSS 19.0) for Windows. The values of Pearson coefficient of correlation (*r*) are depicted in the Table 13.

Table 13

Pearson's Product-Moment Correlation Coefficient (N = 48)

Dependent variable	Correlated variable	*r* – value
Score in Number	Score in Algebra	0.936
	Score in Geometry	0.957
	Score in Data & Chance	0.947
	Score in Knowing	0.983
	Score in Applying	0.990
	Score in Reasoning*	0.962
Score in Algebra	Score in Number	0.936
	Score in Geometry	0.930
	Score in Data & Chance	0.871
	Score in Knowing	0.977
	Score in Applying	0.944
	Score in Reasoning*	0.927
Score in Geometry	Score in Number	0.957
	Score in Algebra	0.930
	Score in Data & Chance	0.909
	Score in Knowing	0.960
	Score in Applying	0.981
	Score in Reasoning*	0.957
Score in Data & Chance	Score in Number	0.947
	Score in Algebra	0.871
	Score in Geometry	0.909
	Score in Knowing	0.923
	Score in Applying	0.954
	Score in Reasoning*	0.952
Score in Knowing	Score in Number	0.983
	Score in Algebra	0.977
	Score in Geometry	0.960
	Score in Data & Chance	0.923
	Score in Applying	0.980
	Score in Reasoning*	0.951
Score in Applying	Score in Number	0.990
	Score in Algebra	0.944
	Score in Geometry	0.981
	Score in Data & Chance	0.954
	Score in Knowing	0.980
	Score in Reasoning*	0.980
Score in Reasoning	Score in Number*	0.962
	Score in Algebra*	0.927
	Score in Geometry*	0.957
	Score in Data & Chance*	0.952
	Score in Knowing*	0.951
	Score in Applying*	0.980

Note. * indicates *N* = 41. In 7 countries, the average achievement in *Reasoning* could not be accurately estimated.

The average scale scores of all the countries in all the domains at the 10th, 25th, 50th, 75th, and 90th percentile levels are shown in Table 35, Table 37, Table 39, Table 41, and Table 43, while their respective correlations are depicted in Table 13, Table 36, Table 38, Table 40, Table 42, and Table 44. These data also revealed a total of 42 correlations. The correlations were identified with $r \geq 0.854$, $p < 0.001$ for the data at 10th percentile level, with $r \geq 0.878$, $p < 0.001$ for the data at the 25th percentile level, with $r \geq 0.882$, $p < 0.001$ for the data at the 50th percentile level, with $r \geq 0.875$, $p < 0.001$ for the data at the 75th percentile level, and with $r \geq 0.859$, $p < 0.001$ at the 90th percentile level.

Findings about correlations. From the TIMSS data, the actual scores of the students cannot be known individually and hence the correlations of student scores in the various domains cannot be exactly computed. However, the correlational trends can be assessed by finding the correlations among the average scores of all the participating countries in the seven domains. As a step towards possible further refinement, the correlations among the average scores in each of the domains of *Number*, *Algebra*, *Geometry*, *Data & Chance*, *Knowing*, *Applying*, and *Reasoning* at the different percentile levels of 10%, 25%, 50%, 75%, and 90% are computed. This gives a closer approximation of the correlational trends among the scores in the seven domains of TIMSS (2007) assessment in mathematics at the eighth grade. Table 14 shows a compilation (from Table 13, Table 36, Table 38, Table 40, Table 42, and Table 44) of the minimum and maximum values of Pearson's product-moment coefficient (r) among the average scores in all the seven domains at different percentile levels.

Table 14

Range of Values of 'r' at Different Percentile Levels of Average Scores

Level of Scores	Range of Minimum	Values of '*r*' Maximum	Level of Scores	Range of Minimum	Values of '*r*' Maximum
Overall Scores	0.871	0.990	50% Percentile Scores	0.882	0.991
10% Percentile Scores	0.854	0.988	75% Percentile Scores	0.875	0.992
25% Percentile Scores	0.878	0.989	90% Percentile Scores	0.859	0.992

Note. p < 0001.

In the process of comparing, one variable is taken as a reference variable and correlation to each of the other variables is found. The findings show the existence of a strong positive correlation at 0.01 level of significance (two-tailed) between a specific reference variable and each of all the other variables, thereby indicating that the data suggests a correlational tendency that students who scored high in one subject tend to score high in other subjects also, whereas students who scored low in a subject tend to score low in other domains also. The correlation between the content subject of *Number* and the cognitive domain subject of *Applying* was found to be most pronounced, followed by the correlation between *Number* and *Knowing*, *Geometry* and *Applying*, *Knowing* and *Applying*, and *Reasoning* and *Applying*. The correlation between *Algebra* and *Data & Chance* has been found to be the lowest among all the correlations. The above correlations are depicted diagrammatically by means of a Bar Chart in *Figure 1*.

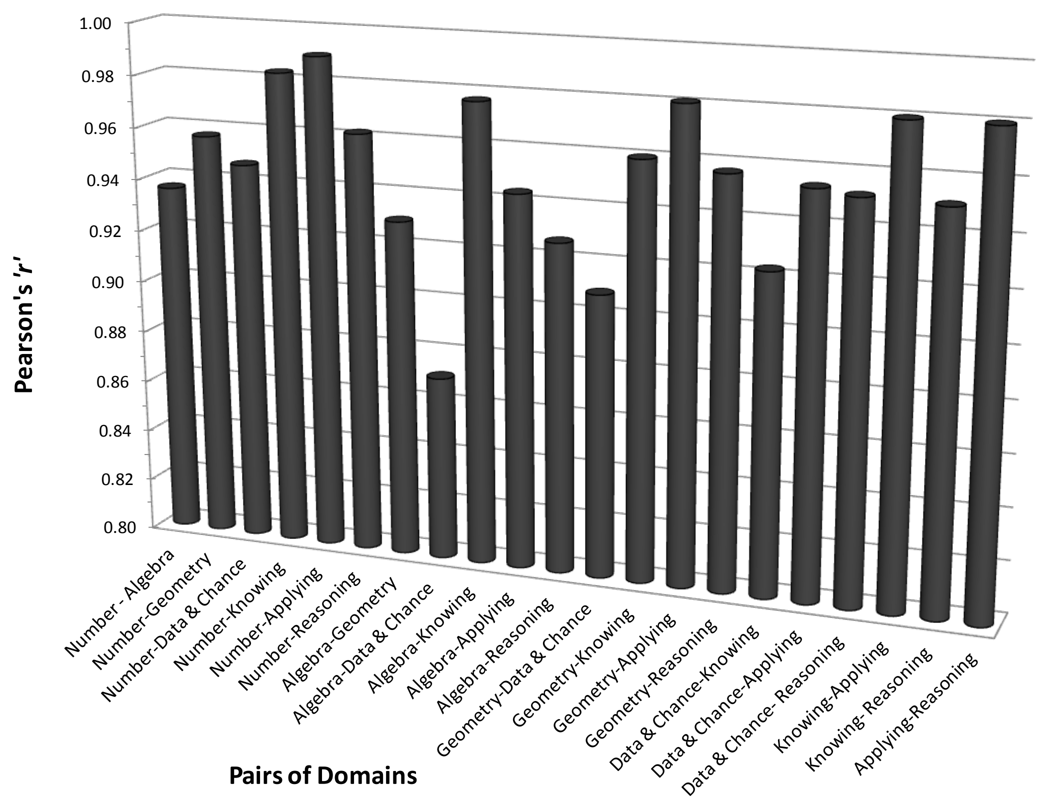

Figure 1. Bar Chart of Correlations among Pairs of Domains for Average Scores

Inferential data analyses procedure. Inferential statistics were used to draw conclusions from the sample tested to the population. The SPSS 19.0 was used to code and tabulate scores collected from the TIMSS (2007) report and provide summarized values. In addition, demographic data were presented using frequency statistics. Then, Multiple Regression analyses were used to detect the amount of shared variance and strength of relationship between the variables of interest.

Multiple regression analysis. While simple regression shows the mutual relationship between two variables, multiple regression analysis shows the relationship of two or more variables on a designated dependent variable. R^2 is a measure of how much of the variability in the dependent variable is accounted by the independent variables. The adjusted R^2 is a measure of how well the model generalizes over the population.

Hypotheses testing. In order to determine whether the five predictors, viz., *Homework Tim*e (Student background), *Classroom Instruction*, *Overall Home Involvement* (with five variant attributes in different contexts), *School Resources*, and *Overall Math Teacher P.D.* (with six variant attributes in different contexts) together account for a statistically significant proportion of the variance in *Y* (Overall Score in each of the content domains *Number*, *Algebra*, *Geometry*, *Data & Chance*), we consider the null hypothesis, which states that "the coefficient of determination for the population, $R^2_{pop} = 0$". The descriptions and codes of the predictors are enumerated in Table 15:

Table 15

Predictor Variables

Predictor	Description	Code
Homework Time (Student background)	Amount of time spent on math homework	BSDMTMH
Classroom Instruction	Frequency of applying facts, concepts and procedures to solve routine problems	BT4MASAC
Home Involvement	1. to ensure that their child completes his/her homework	BC4GAPCH
	2. to raise funds for the school	BC4GAPRF
	3. to serve on school committees	BC4GAPSC
	4. to attend special events	BC4GAPSE
	5. to volunteer for school projects, programs, and trips	BC4GAPVO
School Resources	For math Instruction	BCDSRMI
Math Teacher P.D.	Participation in P.D. in the past two years	
	1. for integrating I.T. into mathematics	BT4MPDIT
	2. for improving students' critical thinking or problem solving skills	BT4GPDCT BT4MPDMC
	3. for mathematics curriculum development	BT4MPDMP
	4. for in mathematics pedagogy/instruction	BT4MPDMT
	5. for mathematics content	BT4MPDMA
	6. for mathematics assessment	

For independent variable 1, i.e. *Homework Time* (amount of time spent on math homework in the Student Activity area) --Index of time spent on math homework (ID: BSDMTMH) has three values: *High*, *Medium*, and *Low*. They are given weights 3, 2, and 1 respectively. The linear product of the respective percentages with the above weights is divided by 100 to arrive at the value of the *Homework Time* in (Student Background) variable BSDMTMH, as shown below.

((HIGH_PERCENT *3) + (MED_PERCENT *2) + (LOW_PERCENT)) /100.

For independent variable 2, i.e., *Classroom Instruction* (frequency of applying facts, concepts and procedures to solve routine problems) with ID: BT4MASAC, there are four values '*Every or almost every lesson*', '*About half the lessons*', '*Some lessons*', and '*Never*'. These four alternative values are given is given weights 4, 3, 2, and 1 respectively. The linear product of the respective percentages with the above weights is divided by 100 to arrive at the value of the *Class Instruction* (facts, concepts, and procedures) variable BT4MASAC, as shown below.

((EVERY percent* 4) + (HALF percent* 3) + (SOME percent* 2) + (NEVER percent)) /100.

For independent variable 3, i.e., *Home Involvement* (Does your school ask parents to ensure that their child completes his/her homework?) ID: BC4GAPCH; in which, the value Values '*Yes'* and '*No'* are given weights 1 and 0, respectively. The percentages corresponding to the value '*Yes'* are divided by 100 to arrive at the value of the Home Involvement variable BC4GAPCH. **(YES percent)/100**. Similarly for the other four attributes BC4GAPRF, BC4GAPSC, BC4GAPSE, and BC4GAPVO of this variable

Home Involvement. The average of these five measures gives the overall score for the independent variable *Home Involvement*.

For independent variable 4, *School Resources*, the procedure is the same as that of variable 1, and for independent variable 5, the overall score for the independent variable *Math Teacher Professional Development* (with 6 attributes), the procedure is the same as that of variable 3. (Please refer to SPSS Output in Appendix 9). The compiled weightages for all the attributes of the predictor variables are depicted in Table 16 and Table 17.

Table 16

Data Used from TIMSS (2007) for Predictor Variables 1, 2, and 3

Variable	S.B.	S.B.	S.B.	C. I.	C. I.	C. I.	C. I.	H. I.	H. I.	H. I.	H. I.	H. I.
*Code**	BS DM TMH	BS DM TMH	BS DM TMH	BT4 MA SAC	BT4 MA SAC	BT4 MA SAC	BT4 MA SAC	BC4 GA PCH	BC4 GA PRF	BC4 GA PSC	BC4 GA PSE	BC4 GA PVO
Value	%High	%Med	%Low	%Every	%Half	%Some	%Nev	%Yes	%Yes	%Yes	%Yes	%Yes
Country/ Weightage	3	2	1	4	3	2	1	1	1	1	1	1
Algeria	‡	‡	‡	53	13	30	4	85	37	48	84	56
Armenia	32	64	4	24	26	28	22	91	53	90	91	84
Australia	15	44	42	28	27	45	1	97	71	97	96	77
Bahrain	15	67	18	34	33	31	1	97	31	32	92	64
Bosnia	24	51	25	39	20	41	‡	92	52	91	84	92
Botswana	29	50	20	38	21	40	1	88	99	89	82	76
Bulgaria	36	48	15	55	26	19	#	83	62	63	95	70
Chinese Taipei	31	46	23	28	31	40	‡	98	38	83	90	77
Colombia	36	48	16	48	35	16	1	98	31	63	93	90
Cyprus	20	70	11	42	33	25	‡	95	74	79	93	51
Czech Republic	5	46	49	34	35	31	1	95	40	70	58	76
Egypt	30	58	13	36	26	38	#	94	56	65	94	81
El Salvador	46	45	9	40	28	30	3	93	44	81	94	89
England	5	31	65	24	32	44	#	99	67	71	99	61
Georgia	34	62	4	55	12	33	‡	99	64	90	89	89
Ghana	28	55	16	63	17	19	1	79	66	95	82	62
Hong Kong	34	48	18	14	29	55	1	91	66	60	92	83
Hungary	16	78	6	41	39	19	1	94	77	62	75	91
Indonesia	29	53	18	30	35	32	2	97	71	80	77	54
Iran	19	55	26	43	28	27	2	89	70	63	72	77
Israel	34	53	13	32	27	39	2	86	33	56	91	83
Italy	45	47	7	51	25	21	3	96	27	51	96	47
Japan	8	36	57	26	40	30	4	78	13	29	100	74
Jordan	26	62	12	50	33	17	#	95	33	46	96	78
Korea	6	31	62	41	47	12	#	60	11	92	93	51
Kuwait	14	58	27	37	28	32	4	90	9	28	79	65
Lebanon	25	67	8	37	28	33	2	91	46	73	79	52
Lithuania	27	69	4	51	23	25	2	97	74	85	99	98
Malaysia	41	47	12	43	22	35	‡	92	85	57	98	77
Malta	24	71	5	29	46	24	#	100	74	75	99	58
Norway	25	53	22	11	28	60	1	92	18	91	90	90

Oman	12	73	15	57	24	19	‡	94	24	21	98	85
Palestine	24	68	7	47	30	21	1	99	38	19	100	80
Qatar	16	67	17	36	40	24	‡	94	28	30	91	75
Romania	66	29	5	40	30	28	1	99	49	68	78	85
Russian Fed	50	49	2	70	22	7	1	88	69	92	98	95
Saudi Arabia	13	61	26	34	31	33	2	97	16	93	96	44
Scotland	8	41	51	34	27	39	#	99	79	85	99	53
Serbia	31	40	28	42	25	31	1	97	72	96	77	83
Singapore	42	43	16	28	36	35	‡	91	69	63	98	96
Slovenia	20	64	16	37	38	24	#	96	44	38	98	70
Sweden	3	35	62	18	26	55	1	96	10	68	85	74
Syria	44	48	8	59	22	17	2	98	14	52	91	80
Thailand	39	45	15	24	40	35	1	89	92	77	95	78
Tunisia	45	44	11	35	24	39	2	97	36	21	79	60
Turkey	22	49	29	34	28	36	1	59	81	62	80	80
Ukraine	40	53	7	72	20	8	‡	93	91	90	97	86
United States	26	62	12	50	31	19	‡	98	82	89	99	97

(# = Rounds to zero; ‡ = Reporting standards not met)

Table 17

Data Used from TIMSS (2007) for Predictor Variables 4 and 5

Variable	S.R.	S.R.	S.R.	PD	PD	PD	PD	PD	PD
*Code**	BSDS RMI	BSDS RMI	BSDS RMI	BT4 MP DIT	BT4 GP DCT	BT4 MP DMC	BT4 MP DMP	BT4 MP DMT	BT4 MP DMA
Value	%High	%Med	%Low	%Yes	%Yes	%Yes	%Yes	%Yes	%Yes
Country/ Weightage	3	2	1	1	1	1	1	1	1
Algeria	11	80	9	27	60	51	66	51	51
Armenia	19	73	8	32	38	69	67	56	45
Australia	55	43	2	57	45	69	61	69	59
Bahrain	24	72	4	70	57	28	49	33	42
Bosnia	6	74	20	39	43	56	60	66	46
Botswana	4	65	30	13	27	11	12	20	27
Bulgaria	29	65	6	69	25	60	42	59	44
Chinese Taipei	36	58	6	73	40	84	79	84	52
Colombia	16	52	31	51	60	67	64	70	53
Cyprus	12	79	9	59	46	55	70	69	48

Czech Republic	62	38	‡	49	28	35	45	47	22
Egypt	27	68	6	54	77	34	66	46	51
El Salvador	13	63	24	26	45	26	42	49	38
England	34	61	5	62	40	61	79	66	59
Georgia	7	77	17	26	59	52	49	30	64
Ghana	11	77	12	13	44	44	38	60	46
Hong Kong	70	30	‡	63	60	72	71	78	56
Hungary	49	48	3	26	34	28	53	51	32
Indonesia	6	61	33	29	57	77	69	71	69
Iran	11	72	18	28	52	47	78	57	44
Israel	36	59	5	35	45	50	63	59	33
Italy	25	73	3	43	9	15	34	16	17
Japan	51	49	‡	27	38	31	76	74	39
Jordan	21	70	9	65	67	62	78	57	53
Korea	30	69	1	31	22	41	50	48	33
Kuwait	14	79	7	45	69	30	62	45	43
Lebanon	37	60	3	50	68	54	67	68	70
Lithuania	22	76	2	69	51	71	81	85	65
Malaysia	42	45	13	61	27	52	46	57	38
Malta	54	42	4	83	30	60	71	47	68
Norway	22	76	2	35	18	44	38	39	22
Oman	16	65	19	24	37	58	42	54	48
Palestine	19	67	14	26	45	34	47	44	35
Qatar	28	70	3	53	50	37	55	41	44
Romania	19	75	6	57	56	53	55	71	69
Russian Fed	28	67	5	67	62	74	73	84	60
Saudi Arabia	8	77	15	24	34	19	47	26	24
Scotland	48	51	1	79	56	74	93	80	71
Serbia	15	70	15	33	37	45	50	72	46
Singapore	91	9	‡	74	63	65	88	82	61
Slovenia	63	37	‡	62	37	66	65	70	72
Sweden	49	50	1	8	28	38	48	41	45
Syria	12	82	6	15	49	17	20	13	32
Thailand	13	66	21	73	82	79	80	82	83
Tunisia	6	73	21	22	36	26	35	24	32
Turkey	8	67	25	18	24	69	48	47	27
Ukraine	13	77	11	75	80	81	82	79	83
United States	51	45	4	60	65	80	76	81	69

Results of Multiple Regression Analysis

These results are discussed separately for each of the four hypotheses related to the second research question.

Hypothesis 2.1 analysis. Null Hypothesis *($H_{0, 2.1}$)*: The average score in *Number* does not change with changes in the variables connected with learning situations. The criterion variable is the average score in *Number* and the predictor variables are *Homework Time* (in Student Activity area), *Class Instruction*, *Home Involvement*, *School Resources,* and *Math Teacher Professional Development*.

The results from the *F*-test related to this hypothesis are reported in *Figure 4*. It can be seen that the *F*-statistic is statistically significant, $F(5, 42) = 11.944$, $p < 0.001$, thus providing evidence that the variance in *Y* accounted for by the five predictors does not equal zero for the population. Specifically, the coefficient of determination in the **Model Summary** Table, $R^2 = 0.587$, indicates that 58.7% of the differences students' differences in *Overall Score in NUMBER* are accounted for by their differences in *Homework Time* (in Student Activity area), *Classroom Instruction, Home Involvement, School Resources, and Math Teacher Professional Development*.

The coefficients for the predictor variables are listed in the column labeled *B* in Table 18. The multiple regression equation for predicting the dependent variable, *Y* (Overall Number Score) is:

$\hat{Y}$ = - 8.133 (Homework Time) – 39.910 (Classroom Instruction) + 175.884 (Home Involvement) + 172.165 (School Resources) + 34.344 (Math Teacher Professional Development) + 64.643.

But, from the SPSS output in *Figure 4*, we observe that only one predictor, i.e., *School Resources for mathematics instruction*, is statistically significant ($p < 0.001$).

However, one more predictor, *Overall Home Involvement* is significant at 10% significance level (p = 0.063). Moreover, the part correlation between Y and $X_{1,}$ partialling out the other predictors, is: $r_{Y(1.2345)}$ = -0.042. As $(-0.042)^2$ = 0.001764, 0.18% of the variance in *Overall Score in Number* is uniquely accounted for by the variance in *Homework Time* (in Student Activity area). The part correlation between Y and $X_{2,}$ partialling out the other predictors is: $r_{Y(2.1345)}$ = -0.133. As $(-0.133)^2$ = 0.017689, 1.77% of the variance in *Overall Score in Number* is uniquely accounted for by the variance in *Classroom Instruction –frequency of facts, concepts, and procedures.* Thirdly, the part correlation between Y and $X_{3,}$ partialling out the other predictors, is: $r_{Y(3.1245)}$ = 0.189, and since $(0.189)^2$ = 0.035721, it shows that 3.57% of the variance in *Overall Score in Number* is uniquely accounted for by the variance in *Overall Home Involvement.* Again, the part correlation between Y and $X_{4,}$ partialling out the other predictors is $r_{Y(4.1235)}$ = 0.589. Since $(0.589)^2$ = 0.346921, it indicates that 34.69% of the variance in *Overall Score in Number* is uniquely accounted for by the variance in *School Resources--for mathematics instruction.* Lastly, the part correlation between Y and $X_{5,}$ partialling out the other predictors is: $r_{Y(5.1234)}$ = 0.059, and since $(0.059)^2$ = 0.003481, 0.35% of the variance in *Overall Number Score* is uniquely accounted for by the variance in *Overall Math Teacher Professional Development.*

Table 18

Multiple Regression Analysis for Variables Predicting NUMBER Score

Variable	*B*	SE *B*	Beta	R^2
				0.59*
Homework Time	-8.13	19.17	-0.05	
Class Instruction	-39.91	29.77	-0.14	
Home Involvement	175.88	92.20	0.24**	
School Resources	172.16	28.96	0.65*	
Math Teacher P.D.	34.34	57.71	0.07	

Note. $N = 48$. $*p < .001$, $**.05 < p < .10$

The test for the null hypothesis, which is based on the ratio of the regression mean square to the residual mean square, is depicted in the ANOVA Table 19. The ratio of the two mean squares, labeled F, is 11.944. Since the observed significance level is less than .001 and the *F*-ratio is large, the null hypothesis (that an average score in *Number* does not significantly change with change in the learning situations) can be rejected because at least one of the coefficients is not 0.

Table 19

Analysis of Variance for NUMBER Score

Model		SS	*df*	MS	*F*	*p-value*
1	Regression	142458.85	5	28491.77	11.95	.000
	Residual	100185.07	42	2385.36		
	Total	242643.92	47			

Note.

a. Predictors: (Constant), Overall Math Teacher Prof Dev, Class Instruction BT4MASAC Facts Concepts, *Homework Time* (in Student Activity area) BSDMTMH Time spent on HW, School Resources BCDSRMI for math instruction, Overall Home Involvement.
b. Dependent Variable: Overall Score in Number.
c. SS = sum of scores; MS = mean score.

Hypothesis 2.2 analysis. Null Hypothesis *($H_{0, 2.2}$)*: The average score in *Algebra* does not change with changes in the variables connected with learning situations. The criterion variable is the average score in *Algebra* and the predictor variables are *Homework Time* (in Student Activity area), *Class Instruction*, *Home Involvement*, *School Resources,* and *Math Teacher Professional Development*.

The results from the *F*-test related to this hypothesis are reported in Table 20. It can be seen that the *F*-statistic is statistically significant, $F(5, 42) = 6.476$, $p < 0.001$, thus providing evidence that the variance in *Y* accounted for by the five predictors does not equal zero for the population. Specifically, the coefficient of determination in the **Model Summary** Table, $R^2 = 0.435$, indicates that 43.5% of the differences students' differences in *Overall Score in ALGEBRA* are accounted for by their differences in *Homework Time* (in Student Activity area), *Classroom Instruction, Home Involvement, School Resources, and Math Teacher Professional Development.*

The coefficients for the predictor variables are listed in the column labeled *B* in Table 20. The multiple regression equation for predicting the dependent variable, *Y* (Overall Algebra Score) is:

$\hat{Y}$ = 15.085 (Homework Time) – 27.832 (Classroom Instruction) + 107.395 (Home Involvement) + 134.747 (School Resources) + 64.955 (Math Teacher Professional Development) + 98.766.

However, from the SPSS output, shown in Figure 5, only one predictor, i.e., *School Resources for mathematics instruction,* is statistically significant ($p < 0.001$). Also, the part correlation between *Y* and $X_{1,}$ partialling out the other predictors, is: $r_{Y(1.2345)} = 0.082$. As $(0.082)^2 = 0.007225$, 0.72% of the variance in *Overall Score in Algebra* is uniquely accounted for by the variance in *Homework Time* (in Student Activity area). The part correlation between *Y* and $X_{2,}$ partialling out the other predictors is: $r_{Y(2.1345)} = -0.097$. As $(-0.097)^2 = 0.009409$, 0.94% of the variance in *Overall Score in Algebra* is uniquely accounted for by the variance in *Classroom Instruction –frequency of facts, concepts, and procedures.* Thirdly, the part correlation between *Y* and $X_{3,}$ partialling out the other predictors, is: $r_{Y(3.1245)} = 0.121$, and since $(0.121)^2 = 0.014641$, it shows that 1.46% of the variance in *Overall Score in Algebra* is uniquely accounted for by the variance in *Overall Home Involvement.* Again, the part correlation between *Y* and $X_{4,}$ partialling out the other predictors is $r_{Y(4.1235)} = 0.485$. Since $(0.485)^2 = 0.235225$, it indicates that 23.52% of the variance in *Overall Score in Algebra* is uniquely accounted for by the variance in *School Resources--for mathematics instruction.* Lastly, the part correlation between *Y* and $X_{5,}$ partialling out the other predictors is: $r_{Y(5.1234)} = 0.117$, and since $(0.117)^2 = 0.013689$, 1.37% of the variance in *Overall Algebra Score* is uniquely accounted for by the variance in *Overall Math Teacher Professional Development.*

Table 20

Multiple Regression Analysis for Variables Predicting ALGEBRA Score

Variable	*B*	SE *B*	Beta	R^2
				0.43*
Homework Time	15.08	21.32	0.09	
Class Instruction	-27.83	33.10	-0.10	
Home Involvement	107.39	102.53	0.15	
School Resources	134.75	32.20	0.53*	
Math Teacher P. D.	64.95	64.17	0.14	

Note. $N = 48$. *$p < .001$, **$.05 < p < .10$

The test for the null hypothesis, which is based on the ratio of the regression mean square to the residual mean square, is depicted in ANOVA Table 21. The ratio of the two mean squares, labeled F, is 6.476. Since the observed significance level is less than .001 and the *F*-ratio is large, the null hypothesis (that an average score in *Algebra* does not significantly change with change in the learning situations) can be rejected because at least one of the coefficients is not 0.

Table 21

Analysis of Variance for ALGEBRA Score

Model		SS	*df*	MS	*F*	*p-value*
1	Regression	95496.01	5	19099.20	6.48	.000
	Residual	123873.98	42	2949.38		
	Total	219370.00	47			

Note.

a. Predictors: (Constant), Overall Math Teacher Prof Dev, Class Instruction BT4MASAC Facts Concepts, *Homework Time* (in Student Activity area) BSDMTMH Time spent on HW, School Resources BCDSRMI for math instruction, Overall Home Involvement.
b. Dependent Variable: Overall Score in Number.
c. SS = sum of scores; MS = mean score.

Hypothesis 2.3 analysis. Null Hypothesis *($H_{0, 2.3}$)*: The average score in *Geometry* does not change with changes in the variables connected with learning situations. The criterion variable is the average score in *Geometry* and the predictor variables are *Homework Time* (in Student Activity area), *Class Instruction, Home Involvement, School Resources,* and *Math Teacher Professional Development*.

The results from the *F*-test related to this hypothesis are reported in Table 22. It can be seen that the *F*-statistic is statistically significant, $F(5, 42) = 8.186$, $p < 0.001$, thus providing evidence that the variance in *Y* accounted for by the five predictors does not equal zero for the population. Specifically, the coefficient of determination in the **Model Summary** Table, $R^2 = 0.494$, indicates that 49.4% of the differences students' differences in *Overall Score in Geometry* are accounted for by their differences in *Homework Time* (in Student Activity area), *Classroom Instruction, Home Involvement, School Resources, and Math Teacher Professional Development*.

The coefficients for the predictor variables are listed in the column labeled B in Table 22. The multiple regression equation for predicting the dependent variable, Y (Overall Geometry Score) is:

$\hat{Y}$ = -8.905 (Homework Time) – 27.551 (Classroom Instruction) + 39.005 (Home Involvement) + 153.679 (School Resources) + 91.764 (Math Teacher Professional Development) + 142.113.

But, by observing the SPSS output in Figure 6, we notice that one predictor, i.e., School Resources for mathematics instruction, is statistically significant ($p < 0.001$).

Moreover, the part correlation between Y and $X_{1,}$ partialling out the other predictors, is: $r_{Y(1.2345)} = 0.046$. As $(0.046)^2 = 0.002116$, 0.21% of the variance in *Overall Score in Geometry* is uniquely accounted for by the variance in *Homework Time* (in Student Activity area). The part correlation between Y and $X_{2,}$ partialling out the other predictors is: $r_{Y(2.1345)} = -.092$. As $(-.092)^2 = 0.008464$, 0.85% of the variance in *Overall Score in Geometry* is uniquely accounted for by the variance in *Classroom Instruction – frequency of facts, concepts, and procedures.* Thirdly, the part correlation between Y and $X_{3,}$ partialling out the other predictors, is: $r_{Y(3.1245)} = 0.042$, and since $(0.042)^2 = 0.01764$, it shows that 1.76% of the variance in *Overall Score in Geometry* is uniquely accounted for by the variance in *Overall Home Involvement.* Again, the part correlation between Y and $X_{4,}$ partialling out the other predictors is $r_{Y(4.1235)} = 0.528$. Since $(0.528)^2 = 0.278784$, it indicates that 27.88% of the variance in *Overall Score in Geometry* is uniquely accounted for by the variance in *School Resources--for mathematics instruction.* Lastly, the part correlation between Y and $X_{5,}$ partialling out the other predictors is: $r_{Y(5.1234)} =$ 0.158, and since $(0.158)^2 = 0.024964$, 2.50% of the variance in *Overall Geometry Score* is uniquely accounted for by the variance in *Overall Math Teacher Professional Development.*

Table 22

Multiple Regression Analysis for Variables Predicting GEOMETRY

Variable	*B*	SE *B*	Beta	R^2
				0.49*
Homework Time	-8.90	21.14	-0.05	
Class Instruction	-27.55	32.83	-0.10	
Home Involvement	39.00	181.69	0.05	
School Resources	153.68	31.94	0.58*	
Math Teacher P. D.	91.76	63.65	0.20	

Note. $N = 48$. $*p < .001$, $**.05 < p < .10$

The test for the null hypothesis, which is based on the ratio of the regression mean square to the residual mean square, is depicted in the ANOVA Table 23. The ratio of the two mean squares, labeled F, is 8.186. Since the observed significance level is less than .001 and the *F*-ratio is large, the null hypothesis (that an average score in *Geometry* does not significantly change with change in the learning situations) can be rejected because at least one of the coefficients is not 0.

Table 23

Analysis of Variance for GEOMETRY Score

Model		SS	*df*	MS	*F*	*p-value*
1	Regression	118760.35	5	23752.07	8.19	.000
	Residual	121868.46	42	2901.63		
	Total	240628.81	47			

Note.

a. Predictors: (Constant), Overall Math Teacher Prof Dev, Class Instruction BT4MASAC Facts Concepts, *Homework Time* (in Student Activity area) BSDMTMH Time spent on HW, School Resources BCDSRMI for math instruction, Overall Home Involvement.
b. Dependent Variable: Overall Score in Number.
c. SS = sum of scores; MS = mean score.

Hypothesis 2.4 analysis. Null Hypothesis *($H_{0, 2.4}$)*: The average score in *Data & Chance* does not change with changes in the variables connected with learning situations. The criterion variable is the average score in *Data & Chance* and the predictor variables are *Homework Time* (in Student Activity area), *Class Instruction, Home Involvement, School Resources,* and *Math Teacher Professional Development*.

The results from the *F*-test related to this hypothesis are reported in Table 24. It can be seen that the *F*-statistic is statistically significant, $F(5, 42) = 10.502$, $p < 0.001$, thus providing evidence that the variance in *Y* accounted for by the five predictors does not equal zero for the population. Specifically, the coefficient of determination in the **Model Summary** Table, $R^2 = 0.556$, indicates that 55.6% of the differences students' differences in *Overall Score in DATA & CHANCE* are accounted for by their differences in *Homework Time* (in Student Activity area), *Classroom Instruction, Home Involvement, School Resources, and Math Teacher Professional Development*.

The coefficients for the predictor variables are listed in the column labeled *B* in Table 24. The multiple regression equation for predicting the dependent variable, *Y* (Overall Data & Chance Score) is:

$\hat{Y}$ = -18.612 (Homework Time) – 34.012 (Classroom Instruction) + 176.235 (Home Involvement) + 168.385 (School Resources) + 20.952 (Math Teacher Professional Development) + 83.017.

The SPSS output in Figure 7, shows that only one predictor, i.e., *School Resources for mathematics instruction*, is statistically significant ($p < 0.001$). However, one more predictor, *Overall Home Involvement* is significant at 10% significance level ($p = 0.070$). Again, the part correlation between *Y* and X_1, partialling out the other predictors, is: $r_{Y(1.2345)} = -0.097$. As $(-0.097)^2 = 0.009409$, 0.94% of the variance in *Overall Score in Data & Chance* is uniquely accounted for by the variance in *Homework Time* (in Student Activity area). The part correlation between *Y* and X_2, partialling out the other predictors is: $r_{Y(2.1345)} = -0.114$. As $(-0.114)^2 = 0.012996$, 1.30% of the variance in *Overall Score in Data & Chance* is uniquely accounted for by the variance in *Classroom Instruction –frequency of facts, concepts, and procedures.* Thirdly, the part correlation between *Y* and X_3, partialling out the other predictors, is: $r_{Y(3.1245)} = 0.191$, and since $(0.191)^2 = 0.036481$, it shows that 3.65% of the variance in *Overall Score in Data & Chance* is uniquely accounted for by the variance in *Overall Home Involvement.* Again, the part correlation between *Y* and X_4, partialling out the other predictors is $r_{Y(4.1235)} = 0.581$. Since $(0.581)^2 = 0.337561$, it indicates that 33.76% of the variance in *Overall Score in Data & Chance* is uniquely accounted for by the variance in *School Resources--for mathematics instruction.* Lastly, the part correlation between *Y* and X_5, partialling out the other predictors is: $r_{Y(5.1234)} = 0.036$, and since $(0.036)^2 = 0.001296$, 0.13% of the variance in *Overall Data & Chance Score* is uniquely accounted for by the variance in *Overall Math Teacher Professional Development.*

Table 24

Multiple Regression Analysis for Variables Predicting DATA & CHANCE

Variable	*B*	SE *B*	Beta	R^2
				0.56*
Homework Time	-18.61	19.72	-0.11	
Class Instruction	-34.01	30.62	-0.12	
Home Involvement	176.23	94.84	0.24**	
School Resources	168.38	29.79	0.64*	
Math Teacher P. D.	20.95	59.36	0.04	

Note. $N = 48$. *$p < .001$, **$.05 < p < .10$

The test for the null hypothesis, which is based on the ratio of the regression mean square to the residual mean square, is depicted in the ANOVA Table 25. The ratio of the two mean squares, labeled *F*, is 10.502. Since the observed significance level is less than .001 and the *F*-ratio is large, the null hypothesis (that an average score in *Data & Chance* does not significantly change with change in the learning situations) can be rejected because at least one of the coefficients is not 0.

Table 25

Analysis of Variance for DATA & CHANCE Score

Model		SS	*df*	MS	*F*	*p-value*
1	Regression	132505.71	5	26501.14	10.50	.000
	Residual	105984.10	42	2523.43		
	Total	238489.81	47			

Note.

a. Predictors: (Constant), Overall Math Teacher Prof Dev, Class Instruction BT4MASAC Facts Concepts, *Homework Time* (in Student Activity area) BSDMTMH Time spent on HW, School Resources BCDSRMI for math instruction, Overall Home Involvement.
b. Dependent Variable: Overall Score in Number.
c. SS = sum of scores; MS = mean score.

Comparing student achievement of the U.S. and five other countries. This section reports the descriptive information obtained for the sample, and will address the analysis and results for each of the null hypotheses established for this study. Archived data (vide Table 9 and Table 10) of the average scores from TIMSS (2007) in respect of the eighth grade middle school students of the six countries intended to be compared, viz., Bulgaria, Chinese Taipei, Korea, Singapore, Thailand and the United States in all the seven domains were compiled and entered into the Statistical Program for the Social Sciences, Version 19.0 (SPSS 19.0) for Windows.

Descriptive statistics. The means and standard deviations of the average scale scores of Bulgaria, Chinese Taipei, Korea, Singapore, and Thailand in each of the content and cognitive domains are shown in Table 9 and Table 10. They are depicted by means of bar diagrams shown in *Figure 2* and *Figure 3*.

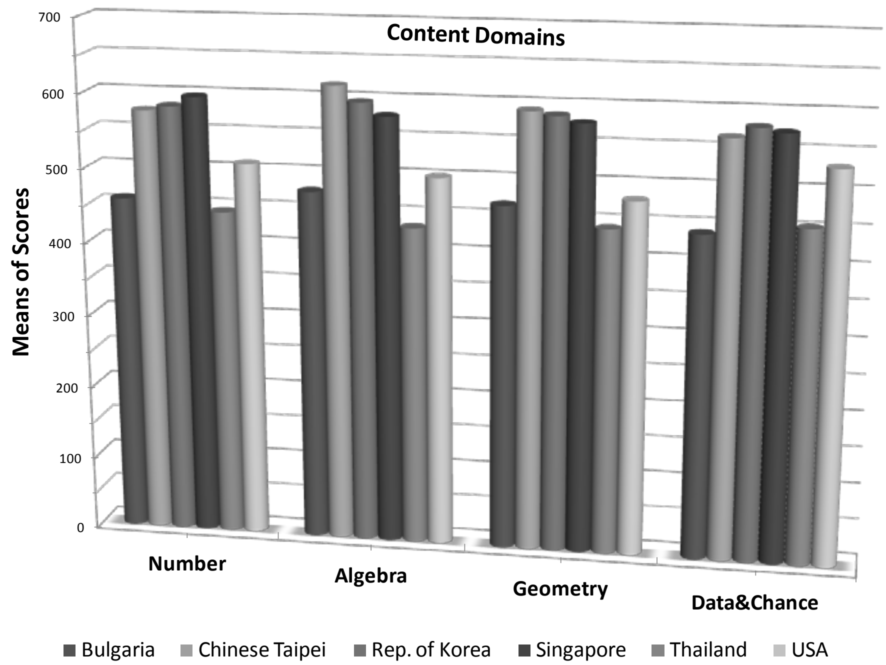

Figure 2. Means of Scores in the Content Domains

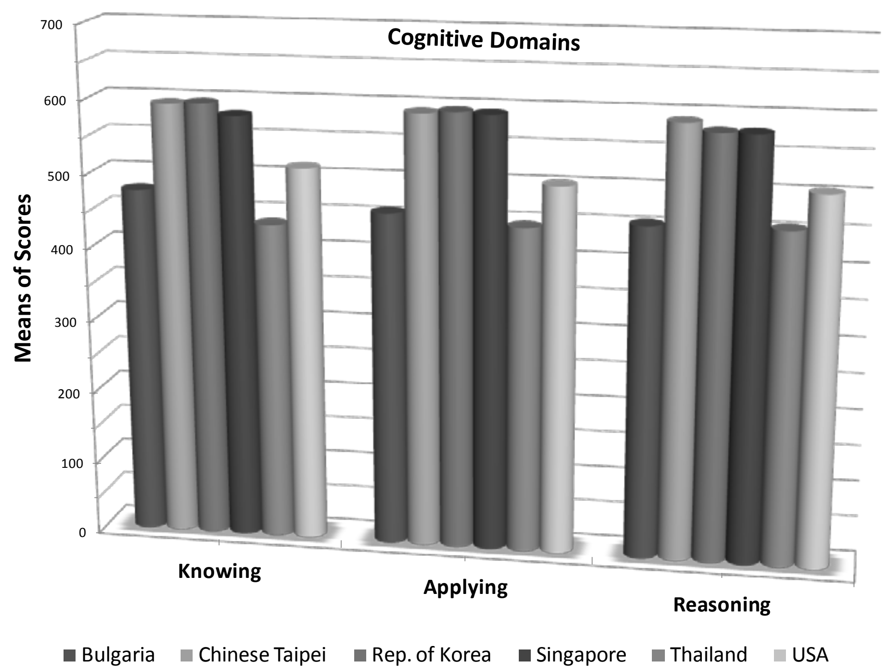

Figure 3. Means of Scores in the Cognitive Domains

There are seven Hypotheses for this section as stated below.

Hypothesis 3.1 $(H_{0,\,3.1})$: $\mu_1 = \mu_2$ (There is no difference between the population means of the United States and any of the other compared countries in *Number*)

Hypothesis 3.2 $(H_{0,\,3.2})$: $\mu_1 = \mu_2$ (There is no difference between the population means of the United States and any of the other compared countries in *Algebra*)

Hypothesis 3.3 $(H_{0,\,3.3})$: $\mu_1 = \mu_2$ (There is no difference between the population means of the United States and any of the other compared countries in *Geometry*)

Hypothesis 3.4 $(H_{0,\,3.4})$: $\mu_1 = \mu_2$ (There is no difference between the population means of the United States and any of the other compared countries in *Data & Chance)*

Hypothesis 3.5 $(H_{0,\,3.5})$: $\mu_1 = \mu_2$ (There is no difference between the population means of the United States and any of the other compared countries in *Knowing)*

Hypothesis 3.6 $(H_{0,\,3.6})$: $\mu_1 = \mu_2$ (There is no difference between the population means of the United States and any of the other compared countries in *Applying)*

Hypothesis 3.7 $(H_{0,\,3.7})$: $\mu_1 = \mu_2$ (There is no difference between the population means of the United States and any of the other compared countries in *Reasoning)*

The researcher employed the independent samples *t*-test method for testing the above hypotheses and to identify statistically significant differences between U.S, students and those from each of the countries of Bulgaria, Chinese Taipei, Korea, Singapore, and Thailand in the four content domains and the three cognitive domains.

Conducting a t-test. Using the test statistic (t) with value $(\bar{x}_1 - \bar{x}_2)/(S_{\bar{x}_1-\bar{x}_2})$, the confidence interval (CI) to test H_0: $\mu_1 = \mu_2$ is constructed the difference μ_1 - μ_2 for each dependent variable as: $\mathrm{CI} = \overline{X} \pm (t_{\alpha/2})(S_{\bar{x}_1-\bar{x}_2})$, where, $\bar{x} = \bar{x}_1 - \bar{x}_2$ and $(S_{\bar{x}_1-\bar{x}_2})$, the standard error of the differences of the means is given by $(S_{\bar{x}_1-\bar{x}_2}) = \sqrt{\{(s_1^2/n_1) + (s_2^2/n_2)\}}$ and $t_{\alpha/2}$ is the *t*-critical value associated with the pre-established level of significance, α and the degrees of freedom *(df)*, whose value is given by $\{(A+B)^2\}/\{A^2/(n_1-1) + B^2/(n_2-1)\}$, where $A = s_1^2/n_1$ and $B = s_2^2/n_2$ .The notation $t_{\alpha/2}$ indicates that α is "split in half" in the two tails of the *t*-distribution. We reject the null hypotheses (H_0: $\mu_1 = \mu_2$) when the confidence interval does not contain the number zero. The computed values of $(S_{\bar{x}_1-\bar{x}_2})$ in respect of each pair of comparison variables (US and another from five other countries) for each of the seven domains are shown in Table 34. The findings along with SPSS outputs of confidence intervals are enumerated in Table 26, Table 27, Table 28, Table 29, Table 30, Table 31, and Table 32.

Table 26

Multiple Comparisons of Average Scores in NUMBER

			95% CI for ΔM	
Pairs of Countries	*ΔM*	SE*ΔM*	*LL*	*UL*
USA- Bulgaria	52	0.08053	51.84	52.16
USA- Chinese Taipei	-67	0.07313	-67.14	-66.86
USA- Rep. of Korea	-73	0.04844	-73.09	-72.90
USA- Singapore	-87	0.06043	-87.12	-86.88
USA- Thailand	66	0.07243	65.86	66.14

Note. CI = confidence interval, *LL* = lower limit, *UL* = upper limit.

Table 27

Multiple Comparisons of Average Scores in ALGEBRA

			95% CI for ΔM	
Pairs of Countries	*ΔM*	SE*ΔM*	*LL*	*UL*
USA- Bulgaria	25	0.08637	21.83	25.17
USA- Chinese Taipei	-116	0.09053	-116.18	-115.82
USA- Rep. of Korea	-95	0.05578	-95.11	-94.89
USA- Singapore	-78	0.06297	-78.12	-77.88
USA- Thailand	68	0.07488	67.85	68.15

Note. CI = confidence interval, *LL* = lower limit, *UL* = upper limit.

Table 28

Multiple Comparisons of Average Scores in GEOMETRY

			95% CI for ΔM	
Pairs of Countries	*ΔM*	SE*ΔM*	*LL*	*UL*
USA- Bulgaria	12	0.08407	11.84	12.16
USA- Chinese Taipei	-112	0.07796	-112.15	-111.85
USA- Rep. of Korea	-107	0.04577	-107.09	-106.91
USA- Singapore	-98	0.05797	-98.11	-97.89
USA- Thailand	38	0.0777	37.85	38.15

Note. CI = confidence interval, *LL* = lower limit, *UL* = upper limit.

Table 29

Multiple Comparisons of Average Scores in DATA & CHANCE

			95% CI for ΔM	
Pairs of Countries	*ΔM*	SE*ΔM*	*LL*	*UL*
USA- Bulgaria	91	0.08099	90.84	91.16
USA- Chinese Taipei	-35	0.06531	-35.13	-34.87
USA- Rep. of Korea	-43	0.04479	-43.09	-42.91
USA- Singapore	-43	0.06611	-43.13	-42.87
USA- Thailand	51	0.06457	50.87	51.13

Note. CI = confidence interval, *LL* = lower limit, *UL* = upper limit.

Table 30

Multiple Comparisons of Average Scores in KNOWING

			95% CI for ΔM	
Pairs of Countries	*ΔM*	SE*ΔM*	*LL*	*UL*
USA- Bulgaria	37	0.08008	36.84	37.16
USA- Chinese Taipei	-80	0.07695	-80.15	-79.85
USA- Rep. of Korea	-82	0.04889	-82.1	-81.9
USA- Singapore	-67	0.05857	-67.11	-66.89
USA- Thailand	78	0.07193	77.86	78.14

Note. CI = confidence interval, *LL* = lower limit, *UL* = upper limit.

Table 31

Multiple Comparisons of Average Scores in APPLYING

			95% CI for ΔM	
Pairs of Countries	*ΔM*	SE*ΔM*	*LL*	*UL*
USA- Bulgaria	45	0.0829	45.16	45.84
USA- Chinese Taipei	-89	0.07416	-89.15	-88.85
USA- Rep. of Korea	-92	0.05467	-92.11	-91.89
USA- Singapore	-90	0.07226	-90.14	-89.86
USA- Thailand	57	0.07243	56.86	57.14

Note. CI = confidence interval, *LL* = lower limit, *UL* = upper limit.

Table 32

Multiple Comparisons of Average Scores in REASONING

			95% CI for ΔM	
Pairs of Countries	*ΔM*	SE*ΔM*	*LL*	*UL*
USA- Bulgaria	50	0.07923	49.84	50.16
USA- Chinese Taipei	-86	0.07025	-86.14	-85.86
USA- Rep. of Korea	-74	0.04504	-74.09	-73.91
USA- Singapore	-74	0.0666	-74.13	-73.87
USA- Thailand	49	0.06602	48.87	49.13

Note. CI = confidence interval, *LL* = lower limit, *UL* = upper limit.

Interpretation. With reference to each of the content domains of Number, Algebra, Geometry, Data & Chance, as well as each of the cognitive domains of Knowing, Applying, and Reasoning, we reject the null hypothesis, viz., H_0: $\mu_1 = \mu_2$ (There is no difference between the population means of the United States and any of the other compared countries). However, in respect of Bulgaria and Thailand, we retain the alternative hypothesis H_a: $\mu_1 - \mu_2 > 0$, because the difference $(\mu_1 - \mu_2)$ can take only positive values. On the other hand, in respect of Chinese Taipei, Korea, and Singapore, we retain the other alternative hypothesis H_b: $\mu_1 - \mu_2 < 0$, because the difference $(\mu_1 - \mu_2)$ can take only negative values.

Findings of independent samples t-test for multiple comparisons. Independent samples *t*-tests have been conducted for the scores of the U.S., paired separately with each of the countries Bulgaria, Chinese Taipei, Republic of Korea, Singapore, and Thailand. The confidence intervals for comparison of the performances of each of the pairs in the four content domains are shown in Table 26, Table 27, Table 28, and Table

29 and for the three cognitive domains in Table 30, Table 31, and Table 32**.** As can be seen from these Tables, in the case of each of the domains, the limits of confidence intervals are either both positive, or both negative, so much so that the value ‘zero’ does not fall within the range of the confidence interval in any of the seven cases.

In particular, denoting the population means of the U.S. and the country of comparison by μ_1 and μ_2 respectively, we observe from the Tables that the value of (μ_1 - μ_2) is always positive for each of the seven domains in the case of Bulgaria and Thailand and always negative in the case of Chinese Taipei, Singapore, and Korea.

In the light of the foregoing findings, we reject each one of the seven hypotheses represented by the following statement: *(H_0):* $\mu_1 = \mu_2$ (There is no difference between the population means of the United States and any of the other compared countries in *Number/Algebra/Geometry/Data & Chance/Applying/Reasoning*).

Therefore, we conclude that the population means of the U.S. scores in TIMSS (2007) in all the seven domains were higher than those of two of the countries compared, viz., Bulgaria and Thailand and lower than those of the remaining three countries compared, viz., Chinese Taipei, Republic of Korea, and Singapore. Thus, in the TIMSS (2007) international eighth grade mathematics students’ assessment, the U.S. performed better than Bulgaria and Thailand in all the seven domains, whereas Chinese Taipei, Republic of Korea, and Singapore outperformed the U.S.

Summary

Using the international student achievement from TIMSS 2007 in respect of eighth mathematics, the researcher made a study to explore the connections between the average scale scores in the content and cognitive domains. The study also involved finding the connections between student achievements in the cognitive domains with the independent variables as factors in student activity area, classroom instruction, home involvement, school resources, and teacher professional development. Finally, the study went on to compare the math achievement of the students of the United States with each of five selected countries, viz., Bulgaria, Chinese Taipei, Republic of Korea, Singapore, and Thailand in all the seven domains.

For the first Research Question, Pearson's product-moment correlation co-efficient (*r*) was used as the measure. The existence of a strong positive correlation at 0.01 level of significance (two-tailed) between dependent variables and correlated variables indicate that the data suggests a correlational tendency that students who scored high in one subject tend to score high in other subjects also, whereas students who scored low in a subject tend to score low in other domains also. Among these correlations, the one between the content subject of *Number* and the cognitive domain subject of *Applying* was found to be most pronounced. This was followed by that between *Number* and *Knowing, Geometry* and *Applying, Knowing* and *Applying,* and *Reasoning* and *Applying.* The correlation between *Algebra* and *Data & Chance* has been found to be the lowest among all the correlations.

For Research Question 2, the data comprised the scores (dependent variables) of all the participating countries in each of the four content domains, while the five chosen independent variables were Homework Time (in Student Activity area), Classroom Instruction (frequency of applying facts, concepts and procedures), Overall Home Involvement (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), School Resources (for math instruction), and Overall Math Teacher Professional Development (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy /instruction, content, and assessment). Analysis for this was made using multiple regression method. Finally, for Research Question 3, the data comprised the average student achievement scores of Bulgaria, Chinese Taipei, Republic of Korea, Singapore, and Thailand in each of the seven subject domains. In this case, in order to compare the United States with each of the other five countries, independent samples *t*-test was used as the measure.

A summary of results from testing the four Hypotheses pertaining to the Research Question 2 are presented in Table 33. Specifically, for all the four Hypotheses (2.1, 2.2, 2.3, and 2.4), a **significant** relationship between each of the criterion variables *Number, Algebra, Geometry, and Data & Chance* and the predictor variable, *School Resources* was found ($.660 < R < .766$, $p < .001$). For the Hypothesis (2.1), a **moderately significant** relationship between the criterion variable *Number* and the predictor variable *Home Involvement* was found ($R = .766$, $p = .063$). Again, for the Hypothesis (2.4) also, a

moderately significant relationship between the criterion variable *Data & Chance* and the predictor variable, *Home Involvement* was found (R =.745, $p < .070$). However, for the other two hypotheses (2.2 and 2.3), **no significant** relationship between the criterion variables *Algebra and Geometry* and the predictor variable *Home Involvement* was found. In respect of all the four hypotheses pertaining to the Research Question 2, **no significant** relationship between the criterion variables and any of the predictor variables *Homework Time (Student Background), Class Instruction, and Overall Math Teacher Professional Development* was found. These details are depicted in Table 33.

Finally, for Research Question 3, the researcher conducted the independent samples *t*-test for finding statistically significant differences between U.S, students and each of the countries Bulgaria, Chinese Taipei, Korea, Singapore, and Thailand in all of the four content domains and the three cognitive domains. The analysis revealed that there are **significant** differences between the population means of the United States and each of the other compared countries in each of the content and cognitive domains. Whereas the United States performed significantly better than Bulgaria and Thailand in all the four content domains, viz., *Number, Algebra, Geometry, and Data & Chance,* as well as in the three cognitive domains of *Knowing, Applying, and Reasoning,* Chinese Taipei, Republic of Korea, and Singapore outperformed the U.S. in all the domains in the eighth grade mathematics assessment of TIMSS (2007).

Chapter Five: Summary, Conclusions and Recommendations

This chapter discusses the purpose of the study and reviews the questions addressed and the procedures adopted, mentions the limitations of the study, and summarizes the findings in respect of each hypothesis. It provides the interpretations of the findings and gives an overall discussion about the results and points out implications for change in educational practice. The chapter concludes with recommendations for further research.

Nations all over the world have recognized that an effective educational system provides a competitive advantage in global economy. Further, it helps in improving the personal development of individuals, in streamlining social coherence and in bridging the economic inequities. As such, regular monitoring of educational performance has assumed intense importance, according to the policy statement of TIMSS (2007).

The data from the report of TIMSS (2007) on the international assessment in mathematics at the eighth grade level consists of the scale scores of the students from the participating countries in the four content subscale domains of *Number, Algebra, Geometry, and Data & Chance* and the three cognitive subscale domains of *Knowing, Applying, and Reasoning,* and also the subscale of overall scores in mathematics. The data also has details of the values of the independent variables under 15 categories—(i) student and family characteristics; (ii) student computer use; (iii) student activities outside of school; (iv) student perception/valuing of mathematics; (v) teacher background characteristics, formal education, and training; (vi) teacher perception of mathematics teaching/learning; (vii) teacher preparation and collaboration; (viii) teacher activities outside of school; (ix) classroom characteristics; (x) classroom instruction; (xi) role of homework; (xii) school characteristics; (xiii) school resources; (xiv) home involvement; and (xv) school climate and safety (TIMSS, 2007).

Review of the Purpose and the Methods

The first research question of this study focuses on finding the connections among the student achievement scores in the seven domains. The researcher has used Pearson's correlation coefficient for data derived from TIMSS (2007) report as the measure to find these connections.In this study, the second research question aims to find if the student achievements in each of the content domains have any connection to any of the 14 independent variables, chosen from five groups of variables, viz., Student Activity, Classroom Instruction, Home Involvement, School Resources, and Math Teacher Professional Development. For the purpose of analyzing these connections, five predictor variables have been compiled: (1) Homework Time, (amount of time spent on math homework in the Student activities area), (2) Classroom instruction (frequency of applying facts, concepts and procedures to solve routine problems), (3) Overall Home Involvement (combining the five independent variables corresponding to the attributes—i) to ensure that their child completes his/her homework, ii) to raise funds for the school, iii) to serve on school, committees, iv) to attend special events, and v) to volunteer for school projects, programs, and trips), (4) School Resources (for mathematics instruction), and (5) Overall Math Teacher P.D. (combining the six variables corresponding to the attributes—i) for integrating I.T. into mathematics, ii) for improving students' critical thinking or problem solving skills, iii) for mathematics curriculum development, iv) for mathematics pedagogy/instruction, v) for mathematics content, and vi) for mathematics assessment). The study employed the multiple regression method to analyze the connections between the dependent variables of eighth grade mathematics scores and the above five independent predictor variables.

The last research question aims to compare the mathematics achievement of the United States in all the seven domains with those of five selected countries -- Bulgaria, Chinese Taipei, Republic of Korea, Singapore, and Thailand. In order to make these comparisons, the researcher used independent samples *t*-test.the null hypotheses that the performance of the United States in the TIMSS (2007) was not different from that of any of the countries compared in any of the seven domains.

Limitations of the Study

Although it is a common practice to use secondary data for research due to the ready availability of samples, for which, collection of data and preparation of statistical packages have already been done, secondary data are not the outcome of an experimental study. Non-sampling errors in collection of data and sampling errors in the form of standard errors are likely to lead to incorrect interpretations (Aud & Hannes, 2009b). The absence of control and treatment groups impairs the amenability of the data to correct interpretation of the connections among the variables and causal relationships. Also, the use of teachers' self-reported instructional practices to construct variables, the non-response items in TIMSS 2007, and the possible bias from regional and community-type differences lessen the authenticity of the data (Rosenberg, et al., 2006). Moreover, some minority students were less likely to participate, although, substitute schools were included, thereby rendering the measurable differences very small. Further, the school entrance requirements and tracking practices in different countries can also impact the population tested. Again, by design, there is a limitation on learning from surveys and multiple choice tests (TIMSS, 2007).

In the analysis of this study, the Pearson's product moment coefficient of correlation (r) has been used. A limitation in using this as a measure is that when $r = 0$, it indicates that there is no linear relationship between two variables. But, there could still be some non-linear relationship. Moreover, there could be a positive linear relationship between some subgroups of the variables, while other subgroups might have a negative linear relationship. However, this singular case of $r = 0$ has not occurred in this analysis.

For multiple regression, there are four assumptions -- linearity, independence, normality, and homoscedasticity i.e., the assumption that the variance of the dependent variable is the same for all the data. However, all these are supposed to be embedded in TIMSS data. Moreover, multicollinearity is a factor that can mutilate the predictive inferences of the variables. However, as the independent variables in this study do not contain similar information, they will not have high inter-correlations and hence, multicollinearity is not likely to be a matter of concern.

Findings

For the first Research Question 1, Pearson's product-moment correlation co-efficient (*r*) was used as the measure. But, correlations among the student scores in different domains cannot be correctly known since the actual student scores are not revealed in TIMSS data. However, to know the correlation trends to a greater degree of approximation, the average student scores in all the domains at different percentile levels i.e., at 10%, 25%, 50%, 75%, and 90%, have also been considered for finding correlations.

The minimum and maximum values of Pearson's product-moment coefficient (*r*) among the scores in all the seven domains at different percentile levels are shown in Table 14.

In respect of Research Question 2, the data consisted of five independent variables, pertaining to 14 factors chosen from five areas; *Student Activities* (Homework Time), *Classroom Instruction* (frequency of applying facts, concepts and procedures), *Overall Home Involvement* (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), *School Resources* (for math instruction), and math teacher's *Overall Professional Development* (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy/instruction, content, and assessment). The dependent variables were *Student Scores* in *Number* for Hypothesis 2.1, *Student Scores* in *Algebra* for

Hypothesis 2.2, *Student Scores* in *Geometry* for Hypothesis 2.3, and *Student Scores* in *Data & Chance* for Hypothesis 2.4.

Multiple Regression analysis revealed a **significant relationship** between each of the criterion variables of *Student Scores* in *Number, Algebra, Geometry, and Data & Chance* and the predictor variable, *School Resources* for all these four Hypotheses, whereas only **a moderately significant relationship** between only two criterion variables *Number* and *Data & Chance* and the predictor variable *Home Involvement* was found. Again, **no significant relationship** was found between any of the criterion variables and any of the other predictor variables *Homework Time (Student Activity), Class Instruction, and Overall Math Teacher Professional Development* was found.

Finally, for Research Question 3, in order to compare the performance of the United States with five other countries, independent samples *t*-test was used as the measure. The data consisted of the average student achievement scores of Bulgaria, Chinese Taipei, Republic of Korea, Singapore, and Thailand in each of the seven subject domains. The findings of the analysis revealed that the population means of the U.S. average scores in each of the seven domains exceeded those of Bulgaria and Thailand and less than those of Chinese Taipei, Republic of Korea, and Singapore

Conclusions

This research study arrives at the following conclusions by interpreting the above findings. With reference to the R.Q. 1, the students who scored high in one domain tend to score high in the other domains also. So also, the students who scored low in a domain tend to score low in the other domains, too. The scores in the content domain of *Number* and those in the cognitive domain of *Applying* were most correlated. The correlations among *Number* and *Knowing*, *Geometry* and *Applying*, *Knowing* and *Applying*, and *Reasoning* and *Applying* were also high in that order. In comparison with others, the correlation between *Algebra* and *Data & Chance* was the lowest.

With reference to RQ 2, *School Resources* (for math instruction) appear to have a strong bearing on student achievement in each of the content domains of *Number*, *Algebra*, *Geometry*, and *Data & Chance*. A moderately strong connection seems to exist between *Homework Time* (the amount of time, the student spends on math homework) and the scores in *Number* and *Data & Chance*, but not with the scores in *Algebra* and *Geometry*. There is no indication of significant connections between student achievement in any of the four cognitive domains and any of the remaining variables, viz., *Classroom Instruction* (frequency of applying facts, concepts and procedures), *Overall Home Involvement*, and math teacher's *Overall Professional Development*. Lastly, with reference to RQ 3, the performance of the U.S. has been surpassed in each of the seven domains by Chinese Taipei, Korea, and Singapore, whereas, the U.S. performed better than Bulgaria and Thailand in all the domains.

Implications and Recommendations

The findings of the study implicitly provide mathematics teachers, educators, and policy makers with some important cues. These are discussed below with reference to each research question separately.

Research question 1. The first finding of the study is that a student who performs well in one domain of mathematics tends to peform well in others also. This implies that teachers would do well to encourage and help the students to reach mastery levels in the arenas in which they fare well in comparison to other arenas. When near-mastery levels are reached in some domains, it can be expected that their self-efficacy gets a boost and has a salutary effect of spilling over to the other domains also. The adage, "Success breeds Success" will possibly be enacted in students' learning efforts. Conversely, a student who lingers at lower stages of learning in some areas is apt to be likewise in other areas also. This implies that the teacher ought to address the situation and take quick remedial steps for bringing the student to higher conceptual and operational levels. Otherwise, its ramifications could have a multiplier effect and percolate to the other areas of mathematics learning. Another implication from the first finding is that teachers should lay more emphasis on student learning in the content domain *Number*, since it is highly correlated to achievement with the cognitive domains of *Knowing* and *Applying*. Next, an emphasis on student learning in the content domain of *Geometry* is also implicitly called for, since it is highly correlated to student achievements in the cognitive domain of *Applying* and thereby, to *Knowing* and *Reasoning*. It is not to be construed in the light of the above implications that the remaining two content domains viz., *Algebra* and *Data & Chance* need less focus. Rather, it should be understood that, as a matter of priority, teachers need to bestow more focus on the domains *Number* and *Geometry*, since they have a rippling effect over other areas of learning.

Research question 2. The second finding gives a clarion call to educational administrators and policy makers. The finding emphatically shows that providing adequate resources in the school for math learning is highly essential to forge student achievement in mathematics in all the content domains. The transition of the educational scenario in several countries from traditional teaching to reformed teaching requires teachers to incorporate innovative practices like problem solving. The student, on the other hand, is geared to play an active role in self-learning and by interacting with peers, instead of being a passive recipient of knowledge. This can be possible only when adequate learning resources are available in the school, even as a study by Ritzhaupt et al. (2010) has indicated that teachers need to be supported with technical support and modern educational games on a large scale for embedding innovative activities in their instruction. Again, as the study of Larson et al. (2012) points out, math learning at the middle school leads students to progress from concrete thinking to abstract thinking and that conceptual learning occurs when students are exposed to challenging tasks and complex topics. Further, teacher preparedness to teach mathematics content with emphasis on students' conceptual learning requires teachers to adopt reform-oriented strategies that involve the use of several resources, according to the study of Smith et al. (2005), who examined the teacher quality as mentioned in NCLB and reform-oriented instruction. The *School Resources* available in different countries for math instruction may be reviewed for standardization to the extent possible.

Another implication of the second finding is that students need to spend more time on doing homework. Teachers will do well to ensure that the students get balanced amount of homework on a regular basis. It will be more productive if the homework assignments are aligned to individual student needs to the extent possible, without being redundant and insipid exercises for some, while being too hard nuts to crack for others.

Guided exercises, which are challenging enough, might keep the students in moderate tension, so that by answering them, students derive satisfaction and enhanced self-confidence. For this purpose, parental help and encouragement in providing a suitable learning environment at home will also be necessary. In this regard, it may be pertinent to note that according to a study by Balli et al., (1997), the ability of parents to provide help was found to be related to student achievement and not merely their involvement. In a later study, Balli (1998) also found that students believed that when they had help from their parents, their performance was better, although not all students enjoyed such help. Further, in a later significant study involving middle grade students of New York City, Minotti (2005) used multivariate analyses of variance and pair-wise comparisons on two groups – one with traditional type of homework and the other based on individualized learning-style homework, derived from the "Dunn and Dunn Learning-Style Model [which] consists of five strands of learning-style elements: environmental, emotional, sociological, physiological, and psychological" (p. 68). The results of the study showed that students who were involved in individualized learning-style based homework showed significantly better performance.

Again, in a recent research, Kaur (2011) examined 8 teachers of mathematics and 115 eighth grade students to find out their perspectives on homework by using field notes and conducting interviews. The findings of the study indicated that homework provided an opportunity to revise the topics and learn from their mistakes and served to enhance the understanding of mathematics concepts and problem-solving skills, and also helped in preparing for test. In an article on the same topic, Kalchman (2011) states that whereas on one hand, homework helps the students in preparing for tests by toning up their abilities in performing mathematical tasks and also in improving their self-confidence, on the other hand, it helps teachers also in planning the curriculum transaction. In the same vein, a study of the cultural aspect of teaching geometry in China by Fang (2010) revealed that errors serve as resources for not only improving student understanding, but also, by providing feedback, help teachers in helping student learning. In an interesting article on a similar issue regarding the role of homework, Coates and Mayfield (2010) state that parental involvement in student homework can help the students in their classrooms also. However, they suggested that parents need to refrain from overacting their involvement, so that their children develop independence, confidence, and own the consequences of their choices.

Although according to the findings of this study, factors related to classroom instruction, professional development of math teacher, and home involvement have no significant relation to student achievement in the content domains, I think, that the implication is only that the earlier mentioned factors viz., ‘school resources’ and ‘homework time’ need to be prioritized even more than these, since they seem to have more bearing on student achievement. With respect to the variable Classroom Instruction, since only one aspect was considered for this study, viz., the measure concerning the frequency of applying facts, concepts and procedures, it might not have been adequate to show significant relationship with student achievement. As regards the variable factor *Homework Time* also, the insignificant relation with student achievement could be due to the fact that students engaged themselves more on doing routine problems, rather than individually customized and technology-driven practice problems. Again, the finding that math teacher professional development is insignificantly related to student achievement throws many questions, such as, whether the professional development of math teacher has been effective in enhancing the quality of instruction, the teacher has been able to translate the gains of professional development courses in actual instructional practice, and whether there is adequate support from the institution for supporting the teachers' ideas for change in instruction, based on training in professional development courses. An in depth experimental study for gauzing the effectiveness of the different facets of math professional development courses in improving teacher instruction and student achievement might be necessary.

Research question 3. The third finding points out that there is a clear difference in the performances of the U.S. and the other five compared countries. Whereas Chinese Taipei, Korea, and Singapore performed better in all the subject domains, the achievement of the U.S. students was better than Bulgaria and Thailand in all the domains. A closer look at the various factors concerning the learning situations of these countries would probably lend more insights about their connections with the differences in student performances; this is however, beyond the scope of this dissertation study. At any rate, this study is a pointer towards taking up such studies that could further help the policy makers and administrators to lay priorities in shifting their focus on the variables that govern the learning situations in U.S. schools.

Future Research

This study generates some ideas and questions for further research. First, as an extension of this study, the variations in all the variables in student, classroom, and school backgrounds among the countries considered for comparison in Research Question 3 of this dissertation study, i.e., U.S., Bulgaria, Chinese Taipei, Korea, Singapore, and Thailand could be studied and their connections to student achievements may be probed in order to find out some tangible reasons for the differences in student performances. For this purpose, the connections of all the variables, which have not been already considered for Research Question 2 of this study, may also be examined for finding significant relationships with student performance in all the four content domains, in order to arrive at some clearer ideas to explain the differences in student performances in the compared countries.

Second, the three research questions of this study, with the addition of the above extension, may be applied to the data of TIMSS (2011) report, which has been published recently, to see if these results persist for the new data set also. This may then lead to some clarity on the reasons for the performance of the U.S. and provide some cues for teachers, educational administrators, and policy makers for inducting some changes in the current educational practice for better performance by the students of the U.S.

Third, a study could be made considering the variables of gender, race, and socio-economic status. This will help to understand the details of the differences in student performance and further help the policy makers to plug in the gaps in the present student learning conditions that are found to be related to student self-efficacy and achievement, by making equitable provisions of enhanced facilities. Last, a further study may be undertaken on the same lines as above, but with the data obtained from national educational reports, to tally them with the results obtained from TIMSS reports.

Appendices

Appendix 1a

Average mathematics scores of fourth- and eighth-grade students, by country: 2007

Country	Average Score	Country	Average Score
Chinese Taipei	598	Ukraine	462
Korea, Rep. of	597	Romania	461
Singapore	593	Bosnia and Herzegovina	456
Hong Kong SAR	572	Lebanon	449
Japan	570	Thailand	441
Hungary	517	Turkey	432
England4	513	Jordan	427
Russian Federation	512	Tunisia	420
United States	508	Georgia	410
Lithuania	506	Iran, Islamic Rep. of	403
Czech Republic	504	Bahrain	398
Slovenia	501	Indonesia	397
Armenia	499	Syrian Arab Republic	395
Australia	496	Egypt	391
Sweden	491	Algeria	387
Malta	488	Colombia	380
Scotland4	487	Oman	372
Serbia	486	Palestinian Nat'l Auth.	367
Italy	480	Botswana	364
Malaysia	474	Kuwait	354
Norway	469	El Salvador	340
Cyprus	465	Saudi Arabia	329
Bulgaria	464	Ghana	309
Israel	463	Qatar	307

Note. Adapted from Exhibit 3 (p. 7), Gonzales, P. (2009). *Highlights from TIMSS 2007: Mathematics and Science Achievement of U.S. Fourth-and Eighth-Grade Students in an International Context:* Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Appendix 1b

Trends in average mathematics scores of fourth- and eighth-grade students, by country: 1995 to 2007

Country	Average Score 1995	2007	Difference 2007-1995
Colombia	332	380	47*
Lithuania[3]	472	506	34*
Korea, Rep. of	581	597	17*
United States[4,5]	492	508	16*
England[4]	498	513	16*
Slovenia	494	501	7*
Hong Kong SAR[2,4]	569	572	4
Cyprus	468	465	-2
Scotland[4]	493	487	-6
Hungary	527	517	-10*
Japan	581	570	-11*
Russian Federation	524	512	-12
Romania	474	461	-12*
Australia	509	496	-13*
Iran, Islamic Rep. of	418	403	-15*
Singapore	609	593	-16*
Norway	498	469	-29*
Czech Republic	546	504	-42*
Sweden	540	491	-48*
Bulgaria	527	464	-63*

Note. *$p < .05$. Within-country difference between 1995 and 2007 average scores is significant. 1. Difference calculated by subtracting 1995 from 2007 estimate using unrounded numbers. 2. Hong Kong is a Special Administrative Region (SAR) of the People's Republic of China. 3. In 2007, National Target Population did not include all of the International Target Population defined by the Trends in International Mathematics and Science Study (TIMSS). 4. In 2007, met guidelines for sample participation rates only after substitute schools were included. 5. In 2007, National Defined Population covered 90 percent to 95 percent of National Target Population. 6. In 2007, nearly satisfied guidelines for sample participation rates only after substitute schools were included. Adapted from Exhibit 3 (p. 7), Gonzales, P. (2009). *Highlights from TIMSS 2007: Mathematics and Science Achievement of U.S. Fourth-and Eighth-Grade Students in an International Context:* Washington, D.C.: U.S. Department of Education, National Center for Education Statistics.

Appendix 2

Distribution of New and Trend Items in the TIMSS 2007 for Mathematics 8th Grade Item Format

Item Format	Number of Items				
	New Items	Trend Items	Total (New+Trend)	Total Score Points	Percentage of Score Points
	Mathematics Items				
Multiple Choice	61	56	117	117	49%
Constructed Response	59	39	98	121	51%
Total Mathematics Items	**120**	**95**	**215**	**238**	

Note. Adapted from Exhibit 2.9 (p. 28), TIMSS 2007 *Technical Report:* Olson, J.F., Martin, M.O., & Mullis, I.V.S. (Eds.). (2008). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Appendix 3

TIMSS Generalized Scoring Guidelines for Mathematics Constructed-response Items

Type of Item	Score Points	Clarifications
1-point Item	1 Point	A 1-point response is correct. The response indicates that the student has completed the task correctly.
	0 Points	A 0-point response is completely incorrect, irrelevant, or incoherent.
2-point Items	2 Points	A 2-point response is complete and correct. The response demonstrates a thorough understanding of the mathematical concepts and/or procedures embodied in the task. • Indicates that the student has completed the task, showing mathematically sound procedures • Contains clear, complete explanations and/or adequate work when required
	1 Point	A 1-point response is only partially correct. The response demonstrates only a partial understanding of the mathematical concepts and/or procedures embodied in the task. • Addresses some elements of the task correctly but may be incomplete or contain some procedural or conceptual flaws • May contain a correct solution with incorrect, unrelated, or no work and/or explanation when required • May contain an incorrect solution but applies a mathematically appropriate process
	0 Points	A 0-point response is completely incorrect, irrelevant, or incoherent.

Note. Adapted from Exhibit 2.10 (p. 30), TIMSS 2007 *Technical Report:* Olson, J.F., Martin, M.O., & Mullis, I.V.S. (Eds.). (2008). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Appendix 4

General Design of the TIMSS Matrix-sampling Blocks

Mathematics Blocks	Source of Items
M01	Block M05 from TIMSS 2003
M02	New items for TIMSS 2007
M03	Block M06 from TIMSS 2003
M04	New items for TIMSS 2007
M05	Block M07 from TIMSS 2003
M06	New items for TIMSS 2007
M07	Block M08 from TIMSS 2003
M08	New items for TIMSS 2007
M09	Block M11 from TIMSS 2003
M10	New items for TIMSS 2007
M11	Block M12 from TIMSS 2003
M12	New items for TIMSS 2007
M13	Block M14 from TIMSS 2003
M14	New items for TIMSS 2007

Note. Adapted from Exhibit 2.12 (p. 33), TIMSS 2007 *Technical Report:* Olson, J.F., Martin, M.O., & Mullis, I.V.S. (Eds.). (2008). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Appendix 5

Booklet Design for TIMSS 2007 – Fourth Grade and Eighth Grade

Student Achievement Booklet	Assessment Blocks			
	Part 1		Part 2	
Booklet 1	M01	M02	S01	S02
Booklet 2	S02	S03	M02	M03
Booklet 3	M03	M04	S03	S04
Booklet 4	S04	S05	M04	M05
Booklet 5	M05	M06	S05	S06
Booklet 6	S06	S07	M06	M07
Booklet 7	M07	M08	S07	S08
Booklet 8	S08	S09	M08	M09
Booklet 9	M09	M10	S09	S10
Booklet 10	S10	S11	M10	M11
Booklet 11	M11	M12	S11	S12
Booklet 12	S12	S13	M12	M13
Booklet 13	M13	M14	S13	S14
Booklet 14	S14	S01	M14	M01

Note. Adapted from Exhibit 2.13 (p. 34), TIMSS 2007 *Technical Report:* Olson, J.F., Martin, M.O., & Mullis, I.V.S. (Eds.). (2008). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Appendix 6

Number of Score Points in TIMSS 2007 in Each Booklet for Mathematics Content Domain

Content Domain	Booklet													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	Mathematics													
Number	13	14	17	11	8	9	9	8	7	5	6	8	12	17
Algebra	6	5	5	9	10	9	11	11	9	13	13	15	15	7
Geometry	9	7	7	6	7	7	7	9	9	7	7	5	4	8
Data and Chance	6	7	6	10	11	9	6	5	7	8	7	4	4	4
Total in Mathematics	**34**	**33**	**35**	**36**	**36**	**34**	**33**	**33**	**32**	**33**	**33**	**32**	**35**	**36**

Note. Adapted from Exhibit 2.25 (p. 42), TIMSS 2007 *Technical Report:* Olson, J.F., Martin, M.O., & Mullis, I.V.S. (Eds.). (2008). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Appendix 7

Content of the TIMSS 2007 Mathematics Teacher Questionnaire at the 8th Grade

Item Content	Description
1. Age	Teacher's age
2. Gender	Teacher's gender
3. Teaching experience	Number of years as a teacher
4. Formal education	Highest level of formal education completed by the teacher
5. Major area of study	Teacher's major area of study during postsecondary education
6. Teaching license	Whether or not the teacher has a teaching license or certificate
7. Preparation to teach	How well prepared the teacher feels to teach the topics included in the TIMSS mathematics test
8. Teacher interactions	Frequency of various types of interactions the teacher has with colleagues
9. Professional development	Whether the teacher participated in various types of professional development activities
10. School safety	Teacher's perception about school safety
11. School facility	Teacher's perception about the adequacy of the school facility
12. School climate	Teacher's perception of job satisfaction, parental support and involvement, expectations for student achievement, students' desire to do well in school and their regard for school property
13. Class size	Number of students in the sampled class
14. Time spent teaching subject	Minutes per week the teacher teaches mathematics to the sampled class
15. Textbook	Whether or not a textbook(s) is used as a primary or supplementary resource
16. Student learning activities	Percentage of time students spend doing various learning activities in a typical week
17. Content-related activities	Frequency with which the teacher asks students to do various content-related activities in mathematics
18. Factors limiting teaching	Extent to which the teacher perceives various student and resource factors to limit teaching
19. Emphasis on content areas	Percentage of time spent on mathematics content areas over the course of the year
20. Topic coverage	When students were taught the TIMSS mathematics topics, by content area
21. Calculator use policy	Whether or not the students are permitted to use calculators during mathematics lessons
22. Calculator use	Frequency with which the students use calculators for various learning activities in mathematics
23. Computer availability	Whether or not the students have access to computers during mathematics lessons and whether or not computers have access to the Internet
24. Computer use	Frequency with which the students use computers for various learning activities in mathematics
25. Homework	Whether or not the teacher assigns mathematics homework
26. Frequency of homework	How often the teacher assigns mathematics homework
27. Amount of homework	Number of minutes it would take an average student to complete a mathematics homework assignment
28. Type of homework	Frequency with which the teacher assigns various types of homework
29. Use of homework	How often the teacher uses mathematics homework for various purposes
30. Sources to monitor progress	Emphasis teacher places on sources to monitor students' progress in mathematics
31. Assessment	Frequency with which the teacher gives a mathematics test or examination
32. Question format	Item formats the teacher typically uses in mathematics tests or examinations
33. Type of questions	Types of questions the teacher uses in mathematics tests or examinations

Note. Adapted from Exhibit 3.4 (p. 58), TIMSS 2007 *Technical Report:* Olson, J.F., Martin, M.O., & Mullis, I.V.S. (Eds.). (2008). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Appendix 8

Content of the TIMSS 2007 Student Questionnaire at the 8th Grade

Item Number	Item Content	Description
1	Age	Month and year of student's birth
2	Gender	Student's gender
3	Language	Student's frequency of use of the language of the test at home
4	Books in the home	Number of books in the student's home
5	Home possessions	Educational resources and general possessions in the student's home
6	Parents' education	Highest level of education completed by mother and father
7	Educational expectations	Level of education the student expects to complete
8	Liking mathematics	How much the student likes and feels competent at mathematics
9	Valuing mathematics	Importance and value the student attributes to mathematics
10	Learning activities in mathematics	Frequency with which student does various learning activities in mathematics lessons
11	Computers	Whether or not student uses a computer, where student uses it, and frequency with which student uses a computer in mathematics and science
12	School climate	Student's affinity for school, perception of other students' motivation in school, and teachers' expectations
13	Safety in school	Whether or not the student experienced being the object of problematic behaviors by other students
14	Out-of-school activities	Frequency with which student does various nonacademic activities and homework outside of school
15	Mathematics homework	Frequency and amount of mathematics homework
16	Parents born in country	Whether or not mother and father were born in country
17	Student born in country	Whether or not student was born in country and if not, the age at which the student emigrated

Note. Adapted from Exhibit 3.6 (p. 61), TIMSS 2007 *Technical Report:* Olson, J.F., Martin, M.O., & Mullis, I.V.S. (Eds.). (2008). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Appendix 9 (a)

Figure 4 shows the SPSS output of multiple regression of Overall Score in *Number* on Homework Time (in Student Activity area), Class Instruction (frequency of applying facts, concepts and procedures), Overall Home Involvement (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), School Resources (for math instruction), and Overall Math Teacher Professional Development (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy /instruction, content, and assessment).

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.766[a]	.587	.538	48.840

a. Predictors: (Constant), Average Math Teacher Prof Dev, Class Instr BT4MASAC Facts Concepts, St Bckg BSDMTMH Time spent on HW, Sch Resources BCDSRMI for math instruction, Average Home Involvement

$R^2 = 0.587$

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	142458.846	5	28491.769	11.944	.000[a]
	Residual	100185.070	42	2385.359		
	Total	242643.917	47			

a. Predictors: (Constant), Average Math Teacher Prof Dev, Class Instr BT4MASAC Facts Concepts, St Bckg BSDMTMH Time spent on HW, Sch Resources BCDSRMI for math instruction, Average Home Involvement
b. Dependent Variable: Overall Score in NUMBER

$F(5, 42) = 11.944, p < .001$

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	64.643	132.070		.489	.627			
	St Bckg BSDMTMH Time spent on HW	-8.133	19.172	-.046	-.424	.674	-.051	-.065	-.042
	Class Instr BT4MASAC Facts Concepts	-39.910	29.767	-.140	-1.341	.187	-.270	-.203	-.133
	Average Home Involvement	175.884	92.205	.236	1.908	.063	.261	.282	.189
	Sch Resources BCDSRMI for math instruction	172.165	28.960	.647	5.945	.000	.710	.676	.589
	Average Math Teacher Prof Dev	34.344	57.714	.073	.595	.555	.367	.091	.059

a. Dependent Variable: Overall Score in NUMBER

b_1 = -8.133 -- statistically *not* significant ($p = .674$); b_2 = -39.910 -- statistically *not* significant ($p = .187$);
b_3 = 175.884 -- statistically not significant ($p = .063$); b_4 = 172.165 -- statistically significant ($p < .001$);
b_5 = 34.344 -- statistically *not* significant ($p = .555$); $a = 64.643$

$r_{Y(1.2345)} = -.042$; $r_{Y(2.1345)} = -.133$; $r_{Y(3.1245)} = .189$; $r_{Y(4.1235)} = .589$; $r_{Y(5.1234)} = .059$

***Overall score in NUMBER* = -8.133 (*Student Background-Time spent on Home work*) – 39.910 (*Class Instruction- Facts and concepts* + 175.884 (*Average Home Involvement*) + 172.165 (*School Resources-for math instruction*) + 34.344 (*Average Math Teacher Prof Dev*) + 64.643**

Figure 4. SPSS Output of Multiple Regression of Number on Predictor Variables

Appendix 9 (b)

Figure 5 shows the SPSS output of multiple regression of Overall Score in *Algebra* on Homework Time (in Student Activity area), Class Instruction (frequency of applying facts, concepts and procedures), Overall Home Involvement (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), School Resources (for math instruction), and Overall Math Teacher Professional Development (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy /instruction, content, and assessment).

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.660[a]	.435	.368	54.308

a. Predictors: (Constant), Average Math Teacher Prof Dev, Class Instr BT4MASAC Facts Concepts, St Bckg BSDMTMH Time spent on HW, Sch Resources BCDSRMI for math instruction, Average Home Involvement

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	95496.013	5	19099.203	6.476	.000[a]
	Residual	123873.987	42	2949.381		
	Total	219370.000	47			

a. Predictors: (Constant), Average Math Teacher Prof Dev, Class Instr BT4MASAC Facts Concepts, St Bckg BSDMTMH Time spent on HW, Sch Resources BCDSRMI for math instruction, Average Home Involvement
b. Dependent Variable: Overall Score in Algebra

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	98.776	146.857		.673	.505			
	St Bckg BSDMTMH Time spent on HW	15.085	21.318	.090	.708	.483	.081	.109	.082
	Class Instr BT4MASAC Facts Concepts	-27.832	33.100	-.102	-.841	.405	-.184	-.129	-.097
	Average Home Involvement	107.395	102.529	.151	1.047	.301	.261	.160	.121
	Sch Resources BCDSRMI for math instruction	134.747	32.203	.532	4.184	.000	.592	.542	.485
	Average Math Teacher Prof Dev	64.955	64.175	.145	1.012	.317	.378	.154	.117

a. Dependent Variable: Overall Score in Algebra

Figure 5. SPSS Output of Multiple Regression of Algebra on Predictor Variables

Appendix 9 (c)

Figure 6 shows the SPSS output of multiple regression of Overall Score in *Geometry* on Homework Time (in Student Activity area), Class Instruction (frequency of applying facts, concepts and procedures), Overall Home Involvement (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), School Resources (for math instruction), and Overall Math Teacher Professional Development (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy /instruction, content, and assessment)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.703[a]	.494	.433	53.867

a. Predictors: (Constant), Average Math Teacher Prof Dev, Class Instr BT4MASAC Facts Concepts, St Bckg BSDMTMH Time spent on HW, Sch Resources BCDSRMI for math instruction, Average Home Involvement

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	118760.351	5	23752.070	8.186	.000[a]
	Residual	121868.461	42	2901.630		
	Total	240628.813	47			

a. Predictors: (Constant), Average Math Teacher Prof Dev, Class Instr BT4MASAC Facts Concepts, St Bckg BSDMTMH Time spent on HW, Sch Resources BCDSRMI for math instruction, Average Home Involvement
b. Dependent Variable: Overall Score in Geometry

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	142.113	145.663		.976	.335			
	St Bckg BSDMTMH Time spent on HW	-8.905	21.145	-.051	-.421	.676	-.094	-.065	-.046
	Class Instr BT4MASAC Facts Concepts	-27.551	32.831	-.097	-.839	.406	-.217	-.128	-.092
	Average Home Involvement	39.005	101.695	.052	.384	.703	.139	.059	.042
	Sch Resources BCDSRMI for math instruction	153.679	31.941	.580	4.811	.000	.666	.596	.528
	Average Math Teacher Prof Dev	91.764	63.654	.196	1.442	.157	.381	.217	.158

a. Dependent Variable: Overall Score in Geometry

Figure 6. SPSS Output of Multiple Regression of Geometry on Predictor Variables

Appendix 9 (d)

Figure 7 shows the SPSS output of multiple regression of Overall Score in *Data & Chance* on Homework Time (in Student Activity area), Class Instruction (frequency of applying facts, concepts and procedures), Overall Home Involvement (to ensure that student completes homework, to raise funds for school, to serve on school committees, to attend special events, and to volunteer for school projects and programs), School Resources (for math instruction), and Overall Math Teacher Professional Development (in integrating information technology into mathematics, improving students' critical thinking or problem solving skills, mathematics curriculum, pedagogy /instruction, content, and assessment)

Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.745[a]	.556	.503	50.234

a. Predictors: (Constant), Average Math Teacher Prof Dev, Class Instr BT4MASAC Facts Concepts, St Bckg BSDMTMH Time spent on HW, Sch Resources BCDSRMI for math instruction, Average Home Involvement

ANOVA[b]

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	132505.712	5	26501.142	10.502	.000[a]
	Residual	105984.101	42	2523.431		
	Total	238489.813	47			

a. Predictors: (Constant), Average Math Teacher Prof Dev, Class Instr BT4MASAC Facts Concepts, St Bckg BSDMTMH Time spent on HW, Sch Resources BCDSRMI for math instruction, Average Home Involvement
b. Dependent Variable: Overall Score in Data & Chance

Coefficients[a]

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.	Correlations		
		B	Std. Error	Beta			Zero-order	Partial	Part
1	(Constant)	83.017	135.839		.611	.544			
	St Bckg BSDMTMH Time spent on HW	-18.612	19.719	-.106	-.944	.351	-.108	-.144	-.097
	Class Instr BT4MASAC Facts Concepts	-34.012	30.617	-.120	-1.111	.273	-.261	-.169	-.114
	Average Home Involvement	176.235	94.836	.238	1.858	.070	.229	.276	.191
	Sch Resources BCDSRMI for math instruction	168.385	29.787	.638	5.653	.000	.695	.657	.581
	Average Math Teacher Prof Dev	20.952	59.361	.045	.353	.726	.334	.054	.036

a. Dependent Variable: Overall Score in Data & Chance

Figure 7. SPSS Output of Multiple Regression of Geometry on Predictor Variables

Appendix 10

Table 33

Statistical Significant Differences for Multiple Regression Analysis

Hypothesis	Criterion Variable	Predictor Variable		Unstandardized Coefficients	Sig.
Hypothesis 2.1	Number	1.	Homework Time	-8.133	0.674
		2.	Class Instruction	-39.910	0.187
		3.	Overall Home Involvement	175.884	0.063
		4.	School Resources	172.165	0.000
		5.	Overall Math Teacher Professional Development	34.344	0.555
Hypothesis 2.2	Algebra	1.	Homework Time	15.085	0.483
		2.	Class Instruction	-27.832	0.405
		3.	Overall Home Involvement	107.395	0.301
		4.	School Resources	134.747	0.000
		5.	Overall Math Teacher Professional Development	64.955	0.317
Hypothesis 2.3	Geometry	1.	Homework Time	-8.905	0.676
		2.	Class Instruction	-27.551	0.406
		3.	Overall Home Involvement	39.005	0.703
		4.	School Resources	153.679	0.000
		5.	Overall Math Teacher Professional Development	91.764	0.157
Hypothesis 2.4	Data & Chance	1.	Homework Time	-18.612	0.351
		2.	Class Instruction	-34.012	0.273
		3.	Overall Home Involvement	176.235	0.070
		4.	School Resources	168.385	0.000
		5.	Overall Math Teacher Professional Development	20.952	0.726

Note. Results indicate that four hypotheses, comprising twenty sub-variant hypotheses reached statistical significant differences at 0.05 % level and two sub-variant hypotheses at 0.10 % level.

Appendix 11

Table 34

Standard Errors of Mean Differences and Degrees of Freedom

DOMAIN	COMPARISON PAIRS	Std. Error of Difference of Means $(S_{\bar{x}_1-\bar{x}_2})$	Degrees of Freedom (*df*)
NUMBER	USA- Bulgaria	0.080527	5497
	USA- Chinese Taipei	0.073130	5921
	USA- Rep. of Korea	0.048443	9701
	USA- Singapore	0.060430	7961
	USA- Thailand	0.072425	7903
ALGEBRA	USA- Bulgaria	0.086371	5273
	USA- Chinese Taipei	0.090528	5178
	USA- Rep. of Korea	0.055775	8098
	USA- Singapore	0.062968	7634
	USA- Thailand	0.074884	7716
GEOMETRY	USA- Bulgaria	0.084069	5324
	USA- Chinese Taipei	0.077956	5609
	USA- Rep. of Korea	0.045770	10007
	USA- Singapore	0.057973	8141
	USA- Thailand	0.077702	7470
DATA & CHANCE	USA- Bulgaria	0.080989	5497
	USA- Chinese Taipei	0.065314	6583
	USA- Rep. of Korea	0.044790	10900
	USA- Singapore	0.066106	7348
	USA- Thailand	0.064566	8753
KNOWING	USA- Bulgaria	0.080080	5497
	USA- Chinese Taipei	0.076950	5680
	USA- Rep. of Korea	0.048892	9401
	USA- Singapore	0.058566	8141
	USA- Thailand	0.071928	7903
APPLYING	USA- Bulgaria	0.082902	5435
	USA- Chinese Taipei	0.074164	5921
	USA- Rep. of Korea	0.054673	8575
	USA- Singapore	0.072261	7792
	USA- Thailand	0.072425	8006
REASONING	USA- Bulgaria	0.079229	5497
	USA- Chinese Taipei	0.070253	6013
	USA- Rep. of Korea	0.045038	10007
	USA- Singapore	0.066603	7098
	USA- Thailand	0.066015	8347

Appendix 12

Table 35

Average Scale Scores of All Countries at the 10th Percentile

Country	*Number*	*Algebra*	*Geometry*	*Data&Chance*	*Knowing*	*Applying*	*Reasoing*
Algeria	307	213	337	258	276	331	251
Armenia	378	426	364	253	395	371	359
Australia	393	365	384	396	389	396	394
Bahrain	268	269	278	298	272	297	292
Bosnia and Herzegovina	347	367	335	298	363	330	331
Botswana	239	300	185	254	268	243	250
Bulgaria	317	321	318	255	322	310	293
Chinese Taipei	442	441	445	436	438	449	450
Colombia	242	279	259	271	258	279	299
Cyprus	341	354	319	318	350	340	323
Czech Republic	412	386	392	388	406	404	395
Egypt	255	274	257	233	251	256	258
El Salvador	251	198	197	229	230	251	241
England	393	378	400	401	401	397	401
Georgia	292	258	252	188	282	260	235
Ghana	163	227	124	163	180	174	181
Hong Kong SAR	441	430	441	425	442	435	424
Hungary	400	390	384	397	404	400	393
Indonesia	274	279	254	252	266	282	273
Iran, Islamic Rep. of	267	294	289	285	288	285	306
Israel	338	338	287	277	336	309	316
Italy	379	351	381	357	371	384	373
Japan	440	438	478	470	456	456	457
Jordan	270	307	288	263	283	282	297
Korea, Rep. of	466	459	472	480	472	470	472
Kuwait	225	218	270	228	235	253	222
Lebanon	356	369	345	263	358	349	292
Lithuania	398	365	394	405	394	405	368
Malaysia	380	353	349	353	366	371	365
Malta	362	353	360	316	355	357	344
Norway	390	327	350	361	378	382	366

Oman	231	244	241	234	230	236	265
Palestinian Nat'l Auth.	212	233	237	206	217	239	233
Qatar	207	157	155	132	182	179	155
Romania	323	333	318	263	320	327	289
Russian Federation	398	398	398	352	401	400	373
Saudi Arabia	180	219	239	215	191	229	247
Scotland	377	356	376	387	382	378	379
Serbia	361	373	355	306	375	352	341
Singapore	478	454	455	432	464	467	445
Slovenia	403	385	398	399	404	410	386
Sweden	415	354	360	379	398	400	374
Syrian Arab Republic	269	274	287	250	267	288	269
Thailand	318	304	307	321	315	331	334
Tunisia	326	329	336	287	324	331	329
Turkey	290	286	255	285	292	287	292
Ukraine	340	329	345	316	340	344	306
United States	408	404	380	399	416	394	405

Table 36

Pearson's Correlation Coefficient of 10th Percentile Scores (N = 48)

Dependent variable	Correlated variable	*r* – value
Score in Number	Score in Algebra	0.918
	Score in Geometry	0.945
	Score in Data & Chance	0.905
	Score in Knowing	0.988
	Score in Applying	0.980
	Score in Reasoning	0.937
Score in Algebra	Score in Number	0.918
	Score in Geometry	0.892
	Score in Data & Chance	0.854
	Score in Knowing	0.960
	Score in Applying	0.914
	Score in Reasoning	0.929
Score in Geometry	Score in Number	0.945
	Score in Algebra	0.892
	Score in Data & Chance	0.910
	Score in Knowing	0.948
	Score in Applying	0.977
	Score in Reasoning	0.939
Score in Data & Chance	Score in Number	0.905
	Score in Algebra	0.854
	Score in Geometry	0.910
	Score in Knowing	0.916
	Score in Applying	0.945
	Score in Reasoning	0.965
Score in Knowing	Score in Number	0.988
	Score in Algebra	0.960
	Score in Geometry	0.948
	Score in Data & Chance	0.916
	Score in Applying	0.977
	Score in Reasoning	0.955
Score in Applying	Score in Number	0.980
	Score in Algebra	0.914
	Score in Geometry	0.977
	Score in Data & Chance	0.945
	Score in Knowing	0.977
	Score in Reasoning	0.962
Score in Reasoning	Score in Number	0.937
	Score in Algebra	0.929
	Score in Geometry	0.939
	Score in Data & Chance	0.965
	Score in Knowing	0.955
	Score in Applying	0.962

Table 37

Average Scale Scores of All Countries at the 25th Percentile

Country	*Number*	*Algebra*	*Geometry*	*Data&Chance*	*Knowing*	*Applying*	*Reasoning*
Algeria	347	270	377	302	318	367	299
Armenia	437	483	430	332	450	435	424
Australia	445	415	434	457	435	444	444
Bahrain	320	328	338	347	324	345	345
Bosnia and Herzegovina	398	424	391	362	417	383	388
Botswana	295	343	250	306	317	293	307
Bulgaria	389	400	396	347	401	386	375
Chinese Taipei	522	545	534	513	533	533	534
Colombia	300	329	307	327	306	328	350
Cyprus	403	415	387	389	409	402	393
Czech Republic	460	434	442	448	451	450	447
Egypt	314	333	319	291	311	317	317
El Salvador	296	253	253	282	277	294	292
England	452	438	455	480	454	456	460
Georgia	351	334	325	271	347	327	304
Ghana	225	283	191	223	241	231	240
Hong Kong SAR	518	512	521	502	525	515	503
Hungary	461	449	446	462	462	455	452
Indonesia	329	338	317	317	324	336	333
Iran, Islamic Rep. of	321	343	345	338	338	337	357
Israel	402	405	359	370	405	387	389
Italy	428	404	433	420	422	431	427
Japan	500	503	532	531	511	512	516
Jordan	337	373	356	334	350	349	363
Korea, Rep. of	529	535	538	541	542	535	532
Kuwait	279	277	321	281	286	304	281
Lebanon	399	413	398	322	402	393	353
Lithuania	451	425	447	467	450	455	427
Malaysia	434	401	406	406	416	421	412
Malta	436	416	431	408	429	431	413
Norway	437	375	401	430	417	428	417
Oman	291	311	308	300	297	300	326
Palestinian Nat'l Auth.	278	297	304	276	284	300	298
Qatar	263	219	220	197	240	237	220

Romania	388	404	390	341	390	392	365
Russian Federation	451	463	452	415	459	452	435
Saudi Arabia	236	271	291	268	246	277	295
Scotland	429	409	427	447	429	429	434
Serbia	420	437	419	379	434	414	407
Singapore	547	526	526	515	531	537	521
Slovenia	450	436	447	452	450	452	439
Sweden	458	405	414	451	438	446	431
Syrian Arab Republic	323	333	346	305	323	340	325
Thailand	373	360	364	377	366	381	387
Tunisia	369	369	380	337	365	371	372
Turkey	351	351	321	351	352	345	358
Ukraine	398	398	404	384	403	403	375
United States	457	451	427	463	463	444	453

Table 38

Pearson's Correlation Coefficient of 25th Percentile Scores (N = 48)

Dependent variable	Correlated variable	*r* – value
Score in Number	Score in Algebra	0.935
	Score in Geometry	0.956
	Score in Data & Chance	0.934
	Score in Knowing	0.989
	Score in Applying	0.987
	Score in Reasoning	0.957
Score in Algebra	Score in Number	0.935
	Score in Geometry	0.920
	Score in Data & Chance	0.878
	Score in Knowing	0.972
	Score in Applying	0.937
	Score in Reasoning	0.947
Score in Geometry	Score in Number	0.956
	Score in Algebra	0.920
	Score in Data & Chance	0.920
	Score in Knowing	0.960
	Score in Applying	0.983
	Score in Reasoning	0.952
Score in Data & Chance	Score in Number	0.934
	Score in Algebra	0.878
	Score in Geometry	0.920
	Score in Knowing	0.934
	Score in Applying	0.954
	Score in Reasoning	0.972
Score in Knowing	Score in Number	0.989
	Score in Algebra	0.972
	Score in Geometry	0.960
	Score in Data & Chance	0.934
	Score in Applying	0.983
	Score in Reasoning	0.968
Score in Applying	Score in Number	0.987
	Score in Algebra	0.937
	Score in Geometry	0.983
	Score in Data & Chance	0.954
	Score in Knowing	0.983
	Score in Reasoning	0.975
Score in Reasoning	Score in Number	0.957
	Score in Algebra	0.947
	Score in Geometry	0.952
	Score in Data & Chance	0.972
	Score in Knowing	0.968
	Score in Applying	0.975

Table 39

Average Scale Scores of All Countries at the 50th Percentile

Country	*Number*	*Algebra*	*Geometry*	*Data&Chance*	*Knowing*	*Applying*	*Reasoing*
Algeria	394	333	423	350	363	410	352
Armenia	498	540	495	415	505	498	492
Australia	503	473	487	525	488	497	502
Bahrain	380	397	404	400	387	400	405
Bosnia and Herzegovina	453	481	450	429	476	441	449
Botswana	357	391	318	361	370	350	365
Bulgaria	465	487	473	439	481	466	458
Chinese Taipei	598	644	619	587	622	611	615
Colombia	363	383	362	386	360	381	406
Cyprus	469	478	460	461	470	470	464
Czech Republic	513	488	499	511	503	502	504
Egypt	384	404	396	360	385	392	385
El Salvador	346	313	312	338	329	341	347
England	514	499	515	558	511	518	522
Georgia	419	421	406	355	422	403	385
Ghana	292	346	265	288	309	293	304
Hong Kong SAR	587	586	592	572	595	584	578
Hungary	522	511	512	531	523	514	517
Indonesia	392	400	386	385	391	394	398
Iran, Islamic Rep. of	384	402	412	396	395	396	415
Israel	473	478	437	466	477	462	467
Italy	482	463	492	487	476	485	485
Japan	561	569	586	593	571	570	578
Jordan	415	450	433	410	429	424	440
Korea, Rep. of	598	613	606	607	615	606	596
Kuwait	338	343	377	341	343	359	345
Lebanon	451	467	455	388	454	446	424
Lithuania	508	489	508	528	510	512	489
Malaysia	494	455	472	461	472	476	467
Malta	508	483	503	492	500	500	481
Norway	487	425	460	505	460	478	477
Oman	356	387	380	371	369	369	391
Palestinian Nat'l Auth.	355	371	379	349	360	369	372
Qatar	324	290	293	271	305	303	293

Romania	459	484	467	421	470	466	449
Russian Federation	511	530	513	485	523	512	502
Saudi Arabia	300	329	349	323	305	332	347
Scotland	488	471	487	516	483	489	495
Serbia	484	509	489	454	501	481	477
Singapore	614	599	597	597	599	606	599
Slovenia	504	491	500	510	501	501	498
Sweden	507	461	475	528	481	497	496
Syrian Arab Republic	384	397	409	365	385	397	388
Thailand	438	424	432	437	426	440	449
Tunisia	419	417	431	392	412	421	419
Turkey	421	430	398	428	426	416	431
Ukraine	462	468	468	453	470	465	445
United States	514	506	480	534	517	502	506

Table 40

Pearson's Correlation Coefficient of 50th Percentile Scores (N = 48)

Dependent variable	Correlated variable	*r* – value
Score in Number	Score in Algebra	0.942
	Score in Geometry	0.965
	Score in Data & Chance	0.950
	Score in Knowing	0.987
	Score in Applying	0.991
	Score in Reasoning	0.972
Score in Algebra	Score in Number	0.942
	Score in Geometry	0.938
	Score in Data & Chance	0.882
	Score in Knowing	0.980
	Score in Applying	0.950
	Score in Reasoning	0.955
Score in Geometry	Score in Number	0.965
	Score in Algebra	0.938
	Score in Data & Chance	0.922
	Score in Knowing	0.967
	Score in Applying	0.986
	Score in Reasoning	0.962
Score in Data & Chance	Score in Number	0.950
	Score in Algebra	0.882
	Score in Geometry	0.922
	Score in Knowing	0.937
	Score in Applying	0.956
	Score in Reasoning	0.971
Score in Knowing	Score in Number	0.987
	Score in Algebra	0.980
	Score in Geometry	0.967
	Score in Data & Chance	0.937
	Score in Applying	0.985
	Score in Reasoning	0.976
Score in Applying	Score in Number	0.991
	Score in Algebra	0.950
	Score in Geometry	0.986
	Score in Data & Chance	0.956
	Score in Knowing	0.985
	Score in Reasoning	0.983
Score in Reasoning	Score in Number	0.972
	Score in Algebra	0.955
	Score in Geometry	0.962
	Score in Data & Chance	0.971
	Score in Knowing	0.976
	Score in Applying	0.983

Table 41

Average Scale Scores of All Countries at the 75th Percentile

Country	*Number*	*Algebra*	*Geometry*	*Data&Chance*	*Knowing*	*Applying*	*Reasoing*
Algeria	441	394	470	396	410	453	403
Armenia	556	592	557	494	559	555	556
Australia	560	531	542	593	541	551	559
Bahrain	441	467	470	453	452	456	468
Bosnia and Herzegovina	505	533	506	491	530	497	507
Botswana	417	437	385	413	423	405	421
Bulgaria	531	558	542	517	548	533	533
Chinese Taipei	659	725	686	652	687	671	679
Colombia	428	436	415	444	411	434	461
Cyprus	529	534	528	529	524	532	529
Czech Republic	567	540	556	576	554	555	558
Egypt	458	476	475	426	460	467	454
El Salvador	400	373	370	394	382	390	402
England	575	562	577	633	566	577	581
Georgia	482	502	482	435	493	472	464
Ghana	363	410	339	350	378	358	370
Hong Kong SAR	642	648	649	628	651	639	639
Hungary	583	570	577	596	583	572	579
Indonesia	456	462	457	450	458	457	458
Iran, Islamic Rep. of	451	461	480	453	453	458	474
Israel	537	544	507	551	538	530	535
Italy	533	517	548	552	527	534	540
Japan	620	633	638	655	629	627	639
Jordan	491	521	506	483	505	498	509
Korea, Rep. of	659	686	666	666	680	669	656
Kuwait	398	407	436	397	402	414	406
Lebanon	506	522	514	454	512	500	493
Lithuania	565	551	571	588	569	567	550
Malaysia	556	512	542	515	533	533	521
Malta	568	539	564	564	559	559	542
Norway	536	475	516	577	499	525	532
Oman	420	460	450	435	440	434	454
Palestinian Nat'l Auth.	432	444	453	417	436	438	444
Qatar	386	363	364	346	370	368	366

Romania	526	561	539	498	542	533	528
Russian Federation	569	591	571	552	585	569	565
Saudi Arabia	364	387	407	374	366	389	400
Scotland	548	531	545	580	536	546	555
Serbia	542	574	555	526	561	546	541
Singapore	671	664	662	672	661	666	666
Slovenia	558	546	552	567	552	550	555
Sweden	552	515	533	603	523	546	556
Syrian Arab Republic	447	464	473	424	451	457	451
Thailand	508	497	505	499	492	503	514
Tunisia	470	467	482	447	461	470	466
Turkey	499	520	485	509	510	495	512
Ukraine	520	535	526	518	534	525	511
United States	571	563	534	605	572	560	560

Table 42

Pearson's Correlation Coefficient of 75th Percentile Scores (N = 48)

Dependent variable	Correlated variable	*r* – value
Score in Number	Score in Algebra	0.943
	Score in Geometry	0.969
	Score in Data & Chance	0.956
	Score in Knowing	0.982
	Score in Applying	0.992
	Score in Reasoning	0.980
Score in Algebra	Score in Number	0.943
	Score in Geometry	0.945
	Score in Data & Chance	0.875
	Score in Knowing	0.984
	Score in Applying	0.956
	Score in Reasoning	0.959
Score in Geometry	Score in Number	0.969
	Score in Algebra	0.945
	Score in Data & Chance	0.915
	Score in Knowing	0.970
	Score in Applying	0.987
	Score in Reasoning	0.966
Score in Data & Chance	Score in Number	0.956
	Score in Algebra	0.875
	Score in Geometry	0.915
	Score in Knowing	0.926
	Score in Applying	0.953
	Score in Reasoning	0.965
Score in Knowing	Score in Number	0.982
	Score in Algebra	0.984
	Score in Geometry	0.970
	Score in Data & Chance	0.926
	Score in Applying	0.984
	Score in Reasoning	0.976
Score in Applying	Score in Number	0.992
	Score in Algebra	0.956
	Score in Geometry	0.987
	Score in Data & Chance	0.953
	Score in Knowing	0.984
	Score in Reasoning	0.986
Score in Reasoning	Score in Number	0.980
	Score in Algebra	0.959
	Score in Geometry	0.966
	Score in Data & Chance	0.965
	Score in Knowing	0.976
	Score in Applying	0.986

Table 43

Average Scale Scores of All Countries at the 90th Percentile

Country	*Number*	*Algebra*	*Geometry*	*Data&Chance*	*Knowing*	*Applying*	*Reasoing*
Algeria	482	448	511	437	452	491	448
Armenia	610	640	616	565	607	607	615
Australia	616	588	596	660	593	604	614
Bahrain	496	524	525	500	508	505	523
Bosnia and Herzegovina	550	580	556	546	576	545	558
Botswana	474	478	445	458	470	453	470
Bulgaria	586	619	600	583	604	587	593
Chinese Taipei	709	789	742	708	741	723	733
Colombia	487	484	462	499	459	481	512
Cyprus	580	580	585	585	570	582	585
Czech Republic	618	589	609	632	602	605	608
Egypt	517	537	538	482	521	527	514
El Salvador	450	429	424	445	431	435	452
England	622	610	625	691	609	623	632
Georgia	538	566	547	503	551	531	529
Ghana	430	466	408	407	441	419	430
Hong Kong SAR	688	696	696	677	698	685	689
Hungary	634	622	636	650	637	626	634
Indonesia	513	519	523	509	517	512	511
Iran, Islamic Rep. of	517	519	545	507	509	517	533
Israel	591	598	567	621	591	585	592
Italy	579	568	598	609	573	578	588
Japan	674	692	687	709	678	678	694
Jordan	551	577	563	543	562	554	564
Korea, Rep. of	709	749	718	715	733	721	709
Kuwait	452	463	484	446	453	462	459
Lebanon	553	570	565	513	561	550	553
Lithuania	614	607	624	644	623	616	604
Malaysia	608	560	606	563	584	585	568
Malta	619	583	614	626	605	607	591
Norway	575	517	563	639	530	562	577
Oman	474	519	508	487	496	488	509
Palestinian Nat'l Auth.	499	508	516	475	500	497	506
Qatar	443	428	428	410	431	427	429

Romania	582	620	602	559	600	588	596
Russian Federation	620	644	620	611	638	618	620
Saudi Arabia	425	439	458	418	421	439	447
Scotland	598	581	594	638	580	593	607
Serbia	593	629	610	589	613	599	598
Singapore	717	714	711	733	708	711	717
Slovenia	607	594	603	619	599	596	607
Sweden	594	563	582	672	560	588	608
Syrian Arab Republic	501	521	528	475	505	509	504
Thailand	574	566	576	557	556	563	573
Tunisia	516	512	527	496	505	516	508
Turkey	572	608	572	583	596	574	588
Ukraine	572	592	578	575	588	576	568
United States	620	611	580	665	619	609	607

Table 44

Pearson's Correlation Coefficient of 90th Percentile Scores (N = 48)

Dependent variable	Correlated variable	*r* – value
Score in Number	Score in Algebra	0.940
	Score in Geometry	0.970
	Score in Data & Chance	0.953
	Score in Knowing	0.976
	Score in Applying	0.992
	Score in Reasoning	0.981
Score in Algebra	Score in Number	0.940
	Score in Geometry	0.950
	Score in Data & Chance	0.859
	Score in Knowing	0.985
	Score in Applying	0.959
	Score in Reasoning	0.959
Score in Geometry	Score in Number	0.970
	Score in Algebra	0.950
	Score in Data & Chance	0.901
	Score in Knowing	0.970
	Score in Applying	0.987
	Score in Reasoning	0.967
Score in Data & Chance	Score in Number	0.953
	Score in Algebra	0.859
	Score in Geometry	0.901
	Score in Knowing	0.906
	Score in Applying	0.943
	Score in Reasoning	0.956
Score in Knowing	Score in Number	0.976
	Score in Algebra	0.985
	Score in Geometry	0.970
	Score in Data & Chance	0.906
	Score in Applying	0.983
	Score in Reasoning	0.973
Score in Applying	Score in Number	0.992
	Score in Algebra	0.959
	Score in Geometry	0.987
	Score in Data & Chance	0.943
	Score in Knowing	0.983
	Score in Reasoning	0.986
Score in Reasoning	Score in Number	0.981
	Score in Algebra	0.959
	Score in Geometry	0.967
	Score in Data & Chance	0.956
	Score in Knowing	0.973
	Score in Applying	0.986

References

Aczel, A. D., & Sounderpandian, J. (2006). *Complete business statistics (6th ed.).* New York, NY: McGraw-Hill.

Akiba, M., LeTendre, G., & Baker, D. (2002). Student victimization: National and school system effects on school violence in 37 nations. *American Educational Research Journal*, *39*, 829-853.

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

Angloinfo: Retrieved from http://bulgaria.angloinfo.com/family/schooling-education/the-school-year/

Asia Society: South Korean Education. Retrieved from http://asiasociety.org/education/learning-world/south-korean-education

Aubrecht, G. (1999). TIMSS: What lessons to learn? *The Science Teacher, 66*(1), 10.

Aud, S., & Hannes, G. (2009a). *The Condition of Education 2011*. NCES 2011-034. U.S. Department of Education. Retrieved from http://nces.ed.gov

Aud, S., & Hannes, G. (2009b). *The Condition of Education 2011*. U.S. performance across international assessments of student achievement. Retrieved from http://nces.ed.gov/programs/coe/analysis/2009-appa1.asp

Baker, D. (1997). Good news, bad news, and international comparisons: Comment on Bracey. *Educational Researcher, 26*, 16-17.

Baker, R., & Jones, A. (2005). How can international studies such as the international mathematics and science study and the programme for international student assessment be used to inform practice, policy and future research in science education in New Zealand? *International Journal of Science Education*, *27*, 145-157.

Balli, S., Wedman, J., & Demo, D. (1997). Family involvement with middle-grades homework: Effects of differential prompting. *Journal of Experimental Education, 66*(1), 31-48.

Balli, S. (1998). When mom and dad help: Student reflections on parent involvement with homework. *Journal of Research & Development in Education*, *31*,142-146.

Battista, M. (1999). Geometry results from the third international mathematics and science study. *Teaching Children Mathematics*, *5*, 367-373.

Baumert, J., & Köller, O. (2000). National and international school performance studies. *European Education*, *32*(3), 37-49.

Beaton, A. E., Mullis, I. V. S., Martin, M. O. Gonzalez, E. J., Kelly, D. L., & Smith, T. A. (1996). *Mathematics achievement in the middle school years: IEA's third international mathematics and science study (TIMSS).* Chestnut Hill, MA: Boston College.

Bennett, W. (1987). "Looking at Japanese education: Implications for American education." *NASSP Bulletin 71*(499), 102-108.

Bennett, G. K., Seashore, H. G., & Wesman, A. G. (1990). *Differential aptitude tests (5th ed.).* San Antonio, TX: Psychological Corporation.

Berliner, D., & Biddle, B. (1995). *The manufactured crisis*. New York, NY: Addison-Wesley Publishing Company.

Bielinski, J., & Davison, M. (2001). A sex difference by item difficulty interaction in multiple-choice mathematics items administered to national probability samples. *Journal of Educational Measurement, 38*(1), 51-77.

Birenbaum, M., Tatsuoka, C., & Yamada, T. (2004). Diagnostic assessment in TIMSS-R: Between-countries and within-country comparisons of eighth graders' mathematics performance. *Studies in Educational Evaluation, 30*, 151-73.

Bloom, B. S. (1981). Human characteristics and school learning. *Educational Researcher*, *10*, 29-30.

Borko, H . (2004). Professional development and teacher learning: Mapping the terrain. *Educational Researcher, 33*(8), 3-15.

Bracey, G. (1997). On comparing the incomparable: A response to Baker and Stedman. *Educational Researcher, 26*, 19-26.

Bruner, J. (1977). *The process of education*. Cambridge, MA: Harvard University Press. Bulgaria - Educational System—Overview. Retrieved from http://education.stateuniversity.com/pages/210/Bulgaria

Carroll, J. B. (1963). A model of school learning. *Teachers College Record, 64,* 722–733.

Casey, M., Nuttall, R., & Pezaris, E. (2001). Spatial-mechanical reasoning skills versus mathematics self-confidence as mediators of gender differences on mathematics subtests using crossnational gender-based Items. *Journal for Research in Mathematics Education, 32(*1), 28-57.

Cheng, Q. (2012). *Reformed and traditional mathematics teaching approaches: Are they related to the mathematics achievements of U.S. students across racial groups?* University of Nevada, Las Vegas. ProQuest Dissertations and Theses, 159. Retrieved from http://search.proquest.com/docview /1234288314? account id=14541 (1234288314).

Choi, K., Choi, T., & McAninch, M. (2012). A comparative investigation of the presence of psychological conditions in high achieving eighth graders from TIMSS 2007 mathematics. ZDM: *International Reviews on Mathematical Education, 44*, 189-199.

CIA: *The World Factbook.* Retrieved from https://www.cia.gov/library/publications/the-world-factbook/index.html

Coates, G. D., & Mayfield, K. (2009). Homework: How much help is enough? *Mathematics Teaching in the Middle School, 15*, 292-293.

Cogan, L., Schmidt, W., & Wiley, D. (2001). Who takes what math and in which track? Using TIMSS to characterize U.S. students' eighth-grade mathematics learning opportunities. *Educational Evaluation & Policy Analysis, 23*, 323-341.

Dewey, J. (1910). *How we think*. Boston, MA: Heath.

Dewey, J. (1929). *The quest for certainty*. New York, NY: Minton, Balch & Co.

Dimitrov, D. M. (2008). *Quantitative research in education. Intermediate & advanced methods.* Oceanside, NY: Whitter Publications Inc.

Diversity in Mathematics Education Center (2007). Culture, race, power, and mathematics education. In F. K. Lester, (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 405-433). Charlotte, NC: Information Age Publishing.

Encyclopedia of Educational Research (1960). London: Collier-Macmillan Limited.

Fang, Y. (2010). The cultural pedagogy of errors: Teacher Wang's homework practice in teaching geometric proofs. *Journal of Curriculum Studies, 42*, 597-619.

Ferrini-Mundy, J., & Schmidt, W. (2005). International comparative studies in mathematics education: Opportunities for collaboration and challenges for researchers. *Journal for Research in Mathematics Education, 36*, 164-74.

Finn, C. (1992). More like them. *The Public Interest, 107*, 106-112.

Forshay, A. W., Thorndike, R. L., Hoytat, F., Pidgeon, D. A., & Walker, D. A. (1962). *Educational achievement of thirteen-year-olds.* Hamburg: UNESCO Institute for Education.

Gonzales, P. (2001). *Using TIMSS to analyze correlates of performance variation in mathematics.* Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Gonzales, P. (2009). *Highlights from TIMSS 2007: Mathematics and science achievement of U.S. fourth-and eighth-grade students in an international context.* Washington, DC: U.S. Department of Education, National Center for Education Statistics.

Hammouri, H. (2004). Attitudinal and motivational variables related to mathematics achievement in Jordan: Findings from the third international mathematics and science study (TIMSS). *Educational Research, 46*, 241-257.

Hiebert, J., & Grouws, D. A. (2007). The effects of classroom mathematics teaching on students' learning. In F. K. Lester, (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 371-404). Charlotte, NC: Information Age Publishing.

Holliday, W. G. (1999). Questioning the TIMSS. *The Science Teacher, 66*, 34-37.

House, J. (2000). Student self-beliefs and science achievement in Ireland: Findings from the third international mathematics and science study (TIMSS). *International Journal of Instructional Media, 27*, 107-115.

House, J. (2001). Relationships between instructional activities and mathematics achievement of adolescent students in Japan: Findings from the third international mathematics and science study (TIMSS). *International Journal of Instructional Media, 28*, 93-105.

House, J. (2002a). The use of computers in a mathematics lesson in Japan: A Case analysis from the TIMSS videotape classroom study. *International Journal of Instructional Media, 29*, 113-124.

House, J. (2002b). Instructional practices and mathematics achievement of adolescent students in Chinese Taipei: Results from the TIMSS 1999 assessment. *Child Study Journal, 32*, 157-178.

House, J. (2003). Self-beliefs and science and mathematics achievement of adolescent students in Hong Kong: Findings from the third international mathematics and science study (TIMSS). *International Journal of Instructional Media, 30,* 195-212.

House, J. (2004). The effects of homework activities and teaching strategies for new mathematics topics on achievement of adolescent students in Japan: Results from the TIMSS 1999 assessment. *International Journal of Instructional Media, 31*, 199-210.

House, J. (2005). Mathematics beliefs and achievement of adolescent students in Japan: Results from the TIMSS 1999 assessment. *Psychological Reports, 97*, 717-720.

House, J. (2006a). Mathematics beliefs, instructional strategies, and algebra achievement of adolescent students in Japan: Results from the TIMSS 1999 assessment. *International Journal of Instructional Media, 33,* 443-462.

House, J. (2006b). Cognitive-motivational characteristics and geometry knowledge of adolescent students in Japan: Results from the TIMSS 1999 assessment. *International Journal of Instructional Media, 33,* 95-111.

House, J. (2009). Mathematics beliefs and achievement of a national sample of native American students: Results from the trends in international mathematics and science study (TIMSS) 2003 United States assessment. *Psychological Reports, 104*, 439-446.

House, J., & Telese, J. (2011). Effects of computer activities and classroom lesson strategies on motivation for mathematics learning for eighth-grade students in the

United States and Korea. *International Journal of Instructional Media*, *38*, 295-306.

House, J., & Telese, J. (2012). Effects of mathematics lesson activities and computer use on algebra achievement of eighth-grade students in the United States and Japan: Findings from the TIMSS 2007 assessment. *International Journal of Instructional Media*, *39*, 69-81.

Howie, S. (2003). Conditions of schooling in South Africa and the effects on mathematics achievement. *Studies in Educational Evaluation, 29*, 227-241.

Innabi, H., & Dodeen, H. (2006). Content analysis of gender-related differential item functioning TIMSS items in mathematics in Jordan. *School Science and Mathematics, 106*, 328-337.

International association for the evaluation of educational achievement. (2007): Completed studies. Retrieved from http://www.iea.nl/completed_studies.html

Jackson, S. (2005). *Research methods and statistics: A critical thinking approach (2nd ed.).* Belmont, CA: Thomas Higher Education.

Kalchman, M. (2011). Using the math in everyday life to improve student learning. *Middle School Journal, 43*, 24-31.

Kaur, B. (2011). Mathematics homework: A study of three grade eight classrooms in Singapore. *International Journal of Science and Mathematics Education*, *9*, 187-206.

Kellaghan, T. (1996). IEA Studies and educational policy. *Assessment in Education, 3*, 143-160.

Kiecolt, K. J., Nathan, L. E. (1985). *Secondary Analysis of Survey Data*. Beverly Hills, CA: Sage.

Kieran, C. (2007). Learning and teaching algebra at the middle school through college levels: Building meaning for symbols and their manipulation. In F. K. Lester, (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 707-762). Charlotte, NC: Information Age Publishing.

Kupari, P. (2003). Instructional practices and teachers' beliefs in Finnish mathematics education. *Studies in Educational Evaluation, 29*, 243-257.

Lagrange, J (2007). Large-scale assessment and mathematics education. In F. K. Lester, (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 1111-1167). Charlotte, NC: Information Age Publishing.

Larson, J. M., Jacobson, M.S., DenOuden, K. M., & Basile, C. G. (2012). *Mathematics teaching in the middle school, 17*, 346-351.

LeTendre, G., Baker, D., & Akiba, M. (2001). Teachers' work: Institutional isomorphism and cultural variation in the U.S., Germany, and Japan. *Educational Researcher, 30*(6), 3-15.

Lindquist, M. (2001). NAEP, TIMSS, and PSSM: Entangled influences. *School Science and Mathematics, 101*, 286-291.

Liou, P. (2010). *Cross-national comparisons of the association between student motivation for learning mathematics and achievement linked with school contexts: Results from TIMSS 2007.* University of Minnesota. ProQuest

Dissertations and Theses, 247. Retrieved from http://search.proquest.com/docview/738330890? accountid=14541. (738330890).

Lomos, C., Hofman, R., & Bosker, R. (2011). The relationship between departments as professional communities and student achievement in secondary schools. *Teaching and Teacher Education, 27*, 722-731.

Louis, R., & Mistele, J. (2012). The differences in scores and self-efficacy by student gender in mathematics and science. *International Journal of Science & Mathematics Education, 10*, 1163-1190.

Luyten, H. (2006). An empirical assessment of the absolute effect of schooling: Regression-discontinuity applied to TIMSS-95. *Oxford Review of Education, 32*, 397-429.

Macnab, D. (2000). Raising standards in mathematics education: Values, vision, and TIMSS. *Educational Studies in Mathematics, 42*, 61-80.

Meelissen, M., & Luyten, H. (2008). The Dutch gender gap in mathematics: Small for achievement, substantial for beliefs and attitudes. *Studies in Educational Evaluation, 34*, 82-93.

Menon, R. (2000). Should the United States emulate Singapore's education system to achieve Singapore's success in the TIMSS? *Mathematics Teaching in the Middle School*, *5*, 345-347.

Ministry of Education, Singapore: MOE *Seminars for Parents.* Retrieved from http://www.moe.gov.sg/events/seminars-for-parents/

Ministry of Education, Republic of China (Taiwan): *An Educational Overview*. Retrieved from http://english.moe.gov.tw/mp.asp?mp=1

Minotti, J. (2005). Effects of learning-style-based homework prescriptions on the achievement and attitudes of middle school students. *NASSP Bulletin 89*, 67-89.

Moore, J. (1998). Are all of the children below average? *Journal of Chemical Education*, *75*, 655.

Moses, R. P. (2001). *Radical equations: Math literacy and civil rights.* Boston, MA: Beacon Press.

Mullis, I.V.S., Martin, M.O., Ruddock, G.J., O'Sullivan, C.O., Arora, A., & Erberber, E. (2005). *TIMSS 2007 Assessment Framework.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College. National Assessment of Educational Progress, (2009). *Mathematics Grade 8: National Results.* Retrieved from http://nationsreportcard.gov/math_2009/gr8_national.asp

National Council of Teachers of Mathematics (1989). *Curriculum and Evaluation Standards for School Mathematics.* Reston, VA: Author.

National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: Author.

National Research Council. (2004). *On evaluating curricular effects: Judging the quality of K-12 mathematics evaluations*. J.Coudrey & V. Sohl (Eds.). Committee for a review of the evaluation data on the effectiveness of NSF-supported and commercially generated mathematics curricular materials. Washington, DC: National Academy Press.

Neuschmidt, O., Barth, J., & Hastedt, D. (2008). Trends in gender differences in mathematics and science (TIMSS 1995-2003). *Studies in Educational Evaluation, 34*(2), 56-72.

Parsons, J . S., Adler, T. F., Futterman R., Goff, S. R., Kaczala, C . M., Meece, J. L., & Midgley, C. (1980). *Self-perceptions, task perceptions, and academic choice: Origins and change (Final Report).* Washington, DC: National Institute of Education. (ERIC Document Reproduction Service No. ED 186477).

Patterson, M., Perry, E., & Decker, C. (2003). Factors associated with high school Mathematics Performance in the United States. *Studies in Educational Evaluation, 29*, 91-108.

Penner, A. (2003). International gender x item difficulty interactions in mathematics and science achievement tests. *Journal of Educational Psychology, 95*, 650-655.

Peterman, L., & Linder-Scholer, B. (1999). The value of an A. *The Science Teacher, 66*, 42-45.

Phan, H. (2008). *Correlates of mathematics achievement in developed and developing countries: An HLM analysis of TIMSS 2003 eighth-grade mathematics scores.* (Ed. D. Dissertation, University of South Florida, United States – Florida). Retrieved from Dissertations & Theses: Full Text. (Publication No. AAT 0389029).

Polya, G. (1957). *How to solve it*. Garden City, NY: Doubleday Anchor Books.

Pong, S., & Pallas, A. (2001). Class size and eighth-grade math achievement in the United States and abroad. *Educational Evaluation & Policy Analysis, 23*, 251-273.

Post, D., & Pong, S. (2000). Employment during middle school: The effects on academic achievement in the U.S. and abroad. *Educational Evaluation & Policy Analysis, 22*, 273-298.

RAND Mathematics Study Panel. (2003). *Mathematical proficiency for all students: Toward a strategic research and development program in mathematics education* (No. 083303331X). Santa Monica, CA: RAND.

Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., et al. (2000). *A User's Guide to Mlwin.* London: Multilevel Models Project, Institute of Education, University of London.

Raudenbush, S. W. (2004).What are value-added models estimating and what does this imply for statistical practice? *Journal of Educational and Behavioural Statistics, 29*, 121-129.

Ravitch, D. (1986). Japan's smart schools. *The New Republic, 194*, 13-15.

Ritzhaupt A, Higgins H, Allred B. (2010). Teacher experiences on the integration of modern educational games in the middle school mathematics classroom. *The Journal of Computers in Mathematics and Science Teaching, 29*, 189-216.

Rodriguez, M. (2004). The role of classroom assessment in student performance on TIMSS. *Applied Measurement in Education, 17*, 1-24.

Rosenberg, A.L., Greenfield, M.V.H., & Dimick, J.B. (2006). Secondary data analysis: Using existing data to answer clinical questions. In J.T. Wei (Ed.) *Clinical Research Methods for the Surgeons.* Totowa, NJ: Humana Press.

Rowan, B ., Correnti, R., & Miller, R . J. (2002). What large-scale survey research tells us about teacher effects on student achievement: Insights from the prospects study of elementary schools. *Teachers College Record, 104*, 1525-1567.

Schmidt, W., Cogan, L., Houang, R., & McKnight, C. (2011). Content coverage differences across districts/states: A persisting challenge for U.S. education policy. *American Journal of Education, 117*, 399-427.

Schmidt, W., Wang, H., & McKnight, C. (2005). Curriculum Coherence: An examination of US mathematics and science content standards from an international perspective. *Journal of Curriculum Studies, 37,* 525-559.

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

Schümer, G. (1999). Mathematics education in Japan. *Journal of Curriculum Studies, 31*, 399-427.

Seidel, T., & Shavelson, R.J. (2007). Teaching effectiveness research in the past decade: The role of theory and research design in disentangling meta-analysis results. *Review of Educational Research, 77,* 454-499.

Sfard, A. (2005). What could be more practical than good research? *Educational Studies in Mathematics, 58*, 393-413.

Sholl, M. J. (1989). The relation between horizontality and rod-and-frame and vestibular navigational performance. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,* 110-125.

Smith, T. M., Desimone, L. M., & Ueno, K. (2005). "Highly qualified" to do what? The relationship between NCLB teacher quality mandates and the use of reform-oriented instruction in middle school mathematics. *Educational Evaluation & Policy Analysis, 27*, 75-109.

Snow, R. E., Frederico, P.A., & Montague, W. E. (1980). *Aptitude, learning, and instruction.* Hillsdale, N J: Lawrence Erlbaum.

Stein, M.K., Remillard, J., and Smith, M.S. (2007). How curriculum influences student learning. In F. K. Lester, (Ed.), *Second Handbook of Research on Mathematics Teaching and Learning* (pp. 319-369). Charlotte, NC: Information Age Publishing.

Stevenson, H . W., Lee, S-Y., & Stigler, J . W. (1986). Mathematics achievement of Chinese, Japanese, and American children. *Science, 231*, 693-699.

Stigler, J., Gallimore, R., & Hiebert, J. (2000). Using video surveys to compare classrooms and teaching across cultures: Examples and lessons from the TIMSS video studies. *Educational Psychologist, 35*, 87-100.

Straus, M. (2003). Some strengths and weaknesses in algebra achievement of Slovene eighth-grade students. *Studies in Educational Evaluation, 29*, 207-226.

Stedman, L. (1997a). Deep achievement problems: The case for reform still stands. *Educational Researcher, 26,* 27-29.

Stedman, L. (1997b). International achievement differences: An assessment of a new perspective. *Educational Researcher, 26*, 4-15.

Tao Xin. , Xu, Z., & Tatsuoka, K. (2004). Linkage between teacher quality, student achievement, and cognitive skills: A rule-space model. *Studies in Educational Evaluation, 30*, 205-23.

Tarr, J., Mittag, K., & Uekawa, K. (2000). A comparison of calculator use in eighth-grade mathematics classrooms in the United States, Japan, and Portugal: Results from the third international mathematics and science study. *School Science and Mathematics, 100*, 139-150.

The Economist (2012). *Exams in South Korea*. Retrieved from http://www.economist.com/node/21541713

Thompson, T., & Preston, R. (2004). Measurement in the middle grades: Insights from NAEP and TIMSS. *Mathematics Teaching in the Middle School, 9*, 514-19.

TIMSS 2007 *Encyclopedia: A Guide to Mathematics and Science Education Around the World, Volumes 1 and 2.* Mullis, I.V.S., et al. (Eds.). (2008). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS. (1999). The Trends in International Mathematics and Science Study (TIMSS 1999). Retrieved from http://nces.ed.gov/timss/results.asp

TIMSS. (2003). The Trends in International Mathematics and Science Study (TIMSS 2003). Retrieved from http://nces.ed.gov/timss/results03.asp

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

TIMSS. (2007). The Trends in International Mathematics and Science Study (TIMSS 2007). Retrieved from http://timss.bc.edu/TIMSS2007/index.html

US Department of Education. (2006). *The Secretary's fifth annual report on teacher quality: A highly qualified teacher in every classroom, Washington, DC:* (U.S. Department of Education).

U.S. TIMSS. (1996). *Summary of findings*: U.S. National Research Center. Retrieved from http://ustimss.msu.edu/

Valverde, G., & Schmidt, W. (2000). Greater expectations: Learning from other nations in the quest for 'World-class standards' in US school mathematics and science. *Journal of Curriculum Studies, 32*, 651-687.

Vandenberg,S . G., & Kuse, A. R. (1978). Mental rotations, a group test of three-dimensional spatial visualization. *Perceptual and Motor Skills, 47*, 599-604.

Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes.*M. Cole, V. John-Steiner, S. Scribner, & E. Souberman (Eds.). Cambridge, MA: Harvard University Press.

Wang, J. (2001). TIMSS primary and middle school data: Some technical concerns. *Educational Researcher, 30*(6), 17-21.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993a). Towards a knowledge base for school learning. *Review of Educational Research, 63*, 249-294.

Wang, M. C., Haertel, G. D., & Walberg, H. J. (1993b). Towards a knowledge base: Why, how, and for whom? *Review of Educational Research, 63,* 365-376.

Wang, Y., & O'Dwyer, L. (2011). Teacher-directed student-use of technology and mathematics achievement: Examining trends in international patterns. *The Journal of Computers in Mathematics and Science Teaching, 30*, 79 - 135.

Wittrock, M.C. (1978). Development processes in learning from instruction. *Journal of Genetic Psychology, 132*, 37-54.

Wobmann, L., & West, M. (2006). Class-size effects in school systems around the world: Evidence from between-grade variation in TIMSS. *European Economic Review, 50*, 695-736.

Xin. T, Xu. Z., & Tatsuoka, K. (2004). Linkage between teacher quality, student achievement, and cognitive skills: A rule-space model. *Studies in Educational Evaluation, 30*, 205-223.

Yoshino, A. (2012). The relationship between self-concept and achievement in TIMSS 2007: A comparison between American and Japanese Students. *International Review of Education*, *58*, 199-219.

Zinth, K., & Dounay, J. (2006). Mathematics and science education in the States: State notes. Denver, CO: Education Commission of the States.

Zuzovsky, R. (2002). International comparative studies in education—what can local policy makers learn and use? The 3rd international mathematics and science study in Israel. *Studies in Educational Evaluation, 28*, 235-252.

Zuzovsky, R. (2008). Closing achievement gaps between Hebrew-speaking and Arabic-speaking students in Israel: Findings from TIMSS-2003. *Studies in Educational Evaluation, 34*, 105-117.

Biography

Venkata Subbaiah Patnam received his Bachelor's Degree in Mathematics at Andhra University in 1959 and Master's Degree in Mathematics at Patna University in 1961, after which he obtained a Postgraduate certificate in Pedagogy. Later, after teaching mathematics for eight years in higher secondary schools in the public sector and four years in colleges, he took a course of study at Mysore University and received a Bachelor's Degree in Education.

His passion for innovation in teaching and learning saw him as an organizer of mathematics clubs, Olympiads, exhibitions, workshops, and seminars. For several years, he was involved as a Resource Person and Course Director of professional development programs, conducted by the Central Schools Organization in India for teachers at the middle and primary grades. He was a member of the Book Review committee of the National Council of Educational Research and Training (NCERT) and contributed in its workshops on innovative practices and curricular issues. The NCERT sponsored him for a Mathematics Education Project at the Center for the Advancement in Mathematics Education in Technology (CAMET), University of Technology, Loughborough, England, on a British Oversees Development Fellowship, during 1982-83.

In 1987, he left service as Principal in the Central School Organization, to join a new Public Sector Educational system in India with a new vision, Navodaya Vidyalaya Samiti (NVS), running a network of pace-setting schools for talented students. As an Assistant Director, and later as Deputy Director, in-charge of a Region, comprising a few States, he planned and executed several innovative practices in school administration, academic activities, and teacher professional development. At the turn of the century, he superannuated from public service and worked as the academic director of a few high schools and later, as the Principal of a college. His Telugu translation of a book for high school mathematics is followed as the prescribed text book in the Open Schools of Andhra Pradesh state.

Resuscitating his lifelong ambition of pursuing research studies, he enrolled in the Ph. D. in Education program of George Mason University in 2007. He intends to engage himself in the professional development of teachers and in researches for improving the quality of teaching and learning in schools, particularly in mathematics.