MODELING OF THE FACTORS AFFECTING SCIENCE ACHIEVEMENT OF EIGHTH GRADE TURKISH STUDENTS BASED ON THE THIRD INTERNATIONAL MATHEMATICS AND SCIENCE STUDY – REPEAT (TIMSS - R) DATA

A THESIS SUBMITTED TO
THE GRADUATE SCHOOL OF NATURAL AND APPLIED SCIENCES
OF
THE MIDDLE EAST TECHNICAL UNIVERSITY

BY

ERTUĞRUL ÖZDEMİR

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF
MASTER OF SCIENCE
IN
THE DEPARTMENT OF SECONDARY SCIENCE AND MATHEMATICS EDUCATION

JULY 2003

Approval of the Graduate School of Natural and Applied Sciences

Prof. Dr. Canan Özgen

Director

I certify that this thesis satisfies all requirements as a thesis for the degree of Master of Science.

Prof. Dr. Ömer Geban

Head of Department

This is to certify that we have read this thesis and that in our opinion it is fully adequate, in scope and quality, as a thesis for the degree of Master of Science.

Prof. Dr. Giray Berberoğlu

Supervisor

Examining Committee Members

Prof. Dr. Giray Berberoğlu

Prof. Dr. Hamide Ertepınar

Doç. Dr. Fitnat Kaptan

Assist. Prof. Dr. Jale Çakıroğlu

Assist. Prof. Dr. Ali Eryılmaz

ABSTRACT

MODELING OF THE FACTORS AFFECTING SCIENCE ACHIEVEMENT OF EIGHTH GRADE TURKISH STUDENTS BASED ON THE THIRD INTERNATIONAL MATHEMATICS AND SCIENCE STUDY – REPEAT (TIMSS - R) DATA

ÖZDEMİR, Ertuğrul

M.S., Department of Secondary Science and Mathematics Education

Supervisor: Prof. Dr. Giray Berberoğlu

July 2003, 86 Pages

The purpose of this study is to investigate the factors that are related to students' science achievement in TIMSS-R. Basically instructional activities, affective characteristics of students and socioeconomic status (SES) were taken as the variables of the model proposed within the Linear Structural Modeling (LSM) framework. This study examined the TIMSS data for Turkish students with the sample size of 7841 through the analysis of Structural Equation Modeling (SEM). Resulting path diagram showed that the largest relationship existed between science achievement and SES of students. It was also observed that students' enjoyment of science did not seem to have a significant contribution on science achievement. In addition, science achievement had a negative relationship with the classroom activities considered as student-centered. On the other hand, the activities considered as teacher-centered had a positive impact on the science achievement scores of the TIMSS tests. It was also observed that science achievement and perception of success/failure in science were highly related with each other. The results were interpreted within the framework of Turkish educational system, and some suggestions for future research studies were proposed.

Keywords: Science Achievement, TIMSS, Structural Equation Modeling, Instructional Activities, Attitude toward Science, Science Education.

ÖZ

SEKİZİNCİ SINIF ÖĞRENCİLERİN FEN BAŞARISINA ETKİ EDEN FAKTÖRLERİN ÜÇÜNCÜ ULUSLARARASI MATEMATİK VE FEN ÇALIŞMASI TÜRKİYE VERİLERİNE DAYALI MODELLENMESİ

ÖZDEMİR, Ertuğrul

Yüksek Lisans, Ortaöğretim Fen ve Matematik Alanları Eğitimi Bölümü

Tez Danışmanı: Prof. Dr. Giray Berberoğlu

Temmuz 2003, 86 Sayfa

Bu çalışmanın amacı sekizinci sınıf öğrencilerin fen başarısının, sınıf içi etkinlikler, ailenin sosyo-ekonomik durumu ve öğrencilerin fen bilimlerine karşı tutumları gibi fen başarısına etki eden faktörler ile ilişkilerinin modellenmesidir. Bu çalışma örneklemi 7841 büyüklüğünde olan TIMSS Türkiye verilerini Yapısal Denklem Modellemesi (SEM) tekniğini kullanarak analiz etmiştir. Bu çalışmanın sonunda ortaya çıkan ilişki modeli göstermiştir ki, öğrencilerin fen başarısı en çok ailenin sosyo-ekonomik durumuyla ilgilidir. Ayrıca, bu model incelendiğinde öğrencilerin fen bilimlerinden hoşlanmaları onların fen başarıları ile ilgili görünmemektedir. Buna ek olarak, öğrencilerin fen başarıları ile öğrenci merkezli sınıf içi etkinlikleri negatif ilişki göstermiştir. Bunun aksine, fen başarısı ile öğretmen merkezli sınıf içi etkinlikleri arasında pozitif bir ilişki gözlenmiştir. Ayrıca öğrencilerin fen başarılarının onların başarı algılarıyla yüksek düzeyde ilişkili olduğu gözlenmiştir. Bu çalışma yukarıda özetlenen sonuçları Türk Eğitim Sistemi çerçevesinde tartışmaktadır.

Anahtar sözcükler: Fen Başarısı, TIMSS, Yapısal Denklem Modellemesi, Sınıf İçi Etkinlikler, Fen Bilimlerine karşı Tutum, Fen Eğitimi.

ACKNOWLEDGMENTS

I wish to express my sincere thanks to my supervisor Prof. Dr. Giray Berberoğlu for his guidance and encouragements during this research. In addition, I would like to thank my wife Demet Özdemir, my parents Mustafa Ahsen Özdemir and Nuran Özdemir, and my sister Ayşegül Gökçakır for their encouragements. I also wish to thank Emel Uysal and Betül Yayan for their very precious contributions for the analysis, and to my colleagues Aslı Aydın, Elvan Alp and Mine Işıksal, who share the same room with me, for their encouragements.

To My Wife Demet

TABLE OF CONTENT

LIST OF FIGURES

Figures

LIST OF TABLES

Tables

CHAPTER 1

INTRODUCTION

1.1 Introduction

The Third International Mathematics and Science Study (TIMSS) is an international survey about science and mathematics education, which implements repetitive data collection process throughout the years. TIMSS was first conducted in 1995 with 42 participant countries. In 1999, TIMSS was repeated in 38 countries including Turkey. This particular one is called the Third International Mathematics and Science Study – Repeat (TIMSS-R). Totally, TIMSS-R database includes responses of approximately 500,000 students, more than 50,000 mathematics and science teachers, and more than 12,000 school principals. Third data collection of TIMSS was carried out in 2003. However, Turkey did not participate in TIMSS 2003. Present study deals with TIMSS-R data for the eighth grade Turkish students with the sample size of 7841.

TIMSS aims at providing feedback to the educational systems of different countries that each country might have a chance to implement some educational reforms, and follow up the impacts throughout the repeated data collection attempts. It seems quite important to participate the TIMSS project regularly for the countries, which are following up the impacts of educational reforms based on the previous observations and evaluations of the TIMSS database. The comprehensive database of the TIMSS may help the researchers to understand the reasons of low achievement in TIMSS (Gonzalez & Miles, 2001). Rankings of the countries with respect to their performances on achievement measures are not the primary concern of the project. However, education policy makers have interests in their countries' ranks among the other participating countries. TIMSS-R results showed that Turkey's science score was significantly below the international average. (Gonzalez & Miles, 2001) (See Appendix E). This result is not completely unexpected for the Turkish educational system, but the project may give an opportunity to investigate the weaknesses and strengths of the students with respect to various content areas and cognitive skills.

Beside the achievement information of the students, TIMSS data also include students' family background information, attitude toward science, and instructional processes based on the items included in the TIMSS student questionnaire. Interpretation of achievement data and the information included in student questionnaire might help researchers to understand the factors that are influential on students' achievement in science.

Thus, in the present research, it is aimed to develop and interpret a linear structural model that incorporates some selected student characteristics, mode of instructional activities, and achievement data to understand possible sources of factors that explain achievement in science.

1.2 Research Question

In the present study, an exploratory study was carried out to select the variables that are going to be included in the model. For this reason, items of the student questionnaire were analyzed through the Principle Component Analysis to form the latent variables included in the model. In proposing the model, previous studies from the literature survey, besides the conceptual framework of the TIMSS student questionnaire data were used. In TIMSS Student Questionnaire students were asked to rate some family background characteristics, instructional activities and students' attitudes toward science. The items in different subcomponent of the Student Questionnaire were further evaluated through the Principle Component Analysis to form meaningful groups for the LISREL model to be studied.

After describing the groups of variables to be included in the study relationships between students' science achievement were examined through the Structural Equation Modeling (SEM).

The research question of this study is; "What linear structural model explains the relationships among a set of latent variables in TIMSS-R data for the Turkish students?"

1.3 Definitions of Important Terms

1.3.1 Principle Component Analysis

Principle component analysis is a technique for determining whether many variables observed in an instrument can be described by a few factors by searching for clusters of variables that are correlated with each other (Fraenkel & Wallen, 1996). This technique will be described in detail in Chapter 3.

1.3.2 Structural Equation Modeling

The idea behind Structural Equation Modeling is to formulate a theory about the relationships among variables. Kelloway (1998) mentioned that a theory could be thought of as an explanation of correlations among variables. It is necessary to note that correlations, in social sciences, give some idea about causal relations, however, correlations can not be used solely to explain the causality between variables (Fraenkel & Wallen, 1996). Structural Equation Modeling technique will be described in detail in Chapter 3.

1.3.3 Observed Variables

Observed variables are, simply, variables directly measured by an instrument. For example, a question in a questionnaire, which is stated as "what is your gender?" can be considered as an observed variable, and it can be named as "gender" (Kelloway, 1998).

1.3.4 Latent Variables

Latent variables are not measured by an instrument directly. Rather, they are formed by combination of observed variables, which can be described under a more general title (Kelloway, 1998). In other words, some questionnaire items can be described all together as a latent variable which are correlated among themselves (Kelloway, 1998).

1.3.5 Endogenous Variables

In studies, there are some variables the researchers want to explain or predict. This type of variables is generally the internal characteristics of individuals (Kelloway, 1998). In this study, ACHV, ENJY, SUCC, and IMPT are endogenous variables.

1.3.6 Exogenous Variables

In studies, there are some variables the researchers want to use in order to explain or predict the endogenous variables. This type of variables is generally independent from individual characteristics, in other words, exogenous variables are external characteristics affecting human behaviors (Kelloway, 1998). In this study, SES, STAC, and TEAC are exogenous variables.

1.3.7 LISREL

LISREL is a computer application for Structural Equation Modeling. It performs Structural Equation Modeling through correlation or covariance matrixes of observed variables. It gives a path diagram and an output document including fit indexes and modification indexes (Kelloway, 1998).

CHAPTER 2

LITERATURE REVIEW

2.1 Introduction

In this study, related literature was reviewed through three subtitles; Science Achievement in TIMSS, Structural Equation Modeling and General Information about TIMSS. Under the first title, it is intended to demonstrate how science achievement is investigated by using TIMSS data for many different countries. Understanding the findings of previous studies on science achievement in TIMSS for many countries is useful to understand and interpret the findings of this study examining science achievement of Turkish Students.

This study mainly uses structural equation modeling (SEM) which is an advanced statistical analysis. Under the second title, it is intended to demonstrate how structural equation modeling is employed in different studies in literature. Understanding this advanced statistical analysis is useful to interpret resulting models of the analysis in this study.

In this study, TIMSS-1999 data is used, therefore, it is very important to understand the general characteristics of this international study. It is also important to compare the findings of this study with the findings of the other studies based on TIMSS, for a better comprehension of location of Turkish Educational System among all participant countries of TIMSS. Under the third title of this section, it is aimed to give some general information about TIMSS.

2.2 Science Achievement in TIMSS

In literature, lots of articles on science achievement based TIMSS data for many countries are available. Some of them focused on students' science achievement in one participant country of TIMSS. On the other hand, some of them are cross-national studies comparing science achievements of students from different countries. In the following part of this section, several studies on science achievement in TIMSS are reviewed, demonstrating their relationships with this study and their inter-relationships.

In this study, many observed variables are spontaneously examined in the analysis. However, the literature includes some studies investigating the science achievement in a narrower frame. House (2000b) aimed to investigate the relationship between students' self-beliefs and their science achievements using data from TIMSS and to examine the generalizability of previous findings for students in a cross-cultural context. In this study, sample was included 5881 13-year-old students from Ireland, and, some variance estimation techniques for complex sampling designs were employed. It was indicated that several specific self-beliefs were associated with higher levels of science achievement. House (2000a), in another study, intended to assess the efficacy of several different instructional activities using data from TIMSS for Hong Kong. Variance estimation techniques for complex sampling designs were employed. The sample size was 6031. As a result, cooperative learning appeared to be related to students' science achievement. This result was consistent with previous research findings on instructional activities. Similarly, Yore, Anderson and Shymansky (2002) aimed to model the relationships of classroom characteristics and students' attributes to students' science achievement. The researchers used hierarchical linear modeling on TIMSS and Science Co-op Local Systemic Change Project data. As results, classroom factors influenced the weightings of significant student attributes, such as, awareness of nature of science, attitude towards science etc.

Some studies in literature dealt not only with the science achievement but also with both science and mathematics achievements. Webster and Fisher (2000) aimed to address the variation in science and mathematics achievement in urban and rural Australia. In Australia, 12852 13 year-old students were asked to complete TIMSS questionnaire and mathematics and science tests. Data were analyzed through multi-level linear models. It was observed that not having the same standards of resources is not a problem.

In the literature, some studies criticized some aspects of TIMSS. They generally discussed validity and reliability of the results of this international study. Especially TIMSS science test is at the focus of the criticisms. Ramseier (1999) aimed to investigate relatively low science achievements of Swiss students and the reasons of difference between science and mathematics achievements in Switzerland. The author stressed the difficult technical language of TIMSS science tests and disharmony between curricular priorities of Switzerland and TIMSS science tests when interpreting the relatively low science scores of Swiss students. Similarly, Fensham (1999) investigated the reasons of unexpected results of TIMSS for Australian students. In Australia, primary students were high achiever in TIMSS science test, which was an unexpected result. According to the author, there were some evidences that the primary education system in Australia was poor in quality. The Author had some doubts about whether TIMSS test measured the same science achievement for each participant. Because, item distributions for each participant country were very much different. Other critics are about science items of TIMSS test. Wang (1998) investigated the items of TIMSS test. According to the author, not all free-response scores reflected students' science achievement, not all science items had only one correct response, not all TIMSS scores were grounded in students' levels of cognitive development, and not all TIMSS items reflected collaboration between mathematics and science educators.

As mentioned above, literature includes many cross-cultural studies based on TIMSS as well as the studies examining only one participant country. Shen (2002) performed a cross-national research study with 38 participant countries. This study examined the relationship between students' achievement and their self-perceptions. Researcher used Pearson correlation for TIMSS data across these countries. For within-country data, positive relationships between students' science and mathematics achievement and three measures of their self-perception: how much they like the two subjects, their self-perceived competence in the subjects, and their perceived easiness of the subjects were found. Shen and Pedulla (2000) intended to examine the relationship between students' self-perception and their science and mathematics achievement using TIMSS data through a cross-national analysis. TIMSS data for 9 and 13 year-old students were used for this study. In this study, 26 countries were included for 9 year-old students and 41 countries were included for 13 year-old students. Within-country data generally showed a positive relationship between students' achievement and self-perceived competence in science and mathematics. However, when investigating between-country data, the opposite relationship was observed. Similarly, Kuiper and Plomp (2001) focused on national findings and analyses from four TIMSS participant countries; Germany, South Africa, Switzerland, and the Netherlands. Their study consisted of some contributions to explain the achievement in TIMSS, covering a variety of topics and research methodologies. They summarized the findings of studies in literature.

Obviously, teacher characteristics play an important role on students' achievement. Vlaardingerbroek and Taylor (2003) intended to overview the teacher education variables associated with science achievement in a cross-national perspective. This study used data for 13 TIMSS participants. According to the authors, this study reinforced the opinion that primary teacher training ideally occurs in a university, and involves a four-year degree program that preferably includes to common standards across institutions.

In TIMSS samples, Population 2 includes seventh and eighth grades, that is, both seventh and eighth grades completed the same questionnaire and the same achievement test. In this instant, a question is raised; are these two grade levels really identical? Wang and Zhu (2003) analyzed middle school data from the Third International Mathematics and Science Study to compare achievement difference between seventh and eighth grades. The authors noted that not all TIMSS items had resulted in a higher mean score at the upper grade level. And this problem existed in all participating nations except for those that gather data from a single grade. The problem of having higher score at lower grade level varies among nations. According to the authors, this problem was due to variations in cross-national curriculum coverage.

TIMSS gave the participant countries a chance of realizing the relative situation of their science and mathematics education. After TIMSS, some of these countries decided to make some reforms in educational systems. McKnight and Schmidt (1998) discussed the TIMSS results for the USA, which are not satisfactory. Science and Mathematics scores of this country were near the international average. The authors noted that these results were related to the US official science and mathematics curricula and the textbooks. According to the authors, official curricula and textbooks include many topics and they are unfocused. Therefore, US science and mathematics education failed to tell the coherent story of these areas. Similarly, Jakwerth (1999) aimed to discuss TIMSS performance assessment results for the United States of America. According to the author, US students did not exhibit the performance that had been expected. And the author added that performance of US students could be explained in part by teacher practices; data from the TIMSS teacher questionnaire showed that too few teachers reported applying the strategies to promote higher-order, complex thinking in the classroom. Similarly, Harris (1999) studied the strengths and weaknesses of students' performance in England. According to the author, the results of the two students achievement components (the written tests in mathematics and science, and the practical tests in these two areas) complement each other. Actually, it is not enough to study one of them in order to understand the students' strengths and weaknesses.

TIMSS was studied by researcher in many aspects. Wang and Schmidt (2001) studied the results of TIMSS about inclusion of History, Philosophy and Sociology of Science in Science Education (HPSS). According to the authors, it is imperative that appreciation of past complexities of science and societies and the nature of scientific knowledge be a part of the education of both scientists and non-scientists. Inclusion of HPSS in science education has been found to be an effective way to reach the goal of enhancing science literacy for all citizens. Another aspect of TIMSS results studied by researchers is effects of class-size on science achievements. Woessmann and West (2002) intended to investigate the classroom size effects in school systems among 18 TIMSS participants. As results, smaller class-sizes exhibit beneficial effects only in countries with relatively low teacher salaries, and while sizeable beneficial effects of smaller classes were found in Greece and Iceland, any effects of smaller classes were rejected for Japan and Singapore. Another aspect of TIMSS studied by researchers is effects of item types on science achievement. Park and Hong (2002) examined the relative effectiveness of item types for estimating science ability based on TIMSS 1999 data. The researchers used Item Response Theory (IRT) to investigate this. The results showed that free response items were appropriate for the students with high science ability and that free response items estimated students' science ability more accurately than multiple choice items. Science achievement in primary school is another aspect of TIMSS studied by researchers. Martin, Mullis, Beaton, Gonzalez, Smith and Kelly (1997) focused science achievement in primary school based on TIMSS data for 26 countries. The authors examined, in detail, the nature of science test, country characteristics, differences in student science achievement, and intended and implemented curricula. In primary years, the students from all participants seemed to have some difficulties in nature of science and there were several differences between intended and implemented science curricula.

2.3 Structural Equation Modeling (SEM)

In order to demonstrate how SEM is applied in research studies and how the resulting path models are interpreted, several articles in the literature are searched in the following paragraphs. Some of these studies used structural equation modeling; however, some of them criticized some studies using this analysis in an incorrect way.

Bos and Kuiper (1999) examined mathematics achievement in a comparative study based on TIMSS data for 10 participant countries with

10 different educational systems. These countries were Belgium-Flemish, Netherlands, Belgium-French, Sweden, Germany, England, Norway, Denmark, Czech Republic, and Lithuania and total sample size was 21635. This study used Partial Least Square Path Analysis on TIMSS Data for 10 countries mentioned above. For 10 education systems, it was founded that the resulting general path model explains 19% or less of the variance in achievement in mathematics. In many systems, it was observed that home educational background and students' attitude towards mathematics have a positive relation with achievement in mathematics, however, out-of-school activities has a negative relation with achievement in mathematics. In literature, it is possible to reach some other studies modeling TIMSS data for different countries. Papanastasiou (2000) explored how predictors related to the family and the school stimulate mathematics outcomes. TIMSS data for Cyprus, Japan and the US were analyzed to reach a path model. According to this model, the factor having the strongest direct influence on attitudes toward mathematics was teaching in Cyprus and Japan, and reinforcement in the US. The model also seemed to indicate that attitudes and self-beliefs could not be used to predict students' achievements in mathematics. Similarly, Welsh, Parke, Widaman and O'Neil (2001) primarily aimed to test a reciprocal model in order to understand the direction of effects between social competence and academic competence over time. Sample of the study was 163 first, second and third grade students. The model indicated that academic achievement directly influenced social competence from both first to second and second

to third grade. Also, social competence was reciprocally related to academic achievement from second to third grade.

Graduate Record Examination (GRE) is an international examination, which has thousands of examinees every year. Therefore, this exam has commonly been the subject of researches. Stricker & Rock (2002) intended to investigate the relationships between examinees' background characteristics and performance on GRE General Test by structural equation modeling technique. According to the results of this study, examinees' initial characteristics (gender, ethnicity, parental education, geographic region, and age) had the modest relationships with their test performance. In addition, among these initial characteristics, parental education had the strongest association with the test score. Also, gender seemed to have a significant association with the test performance. According to the authors, these findings underlined the importance of socioeconomic status, a neglected construct in recent years.

Path models may not include achievement, unlike in the previous study. Attitudes can also be explained by modeling. This study deals both achievement and attitudes. Papanastasiou (2002) aimed to investigate the school, teaching and family influence on students' attitudes toward science in Cyprus using TIMSS 1995 data and structural equation modeling with LISREL. In the resulting path model, the highest correlation among the latent variables of family, reinforcement, teaching, climate was the correlation between attitudes and teaching. On the other hand, the lowest correlation among these latent variables was the correlation between attitudes and family (educational background).

SEM is generally considered as a very complex and incomprehensible statistical analysis. Martin (1987) intended to guide for perplexed researchers about structural equation modeling. It was attempted to explain basic logic behind SEM. The author reminded that structural equation modeling is an extension of familiar techniques, such as, multiple regression and factor analysis.

As mentioned above, some studies in the literature criticized the misuse of structural equation modeling. Biddle and Marlin (1987) intended to examine and discuss the legitimate advances of SEM and credulity on it. In this study, advantages and problems associated with structural equation modeling were discussed. Problems of SEM were discussed through four subtitles: causal modeling problems, path diagram problems, regression analysis problems, and LISREL problems. Under the first subtitle, the author noted that the strongest conclusion when assessing a path model is that it predicts the pattern of observed associational and possibly temporal relations. One cannot say anything about causality between the variables in the model. Under the second subtitle, the author noted that path diagrams are an efficient way of summarizing complex information, but they always leave out information that the reader may need. Under the third subtitle, the author noted that ordinary least-squares regression requires a correlation or a covariance matrix as input, and, in constructing and interpreting such a matrix, it is assumed that all variables are in interval scale, the dependent variables are normally distributed and all variables are linearly related to one another. These assumptions should not be violated when using Structural Equation Modeling. Under the last subtitle, the author noted that LISREL, as a powerful statistical tool, makes some additional assumptions. Actually, LISREL and ordinary least-square regression analysis do not use identical procedures. Therefore, there are some additional assumptions unique to LISREL.

2.4 General Information about TIMSS

In this section of the study, it aimed to introduce the reader to general characteristics of Third International Mathematics and Science Study in both positive and negative aspects. It is possible to consider the studies on TIMSS in two groups; one group is aimed to give descriptive information, however, the other group criticizes TIMSS and other international studies.

Some studies on TIMSS are just summary of this international study. Jones (1999), in his content analysis research, described what TIMSS is, its goals, the study design and focus populations. In this study, it was mentioned that TIMSS has evaluated attitudes towards science, science achievement and background characteristics of 500.000 8^{th} grade students from 47 countries around the world. Similarly, Peak (1996) reported the initial findings of the Third International Mathematics and Science Study for the US. The author reported the results of students' assessment to describe how the US students perform in mathematics and science. In addition to this, it was examined the possible factors which may have an important influence on achievement, such as, educational standards, curriculum, teacher characteristics, and student initial characteristics etc.

One of the most important results of TIMSS is the superior achievement of Hong Kong, Japan, Korea, and Singapore, the only participants of TIMSS from East Asia. Leung (2002) investigated the reasons of high achievement of East Asian Students in TIMSS. Common characteristics of these countries were investigated. High population densities, large class sizes, relatively low attitude toward mathematics and science and common Confucian Culture were reported for these East Asian Countries. The author stated that there is not enough evidence to relate the high achievement of student from East Asia to their common Confucian Culture. However, low attitude of these students can be explained by their unique culture stressing humility and modesty. Similarly, Menon (2000) investigated possible reasons of Singapore’s success in TIMSS. The author mentioned five possible reasons for this situation. First, it was mentioned that Singapore students were exam smart. In other words, they were prepared thoroughly for examinations by constant practice, review of past examinations etc. Second reason is attitude toward education. That is, education is seen as a passport to upward social mobility in Singapore, therefore, parents provide extra tutoring from kindergarten for their pupils. The third reason is public ranking of schools. In Singapore, most school principals exhort teacher to improve their students’ test performance. Forth and fifth reasons are centralized curriculum and high quality teacher education.

Some studies in the literature criticize international assessments, specifically TIMSS. Wolf (1998) was intended to discuss the validity issues in international assessments. The author mentioned two main problems in international assessments; one is that the different curricula of different countries may result in different scores, and, the other problem is that an international assessment study may have a testing procedure that is not appropriate for countries' usual testing practice. The only problem of TIMSS is not validity. Zuzovsky (1999) aimed to discuss the reliability of TIMSS scores of Israeli students. The author stated that some tasks in TIMSS achievement test have quite vague coding system, and, all "why" questions that demand explanations are hard to interpret and have low inter-rater consistency. The author also stated that heavily elaborated coding systems yield lower inter-rater consistency. Another problem about TIMSS and other international assessment studies is translation of achievement tests into different languages. Ercikan (1998) intended to discuss the translation effects in international assessment. The author explored the application of a statistical method to examine the effect of translations on the equivalence of test items and the comparability of test scores in order to identify poorly translated items in TIMSS tests and to examine how the comparability of scores is affected by problems in translations. To detect translation problems, the method of differential item functioning (DIF) was used. The problems of international assessments examined in the lines above also affects surely the results of this study which models TIMSS data for Turkey. There are lots of studies criticizing TIMSS and other cross-national studies in many aspects in the literature. One of these aspects is how different subsets of TIMSS science items could measure the achievement fairly. Beaton (1998) investigated how fair the TIMSS tests were to the various participating countries. The author used the method of Test-Curriculum Matching Analysis (TCMA). This method was used to investigate how results might have changed if different subsets of TIMSS items were considered. The results suggested that the relative position of participating countries changed very little as a result of the item selection. Similarly, Lokan (1999) attempted to investigate issues related to fairness in the TIMSS Performance Assessment tasks by variables such as gender, language and socio-educational background. There were some important results of this study: there was evidence that the gender effects were less overall for the hands-on tasks of TIMSS Performance Assessment than they might be for the paper-pencil tests; there is also evidence that the TIMSS hands-on type of assessment is less prone to equity problems than TIMSS main assessment. Similarly, Cheng and Cheung (1999) focused the challenges to the validity of country ranking of TIMSS. According to the authors, in TIMSS, there were some problems in relevance of TIMSS to countries' curricula, and some methodological limitations, lack of high quality process data at classroom level, lack of contribution to theory building from the analysis of TIMSS data, and limited policy implications. Similarly, Bracey (1998) discussed six problems in cross-cultural assessment. These problems might be summarized as follows; differences in age of students from different participant countries, differences in enrollment percentages of students in different countries, defining the population of a subject (mathematics or science) in different countries, some unbelievable numbers in reports, cultural differences and the relevance of the results. In the same way, Smith (1997) aimed to investigate the generalizability of scoring TIMSS open-ended items. It was reminded that scoring responses of students used a two-digit diagnostic code rubric with the first digit determining the correctness of the response and the second digit used to identify common approaches or misconceptions. The generalizability for an individual's score on a specific item was found to be less stable for some items. Similarly, O'Leary (2001) examined the TIMSS data in order to determine the extent to which the relative standing of countries was consistent across three different item formats; multiple choice, short answer, and extended response. According to the results of this study, Irish students' performance was closed to the international averages for short answer and multiple choice items, however the performance on extended response items was significantly above the international average. The match of items and Irish curriculum was not good and this curriculum judged to encourage higher-order thinking less than in other countries, therefore this high performance of Irish students was very surprising.

TIMSS were studied frequently in gender difference perspective. Wester and Henriksson (2000) examined the interaction between item format and gender differences in mathematics performance based on TIMSS data with sample of 8851 sixth, seventh and eight grade Swedish students. Results showed that no significant changes in gender differences exist when the item format is changed.

2.5 Summary

TIMSS has lots of important results, which are available in the literature. Results about science achievement can be summarized as follows. House (2000a) noted that science achievement of students appeared to be related to cooperative learning. In addition, Woessmann and West (2002) found that smaller class sizes exhibited beneficial effects on science achievement only in countries with relatively low teacher salaries. House (2000b) also stated in another study that several specific self-beliefs were associated with higher levels of science achievement of student. Similarly, Shen (2002) found that a positive relationship between science achievement of students and three measures of their self-perceptions: how much they like science, their self-perceived competence in science, and their perceived easiness of this subject were found. In the same way, Shen and Pedulla (2000) observed that a positive relationship between students' science achievement and their self-perceived competence in science existed. In addition, Yore et al. (2002) studied attitude toward science rather than science achievement. According to the authors, several classroom factors influenced some attributes of students, such as, awareness of nature of science, and attitude toward science. In another perspective, Park and Hong (2002) noted that free response items were appropriate for the students with high science ability, and that free-response items estimated students' science achievement more accurately than multiple choice items.

Results of studies in the literature about the problems of Third International Mathematics and Science Study can be summarized as follows. Ramseier (1999) stressed that difficult technical language of TIMSS science test and disharmony between curricular priorities of participant countries and TIMSS science test are two of the problems of this international study. In addition, Wang (1998) mentioned four problems in TIMSS, these are; not all free-response scores reflected students' science achievement, not all science items had only one correct response, not all TIMSS scores were grounded in students' level of cognitive development, and not TIMSS items reflected collaboration between mathematics and science educators. Similarly, Fensham (1999) had some doubts about whether TIMSS test measured the same science achievement for each participant country, item distributions for each country were very much different. In the same way, Wolf (1998) mentioned that an international assessment study may have a testing procedure that is not appropriate for the countries' usual testing practice. Similarly, Bracey (1998) summarized the problems in TIMSS as follows; differences in age of students from different countries, differences in enrollment percentages of students in different countries, and cultural differences. In addition, Ercikan (1998) discussed possible negative or positive translation effects for some countries in an international assessment study like TIMSS.

When putting an end the literature, to sum up, there are two groups of studies in the literature. First group is the studies that examined TIMSS data and interpreted the results and the other group is the studies that examined some problematic aspects of TIMSS and other international assessment studies. The present study benefits from both groups.

CHAPTER 3

METHODOLOGY

In this chapter, research design, sampling, instruments, statistical analysis, the procedures, validity and reliability of path model were explained. Sampling and instruments sections were explained with reference to TIMSS original study. And, in the statistical analysis section, principle component analysis and the Structural Equation Modeling (SEM) were explained.

3.1 Research Design

Simply, this study is an advanced correlational research, which uses structural equation modeling. Correlational research is carried out in order either to explain human behaviors or to predict likely outcomes (Fraenkel & Wallen, 1996). This study may be regarded as latter. A major purpose of prediction correlational research is to predict a score on a variable using another variable that is known to be highly correlated to this variable (Fraenkel & Wallen, 1996). Specifically, major purpose of the present study is to predict a model best explaining the relationships between science achievement and some factors affecting it. When studying a correlational research, it is very important to discuss the relationship between correlation and causation. Sometimes correlation coefficients may suggest causation. However, only correlation coefficients can never establish causation between two variables. Because, sometimes highly correlated two variables may not have causation with each other, instead, they may have a common cause. Therefore, the results of correlational studies must always be interpreted with caution that they may suggest, but they cannot establish causation (Fraenkel & Wallen, 1996)

3.2 Sampling

In the TIMSS-R, there are three populations defined. Population-1 is third and fourth grade students, population-2 is seventh and eighth grade students and population-3 is the students in their final year of high school. This study deals with the population-2 only. It should be noted that population-2 of Turkey includes only eighth grade students.

In the TIMSS-R, the sampling procedure was carried out through the following steps. Random selection of schools were carried out by Canada Statistics with the information provided by the Ministry of National Education. A stratified sampling method was used in selecting the subjects of the study. Below, the brief explanations of sampling procedures were presented.

3.2.1 Sampling of Schools

The first step of sampling is school sample selection. In this step, a representative sample of schools in Turkey was selected. Selection of schools is based on a systematic probability-proportional-to-size technique, which includes the consideration of some strata such as, geographical region, public/private (Gonzales & Miles, 2001).

3.2.2 Sampling of Classrooms and Students

After school sampling, the second sampling step is classroom and student sampling. Generally, in each school, one classroom was sampled from each target grade (eighth grade in this study). Most participant countries tested all students in selected classrooms. However, some participants with particularly large classrooms in their schools decided to randomly subsample a fixed number of students per selected classroom (Gonzales & Miles, 2001). In the present study, all of the students in the population-2 were included in the analyses. The sample size of population 2 for Turkey was 7841. At this point, it should be noted that the sample size in the analysis is 5297. This difference is due to the list-wise deletion of the cases in the data. In missing data handling, list-wise deletion was chosen, because, in data file, there was a large number of students who did not complete the questionnaire in a regular way. In other words, lots of students did not answer the items carefully. Therefore, these students were excluded in the analysis. Finally it should be noted that, sampling procedures in the TIMSS-R is independent from the present study. In other words, the present study had no control on the sampling procedures in TIMSS-R. Therefore, a detailed explanation of sampling procedures is not available in this thesis.

3.3 Instruments

The TIMSS-R project basically has an achievement test with multiple forms, a student questionnaire, a teacher questionnaire, and a school principal questionnaire. TIMSS-R instruments were prepared in English and translated into 33 languages (Gonzalez & Miles, 2001). In the present study, student achievement tests and student questionnaires were used.

3.3.1 Achievement Tests

The science and mathematics tests were developed for TIMSS in 1995 by a group of science and mathematics educators. The tests are written tests on science and mathematics achievements. Totally TIMSS achievement tests include 146 science and 163 mathematics items. However, not all of the students were asked to respond all of the items. Instead, they were asked to complete a rotated design of booklets, which included both science and mathematics items. TIMSS-R assessment contained eight booklets with 90 minutes response time, and each student was asked to complete only one booklet including both science and mathematics items (Gonzalez & Miles, 2001). The present study deals only with the science achievement of students.

TIMSS achievement tests include totally 146 both open-ended and objective type science items. The science items can be considered in several content subtitles (Gonzalez & Miles, 2001). Table 3.1 shows the number of items in each content subtitle.

Table 3.1 Number of science items in TIMSS Test

Category	Number of Items
Earth Science	22
Life Science	40
Physics	39
Chemistry	20
Environmental issues	13
Nature of Science	12
Total	146

The tasks in the TIMSS achievement tests can also be divided into several categories, which are in Table 3.2.

Table 3.2 Tasks in TIMSS Achievement Tests

Categories of Tasks in TIMSS Achievement Tests
Understanding
Theorizing
Analyzing
Problem solving
Using tools
Performing routine procedures
Performing science process
Investigating the natural world
Communicating

Table 3.2 shows that the TIMSS achievement test includes not only traditional tasks but also some tasks that are suggested by contemporary approaches.

There are several science achievement scores in the TIMSS-R data. These are raw score on science, standardized score on science, Rasch science score and plausible scores. In the present study, plausible values of science achievement scores of each student were analyzed. In the TIMSS-R data, each student has five plausible scores in science. Students were completed different booklets of achievement test, thus their raw scores were not comparable. These plausible science scores were formed to be able to compare all students completed different booklets. Actually these plausible scores are estimated scores, which assume that each student answers each question. These five plausible scores were used in the present study to form the latent variable called science achievement.

Finally, it should be noted that not all questions of TIMSS achievement tests are released. However, some sample items are available on Internet for researchers.

3.3.2 Questionnaires

In the TIMSS-R, three questionnaires were applied. These are, student questionnaire, teacher questionnaire, and school principal questionnaire (Gonzalez & Miles, 2001). This study used the student questionnaire only.

TIMSS student questionnaire includes 39 items. These items consist of background questions, attitude questions, and the questions about teaching/learning processes. The scales of items change through the items. Majority of items has four alternatives. There are also two or five alternatives in some items. In the TIMSS questionnaire, the items measure either students' opinions of agreement on the content of items or frequency of events in teaching/learning processes (Gonzalez & Miles, 2001). It is possible to reach all of the items in the TIMSS questionnaire on Internet.

3.4 Statistical Analysis

In the present study, mainly two statistical analyses were used. These are factor analysis and structural equation modeling.

3.4.1 Factor Analysis

Factor analysis is a technique, which allows researchers to determine whether many variables can be described by a few factors, and the technique involves a search for clusters of variables, all of which are correlated with each other (Fraenkel & Wallen, 1996).

In the present study, this technique was used to describe lots of items of TIMSS questionnaire (i.e. observed variables) under a few categories in order to form latent variables for the LISREL model.

3.4.2 Structural Equation Modeling

A Structural Equation Model is a composition of patterns of relationships among variables. The idea behind structural equation modeling (SEM) is to formulate a theory about relationships among variables. Kelloway (1998) mentioned that a theory could be thought of as an explanation of why variables are correlated or not correlated. It is necessary to note that correlations, in social sciences, give some idea about causal relations, however, only correlations can not be used to explain the causality between variables (Fraenkel & Wallen, 1996). The process of structural equation modeling can be described in five steps;

1. Model specification
2. Identification
3. Estimation
4. Testing fit
5. Modification

3.4.2.1 Model Specification

Structural equation modeling is inherently a confirmatory technique. SEM is not appropriate for the exploratory identification of relationships. The most important requirement of SEM is the a priori specification of a model (Kelloway, 1998).

At this point, the definitions of some components of a model should be reminded, such as observed variable, latent variable, endogenous variable, exogenous variable, path coefficient. Observed variables are characteristics of individuals measured directly by an instrument. Latent variables are variables that are not measured by the instruments directly, however, they are formed by combination of two or more observed variables. Endogenous variables are variables that the researchers want to explain or predict. Exogenous variables are variables that the researchers use to explain or predict the endogenous variables. When the Structural Equation Modeling technique is performed, it gives a path diagram with the path coefficients among variables on it. Actually, a path coefficient is a standardized multiple regression coefficient. In addition, the square of it can be interpreted as what percentage of variance in an endogenous variable is explained by the combination exogenous variables. (Kelloway, 1998; Hinkle Wiersma and Jurs, 1988)

3.4.2.2 Identification

Second step of SEM involves identification of unknown parameters (e.g. factor loadings or path coefficients) based on observed covariances or correlations. The issues of identification deal with whether a unique solution for the model can be obtained. Models may be underidentified, just-identified, or overidentified. A just-identified model will always provide a unique solution that is, a unique set of path coefficient. When the number of unknowns exceeds the number of equations, the model is said to be underidentified. In this case there is no unique solution. Actually, there are an infinite number of solutions (Kelloway, 1998).

3.4.2.3 Estimation and Testing Fit

When the models are over identified, there are a number of unique solutions, and the task is to find the solution that provides the best fit to the data. Therefore, the identification of a structural equation model is just a matter of the number of estimated parameters. LISREL computer package is used to test the fit of the model (Kelloway, 1998). Model-data fit is traditionally evaluated by chi-square technique. Since it is very sensitive to sample sizes, there are couples of different indexes, which are used to evaluate overall fit. LISREL gives a number of fit indexes, and four of them were reported and discussed in this thesis. These are;

- Root mean square error of approximation (RMSEA)
- Standardized root mean square residual (S-RMR)
- Goodness of fit index (GFI)
- Adjusted goodness of fit index (AGFI)

If a model is said to be a working model, chi-square should not be significant, or chi-square and df should be approximately comparable, and the fit indexes should be within a specific range. Table 3.3 shows the conditions fit indexes.

Table 3.3 Acceptable conditions of fit statistics

Fit indexes	Conditions
RMSEA	< 0.05
S – RMR	< 0.05
GFI	> 0.90
AGFI	> 0.90

Finally, Kelloway (1998) noted that interpretation of fit indexes of a model is a study of validity, however, fit of the model is a necessary but not sufficient requirement for the validity of a model.

3.4.2.4 Modification

The goal of model modification is to improve either the parsimony or the fit of the model. There are two forms of model modification. First, one could delete non-significant paths from the model and, second, one could add new paths based on empirical results (Kelloway, 1998). LISREL suggests some modifications in its written output document.

3.5 Procedures

The procedures of this study may be considered in five steps;

1. Downloading and cleaning the data
2. Principle component analysis of the data
3. Description of latent variables
4. Proposing a model
5. Evaluating the fit of the model

3.5.1 Downloading Converting and Cleaning of the Data

The first step of this study was the obtaining the TIMSS-R data from the official web site of Third International Mathematics and Science Study on Internet (http://isc.bc.edu/). Data files for Turkey, other additional files, and user guides were downloaded from the official web site of TIMSS. Data files are simply the computer files including students' responses to items in

TIMSS questionnaire and their score of TIMSS achievement test. As well as data files, some additional files were downloaded. These are program files and codebook files, which are used to modify data files in order to make them analyzable. Finally, user guides, which include information about how one can analyze TIMSS data, were also downloaded from Internet. After downloading these computer files, data files of TIMSS for Turkey were converted in a form that is readable, editable and analyzable by using SPSS program with the help of codebook and program files, which were also downloaded from Internet. After modification of the data, unrelated items were taken out of the data files. Therefore, data file become smaller and easily analyzable.

3.5.2 Factor Analysis of Data

The second step is performing factor analysis on converted TIMSS data by using SPSS statistical package. The aim of this step is to seek if the questions in TIMSS questionnaire can be clustered into some general groups. The questionnaire items about science were represented by 40 observed variables. These 40 observed variables of TIMSS student questionnaire were analyzed through the principle component analysis with varimax rotation. Varimax rotation revealed nine factors with eigenvalues given in the Table 3.4.

Table 3.4 Eigenvalues of Factors

Factors	Eigenvalues
Factor1	3.377
Factor2	2.836
Factor3	2.756
Factor4	2.678
Factor5	2.248
Factor6	1.999
Factor7	1.888
Factor8	1.845
Factor9	1.663

The following factor structure given in Table 3.5 was obtained as a result of the analysis. It was observed from Table 3.5 that the variables grouped under nine factors. In Table 3.5, the values less than .100 were not indicated. These variables were further evaluated in order to describe the latent variables for the LISREL analysis.

Table 3.5 Factor structure of TIMSS questionnaire

Items	F.1	F.2	F.3	F.4	F.5	F.6	F.7	F.8	F.9
Work in small groups-new topic	**.716**	-	-	-	-	-	-	-	-
Work in pairs or small groups	**.694**	-	-	-	-	-	-	-	-
Students do experiment in class	**.626**	-	-	-	-	-	-	-	-
Work from worksheets	**.595**	.131	-	-	-	.185	-	-	-
Discuss practical problem	**.574**	-	-	.122	.338	-.128	-	-	-
Solve everyday life things	**.530**	-	-	.104	.205	-	-	-.137	-
Have a quiz or test	**.499**	-	-	-	-	.135	-	-	-
Teacher gives demonstration	**.481**	-	-	-	.171	.319	-	-	-
Ask what student know	**.421**	-	-	-	.378	.120	-	-	-
I like science	-	**.730**	-.239	-	-	-	-	-	-
I enjoy learning science	.105	**.727**	-.146	.202	.161	-	.130	-	-
Science is an easy subject	.191	**.663**	-	.118	-	-	-	-	-
I like job involving science	.188	**.566**	-	.401	-	-	-	-	-
I usually do well in science	-	**.535**	.330	-	-	-	.182	-	-
Science is boring	-	**-.507**	.337	-	-	-	-	-	.141
I am just not talented	-	-.171	**.793**	-	-	-	-	-	.118
More difficult than classmate	-	-.167	**.793**	-	-	-	-	-	.131
Science is not of my strengths	-	-.163	**.766**	-	-	-	-	-	-
Would more like if less difficult	-	-.216	**.687**	-	-	-	-	-	.206
Do well to please parents	.132	-	-	**.786**	-	-	-	-	.149
Do well to please self	-	-	-	**.766**	-	-	-	-	-
Do well to get desired job	.104	.201	-	**.691**	-	-	-	-	-
Do well to enter desired school	-	.149	-	**.658**	.174	-	.172	-	-
Science is important in life	.127	.381	-	**.444**	-	-	-	-	-
Teacher explains rules	.119	-	-	-	**.723**	-	-	-	-
Shows how to do problems	-	.146	-	-	**.719**	.139	-	-	-
Copy notes from the board	-	-	-	-	**.575**	.348	-	-	-
Teacher solves examples	.243	-	-	.114	**.547**	.116	-	-	-
Teacher uses board	.107	-	-	-	.228	**.842**	-	-	-
Students use board	-	-	-	-	.195	**.838**	-	-	-
Teacher gives homework	.304	-	-	-	.203	**.445**	-	-	-
Success is important-self	-	.127	-	.125	-	-	**.769**	-	-
Success is important-friends	-	-	-	.122	-	-	**.758**	-	-
Success is important-mother	-	-	-	.112	-	-	**.749**	-	-
Highest education level-father	-	-	-	-	-	-	-	**.834**	-
Highest education level-mother	-	-	-	-	-	-	-	**.794**	-
Number of books at home	-	-	-	-	-	-	-	**.644**	.116
Good luck for success	-	-	.176	-	-	-	-	-	**.797**
Natural talent for success	-	-	.109	-	-	-	-	-	**.776**
Memorize notes for success	-	-	.202	-	-	-	-	-	**.452**

3.5.3 Description of Latent Variables

In the LISREL analysis in order to keep a simple model to evaluate only some of the factors were taken into consideration. Factor selection was carried out through the consideration of eigenvalues, factor loadings and the meanings of the items loadings on the same factor. Moreover, each factor was represented by only three or four observed variables. Selection of observed variables was carried out through the consideration of both factor loadings and the meanings of the items loaded on the same dimension. Consequently, six of the factors with totally 19 observed variables from TIMSS student questionnaire were included to form the latent variables to use in Structural Equation Modeling. These factors were named by examining the common characteristics of the loaded observed variables. Table 3.6 shows the factors that will be used in the Structural Equation Modeling with their eigenvalues and total variance explained.

Table 3.6 Eigenvalues and variances of factors selected to form latent variables

Factor Name	Abbreviation	Eigenvalues	% of Variance
Student-centered classroom activities	STAC	6.287	8.44 %
Teacher-centered classroom activities	TEAC	1.708	5.62 %
Socioeconomic status of students	SES	1.089	4.61 %
Students' perception of success/failure in science	SUCC	2.483	6.89 %
Students' perception of enjoyment of science	ENJY	3.802	7.09 %
Students' perception of importance of success in science	IMPT	1.307	4.72 %

Table 3.7 Descriptions of latent and observed variables

Latent Variables	Observed Variables	Questionnaire Items
STAC	Prac	Students discuss practical problems in the classroom when starting a new topic
	Expr	Students do experiments in science lessons in the classroom
	Smgp	Students work in small groups in the classroom
TEAC	Rule	Teacher starts a new topic by explaining rules and definitions
	Prob	Teacher shows how to solve science problems
	Eg	Teacher solves examples related to new topic
SES	Edfa	Father education level
	Edmo	Mother education level
	Book	Number of books at home
SUCC	Syt1	If science was less difficult, I would like it more
	Syt2	Although I do my best, science is more difficult for me than majority of my friends
	Syt3	No one can be good in every topic, and I am not capable in science
	Syt4	Science is not of my strengths
ENJY	Liks	I like science
	Enjy	I enjoy learning science
	Bore	Science is boring
IMPT	Mip1	My mother thinks that success in science is important
	Fip1	My friends think that success in science is important
	Sip1	I think that success in science is important

Finally, six latent variables were formed by using totally 19 observed variables from the TIMSS questionnaire. Table 3.7 shows the latent variables and corresponding observed variables forming them. Appendix B includes factor loadings of observed variables.

3.5.4 Proposing a LISREL model

After factor analysis, the third step of the study was to perform structural equation modeling analysis. The most crucial sub-step of modeling is to get correlation or covariance matrix. In this study, analyses were done through correlation matrix (see Appendix C). SPSS statistical package were used to get the correlation matrix. Then, a syntax including correlation matrix was written by using LISREL program. Correlation matrix were copied and pasted from SPSS output file to LISREL syntax file. When the LISREL is run, it gives two outputs; a path diagram and an output document. The former is the actual path model, and the latter is a written document including fit indexes and modification indexes for resulting path diagram. Figure 3.1 shows the path diagram, which is going to be tested in the present study.

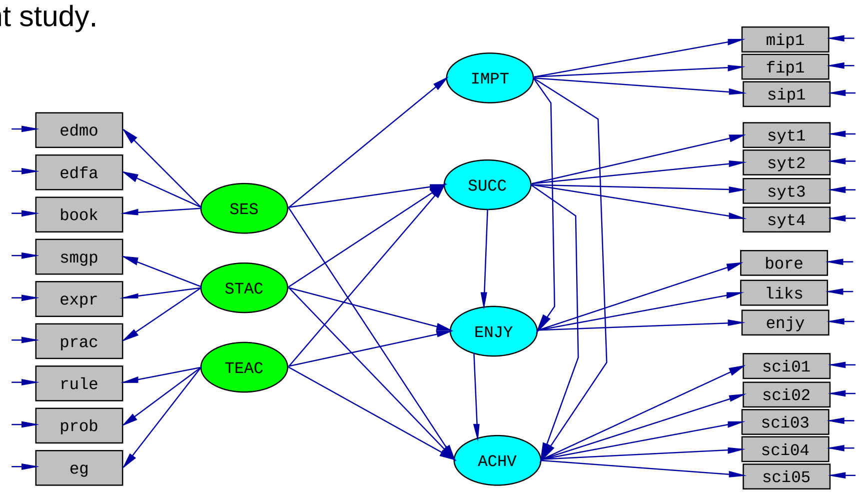

Figure 3.1 Proposed Path Diagram

3.5.5 Evaluating the Fit of the Model

As mentioned above, model-data fit is evaluated by chi-square technique. However, chi-square is very sensitive to sample sizes, thus there are a number of fit indexes that are used to evaluate the overall fit of the model. LISREL gives a number of fit indexes. In the present study, four of them were used to test the fit of the model. These are Root Mean Square Approximation (RMSEA), Standardized Root Mean Square Residual (S-RMR), Goodness of Fit Index (GFI), and Adjusted Goodness of Fit Index (AGFI).

If a model is said to be fitted with data, chi-square and df should be comparable, and fix indexes should be in specific conditions. These conditions were given in Table 3.3.

At this point, it should be noted that modeling is a repetitive process including trial and error. In other words, researchers propose a lot of models in time until a best-fit model is reached. This is done through the modification indexes given in the written output document of LISREL.

3.6 Validity and Reliability

Validity and reliability of the TIMSS data were actually studied by the developers of this international study. Nevertheless, validity and reliability of the LISREL model resulted in the present study should be discussed.

3.6.1 Validity

According to Fraenkel and Wallen (1996), validity is the degree to which correct inferences can be made based on the results from an instrument, and it depends not only on the instrument itself, but also on the instrumentation process and the characteristics of the sample. According to Kelloway (1998), goodness of fit of a model is a validity evidence for the path models. However goodness of fit is necessary but not sufficient requirement for the validity of Structural Equation Modeling. Goodness of fit of the model developed in the present study was discussed in Chapter 4.

The other validity evidence for the model is the factor structure of the data used. In the present study, after getting and modification of the TIMSS-R data, the second step was factor analysis of the TIMSS-R questionnaire data. The main purpose of factor analysis was to create a base for Structural Equation Modeling by describing some factors in order to form latent variables. On the other hand, a second purpose of the factor analysis is to provide an evidence for the construct validity of TIMSS results (Hinkle et al., 1988). Factor structure of the TIMSS questionnaire items was discussed above in this chapter.

3.6.2 Reliability

According to Fraenkel and Wallen (1996), reliability is the degree to which the scores obtained from an instrument are consistent whatever the instrument measures. Reliability study was performed through the internal consistencies of the factors. The internal consistency values (Cronbach-alpha) of factors are shown in Table 3.8.

Table 3.8 Internal consistency of the factors

Factors	Observed Var.	Cronbach-alpha
STAC	Prac, Expr, Smgp	0.6015
TEAC	Rule, Prob, Eg	0.5978
SES	Edfa, Edmo, Book	0.6531
SUCC	Syt1, Syt2, Syt3, Syt4	0.8164
ENJY	Liks, Enjy, Bore	0.7184
IMPT	Mip1, Fip1, Sip1	0.6822

Even though some of the reliabilities are around 0.60 in some of the latent variables, these values were interpreted as satisfactory for such a small number of observed variables.

CHAPTER 4

RESULTS

In this section the relationships among the latent variables, as indicated in the proposed model will be evaluated with respect to fit indexes, significance of the proposed relationships, and amount of explained variances.

4.1 Structural Equation Modeling

Totally 19 observed variables from TIMSS student questionnaire and five plausible scores of TIMSS science test were examined through the Structural Equation Modeling in this study. Descriptive statistics for observed variables are given in Appendix A and Appendix D. Latent variables, which are formed by 19 observed variables from the TIMSS questionnaire, are as follows;

- Student-centered classroom activities (STAC)
- Teacher-centered classroom activities (TEAC)
- Socio-economic status of students' family (SES)
- Perception of success/failure in science (SUCC)
- Enjoyment of science (ENJY)

- Importance of success in science (IMPT)

Among these factors, STAC, TEAC, and SES can be identified as exogenous factors, and SUCC, ENJY, and IMPT can be identified as endogenous factors. Because, in the present study, it was generally intended to explain the students' science achievement (ACHV), SUCC, ENJY, and IMPT by using STAC, TEAC, and SES.

It was mentioned in the method section that it is necessary to get correlation or covariance matrix of observed variables to perform structural equation modeling. In this study, modeling was performed through the correlation matrix formed by using SPSS program. After forming the correlation matrix of observed variables, it was copied on a syntax file of LISREL program. The printed form of this syntax file is in Appendix C. The resulting path diagram with the best fit indexes is in Figure 4.1. In addition, fit indexes of this path model are available in Table 4.1.

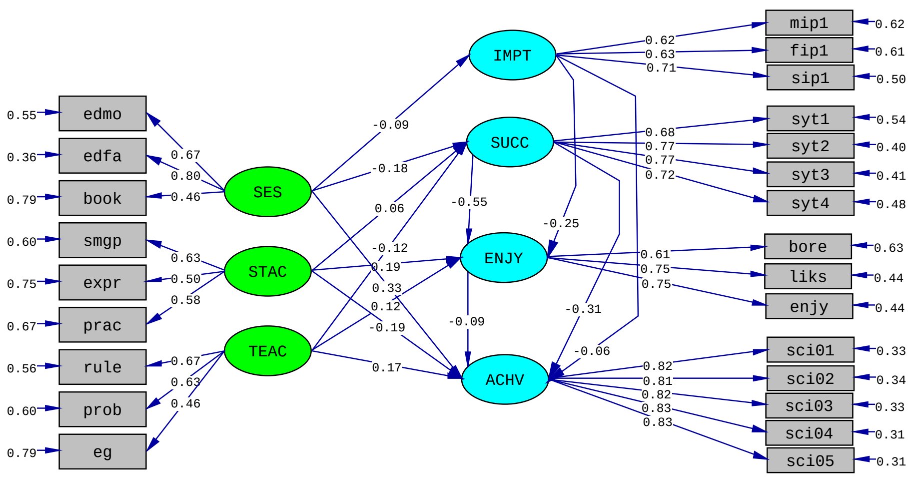

Figure 4.1 Path Diagram

The relationships given in the diagram above were further evaluated with respect to t-statistics. According to Figure 4.2, all the relationships defined in the model were found as significant at 0.05 level of significance.

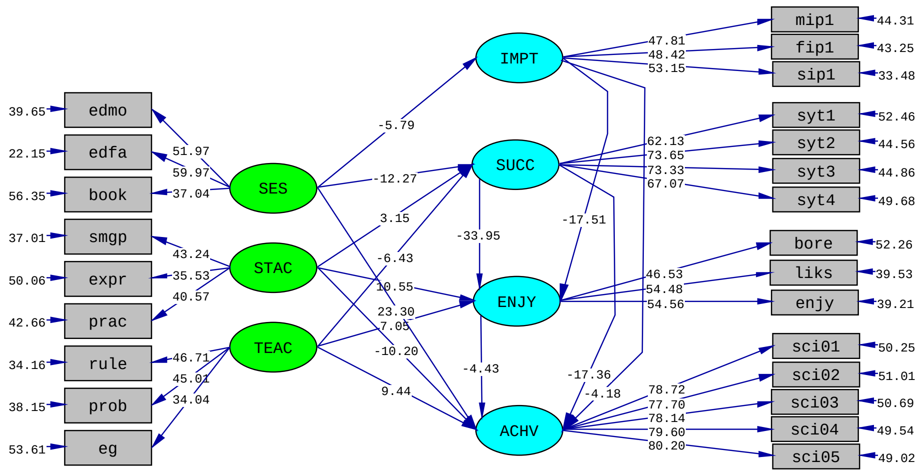

Figure 4.2 Significance of Relationships in the Model (t-values)

In the model, the relationship between socioeconomic status (SES) and students' science achievement (ACHV) is the one of the highest relationships. Path coefficient of this relationship is 0.33 and t-value of this path is 23.30. It means that as SES indicators increase students science achievement increases.

Another highest relationship is between science achievement (ACHV) and the perception of success/failure in science (SUCC) with the path coefficient of –0.31 and t-value of –17.36. It means that if students perceive themselves as successful in science, they are successful in the TIMSS science achievement test.

Also, the relationship between teacher-centered classroom activities (TEAC) and science achievement (ACHV) are significantly related to science achievement. The path coefficient of TEAC-ACHV path is 0.17. And the t-value of this path is 9.44. The path coefficient of STAC-ACHV path is –0.19 with the t-value of –10.20. According to the model, the science achievement (ACHV) has a negative relationship with the student-centered classroom activities (STAC). It means that teacher-centered classroom activities has more impact on science achievement than student-centered classroom activities, and student-centered classroom activities has a negative impact on science achievement of students.

The science achievement seems to have a low or no relationships with enjoyment (ENJY) and importance of success in science (IMPT) according to the model. The path coefficient of ENJY-ACHV is –0.09 with the t-value of –4.43. And the coefficient of IMPT-ACHV is –0.06 with the t-value of –4.18. It means that students can be successful in science even if they do not enjoy science or give no importance to success in science.

Enjoyment (ENJY) of science seems to have higher relationship with the student-centered classroom activities (STAC) than teacher-centered classroom activities (TEAC). The path coefficient of ENJY-STAC is 0.19 with the t-value of 10.55. Also the coefficient of ENJY-TEAC is 0.12 with the t-value of 7.05. It means that student-centered classroom activities have more impact on enjoyment than teacher-centered classroom activities.

According to the LISREL model, enjoyment of science (ENJY) has a very high relationship with the perception of success/failure in science (SUCC). The coefficient of ENJY-SUCC is –0.55 with the t-value of –33.95. It means that if students enjoy science, they also perceive themselves as successful in science. And t-value shows that this relationship is significant.

There seems to be a negative relationship between enjoyment of science (ENJY) and importance of success in science (IMPT). The coefficient of ENJY-IMPT is –0.25 with t-value of –17.51. It means that the students give importance to success in science even if they do not enjoy this subject.

According to the model, perception of success/failure in science (SUCC) has more relationship with teacher-centered classroom activities (TEAC) than student-centered classroom activities (STAC). The coefficient of SUCC-TEAC is –0.12 with the t-value of –6.43. And the coefficient of SUCC-STAC is 0.06 with the t-value of 3.15. The students in a teacher-centered classroom perceive themselves as more successful than those in student-centered classroom.

4.2 Goodness of Fit of the Model

In Chapter 4, fit of the structural models was explained theoretically. Fit of the models is tested through chi-square value and several fit indexes. Actually, chi-square and df should be comparable (Kelloway, 1998). In the present study, chi-square and df values are 2989 and 251 which seem to be problematic. However, it is known that chi-square and df are affected from the sample size. Large sample size of this study may affect these indexes negatively.

In this study, fit indexes presented in Figure 4.1 were used to examine goodness of fit of the model. In Chapter 4, it was mentioned that if a model is a fitted model, its fit indexes should be in some specific ranges. These ranges were given in Table 3.3. According to the Table 4.1, all of fit indexes are in the acceptable criterion.

Table 4.1 Fit indexes of model and acceptable criteria

Fit index	Value	Criterion
RMSEA	0.039	< 0.05
S-RMR	0.045	< 0.05
GFI	0.97	> 0.90
AGFI	0.96	> 0.90

An additional fit measure is R^2 or “coefficient of determination”. Kelloway (1998) noted that the coefficient of determination (R^2) could be reported as an index of overall fit. It is also known that R^2 is a measure of variation in latent variables that is attributed to the combination of observed variables. R^2 values of endogenous variables were given in Table 4.2. It should finally be noted that a well-fitted model could quite possibly explain a modest variance in endogenous variables (Kelloway, 1998).

Table 4.2 Coefficient of determination (R^2) of endogenous variables

Variable	R^2
IMPT	0.01
SUCC	0.04
ENJY	0.46
ACHV	0.28

According to Table 4.2, path model explains 1% of variance in IMPT, and 4% of variance in SUCC, which are small amount of total variance. However, the model explains 46% of variance in ENJY, and 28% of variance in ACHV, which are moderate amounts of total variance (Kelloway, 1998).

CHAPTER 5

DISSCUSSION

The purpose of this study is to model the relationships between science achievement of eighth grade Turkish students and some factors affecting it, such as instructional activities, socio-economic status (SES) of students, and attitude toward science by using TIMSS data.

5.1 Discussion of Results of Structural Equation Modeling

TIMSS has many important results for most of the participating countries. Specifically for Turkey, there are several lessons that should be taken from TIMSS. The relative standing of Turkey in TIMSS-R, which is significantly below the international average, creates some doubts about the quality of science education system in our country. This study aimed to explain the patterns of relationships between students' science achievement and some factors affecting it. The findings could be used to change some critical issues in the system that may cause enhancement in science achievement in the long run.

One of the largest relationships in the model was found between science achievement and socioeconomic status (SES) of students. In other words, although SES of student is independent from science instruction, it has more impact on science achievement than the impact of instructional activities. In the present study parental educational level with the number of books at home were taken as the indicators of SES. Naturally, educationally rich environment has impact on students' academic performances. This strong relation could be used as an advantage to enhance achievement of students throughout the cooperation between parents and the school system. Family support programs may foster academically sound parental activities that could be reflected into students' academic performances.

One of the most striking results of the model is the negative relationship between student science achievement and student-centered classroom activities. Several reasons for this negative relationship can be stated with respect to some known facts about students and science education in Turkey. The first impression about Turkish students is that they are accustomed to teacher-centered instruction that generally base on lecturing. Therefore, Turkish students may not sufficiently benefit from the student-centered activities, which are not familiar for them. Students traditionally are receptive and passive learners. Any activity that requires active participation may not be as effective as in any other western country in the world. Another dimension of this relationship is related to teacher characteristics. Science teachers may have some difficulties in performing student-centered classroom activities efficiently. Even though students claim that they discuss the topics, perform pair works and experiments, in the classroom, the quality of the discussions and group works are always questionable. Another possible reason for the negative relationship between science achievement and student-centered classroom activities is that the TIMSS achievement test did not seem to be measure the outcomes of student-centered instruction such as some higher order thinking skills and social skills. The negative relationship between science achievement and student-centered classroom activities has the agreement with the some studies in the literature. Leung (2002) observed a negative relationship between science achievement and student-centered teaching/learning processes in East Asian countries. Considering the Leung's study and finding of the present study it could be said that in a traditional system a new methodology should be implemented with a special care by considering student expectations and teacher qualifications.

On the contrary, analyses showed that teacher-centered classroom activities had a positive impact on students' science achievement. It means that students with teachers, who solve examples and explain rules, are successful in the TIMSS. The first possible reason for this result is about the TIMSS achievement test. As mentioned above, the TIMSS achievement test seemed to be appropriate for measuring the outcomes of teacher-centered instruction. Another possible reason for this relationship is that teacher-centered instructional environment may be more appropriate for our culture. This result indicates the importance of teacher and the methodology they use in the classroom.

The LISREL model developed in the present study also showed that teacher-centered classroom activities had more positive impact on students' perception of success/failure in science than the impact of student-centered classroom activities. It means that students who perceive themselves as successful in science come more likely from teacher-centered atmosphere. This result emphasizes the importance of teachers once more in the LISREL model tested in the present study.

Another important factor that is influential on science achievement was found as the students perceptions of success/failure in science. It means that the students who perceive themselves as successful in science are really high achiever in science. This result shows agreement with the literature. Shen (2002), and Shen and Pedulla (2000) who noted positive relationships between science achievement and self-perceptions. At this point, it should be added that there seems to be relatively high relationship between students' perception of success/failure in science and SES. It means that students who perceive themselves as successful come more likely from the families with high SES. These results indicate the importance of students' self-perceptions in teaching/learning processes.

Another result, which attracts attention, is no or very low relationship between students' science achievement and their perception of enjoyment of science. In other words, according to the model, students may be successful in science, even if they do not enjoy science. This is an unexpected result, however there may be some reasons which we can address to Turkish Educational System. In Turkey, students have several "high stake" central exams during their education life. These exams are so vital and important for students and parents since it is perceived as a tool for social mobility. Therefore, Turkish students are motivated to be high achievers in many topics in order to be successful in central exams even if they do not enjoy these topics.

When evaluating Turkey's science score in TIMSS, some critical features of the TIMSS international assessment should be taken into consideration. Ramseier (1999) stressed the disharmony between content of the TIMSS science test and science curricula of Switzerland. For the Turkish case, the content of the released items could be evaluated in terms of appropriateness with respect to the content of the school curricula. In addition, the translation effects in international assessment, which Ercikan (1998) discussed, should also be considered for the Turkish versions of the test and questionnaire items. And finally, Bracey (1998) discussed the cultural differences that possibly affect the relative achievement scores of different countries in an international assessment. Since there is no cross-cultural comparison carried out in the present study the findings could be interpreted in the one cultural context only.

5.2 Internal Validity

Like all studies, there are some threats to internal validity of this study. When examining internal validity of a correlational study, the major problem to be considered is that some extraneous variables may explain any results obtained. A researcher who conducts a correlational study should always be alert to alternative explanations for relationships found (Fraenkel and Wallen, 1996). According to this principle, the threats to internal validity of this study are discussed one by one below.

5.2.1 Subject Characteristics

Some characteristics of students, which are not controlled, may explain any relationships found in this study. This threat does not seem to be possibly effective in this study, because sample size of the TIMSS for Turkey is 7841. Such a large sample eliminates this threat.

5.2.2 Location

Non-included characteristics of a specific location where TIMSS applied in Turkey may explain any relationships observed in this study. Actually, there are many locations where the TIMSS applied in Turkey. Therefore, non-included characteristics of a specific location would not have an important effect on the results.

5.2.3 Data Collector Characteristics and Data Collector Bias

Non-included characteristics of a specific data collector may explain any relationships observed in this study. Actually, there are many data collectors in the TIMSS. Therefore, non-included characteristics of a specific data collector would not have an important effect on the results. Like in data collector characteristics, data collector bias was eliminated due to large number of data collectors in the TIMSS.

5.3 Implications

Several implications of this study may be mentioned, and some of them were explained in the lines below.

- Perhaps, the most important implication of this study is for the faculties of education in Turkey. In our country, it is a common fact that the teachers have some difficulties in performing an effective teaching/learning process, which was also showed in the present study. Therefore teacher-training system of our country should be revised, and preferably have more pre-service teaching practice and emphasize the student-centered classroom activities. However, even though the textbooks and curricula are designed with respect to student-centered activities, in practice it is always questionable to implement these methods for enhancing students' achievement. This points out the importance of teacher trainings in the Turkish educational system.
- In our country, it is also possible to claim that the majority of the teachers do not apply the techniques that they learned in their undergraduate education. In the present study, it was found that teachers perform teacher-centered activities in the classroom more frequently than student-centered activities. Teachers should increase the variety of instructional activities they apply.
- The results of the present study showed that students' self-perceptions have an important impact on their science achievement. This finding emphasizes the importance of students' self-perceptions and other affective characteristics in teaching/learning processes. Therefore, teachers should take into consideration the students' self-perceptions and all the other affective characteristics.

- The present study showed that teacher-centered classroom activities had a positive significant impact on students' science achievement. Therefore, teacher may efficiently perform such activities to increase the science achievement of students in their classrooms. This finding emphasizes the importance of teacher trainings and characteristics as mentioned before. However, it does not mean that student-centered activities should be given up.
- This study showed that one of two highest relationships in the model was between science achievement and SES of the students. According to this result, one can state that a good scholarship program may decrease the impact of SES on science achievement along with the increased cooperation between parents and school administration.
- Another impression of science education in our country is about the disharmony between curricula and real teaching/learning practices. When the science curricula and textbooks are evaluated closely, they seem to be somehow student-centered on paper, because science curricula and textbooks include lots of science experiments for students, and they suggest many discussion topics as well. However, school and class atmosphere seem to be still quite authoritarian and teacher-centered, especially in rural areas. In other words, there are some contradictions between science curricula and real teaching/learning processes in schools. In such an atmosphere, an efficient student-centered teaching/learning processes can hardly be performed. Therefore, the Ministry of

National Education should take more precautions in order to decrease the disharmony between science curricula and the real instructional practices.

5.4 Suggestions for Further Research

This study showed the significant impact of SES on science achievement of students once again. Therefore, the researchers should always take into account the effects of SES when studying achievement. Since SES is a very dominant variable to explain achievement models without SES variables could be developed in the further studies in order to understan the impact of school related activities on students achievemnt.

When putting an end to this study, one more suggestion for researchers may be mentioned. TIMSS-R had nearly 40 countries participated, thus the data from TIMSS-R is very appropriate for cross-cultural studies. Researchers should conduct more studies in the cross-cultural context.

REFERENCES

Beaton, A. E. (1998). Comparing Cross-National Student Performance on TIMSS Using Different Test Items. *International Journal of Educational Research, 29*, 529-542.

Biddle, B. J., & Marlin, M. M. (1987). Causality, Confirmation, Credulity, and Structural Equation Modeling. *Child Development, 58*, 4-17.

Bos, K., & Kuiper, W. (1999). Modeling TIMSS Data in a European Comparative Perspective: Exploring Influencing Factors on Achievement in Mathematics in Grade 8. *Educational Research and Evaluation, 5*, 57-179.

Bracey, G. W. (1998). Tinkering with TIMSS. *Phi Delta Kappan, 80*, 32-38.

Cheng, Y. C., & Cheung, W. M. (1999). Lessons from TIMSS in Europe: An Observation from Asia. *Educational Research and Evaluation, 5*, 227-236.

Ercikan, K. (1998). Translation Effects in International Assessments. *International Journal of Educational Research, 29*, *543-553.*

Fensham, P. J. (1999). International Success, But Is It Science? Identifying Strengths and Weaknesses in Australian Primary School Science from TIMSS and Other Data. *Australian Science Teacher Journal, 45*, 39-45.

Fraenkel, J. R., & Wallen, N. E. (1996). *How to Design and Evaluate Research in Education* (3rd ed). New York: McGraw-Hill.

Gonzalez, E. J., & Miles, J. A. (2001). *TIMSS 1999 User Guide for the International Database.* International Association for the Evaluation of Educational Achievement. Boston, MA.

Harris, S. (1999). TIMSS Performance Assessment: The Strengths and Weaknesses of Students in England. *Studies in Educational Evaluation, 25*, 287-295.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1988). *Applied Statistics for the Behavioral Sciences* (2nd ed). Boston: Houghton Mifflin Company.

House, J. D. (2000a). Relationships Between Instructional Activities and Science Achievement of Adolescent Students in Hong Kong: Finding from the Third International Mathematics and Science Study (TIMSS). *Studies in Educational Evaluation, 27*, 275-289.

House, J. D. (2000b). 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.

Jakwerth, P. (1999). TIMSS Performance Assessmnent Result United States. *Studies in Educational Evaluation, 25*, 277-281.

Jones, R. W. (1999). The Third International Mathematics and Science Study: Anatomy of This International Comparative Study in Science Education. *Australian Science Teachers Journal, 45*, 27-36.

Kelloway, E. K. (1998). *Using LISREL for Structural Equation Modeling A Researcher's Guide.* London: SAGE Publications.

Kuiper, W., & Plomp, T. (2001). National and Cross-National Perspective on TIMSS Populations 2 and 3 Findings. *Studies in Educational Evaluation, 27*, 1-6.

Leung, F. K. (2002). Behind the High Achievement of East Asian Students. *Educational Research and Evaluation, 8*, 87-108.

Lokan, J. (1999). Equity Issues in Testing: A Case of TIMSS Performance Assessment. *Studies in Educational Evaluation, 25*, 297-314.

Martin, J. A. (1987). Structural Equation Modeling: A Guide for the Perplexed. *Child Development, 58*, 33-37.

Martin, M. O., Mullis, V. S., Beaton, A. E., Gonzalez, E. J., Smith, T. A., & Kelly, D. L. (1997). *Science Achievement in the Primary School Years. IEA's Third International Mathematics and Science Study (TIMSS).* International Association for the Evaluation of Educational Achievement. Boston, MA.

McKnight, C. C., & Schmidt, W. H. (1998). Facing Facts in U.S. Science and Mathematics Education: Where We Stand, Where We Want to Go. *Journal of Science Education and Technology, 7*, 57-76.

Menon, R. (2000). Should the United States Emulate Singapore's Education System to Achieve Singapore's Success in the TIMSS. *Mathematics Teaching in Middle School, 5*, 345-348.

O'Leary, M. (2001). Item Format as a Factor Affecting the Relative Standing of Countries in the Third International Mathematics and Science Study (TIMSS). *Annual Meeting of the American Educational Research Association,* (Seattle, WA).

Papanastasiou, C. (2000). Effects of Attitudes and Beliefs on Mathematics Achievement. *Studies in Educational Evaluation, 26*, 27-42.

Papanastasiou, C. (2002). School, Teaching and Family Influence on Student Attitudes Toward Science: Based on TIMSS Data for Cyprus. *Studies in Educational Evaluation, 28*, 71-86.

Park, C., & Hong, M. (2002). A Relative Effectiveness of Item Types for Estimating Science Ability. *Journal of the Korean Association for Research in Science Education, 22*, 122-131.

Peak, L. (1996). *Pursuing Excellence: Eight-Grade Mathematics and Science Teaching, Learning, Curriculum, and Achievement in International Context. Initial Findings from the Third International*

Mathematics and Science Study (TIMSS). National Center for Education Statistics, Washington, DC.

Ramseier, E. (1999). Task Difficulty and Curricular Priorities in Science: Analysis of Typical Features of the Swiss Performance in TIMSS. *Educational Research and Evaluation, 5*, 105-126.

Shen, C. (2002). Revisiting the Relationship Between Students Achievement and their Self-perceptions: A Cross-National Analysis Based on TIMSS 1999 Data. *Assessment in Education, 9*, 161-184.

Shen, C., & Pedulla, J. J. (2000). The Relationship Between Students' Achievement and their Self-perception of Competence and Rigour of Mathematics and Science: A Cross-national Analysis. *Assessment in Education, 7*, 237-253.

Smith, T. A. (1997). The Generalizability of Scoring TIMSS Open-Ended Items.). *Annual Meeting of the American Educational Research Association,* (Chicago, IL).

Stricker, L. J., & Rock, D. A. (1995). Examinee Background Characteristics and GRE General Test Performance. *Intelligence, 21*, 49-67.

Wang, H. A., & Schmidt, W. H. (2001). History, Philosophy and Sociology of Science in Science Education: Results from the Third International Mathematics and Science Study. *Science & Education, 10*, 51-70.

Wang, J. (1998). A Content Examination of the TIMSS Items. *Phi Delta Kappan, 80*, 36-39.

Wang, J., & Zhu, C. (2003). An In-Dept Analysis of Achievement Gaps Between Seventh and Eighth Grades in the TIMSS Database. *School Science and mathematics, 103*, 186-193.

Webster, B. J., & Fisher, D. L. (2000). Accounting for Variation in Science and Mathematics Achievement: A Multilevel Analysis of Australian Data Third International Mathematics and Science Study. *School Effectiveness and School Improvement, 11*, 339-360.

Welsh, M., Parke, R. D., Widaman, K., & O'Neil, R. (2001). Linkages Between Childeren's Social and Academic Competence: A Longitudinal Analysis. *Journal of School Psychology, 39*, 463-481.

Wester, A., & Henriksson, W. (2000). The Interaction between Item Format and Gender Differences in Mathematics Performance Based on TIMSS data. *Studies in Educational Evaluation, 26*, 79-90.

Woessmann, L., & West, M. R. (2002). *Class-Size Effects in School Systems around the World: Evidence from Between-Grade Variation in TIMSS.* Harward University, Cambridge, MA.

Wolf, R. M. (1998). Validity Issues in International Assessments. *International Journal of Educational Research, 29*, 491-501.

Vlaardingerbroek, B., & Taylor, T. G. (2003). Teacher Education Variables as Correlates of Primary Science Ratings in Thirteen TIMSS

Systems. *International Journal of Educational Development, 23*, 429-438.

Yore, L. D., Anderson, J. O., & Shymansky, J. A. (2002). Modeling the Relationships of Classroom Characteristics and Student Attributes to Students' Science Achievement. *Annual Meeting of the National Association for Research in Science Teaching,* (New Orleans, LA)

Zuzovsky, R. (1999). Problematic Aspects of the Scoring System of the TIMSS Practical Performance Assessment: Some Examples. *Studies in Educational Evaluation, 25*, 315-323.

APPENDIX A

FREQUENCIES AND PERCENTAGES OF THE ITEMS OF TIMSS QUESTIONNAIRE

Table A.1 Mother Education Level (edmo)

Alternatives	Frequencies	Percentages
Some primary school	2100	26.8
Finish primary school	4026	51.3
Some secondary school	221	2.8
Finish secondary school	639	8.1
Some vocational education	54	0.7
Some university	18	0.2
Finish University	169	2.2
Missing Value	614	7.8
TOTAL	7841	100

Table A.2 Father Education Level (edfa)

Alternatives	Frequencies	Percentages
Some primary school	771	9.8
Finish primary school	3854	49.2
Some secondary school	408	5.2
Finish secondary school	1120	14.3
Some vocational education	171	2.2
Some university	78	1.0
Finish University	635	8.1
Missing	804	10.3
TOTAL	7841	100

Table A.3 Number of books at home (book)

Alternatives	Frequencies	Percentages
0 – 10 books	1643	21.0
11 – 25 books	2816	35.9
26 – 100 books	2150	27.4
101 – 200 books	650	8.3
More than 200 books	469	6.0
Missing	113	1.4
TOTAL	7841	100

Table A.4 My mother thinks that success in science is important (mip1)

Alternatives	Frequencies	Percentages
Strongly agree	4105	52.4
Agree	2774	35.4
Disagree	273	3.5
Strongly disagree	125	1.6
Missing	564	7.2
TOTAL	7841	100

Table A.5 My friends think that success in science is important (fip1)

Alternatives	Frequencies	Percentages
Strongly agree	3617	46.1
Agree	3133	40.0
Disagree	421	5.4
Strongly disagree	120	1.5
Missing	550	7.0
TOTAL	7841	100

Table A.6 I think that success in science is important (sip1)

Alternatives	Frequencies	Percentages
Strongly agree	4967	63.3
Agree	2134	27.2
Disagree	152	1.9
Strongly disagree	96	1.2
Missing	492	6.3
TOTAL	7841	100

Table A.7 If science was less difficult. I would like it more (syt1)

Alternatives	Frequencies	Percentages
Strongly agree	1271	16.2
Agree	2242	28.6
Disagree	2890	36.9
Strongly disagree	971	12.4
Missing	467	6.0
TOTAL	7841	100

Table A.8 Although I do my best. science is more difficult for me than majority of my friends (syt2)

Alternatives	Frequencies	Percentages
Strongly agree	823	10.5
Agree	2174	27.7
Disagree	3198	40.8
Strongly disagree	1047	13.4
Missing	599	7.6
TOTAL	7841	100

Table A.9 No one can be good in every topic and I am not capable in science (syt3)

Alternatives	Frequencies	Percentages
Strongly agree	730	9.3
Agree	1962	25.0
Disagree	3371	43.0
Strongly disagree	1150	14.7
Missing	628	8.0
TOTAL	7841	100

Table A.10 Science is not one of my strengths (syt4)

Alternatives	Frequencies	Percentages
Strongly agree	866	11.0
Agree	2196	28.0
Disagree	2968	37.9
Strongly disagree	1165	14.8
Missing	649	8.3
TOTAL	7841	100

Table A.11 I like science (liks)

Alternatives	Frequencies	Percentages
Like a lot	2449	31.2
Like	4040	51.5
Dislike	744	9.5
Dislike a lot	273	3.5
Missing	335	4.3
TOTAL	7841	100

Table A.12 I enjoy learning science (enjy)

Alternatives	Frequencies	Percentages
Strongly agree	2773	35.4
Agree	3602	45.9
Disagree	831	10.6
Strongly disagree	226	2.9
Missing	409	5.2
TOTAL	7841	100

Table A.13 Science is boring (bore)

Alternatives	Frequencies	Percentages
Strongly agree	360	4.6
Agree	1169	14.9
Disagree	3952	50.4
Strongly disagree	1800	23.0
Missing	560	7.1
TOTAL	7841	100

Table A.14 Students do experiments in science lessons in the classroom (expr)

Alternatives	Frequencies	Percentages
Almost always	990	12.6
Pretty often	1194	15.2
Once in a while	2940	37.5
Never	2122	27.1
Missing	595	7.6
TOTAL	7841	100

Table A.15 Students discuss practical problems in the classroom when starting a new topic (prac)

Alternatives	Frequencies	Percentages
Almost always	1375	17.5
Pretty often	1753	22.4
Once in a while	2817	35.9
Never	1351	17.2
Missing	545	7.0
TOTAL	7841	100

Table A.16 Students work in small groups in the classroom (smgp)

Alternatives	Frequencies	Percentages
Almost always	837	10.7
Pretty often	953	12.2
Once in a while	2743	35.0
Never	2708	34.5
Missing	600	7.7
TOTAL	7841	100

Table A.17 Teacher shows how to solve science problems (prob)

Alternatives	Frequencies	Percentages
Almost always	4409	56.2
Pretty often	2429	31.0
Once in a while	660	8.4
Never	114	1.5
Missing	229	2.9
TOTAL	7841	100

Table A.18 Teacher starts a new topic by explaining rules and definitions (rule)

Alternatives	Frequencies	Percentages
Almost always	4436	56.6
Pretty often	2063	26.3
Once in a while	841	10.7
Never	176	2.2
Missing	325	4.1
TOTAL	7841	100

Table A.19 Teacher solves examples related to new topic (eg)

Alternatives	Frequencies	Percentages
Almost always	2730	34.8
Pretty often	2434	31.0
Once in a while	1715	21.9
Never	465	5.9
Missing	497	6.3
TOTAL	7841	100

APPENDIX B

FACTOR LOADINGS

Table B.1 Student-centered classroom activities (STAC)

Observed Variables	Factor Loadings
Smgp	0.716
Expr	0.616
Prac	0.574

Table B.2 Teacher-centered classroom activities (TEAC)

Observed Variables	Factor Loadings
Rule	0.723
Prob	0.719
Eg	0.547

Table B.3 Socio-economic status of students' family (SES)

Observed Variables	Factor Loadings
Edfa	0.834
Edmo	0.794
Book	0.644

Table B.4 Perception of success/failure in science (SUCC)

Observed Variables	Factor Loadings
syt3	0.793
syt2	0.793
syt4	0.766
syt1	0.687

Table B.5 Enjoyment of science (ENJY)

Observed Variables	Factor Loadings
Liks	0.730
Enjy	0.727
Bore	-0.507

Table B.6 Importance of success in science (IMPT)

Observed Variables	Factor Loadings
sip1	0.769
fip1	0.758
fip1	0.749

APPENDIX C

LISREL SYNTAX OF THE PATH MODEL

```
TIMSS SCIENCE
Observed Variables
edmo edfa book mip1 fip1 sip1 syt1 syt2 syt3 syt4 bore liks enjy smgp expr
prac rule prob eg sci01 sci02 sci03 sci04 sci05
Correlation Matrix
 1.0000  0.5387  1.0000  0.2819  0.3770  1.0000 -0.0703 -0.0483
-0.0750  1.0000 -0.0168  0.0079 -0.0196  0.3952  1.0000 -0.0480
-0.0412 -0.0552  0.4326  0.4418  1.0000 -0.0691 -0.0866 -0.0674
 0.0874  0.0491  0.0750  1.0000 -0.0926 -0.1039 -0.0894  0.1091
 0.0448  0.0970  0.5710  1.0000 -0.1055 -0.1161 -0.1102  0.1153
 0.0668  0.1177  0.4856  0.5822  1.0000 -0.0904 -0.1216 -0.0987
 0.1072  0.0648  0.1228  0.4496  0.5336  0.5978  1.0000  0.0712
 0.0689  0.0763 -0.1231 -0.1146 -0.1493 -0.3576 -0.3429 -0.3548
-0.3434  1.0000  0.0521  0.0597  0.0768 -0.1465 -0.1372 -0.1975
-0.3354 -0.3419 -0.3233 -0.3458  0.4403  1.0000  0.0579  0.0436
 0.0560 -0.1990 -0.1976 -0.2265 -0.2692 -0.2769 -0.2727 -0.2787
 0.4606  0.5858  1.0000 -0.0514 -0.0571  0.0051 -0.0029 -0.0283
 0.0003  0.0233  0.0314  0.0451  0.0143  0.0134  0.0846  0.1181
 1.0000  0.0179 -0.0283  0.0176 -0.0282 -0.0536 -0.0307  0.0008
-0.0036  0.0039 -0.0188  0.0261  0.1246  0.1504  0.3286  1.0000
 0.0548  0.0323  0.0627 -0.0368 -0.0654 -0.0338 -0.0119 -0.0238
-0.0142 -0.0355  0.0758  0.1114  0.1629  0.3696  0.2589  1.0000
 0.0380  0.0638  0.0810 -0.1107 -0.1186 -0.1253 -0.0320 -0.0723
-0.0697 -0.0739  0.0960  0.1095  0.1513  0.1004  0.1160  0.2607
 1.0000  0.0341  0.0507  0.0595 -0.1096 -0.1287 -0.1235 -0.0374
-0.0589 -0.0535 -0.0616  0.1108  0.1583  0.2206  0.0967  0.1421
 0.1894  0.4341  1.0000 -0.0046  0.0004  0.0235 -0.0756 -0.0909
-0.0761 -0.0140 -0.0161 -0.0233 -0.0396  0.0845  0.1115  0.1358
 0.1902  0.1679  0.2032  0.2994  0.2751  1.0000  0.2418  0.2662
 0.1844 -0.0542 -0.0142 -0.0889 -0.2142 -0.2498 -0.2571 -0.1982
 0.1219  0.0963  0.0734 -0.1588 -0.0459  0.0318  0.1035  0.0646
-0.0114  1.0000  0.2267  0.2451  0.1852 -0.0893 -0.0379 -0.0944
-0.1655 -0.2032 -0.2334 -0.1682  0.1200  0.0874  0.0565 -0.1771
-0.0433  0.0495  0.1222  0.0791 -0.0395  0.6566  1.0000  0.2225
 0.2468  0.1873 -0.0674 -0.0148 -0.1014 -0.1921 -0.2300 -0.2632
```

```
-0.1981  0.1116  0.0877  0.0577 -0.1331 -0.0714  0.0071  0.1150
 0.0970  0.0045  0.6800  0.6556  1.0000  0.2402  0.2409  0.1775
-0.1049 -0.0377 -0.1026 -0.1883 -0.2075 -0.2520 -0.1709  0.1279
 0.0976  0.0817 -0.1519 -0.0599  0.0621  0.1285  0.0770 -0.0316
 0.6598  0.6926  0.6805  1.0000  0.2207  0.2467  0.1530 -0.0493
-0.0192 -0.0840 -0.1869 -0.2277 -0.2425 -0.1819  0.1825  0.1015
 0.0745 -0.1512 -0.0823  0.0271  0.0695  0.0603 -0.0346  0.6993
 0.6764  0.6702  0.6897  1.0000
Sample Size 7841
Latent Variables SES IMPT SUCC ENJY STAC TEAC ACHV
Relationships
sci01 sci02 sci03 sci04 sci05 = ACHV
liks enjy bore = ENJY
syt1 syt2 syt3 syt4 = SUCC
mip1 fip1 sip1 = IMPT
smgp prac expr = STAC
rule prob eg = TEAC
edfa edmo book = SES
ACHV = ENJY
ACHV = TEAC
ACHV = STAC
ACHV = SES
ACHV = SUCC
ACHV = IMPT
ENJY = STAC
ENJY = TEAC
ENJY = SUCC
ENJY = IMPT
SUCC = SES
IMPT = SES
Lisrel Output
Path Diagram
End of Problem
```

APPENDIX D

DESCRIPTIVE STATISTICS OF OBSERVED VARIABLES

Table D.1 Descriptive Statistics of Questionnaire Items

Item	N	Min	Max	Mean	Variance	Skew.	Kurtosis
Edmo	7227	1	7	2.07	1.362	2.177	5.974
Edfa	7037	1	7	2.84	2.759	1.391	1.022
Book	7728	1	5	2.42	1.201	0.614	-0.131
Syt1	7374	1	4	2.48	0.858	-0.104	-0.858
Syt2	7242	1	4	2.62	0.753	-0.220	-0.601
Syt3	7213	1	4	2.69	0.737	-0.307	-0.501
Syt4	7192	1	4	2.62	0.801	-0.176	-0.712
Liks	7506	1	4	1.85	0.547	0.793	0.766
Enjy	7432	1	4	1.80	0.566	0.776	0.425
Bore	7281	1	4	2.99	0.605	-0.609	0.232
Mip1	7277	1	4	1.51	0.428	1.294	1.961
Fip1	7291	1	4	1.59	0.455	1.021	1.102
Sip1	7349	1	4	1.37	0.353	1.738	3.645
Smgp	7241	1	4	3.01	0.968	-0.751	-0.452
Expr	7246	1	4	2.85	0.983	-0.548	-0.714
Prac	7296	1	4	2.57	0.993	-0.198	-1.014
Rule	7516	1	4	1.57	0.610	1.216	0.660
Prob	7612	1	4	1.54	0.512	1.192	0.858
Eg	7344	1	4	1.99	0.859	0.501	-0.779
Raw Score of Science	7841	3	40	16.79	29.027	0.474	0.230
Standardized Raw Score	7841	21.387	86.528	49.656	100.027	0.323	-0.171

APPENDIX E

THE RELATIVE STANDINGS OF TIMSS PARTICIPANTS

Table E.1 Science Scores of TIMSS Participants

Country	Score	Standard Error
Australia	540	4.4
Belgium-Flemish	535	3.1
Bulgaria	518	5.4
Canada	533	2.1
Chile	420	3.7
Chinese-Taipei	569	4.4
Cyprus	460	2.4
Czech Republic	539	4.2
England	538	4.8
Finland	535	3.5
Hong Kong	530	3.7
Hungry	552	3.7
Indonesia	434	4.5
Iran	448	3.8
Israel	468	4.9
Italy	493	3.9
Japan	550	2.2
Jordan	450	3.8
Korea (South)	549	2.6

Table E.2 Science Scores of TIMSS Participants (Continued)

Country	Score	Standard Error
Latvia	503	4.8
Lithuania	488	4.1
Macedonia	458	5.2
Malaysia	492	4.4
Moldova	459	4.0
Morocco	323	4.3
Netherlands	545	6.9
New Zealand	510	4.9
Philippines	345	7.9
Romania	472	5.8
Russian Fed.	529	6.4
Singapore	568	8.0
Slovak Rep.	535	3.3
Slovenia	533	3.2
South Africa	243	7.9
Thailand	482	4.0
Tunisia	430	3.4
Turkey	433	4.3
United States	515	4.6
International Average	488	0.7