A CROSS-CULTURAL COMPARISON OF THE EFFECT OF HUMAN AND PHYSICAL RESOURCES ON STUDENTS' SCIENTIFIC LITERACY SKILLS IN THE PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT (PISA) 2006

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

BY

ÖZGÜR ÇELEBİ

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY
IN
SECONDARY SCIENCE AND MATHEMATICS EDUCATION

SEPTEMBER 2010

Approval of the thesis:

A CROSS-CULTURAL COMPARISON OF THE EFFECT OF HUMAN AND PHYSICAL RESOURCES ON STUDENTS' SCIENTIFIC LITERACY SKILLS IN THE PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT (PISA) 2006

submitted by **ÖZGÜR ÇELEBİ** in partial fulfillment of the requirements for the degree of **Doctor of Philosophy in Secondary Science and Mathematics Education Department, Middle East Technical University** by,

Prof. Dr. Canan Özgen ____________________
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Ömer Geban ____________________
Head of Department, **Secondary Science and Mathematics Education**

Prof. Dr. Ömer Geban ____________________
Supervisor, **Secondary Science and Mathematics Education Dept.**

Examining Committee Members:

Prof. Dr. Ayhan Yılmaz ____________________
Secondary Sci. and Math. Edu. Dept., Hacettepe Uni.

Prof. Dr. Ömer Geban ____________________
Secondary Sci. and Math. Edu. Dept., METU

Prof. Dr. Hamide Ertepınar ____________________
Elementary Education Dept., METU

Assoc. Prof. Dr. Yezdan Boz ____________________
Secondary Sci. and Math. Edu. Dept., METU

Assist. Prof. Dr. Ömer Faruk Özdemir ____________________
Secondary Sci. and Math. Edu. Dept., METU

Date: 28.09.2010

I hereby declare that all information in this document has been obtained and presented in accordance with academic rules and ethical conduct. I also declare that, as required by these rules and conduct, I have fully cited and referenced all material and results that are not original to this work.

Name, Last name : Özgür Çelebi

Signature :

ABSTRACT

A CROSS-CULTURAL COMPARISON OF THE EFFECT OF HUMAN AND PHYSICAL RESOURCES ON STUDENTS' SCIENTIFIC LITERACY SKILLS IN THE PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT (PISA) 2006

Çelebi, Özgür
Ph.D., Department of Secondary Science and Mathematics Education
Supervisor : Prof. Dr. Ömer Geban

September 2010, 203 pages

This study investigates the students' characteristics and school characteristics and their influences on scientific literacy skills of 15-year-old students across Turkey, Canada, and Sweden, through the use of data from Organisation for Economic Co-operation and Development's (OECD's) Programme for International Student Assessment (PISA) 2006. The purpose of this study is to gain a more complete understanding of the effect of human and physical resource allocations and their interaction on students' scientific literacy skills using Hierarchical Linear Modeling (HLM) techniques. By PISA 2006 results, in terms of scientific literacy, Canada is a high performing country whereas Turkey is far below the average and Sweden has a rank in the average. For modeling scientific literacy, student-level characteristics determined by student questionnaire, and school-level characteristics determined by school questionnaire were used. Results of the present study indicated that there were significant between-school differences in scientific literacy skills of students for all three countries. Turkey had the highest between-school variance and it was more than half of the total variance whereas in

Canada and Sweden they were far lower. School type and size were common school factors affecting students' scientific literacy skills in Canada and Sweden; however, in Turkey school admittance policies, educational resources, science promotional activities, and teacher qualities were school characteristics which have impact on scientific literacy. Enjoyment of learning science, self-efficacy in science, general value given to science, awareness of environmental issues, responsibility for sustainable development, and confidence in use of information technologies were common student factors affecting development of scientific literacy skills in the three countries. Finally, in all three countries cross-level interactions of student and school characteristics for developing scientific literacy skills were observed.

Keywords: Programme for International Student Assessment (PISA), Scientific Literacy, Science Education, Hierarchical Linear Modeling (HLM), Student-Level Factors, School-Level Factors

ÖZ

PISA 2006 ULUSLARARASI ÖĞRENCİ DEĞERLENDİRME PROGRAMI'NDA İNSAN KAYNAKLARI VE FİZİKSEL KAYNAKLARIN ÖĞRENCİLERİN FEN OKURYAZARLIĞINA OLAN ETKİSİNİN KÜLTÜRLERARASI KARŞILAŞTIRILMASI

Çelebi, Özgür
Doktora, Ortaöğretim Fen ve Matematik Alanları Eğitimi Bölümü
Tez Yöneticisi: Prof. Dr. Ömer Geban

Eylül 2010, 203 sayfa

Bu çalışma, Türkiye, Kanada ve İsveç'te öğrenci ve okul karakteristiklerinin 15 yaş öğrencilerinin Uluslararası Öğrenci Değerlendirme Programındaki (PISA 2006) fen okuryazarlık sonuçlarına etkilerini incelemektedir. Bu çalışmanın amacı, farklı kültürlerde, insan ve fiziksel kaynakların ve bunların etkileşimlerinin fen okuryazarlığına olan etkisini hiyerarşik lineer modelleme (HLM) teknikleri kullanarak belirlemektir. PISA 2006 sonuçlarına gore fen okuryazarlığı bakımından Kanada yüksek puan alırken ve İsveç ortalama seviyesindeyken, Türkiye'nin puanı ortalamanın oldukça altındadır. Fen okuryazarlığını modellerken öğrenci anketiyle belirlenen öğrenci karakteristikleri ve okul anketiyle belirlenen okul karakteristikleri kullanılmıştır. Bu çalışmanın sonuçları üç ülkede de öğrencilerin fen okuryazarlıkları bakımından okullar arasında farklılıklar olduğunu göstermiştir. Türkiye'de toplam varyansın yarıdan fazlası okullar arası ve en yüksek oran olarak belirlenirken Kanada ve İsveç için bu değerler çok daha düşük bulunmuştur. Okulun çeşidi ve büyüklüğü Kanada ve İsveç için öğrencilerin fen okuryazarlığını etkileyen ortak etkenler olarak belirlenirken Türkiye için öğrenci seçim yöntemleri, eğitim malzemeleri, fen öğretimini teşvik edici etkinlikler ve öğretmen kalitesi etkili olan okul karakteristikleri olarak bulunmuştur. Fen okuryazarlığı geliştirmede etkili olan ortak öğrenci seviyesi etkenleri ise fen öğrenmeden hoşlanma, fen özyeterlik algısı, genel anlamıyla bilime verilen değer, çevre sorunları konusunda farkındalık derecesi, sürdürülebilir kalkınma için hissedilen sorumluluk ve bilgi teknolojileri kullanma konusunda kendine duyulan güven olarak belirlenmiştir. Son olarak, her üç ülkede de fen okuryazarlığı geliştirmede etkili öğrenci ve okul karakteristikleri arasında etkileşimler gözlemlenmiştir.

Anahtar Kelimeler: Uluslararası Öğrenci Değerlendirme Programı (PISA), Fen Okuryazarlığı, Fen Eğitimi, Hiyerarşik Lineer Modelleme (HLM), Öğrenciye İlişkin Etkenler, Okula İlişkin Etkenler

To All My Beloved

ACKNOWLEDGMENTS

I would like to express my gratitude to my supervisor Prof. Dr. Ömer Geban for his guidance, encouragement, and support.

I would also thank to Çiğdem İş Güzel and Sevim Sevgi for their support, comments and help.

Last, but not least, my love goes to Pınar, Tolga, my father and my mother for being my real source of strength and for believing in me.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

ACER	Australian Council for Educational Research
ANOVA	Analysis of Variance
CFA	Confirmatory Factor Analysis
CFI	Comparative Fit Index
CHIPCT	Theoretical Values from a Chi-square Distribution of the Number of Random Factors variable in residual file
CITO	National Institute for Educational Measurement
DIF	Differential Item Functioning
DVD	Digital Versatile Disc
EB	Empirical Bayes
GPA	Grade Point Average
HLM	Hierarchical Linear Modeling
HS&B	High School and Beyond Survey
ICT	Information and Communications Technology
IEA	International Association for the Evaluation of Educational Achievement
ILS	Department of Teacher Education and School Development, University of Oslo
IPN	Leibniz Institute for Science Education, University of Kiel
IRT	Item Response Theory
ISCED	International Standard Classification of Education
ISEI	International Socio-Economic Index of Occupational Status
MDIST	Mahalanobis Distance of Empirical Bayes estimates of the fixed effect parameters from their fitted values variable in residual file
MDRSVAR	Natural Logarithm of the Residual Standard Deviation from the Final Fitted Fixed Effect Model variable in residual file
NIER	National Institute for Educational Policy Research

NNFI	Non-Normed Fit Index
NPM	National Project Manager
OECD	Organisation for Economic Co-operation and Development
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PQM	PISA Quality Monitor
PV	Plausible Value
RMR	Root-Mean Square Residual
RMSEA	Root-Mean Square Error of Approximation
SEG	Science Expert Group
SEM	Structural Equation Modeling
SES	Socioeconomic Status
SPSS	Statistics Package for Social Sciences
TIMSS	Trends in International Mathematics and Science Study
TIMSS-R	Third International Mathematics and Science Study-Repeat
UNESCO	United Nations Educational, Scientific and Cultural Organization
VCR	Video Cassette Recording
WLE	Weighted Likelihood Estimate

CHAPTER 1

INTRODUCTION

This study aims to investigate the students' characteristics and school characteristics and their influences on scientific literacy skills of 15-year-old students across Turkey, Canada, and Sweden, using the data from Organisation for Economic Co-operation and Development's (OECD's) Programme for International Student Assessment (PISA) 2006.

1.1 International Assessment Studies and Surveys

With the arrival of globalization and characterization of economies with this notion, educators and policymakers have trended towards international comparisons to assess functioning and performance of education systems. These comparative studies now have been assessing many issues for years, from access to education, equal allocation of resources, optimization of learning processes according to relationship with achievement, to character and quality of schools. Findings of those studies provide educators and governments with benchmarks to assess the performance of education systems and with strategies to improve achievement. Due to the differences in implementation of educational applications across countries and translation effects inherent to tests, there are certain disadvantages of international assessment studies and surveys. Nonetheless, they provide a remarkable and an accepted level comparability between countries on their performances and growth in education systems.

The most well-known, current and inclusive ones of international comparisons in education are Programme for International Student Assessment

(PISA), Trends in International Mathematics and Science Study (TIMSS), and Progress in International Reading Literacy Study (PIRLS). IEA's (International Association for the Evaluation of Educational Achievement) Trends in International Mathematics and Science Study (TIMSS) assesses mathematics and science achievement of students at fourth and eighth grades, and collects data about curricula and instructional methods of science and mathematics from school administrators and teachers in the study (TIMSS & PIRLS International Study Center, 2006b, p.7). Up to now, TIMSS studies were administered in 1995, 1999, 2003 and 2007. IEA's Progress in International Reading Literacy Study (PIRLS) is an international study of the literacy, achievement, behaviors and attitudes about reading of fourth-grade equivalent students in the participating countries (TIMSS & PIRLS International Study Center, 2006a, pp. 1-7). Up to now, PIRLS studies were conducted in 2001 and 2006.

On the other hand, Organisation for Economic Co-operation and Development's (OECD's) Programme for International Student Assessment (PISA) is a system of international assessments that assesses 15-year-olds' reading, mathematics and scientific literacies and capabilities in every assessment, but each application with its own emphasis on one of these domains respectively. Students' functional skills toward the end of compulsory education are investigated and measures of general or cross-curricular competencies are included, resultantly making PISA the most comprehensive international assessment up to now. Together with PISA 2000 emphasizing on reading literacy and PISA 2003 emphasizing on mathematics literacy, PISA 2006 emphasizing on scientific literacy completed the first cycle of assessment and PISA will go through the cycle again in the upcoming nine years.

1.2 Organisation for Economic Co-operation and Development (OECD)

The Organisation for Economic Co-operation and Development is a forum that brings together currently 30 governments to deal with economic, social, and environmental issues (OECD, 2007, p.1). OECD assembles countries committed to democracy and market economy with the tasks of (About OECD, n.d.):

- Maintaining a sustainable growth in economy
- Increasing employment rate and decreasing unemployment
- Improving living standards
- Retaining financial stability
- Assisting other countries' in economic development
- Contributing to expansion of world trade

Since it was established in 1961, OECD has been providing statistical, economic and social data by analyzing and forecasting economic developments, researching social changes and growing patterns in trade, environment, agriculture, technology, and financial policies. The organization mediates to comparisons and sharing of experiences in policy making, seeking for solutions to common problems, and matching of domestic policies with international policies between governments. Also, the Organisation shares and exchanges expertise with more than 100 non-member countries and economies. OECD is a large publisher in the fields of economics and public policy with an average of 250 new titles per year (About OECD, n.d.).

Currently, the 30 member-countries of OECD are Australia, Austria, Belgium, Canada, Czech Republic, Denmark, Finland, France, Germany, Greece, Hungary, Iceland, Ireland, Italy, Japan, Korea, Luxembourg, Mexico, the Netherlands, New Zealand, Norway, Poland, Portugal, Slovak Republic, Spain, Sweden, Switzerland, Turkey, United Kingdom, United States (Members and partners, n.d.).

As of May 2007, Chile, Estonia, Israel, Russia and Slovenia were invited to open discussions for membership of the Organisation, whereas Brazil, China, India, Indonesia and South Africa were offered an enhanced engagement, with a view to possible membership (Members and partners, n.d.).

1.3 Programme for International Student Assessment (PISA)

In 1997, OECD launched the Programme for International Student Assessment (PISA) aiming to gather cross-nationally comparable evidence on student performance (OECD, 2007, p.3). PISA is an internationally standardized assessment that "*represents a commitment by governments to monitor the outcomes of education systems in terms of student achievement on a regular basis and within an internationally agreed common framework*" (OECD, 2007, p.3). Starting in 2000, PISA surveys are administered on three-year periods with cooperation of OECD member countries with volunteering partner countries that are not members of OECD. The target population of PISA is 15-year-old students taking into consideration that they are approaching the end of their compulsory education period in most countries. Regarding this, PISA assesses students' abilities of using knowledge and skills to cope with real-life challenges and situations after school life.

The PISA framework starts with the concept of literacy, which has a much broader meaning than the ability to read and write. PISA not only assesses knowledge, but also studies students' skills to use knowledge and experiences. The term 'literacy' is used to sum up this cooperation and coordination of knowledge and skills.

What PISA aims to measure is reading, mathematics and science literacy developed within a broad framework of four interrelated aspects (OECD, 2007, p.20):

- structure of knowledge that the learner needs to get
- competencies to apply
- contexts for encountering problems and applying relevant knowledge and skills
- attitudes towards the subject

PISA 2000 is the first survey of the first PISA cycle, which focused on reading literacy. The survey was implemented in 43 countries, 32 of which are OECD member countries and the remaining 11 were participating but were not members. Reading was the major domain, where science and mathematics assessment also took place. Key skills for reading literacy were taken as retrieving information from texts, interpreting information, and reflecting. Also a questionnaire on student factors influencing achievement such as family background, home environment, involvement of parents in school affairs, and parental expectations was administered. Besides those materials for students, to determine the impact of school characteristics on achievement, a school questionnaire was given to school principals. PISA 2000 surveyed the relationship between students' interest and engagement in learning, ability to organize the learning process, and performance in learning. Results indicated that the relationships between those characteristics are mutually reinforcing and not simply causal in one direction and learning autonomy can be fostered at school (OECD, 2001, pp. 98-100).

In 2003, second PISA survey was conducted with the main focus of mathematics literacy. A thorough assessment of mathematics and assessments with less detail in problem solving, science, and reading were made (OECD, 2004, p. 20). Students' performance in mathematics was measured in three dimensions: content, processes and the situations in which problems are encountered. Furthermore, mathematical content was examined in four specific areas: "space and shape", "change and relationships", "quantity", and "uncertainty". In PISA 2003, as in 2000, students self-assessed their learning characteristics. Moreover, PISA 2003 initiated the assessment of wider competencies as problem-solving skills (OECD, 2005, pp. 8-9). It included 41 countries, most of which also participated in PISA 2000, and the whole 30 member countries of OECD attended. Multiple examinations of data among countries yielded a common result that there were both high performers and poor performers in each country.

As the last survey of the first cycle, PISA 2006 focused on science but the assessment also included reading and mathematics and student level and school level factors that could affect performance differences were also surveyed. For the first time, attitudinal questions were included within the test, rather than using a separate attitude questionnaire. Around 400,000 students from 57 countries participated in this survey, which represents 20 million 15-year-old students (OECD, 2007, pp. 16-19). These 57 countries represent about 90% of the world economy (OECD, 2007, p.16).

Understanding fundamentals of science and solving scientific problems are key factors in today's society for not only science-oriented proficiency, but also for participating in society and labor market. However, orientation to science and technology in universities has dropped noticeably (OECD, 2007, p.16). Reasons for this are varied but attitudes toward science may play an important role. Therefore, PISA 2006 also assessed students' attitudes toward science, competencies, and their awareness of science-oriented life opportunities.

By the ongoing research on PISA 2006 data, among countries there are different profiles of students who show stronger scientific competencies.

1.4 Concepts and Innovations in PISA

PISA is not the first international study assessing student performance, gathering data on student background and institutional factors that could differentially affect performance, but the most comprehensive and thorough one. Leading experts in their respective fields make decisions on the PISA survey and it is applied by governments in cooperation, therefore, bringing in high reliability and validity for the results of PISA. Across the world, PISA findings are used to measure knowledge and skills of students in comparison with other countries, to help educational improvement, and to understand strengths and weaknesses of education systems.

Key features for the development of PISA have been (OECD, 2007, pp. 16-17):

- policy orientation, which helps identify differences in performance patterns and characteristics of education that has high standards
- innovative “literacy” concept, oriented to application of knowledge and skills in a variety of situations
- relevance to lifelong learning, with motivation to learn, self-beliefs, and learning strategies besides curricular and cross-curricular competencies
- regularity, which enables countries monitor their progress
- geographical coverage that countries in PISA represent more than half of the world’s geography and more than 90% of world’s economy

Other international comparative studies have been conducted over the past 40 years and these surveys have concentrated on outcomes linked directly to the curriculum and only to those parts that are common across participating countries’ curricula. However, PISA takes a different approach in many respects. First, PISA is an initiative taken by governments and the results drive policy makers on educational and labor market policies. Second, it is conducted on a regular basis, which helps countries monitor their progress.

1.4.1 Literacy and PISA's Literacy Concept

The traditional and most basic definition of literacy is considered to be the ability to use language to read and write. By contemporary perspective, the word refers to reading and writing at a level adequate for communication to take part in the society (Literacy, 2009). The United Nations Educational, Scientific and Cultural Organization (UNESCO) made the following definition (UNESCO Education Sector, 2004, p. 13):

> "*'Literacy' is the ability to identify, understand, interpret, create, communicate, compute and use printed and written materials associated with varying contexts. Literacy involves a continuum of learning to enable an individual to achieve his or her goals, to develop his or her knowledge and potential, and to participate fully in the wider society.*"

However, PISA's literacy concept is much broader than simply reading and writing in a language. It is concerned with students' capacity to extrapolate from what they learned, to apply knowledge in several settings, to analyze, to reason and to communicate effectively as they confront, solve and interpret problems in a variety of situations (OECD, 2007, p.20).

PISA's literacy is measured on a continuum but not in an "existent-nonexistent" manner. Certain points on literacy continuum are defined to create ranges of abilities and competences from most basic to higher levels PISA seeks so that students can be categorized as adequate or inadequate at certain areas.

Literacy is an all-time process. Students acquire literacy not only at school, but also outside school through interactions with wider communities. To continue learning in areas such as reading, mathematics and science, and to apply what they learned to real world situations, students should at least have basic understanding of fundamentals and use principles and processes flexibly in various situations.

Therefore, PISA measures the ability to complete tasks related to real life rather than it assesses subject-specific knowledge (OECD, 2007, p.20).

1.4.1.1 Reading Literacy

Reading literacy is defined in PISA studies as the ability to understand, use and reflect on written texts in order to participate effectively in life (OECD, 2001, p.21). Reading literacy is not just decoding written material or literal comprehension; it is also understanding and reflecting on texts, using written information to fulfill goals and to function effectively. As the reader seeks to use and understand what is being read, there is a variety of responses to the given text. This dynamic process has many dimensions and PISA assessments use three of them:

- The form of reading material, that is, PISA includes continuous prose texts and distinguishes between types such as narration, exposition, and argumentation, whereas non-continuous texts such as lists, forms, graphs and diagrams are also included.
- The type of reading task, that is, regarding the age group of students basic reading skills are not of concern but PISA's reading tasks include retrieving information, general understanding of texts, interpretation, reflecting on own knowledge, evaluating and arguing own points of views.
- The use of the document, that is, the context or situation such as private or public use, occupational use and educational use.

1.4.1.2 Mathematical Literacy

Mathematical literacy is defined in PISA as the capacity to identify, understand and engage in mathematics, and to make judgments about the importance of mathematics in life (OECD, 2007, p.21). It indicates the ability to put mathematical knowledge and skills to functional use rather than just mastering them within school curriculum. In order to assess this mathematical literacy, PISA uses three broad dimensions (OECD, 2004, p.25):

- The content or structure of mathematical knowledge: Familiarity with mathematical concepts and mastery of a balanced mathematical curriculum are required. Content is defined in terms of clusters of relevant, connected mathematical concepts that appear in real life situations. These clusters include concepts of quantity, space and shape, change and relationships, and uncertainty.
- The processes that need to be performed: Students are expected to recognize the features of the problem situation, and to use relevant mathematical competencies to solve the problem. That is, they are expected to "mathematizate" the problem and pursue a certain mathematical argument.
- The situations in which students encounter mathematical problems and relevant knowledge and skills are applied: Students are given tasks, sometimes fictional, based on situations which represent problems encountered in real life. In order of closeness to the student, the situations are classified as private life or personal, school life, work and sports, local community and society, and scientific.

1.4.1.3 Scientific Literacy

Today, science and technology shape human lives to a considerable extent. Scientific literacy is not just the charge of scientists; it is a requirement for all, no matter a science-career is chosen or not. Understanding of science and applying a scientific perspective are the basis for scientific literacy.

PISA defines scientific literacy as combination of (OECD, 2007, pp. 34-35):

- Use of scientific knowledge
- Understanding the features of science
- Awareness of the role of science and technology in shaping our lives.
- Willingness to engage with science

PISA 2006 defines scientific literacy and develops its science assessment tasks and questions within a framework of four interrelated aspects: (i) contexts in which tasks are presented, (ii) the competencies that students need to apply, (iii) the knowledge domains involved, and (iv) student attitudes (OECD, 2007, p. 35). Life situations that involve science and technology require people to identify scientific issues, explain phenomena scientifically and use scientific evidence. This process is influenced by the knowledge of science and knowledge about science, and by attitudes toward science. That is to say, PISA's scientific literacy includes both knowledge of the different scientific disciplines and the natural world, and knowledge about science as a type of search for knowledge. Moreover, PISA's science competencies include not just the abilities in science but also attitudes toward science, because attitudes, beliefs, motivations, self-efficacy, and values play an important role in students' decisions to develop their scientific knowledge further, follow careers in science, and use science throughout their lives.

1.5 The Present Study

The purpose of this study is to gain a more complete understanding of the effect of human and physical resource allocations and their interaction on students' scientific literacy skills across Turkey, Sweden, and Canada. By PISA 2006 results, in terms of scientific literacy, Turkey is far below the average whereas Canada is a high performing country and Sweden has a rank in the average. Hierarchical linear modeling (HLM) techniques are used separately for these three different cultural settings using the database of PISA 2006 and to model scientific literacy, student-level characteristics determined by student questionnaire, and school-level characteristics determined by school questionnaire are used.

With its multilevel approach and target populations, the present study aims to benefit from uniting the powers of Hierarchical linear modeling and distinguished data of PISA.

1.5.1 Multilevel Approach

In social research and many other fields, structure of the research data is mostly hierarchical. That is, the individual subjects of study may be classified or arranged in groups which themselves have qualities that influence the study. This feature is called nesting. As for the case of education, students are nested within classes; classes are nested within schools; schools are nested within cities; cities within regions; and so on. Multilevel models allow for the study of these hierarchical relationships at any level in a single analysis, while not ignoring the variability associated with each level of the hierarchy. For the cause of the present study, students are taken as nested within schools to model the real world more realistically and to account for differences in performances with less error within this hierarchical setting.

1.5.2 Significance of the Study

PISA is a comprehensive international study. The countries in which it is applied constitute almost the whole world economy, because many developed and prosperous countries are included. However, even well-developed countries may have significant problems in education, or to say it bluntly, problems in preparing individuals for life. Also they should investigate and monitor the quality of their educational systems. In that sense, PISA databases are the most fruitful sources so far, because other studies have assessed the achievement in particular subjects whereas PISA assessed many factors for obtaining new knowledge and using knowledge at various situations rather than only possessing specific knowledge. PISA provides strong and cross-culturally valid measures of competencies that are relevant to everyday life and information on factors associated with educational success.

Up to date, most of the studies on PISA were either informative about the PISA project and its success or accounts for success levels of individual countries and for what factors they succeeded or not. Very few studies investigated patterns for success and competencies. The present study tries to model the factors and resources affecting scientific literacy of 15-year-old students across different cultural settings. Countries of interest were selected from high, middle, and low levels of scientific literacy from the results PISA 2006 so that the resulting patterns for scientific literacy and science competencies would be independent from country scores and might be applicable to any country provided that data is present. Also, the findings about school and student level factors might give us insight on why Turkey scores low in similar studies and what the deficiencies of Turkish education system are.

CHAPTER 2

LITERATURE REVIEW

2.1 International Assessments and Surveys

International studies of educational achievement began in the early 1960s and recently they received a great recognition and attention by governments and public. The main aim of all those studies has been to collect and analyze valid cross-cultural information about educational achievement. Conceptualizing and conducting such studies are not easy but when done they are greatly beneficial (Beaton et al., 1999; Lie & Linnakylä, 2004). Multiple groups such as researchers, educators, policy-makers, and general public benefit from these information for years now. Experiences show that results from international surveys can have significant impact on some countries' educational policies (Fredriksson, 2001).

Countries differ in their curricula design with varying degrees. Also, within a country, differences may occur between regions and even among schools in the same region. However, such differences do not necessarily imply any deficiency because a country may have its own rationale for adopting a curriculum different from the international pattern. On the other hand, one of the major benefits of participating in an international assessment is it requires the country to scrutinize the curriculum (Beaton et al., 1999). Curricula may become outdated in some respects and lack in including important topics to communities. Therefore, close analyses and comparisons of curricula of several countries is a prerequisite step in deciding which topics and test items should be included to be fair to all countries in a common test. This drives every country to review their curricula thoroughly.

In this way, countries are alerted to differences between their curriculum and that of others in terms of the emphasis given to different subject areas and the content covered. The information obtained from cross-national assessments can be of great value in determining whether to revise or change the curriculum (Beaton et al., 1999). Some countries are involved in more than one international study, as well as conducting own national studies. To compare results of those different studies, they must have sufficient common test items. However, it is often not possible to make such comparisons because there has been little coordination in the design and conduct of different international and national studies (Beaton et al., 1999).

Holding a cross-national structure and multi-aspect views, assessments and surveys so called international have a comparative nature. However, it is a complex issue to conduct, analyze and interpret international comparative studies. They are complicated because of issues of population definition and sampling methodologies, and difficulty in identifying common frame of reference (Mislevy, 1995). Such studies should be seen as means for gaining valuable knowledge about similarities and difference, but not as competitions to see who plays for top. One must keep in mind that educational systems and students' success vary in many aspects from nation to nation.

The set of tasks on which performances are compared has been the center for international assessments. As a result, the tendency is to report single-number indexes determined from the performance in those settings, relying on the notion of the comparability of universal traits such as intelligence, mathematics achievement, and science achievement. Yet, culture, background of students, relevance, and educational experiences are often ignored in those settings. Besides, their inability to communicate the degree and character of differences among nations and within nations is the most serious limitation of single-number indexes (Mislevy, 1995). International assessments convey context, clues, and current conditions as well as they provide useful information to improve educational systems, but not the ways to.

On the other hand, there are basic requirements that international surveys should satisfy (Goldstein, 2004):

- Cultural specificity should be recognized during test development and in the subsequent analysis
- Statistical models used in the analysis should be complex as close to the reality as possible so that multidimensionality is incorporated and country differences are retained rather than eliminated in favor of a common scale
- Multilevel nature of any comparisons should be stressed
- Such studies should be longitudinal
- The design of questionnaires and test items should expose diversity rather than define “untypical” to exclude

Most of the international studies of educational achievement typically reveal substantial differences in the performance levels of student from different countries in many subject areas. In international comparative studies, data analyses are expected to detect differences and similarities between countries (Papanastasiou, 2002). Also, those studies collect information to examine the likely factors to influence achievement. Attention is mostly given to groups of explanatory factors (Beaton et al., 1999):

- Home background: Factors such as education and occupations of parents, socioeconomic status, educational resources at home
- School characteristics: Such as school type, school size, student/teacher ratio
- Teacher characteristics: For example, educational levels of teachers, teaching experience, sex of teacher
- Teaching conditions and practices

- Student motivation

The typical focus of interest for international educational achievement studies have been the negative correlation between poverty and education (Naumann, 2005). However, the main focus in the construction of analytical models should be on the impact of variables related to teaching organization on educational success because those variables are more likely to be influenced by practical policies than social background of the students.

2.2 Studies about PISA

Comparative research in education or student achievement is not a new area. Recently there has been a growing interest in international comparisons between education systems. When compared with earlier international assessments, the new aspect of PISA is its size and scope. Besides, except for a few, earlier projects have generally been single-occasion events where PISA is a cycle of multiple comparisons. In the long run, PISA and other large-scale international studies will add new dimensions to school assessment and it will be easier to compare and to be compared with other countries (Fredriksson, 2001).

2.2.1 PISA as an Assessment

One of the most desired characteristics for an international study to be comparative is that it covers a wide range of cultures and languages. Therefore, at the beginning of test development process, the OECD calls for item submissions from participating countries in accordance with a set of submission guidelines. With the collection of materials from participating countries, it is aimed to be ensured that as many perspectives as possible is represented in the materials (Kirsch et al., 2002).

PISA defined different domains. Unlike other studies, PISA has not focused on comparing and analyzing national curricula to define those domains, instead, but they are defined in terms of knowledge and skills needed in adult life (Fredriksson, 2001). Being not directly tied to the school curriculum PISA is designed to assess the practical outcomes of education systems. However, TIMSS and PISA can be viewed as complementary measures because TIMSS focuses tightly on specific parts of curriculum whereas PISA assesses outcomes attained by a solid education but are obtained by combinations of discrete parts of curriculum (Bybee & Stage, 2005).

Compared to other international studies, PISA focuses on process skills rather than conceptual understanding of items (Kjærnsli & Lie, 2004). The items in PISA are based mostly on authentic texts. Kjærnsli and Lie (2004) categorized all science items from PISA 2000 into two groups as process skills items and conceptual understanding items. They found that as a general pattern for all countries, items focusing on conceptual understanding have higher scores and process skills items appear more difficult for students. Nevertheless, Kjærnsli and Lie suggested that one should focus on the aspect of scientific process skills as competencies that are more general and likely to be used more easily in other than scientific contexts.

Coaching is known to improve student performance in tests relevant to future academic and occupational careers (Becker, 1990). However, large-scale assessments such as PISA do not influence individual student careers and do not have direct implications for the students. Their results are usually taken into consideration when making political decisions. Therefore, assessments affecting policies, such as PISA, should be as immune to coaching as possible. Brunner, Artelt, Krauss, and Baumert (2007) investigated the effects of coaching and pretesting on PISA for the domains of reading and mathematics separately. Both coaching and pretesting had little or no effect on performances of students, however, all increments observed in performances were statistically not significant.

PISA data includes information on assessments, structure of school, administration, educational resources, teacher education, professional development, gender, socioeconomic status, home background, engagement with literature, and underachievement (Topping, 2006; Topping et al., 2003). However, interpreting the data requires great caution because the data (Topping, 2006):

- represent cross-sectional study
- correlational
- do not explain causes and effects
- represent averages
- cannot spot cures
- can define important factors
- can diagnose significant problems
- are important in policymaking

2.2.2 Findings from PISA

Many other assessments examine whether students attain the knowledge and skills described in the curriculum whereas PISA intends to project future use from students' current levels of knowledge, attitudes, and skills. Focusing on scientific literacy, the two contexts used for main assessment and student questionnaire of PISA 2006 were resources and environments. Within the context, a situation is presented and a number of questions are asked about the situation. A typical PISA item demands more complex cognitive processes than typical items from other assessments.

First results from PISA suggest that interest for studying science has declined considerably among young people (OECD, 2007, p.16; Riley &

Torrance, 2003). The decline in numbers choosing to study science and concern for the society has been mentioned numerous times (Jenkins, 1994; Osborne, Simon, & Collins, 2003) and by the earlier ones and the latest PISA study, this loss of interest in science and science related careers was confirmed once again.

Bybee (2008) examined the data from PISA 2006 and using the released units, he evaluated students' science competencies and attitudes relative to resource use and environmental issues. Bybee claimed that scientific literacy includes actions about resource use and environmental quality. On the other hand a scientifically literate person must acquire attitudes that contribute to actions. He defined general characteristics for the knowledge about and attitudes toward resources and environmental issues for all countries as:

- There was an association between scientific knowledge and awareness of issues
- Students were concerned about global issues in general
- Level of concern was not strongly associated with performance on science test
- Disadvantaged students with respect to socio-economic status were more optimistic about solutions to environmental issues and about improvements
- Students with higher levels of understanding of science had higher degrees of responsibility for sustainable development

In a research report of International Reading Association PISA Task Force, Topping, Valtin, Roller, Brozo, and Dionisio (2003) investigated whether PISA's guiding notions are aligned with conceptions of literacy and defined some policy guidelines from findings of PISA 2000. They commented that with regard to quality and equality, PISA results are useful for assessing the relative standings of countries for performance and the variances, for identifying groups at risk, and for identifying characteristics of poorly performing students. The researchers recommended continuous monitoring of educational system because within-country differences was greater than between-countries differences; the impact of overall socioeconomic status (SES) of school population was greater than that of individual students; parental involvement is correlated with achievement; the quality of educational resources is more effective in achievement than that of buildings and physical resources; and systems with fewer types of schools are associated with higher performances and fewer differences in student outcomes.

Investigating the PISA data for relationships between school characteristics and student performance, it is evident that school disciplinary climate, student-teacher relations, and academic record utilization are relatively strongly related to performance even when conditioned on gender, language, and both student and school level socioeconomic status (Anderson et al., 2007). When modeling correlates of learning, a general PISA model would not reflect the nature of those relationships and thus each country should be modeled separately.

When the interactions between student gender, content dimension, item formats, and task demands were explored, there seems to be an advantage of females in open-ended constructed responses and an advantage of males in limited response or selection items (LaFontaine & Monseur, 2006 as cited in Anderson et al., 2007).

White and Smith (2005) conducted an analysis on the views of 5,416 headteachers, principals, and administrators working in secondary schools from 25 countries about employing and retaining teachers. The researchers questioned school administrators' perceptions of the hindering effect of teacher shortage, inadequacy, and turnover on students' learning and whether these problems were related to different characteristics and geographic locations of schools. The school-level data, which was collected by PISA School Questionnaire and complementary to student data, was analyzed. The results showed that there were substantial variations between countries with respect to the views of administrators about the hindering effect of teacher shortages/inadequacy and turnover. Also, differences in those views were rarely related to the size, location, or other characteristics of the schools among administrators from the same country.

Use of information and communication technologies causes significant improvement in science competencies of students (Balım et al., 2009). Generally those students who possess computers at home or at school also have access to Internet and some educational software. That could be among reasons for better scientific competencies of students.

Papanastasiou, Zembylas, and Vrasidas (2005) examined the various use of computers in learning science and its relation with scientific literacy through the data of U.S.A. and Germany from PISA 2000. They claimed that computer use alone cannot explain its effects on student achievement; how it is used is important (Papanastasiou et al., 2003; Papanastasiou et al. 2005). Evaluating the effect of educational technology requires an understanding of how to use in classroom setting, learning goals involved, types of assessments to evaluate improvements due to that technology, and awareness of the complexity of learning process. The researchers performed a series of multivariate regressions controlling the students' socioeconomic status (SES) to investigate the patterns of computer use within countries and variations of those patterns among countries. Main independent variables during analyses related to computers were availability, frequency of use at home, at school and other settings, comfort with use, and activities involving computers. Results suggest that variables of PISA database about computer use can account for a significant proportion of the variance of science literacy.

2.3 Scientific Literacy

The term "scientific literacy" appeared in late 1950s and is a general concept that has a variety of meanings. A clear definition of scientific literacy has not been made, widely accepted and used (Bybee, 1997, p. 63); it is still controversial. However, it implies a broad and functional understanding of science for education and development but not for science or technology related careers (DeBoer, 2000). Recently, the trend is to define scientific literacy as a measurable outcome and to include everything possible in its definition (DeBoer, 2000).

Pella, O'Hearn, and Gale (1966) investigated the referents of scientific literacy in systematically selected papers from 1946 to 1964 to provide an empirical basis for the definition of scientific literacy. From the frequencies of referents of scientific literacy, they concluded that scientific literacy of individuals were characterized by understanding of relationship of science and society, nature of science, science ethics, difference and relation between science and technology, basic concepts in science, and interrelations of science and the humanities.

In 1974, Showalter (as cited in Laugksch, 2000) defined seven dimensions for the scientific literacy:

i. Understanding of nature of scientific knowledge
ii. Applying appropriate science concepts, principles, laws, and theories in interacting with the universe
iii. Use science processes in solving problems, making decisions, and understanding the universe
iv. Acting consistent with the values underlying in science
v. Understanding and appreciation of cooperative enterprises of science and technology, interrelations with each other and with other aspects of society
vi. Developing a richer and more satisfying view of the universe as a result of science education and

vii. Developing manipulative skills associated with science and technology

DeBoer (2000) stated nine contemporary goals for teaching science and their implications for today's society. Those goals were:

i. Teaching and learning science as a cultural force in the modern world
ii. Preparation for the working life
iii. Application of science to daily lives
iv. Teaching students to be informed citizens
v. Learning about science as a particular way of examining the natural world
vi. Understanding science-related matters that appear in the popular media
vii. Learning about science for its aesthetic appeal
viii. Preparing citizens with sympathetic attitude to science
ix. Understanding the nature and importance of technology and the relationship between science and technology

Scientific literacy is about the public's understanding of science and that understanding is open-ended and evolving. It is primarily about the level of scientific understanding in adult population and there are many ways to be scientifically literate (DeBoer, 2000) and many definitions of scientific literacy (Laugksch, 2000).

2.3.1 Scientific Literacy of PISA

PISA's notion of learning is a lifelong process which seeks students' future needs and applications of things they learn. Thus, while knowledge of students is assessed, also their ability to apply knowledge and experiences to real-life situations is inspected (OECD, 2006; OECD, 2007).

Assessment of scientific literacy is of particular importance for PISA 2006 because it is the major domain assessed. In today's society, science and technology plays an important role, therefore to participate in the society, an understanding of science and technology is crucial. Competencies defined for PISA 2006's scientific literacy are the keys to define students' knowledge, values and abilities today and to relate these to the future.

Scientific literacy definition of PISA 2006 refers to an individual's (OECD, 2006, p. 23):

- Use of scientific knowledge to obtain further knowledge, to identify scientific problems, to make scientific explanations, and to draw conclusions based on scientific evidence
- Understanding the nature of science and regarding it as foundation of human knowledge and enquiry
- Awareness of the role of science and technology in shaping our lives
- Compliance to engage with science-related issues, and with ideas of science

Individuals cannot be taken as scientifically literate or scientifically illiterate; scientific literacy is a continuum from less developed to more developed (Bybee, 1997). Students with higher levels of scientific literacy skills will not recall simple factual knowledge only, but will create and use conceptual models to predict and explain, analyze scientific investigations, draw evidence, evaluate alternative explanations of the same phenomena, and communicate conclusions with precision (OECD, 2006, p. 25).

Taking into consideration the "process skills versus the conceptual understanding" dichotomy of science items, PISA has a much stronger emphasis on the process skills aspect compared to other international assessments, (Kjærnsli & Lie, 2004).

PISA project intends to report results in terms of proficiency scales. When developing a scale along which various points can be related to certain levels of performance, the very first decision is defining the number of scales and subscales. However, scientific literacy could be scaled theoretically in many alternative ways (Harlen, 2001).

Using techniques of modern item response modeling, it is possible to construct a scale of scientific performance, to assign each item in the assessment a point score on this scale according to its difficulty, and to assign each student a point score on the same scale representing his or her estimated ability. PISA 2006 constructed such scales for each of the science competencies and for each of the knowledge domains. Also a combined scale that combined the questions from all scales was created. The combined scale was standardized as such that the mean score of science performance across OECD countries was set to 500 score points and the standard deviation was set to 100 score points while the data is weighted in order to provide equivalent contribution of OECD countries. With this transformation, about two-thirds of the scores of students across OECD countries lie between 400 and 600 points. However, this definition of scale was applied only for the combined scale; therefore, individual science scales may differ from this definition with respect to mean score and standard deviation.

Science proficiency levels are defined for describing what competencies students' obtained scores demonstrate. Student scores in science are grouped into six proficiency levels, with Level 6 representing the highest scores; thus the most difficult tasks, and Level 1 representing the lowest scores and thus the easiest tasks. Each of the six levels can be understood as descriptions of the kind of science competency that a student needs to attain them.

Table 2.1 Summary Descriptions of PISA 2006 Proficiency Levels on the Science Scale

Level	Lower Score Limit	What Students Can Typically Do (by definition of PISA)
6	707.9	"Identify, explain and apply scientific knowledge and knowledge about science in a variety of complex life situations consistently; link different information sources and explanations and use evidence from those sources to justify decisions; clearly and consistently demonstrate advanced scientific thinking and reasoning, and demonstrate willingness to use scientific understanding in support of solutions to unfamiliar scientific and technological situations. Students at this level can use scientific knowledge and develop arguments in support of recommendations and decisions that center on personal, social or global situations."
5	633.3	"Students can identify the scientific components of many complex life situations, apply both scientific concepts and knowledge about science to these situations, and can compare, select and evaluate appropriate scientific evidence for responding to life situations. Students at this level can use well-developed inquiry abilities, link knowledge appropriately and bring critical insights to situations. They can construct explanations based on evidence and arguments based on their critical analysis."
4	558.7	"Students can work effectively with situations and issues that may involve explicit phenomena requiring them to make inferences about the role of science or technology. They can select and integrate explanations from different disciplines of science or technology and link those explanations directly to aspects of life situations. Students at this level can reflect on their actions and they can communicate decisions using scientific knowledge and evidence."
3	484.1	"Students can identify clearly described scientific issues in a range of contexts. They can select facts and knowledge to explain phenomena and apply simple models or inquiry strategies. Students at this level can interpret and use scientific concepts from different disciplines and can apply them directly. They can develop short statements using facts and make decisions based on scientific knowledge."
2	409.5	"Students have adequate scientific knowledge to provide possible explanations in familiar contexts or draw conclusions based on simple investigations. They are capable of direct reasoning and making literal interpretations of the results of scientific inquiry or technological problem solving."

Table 2.1 Continued

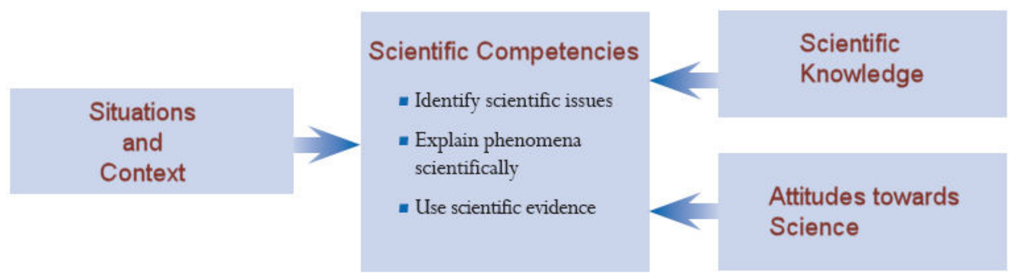
Level	Lower Score Limit	What Students Can Typically Do (by definition of PISA)
1	334.9	"Students have such a limited scientific knowledge that it can only be applied to a few, familiar situations. They can present scientific explanations that are obvious and that follow explicitly from given evidence."

Adapted from *PISA 2006: Science Competencies for Tomorrow's World, Volume 1: Analysis*

PISA 2006 definition of scientific literacy consists of four interrelated aspects (OECD, 2006, pp. 25-36):

- Situations and Context
- Scientific Knowledge
- Scientific Competencies
- Attitudes towards Science

Of these interrelated aspects, context, knowledge and attitudes trigger competencies in PISA's scientific literacy.



Adapted from *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*

Figure 2.1 Science Assessment Framework for PISA 2006

2.3.1.1 Situations and Context Aspect

Engagement with science in a variety of situations is an important aspect of scientific literacy because the choice of methods and representations is often dependent on the situations in which issues are presented. Assessment items of PISA are related to general life situations and not limited to school life. The focus of the items is on personal, social, and global situations. PISA 2006 assesses relevant scientific knowledge that is not restricted to the common aspects of national science education curricula of participating countries. In accordance with PISA's focus on scientific literacy, evidence of the successful use of scientific competencies in important situations reflecting the world is required. PISA is not an assessment of context, but it assesses competencies, knowledge and attitudes as presented or relate to contexts. The contexts used for assessment items are chosen according to the relevance to students' interests and lives where science items are developed keeping linguistic and cultural differences in mind (OECD, 2006, pp. 25-28).

2.3.1.2 Scientific Knowledge Aspect

Scientific knowledge refers to both knowledge of science and knowledge about science. Knowledge of science is the knowledge about the natural world across the major fields of physics, chemistry, biological science, Earth and space science, and science-based technology. It is required for understanding science and for making sense of experiences in personal, social and global contexts. On the other hand, knowledge about science refers to the means and goals of science, which are scientific enquiry and scientific explanations respectively. Scientific explanations can be regarded as the results of scientific enquiry (OECD, 2006, pp. 31-34)

2.3.1.3 Scientific Competencies Aspect

Major competencies investigated in PISA 2006 are students' abilities to identify scientific issues; describe, explain or predict phenomena based on scientific knowledge; interpret evidence and conclusions; and use scientific evidence to make and communicate decisions. Cognitive processes that are of special meaning for scientific literacy of PISA are inductive/deductive reasoning, critical and integrated thinking, transforming representations, constructing and communicating arguments and explanations based on data, thinking in terms of models, and using mathematics. The competency identifying scientific issues includes recognizing questions that could be investigated scientifically, identifying keywords critical to scientific search, and recognizing key features of scientific investigation. The competency explaining phenomena scientifically is prominent with applying appropriate knowledge of science in a given situation and includes describing or interpreting phenomena, predicting changes, and recognizing or identifying appropriate descriptions, explanations, and predictions. Using scientific evidence requires using scientific findings as evidence for claims or conclusions while both knowledge of science and knowledge about science can be used. It includes interpreting scientific evidence, making and communicating conclusions, identifying assumptions, evidence and reasoning behind conclusions, and reflecting on the societal implications of science and technological enhancements (OECD, 2006, pp. 29-31)

2.3.1.4 Attitudes towards Science Aspect

PISA 2006 takes an innovative approach to assessing student attitudes by not only asking about attitudes in a questionnaire, but also asking in the course of the science part of the assessment what their attitudes are towards the issues that they are being tested on. This much attention of PISA to attitudes towards science is based on the belief that a person's scientific literacy includes certain attitudes, beliefs, motivations, sense of self-efficacy, values, and actions. Science education aims to develop attitudes that make students to attend to scientific issues and to acquire and apply scientific and technological knowledge for personal, social, and global benefit. PISA 2006 science assessment evaluated students' attitudes in three areas: i) interest in science; ii) support for scientific enquiry; and iii) responsibility towards resources and environment. Besides engagement in science and valuing science for own and for society, the student questionnaire collects information about students' attitudes in all three areas mentioned in a non-contextualized manner. On the other hand, contextualized items provide data on whether students' attitudes differ when assessed in or out of context, whether they vary between contexts, and whether they correlate with performance at the assessment unit level (OECD, 2006, pp. 35-36)

2.3.2 Factors Affecting Scientific Literacy

2.3.2.1 Student-Level Factors and Scientific Literacy

Socio-economic status is frequently reported to be the most important variable related to science and mathematics achievement (Caldas & Bankston, 1997; Marks et al., 2006; Kalender & Berberoğlu, 2009).

Differences in student performance are cited to be produced by disparities in access to educational resources and materials. Moreover, income and wealth have considerable impact on achievement and education outcomes (Marks et al., 2006). In developing countries where deprivation of basic resources is an important problem for a considerable proportion of students, this lack of material resources is a major cause for socioeconomic inequality in education.

Cultural resources are more effective than material resources at home regarding achievement levels of students (Turmo, 2004; Marks et al., 2006).

However, there is no differential effect of cultural factors across reading, mathematical and scientific literacies. In most instances, effects are similar.

Engagement and achievement in science learning are on many occasions reported to be facilitated by positive attitudes and affective variables (Kupermintz, 2002; Osborne et al., 2003; Singh et al. 2002; Singh et al. 2005). Engagement is raised by opportunities for students to regulate their own learning and giving autonomy (Osborne et al., 2003).

Children's beliefs of self-efficacy lead to behaviors that in turn contribute to achievement. Students with high self-efficacy try different strategies and persist, on the other hand, those who doubt their ability will give up a learning process if early efforts do not result in perceived success (Jinks & Morgan, 1999). Self-efficacy is personal expectation, self-evaluation, and personal judgments about one's ability to succeed, and it ultimately influences motivation and academic performance in the subject area (Andrew, 1998).

2.3.2.2 School-Level Factors and Scientific Literacy

Peer socioeconomic status or school's social status has a significant and independent effect on individual academic achievement and the degree of the effect is close to that of individual social status (Caldas & Bankston, 1997).

Instructional quantity represented by homework load and values that facilitate effective use of homework time, influences science learning, and thus promotes career aspirations depending on science literacy (Wang & Staver, 2001).

In almost every country, there exist various types of schools with respect to administration authority, orientation of students to different careers, and type of education. Thus the type of the school leads to different educational and occupational tracks for individuals. School characteristics such as amount of educational and material resources, disciplinary climates, academic environment, curriculum, teaching practices, teacher qualities, and administration influence students' educational outcomes to varying degrees (Marks et al., 2006).

School systems highly differentiated with respect to student allocation to different academic tracks can increase degrees of socioeconomic inequalities in education (Marks et al., 2006). Therefore, academic allocation of students to diverse tracks should be based on objective evaluation of student performance, rather than socioeconomic criteria.

Motivational differences related to class type and ability are found to impact learning of science (DeBacker & Nelson, 2000). Therefore, class types with respect to instructional domain such as biology and physics, and ability grouping could be among the school factors influencing scientific literacy.

Classroom environment shows a positive correlation with positive attitudes, higher levels of involvement, strong relations with classmates, and use of different teaching strategies and learning activities (Osborne et al., 2003). Teaching quality determined by clarification of learning goals, communication, reviews, supportive social contexts helping students feel accepted, cared and valued, and allowance for different cognitive styles, is much effective in science performance (Cooper & McIntyre, 1996). Students perceive lack of quality teaching and a lack of previous student preparation as the major failure factors and there is significant association between achievement and student perceptions of teacher expectancies, parent involvement, quality of science teaching, a supportive learning environment, and previous preparation (Fonseca & Conboy, 2006).

Most international data almost invariably suggest that there is negative or no relationship between involvement of family in schoolwork and the student achievement (Marks et al., 2006; Okpalla et al., 2001). However, effectiveness of parental involvement depends on type of involvement, family income, and home environment.

2.4 Attitudes towards Science

Students' attitudes towards science have been extensively researched by science educators for the past thirty years. However, definition of an attitude towards science is fuzzy, and is often not well understood. This lack of clarity has been of great interest besides the effects on studying science and thus science educators tried to elaborate the concept. At first place, Klopfer (1971, as cited in Osborne, Simon, & Collins, 2003) put forward some affective behaviors that are indicators of positive attitudes toward science. Those were manifestation of favor for science and scientists, acceptance of scientific enquiry as a way of thought, adoption of scientific attitudes, enjoying science learning, development of interests in science and science-related activities, and development of interest in pursuing careers in science or science-related work. Further, the basic distinction between "attitudes toward science" and "scientific attitudes" was made by Gardner (1975). Scientific attitudes characterize scientific thinking and are cognitive in nature, whereas attitudes towards science are the feelings, beliefs and values held about the enterprise of science, school science, and the impact of science on society (Osborne et al., 2003). Osborne et al. (2003) asserted that attitudes towards science are not composed of single unitary construct, but rather consist of a large number of subconstructs all of which contribute in varying proportions to an individual's attitudes towards science. From the accumulated research, Osborne and fellow researchers exemplified a set of components in measures of attitudes to science:

- perception of science teacher
- anxiety to science
- value of science
- self-esteem at science
- efficacy in science-related actions
- motivation towards science

- enjoyment of science
- attitudes of peers, friends and parents towards science
- nature of classroom environment
- achievement in science
- fear of failure on course

Primarily, attitudes have been measured through the use of questionnaires that commonly consist of Likert-scale items. Such items are derived from the free response answers generated by students, which is the major justification for their validity. However, if there is more than one construct in each scale, giving a summated rating for attitudes is meaningless. Many instruments allegedly called successful suffer from significant problems of internal consistency and unidimensionality (Gardner, 1995).

PISA's attention to attitudes towards science is based on the notion that a person's scientific literacy includes certain attitudes, beliefs, motivational orientations, sense of self-efficacy, and values. PISA 2006 science assessment evaluated students' attitudes in three areas: (i) interest in science, (ii) support for scientific enquiry, and (iii) responsibility towards resources and environments. These three areas were regarded as international indicators of students' positive reception of science, science-specific attitudes and values, and responsibility towards science-related problems with national and international influences. However, the main interest was to gather information about the causes and effects of declining enrolments for science studies among students rather than assessing attitudes toward school science or teachers (OECD, 2007).

2.5 Related Modeling Studies

Hierarchical linear modeling is an approach to analysis of hierarchical or nested data. HLM, also called "multilevel modeling", "random coefficient modeling", or "mixed effects modeling", is a more advanced form of simple linear regression and multiple linear regression. HLM methodology can be viewed as a generalization and synthesis of regression and analysis of variance (ANOVA). The basic idea of this approach is that groups at the lowest level have intercepts and slopes that represent the correlation among variables within each group. The intercepts and slopes for each group vary across groups and the idea is to explain this variability. Multilevel analysis allows variance in outcome variables to be analyzed at multiple hierarchical levels, whereas in simple linear and multiple linear regressions all effects are modeled to occur at a single level. Therefore, HLM is appropriate for use with nested data.

2.5.1 Related Studies using Hierarchical Linear Modeling

Lee and Bryk (1989) used the data from High School & Beyond survey (http://nces.ed.gov/surveys/hsb/) to identify the characteristics of secondary schools that promote high levels of achievement and equitable distribution of achievement across different social, racial, ethnic, and academic backgrounds of students. The data of the study consisted of a subsample of HS&B survey including 10,187 students in 160 high schools. Hierarchical linear modeling techniques were used to investigate the effect of environment and academic organization of high schools on four social distribution parameters related to mathematics achievement. They found that high average achievement in mathematics is related to school's social composition and academic emphasis. That is, organizational differences among schools apply a substantial impact on students' achievement. The study also provided evidence that a student's individual interest in academic activities contributes to individual achievement and that the interest and commitment of teachers contribute to academic achievement.

Lee, Smith and Croninger (1997) investigated the characteristics of high school organization that are positively related to learning in mathematics and science and related to equity during the first and last two years of high school. Their sample was 9,631 students in 789 U.S. high schools and HLM models had a three-level nested structure where multiple test scores were nested in students, and students in turn were nested in schools. Researchers concluded that structural practices that schools engage in influence academic achievement and that these effects on learning persist into the later years of high school, rather than disappear after the early years. Their results also showed that the forms that social relations, the curriculum, and instruction take in individual schools are more important than the number and type of practices that schools adopt.

Ma and Klinger (2000) examined the effect of student and school factors on performance of Grade 6 students in New Brunswick province of Canada in mathematics, science, reading, and writing. The province is officially bilingual and there are separate English and French school systems. In Ma and Klinger’s analysis, data of the Grade 6 student population in the English system participated in the New Brunswick School Climate Study was used. Thus, the sample of the study was 6,883 students from 148 schools. The researchers emphasized that gender and socioeconomic status are significant predictors of academic achievement, and that school size and parental involvement significantly affected only the relationship between mathematics achievement and individual socioeconomic status.

Akyüz (2006) investigated the effects of mathematics teacher and classroom characteristics on students’ mathematics achievement across Turkey, European Union countries, and candidate countries to European Union by analyzing the data from student and teacher background questionnaire and mathematics achievement test in the Third International Mathematics and Science Study-Repeat (TIMSS-R). A sum of 44,806 students from three groups was the subjects of the study and effects of teachers’ background, their instructional practices, and learning environment and activities on mathematics achievement were modeled by multilevel techniques. There were substantial differences between groups with respect to teacher and class characteristics, whereas home educational resources had positive impact in all groups. Teaching practices showed a great variation from country to country, however, a more disciplined classroom environment is associated with greater success.

İş Güzel (2006) examined the impact of human and physical characteristics of education environment on students' mathematical literacy skills across Turkey, European Union countries, and candidate countries for European Union using PISA 2003 data. Subjects of the study were 4,855 students from 159 Turkish schools, 5,129 students from 189 European Union countries, and 4,419 students from 185 schools from candidate countries. In all three groups, higher mathematics literacy is associated with more educational resources at home, higher levels of self-efficacy in mathematics, lower levels of mathematics anxiety, more positive self-concept in mathematics, and more positive disciplinary climate during lessons. On the other hand, schools with higher average self-efficacy of students in mathematics had higher mean school mathematical literacy scores. In Turkey, school size and mathematics student-teacher ratio influenced the disciplinary environment of lessons whereas academic selection procedures influenced the mathematics self-efficacy in the candidate countries.

Çalışkan (2008) examined the effect of school and student characteristics on scientific literacy skills of Turkish students who participated in PISA 2006. The whole extent of Turkish sample was 4,942 students from 10 primary schools, 88 general high schools, and 62 vocational high schools in PISA 2006, and data for them were all used in Çalışkan's study. The researcher used PISA indices as factors and investigated the differences between school types and regions of schools. Results showed that economic and social status, general value of science and self-efficacy in science had significant effect on Turkish students' scientific literacy skills but there were great variations between schools. Students from general high schools obtained higher scientific literacy scores than those from vocational high schools.

Young, Reynolds and Walberg (1996) investigated the relative importance of school and individual factors in the determination of science learning. The researchers analyzed the achievement test scores in relation to student and school factors in a sample of 2,535 tenth grade students from 51 public schools. Hierarchical linear modeling analyses showed that individual measures accounted for most of the variance where initial science attitude, instructional time, home environment, and exposure to mass media were significant student-level factors influencing science achievement. Results also indicated that most of the variance accounted for in science achievement was at the individual level rather than at the school level. Moreover, the researchers concluded that educational productivity is driven largely by individual-level psychological factors.

2.5.2 Related Studies using Other Modeling Techniques

Ferry, Fouad and Smith (2000) examined the effects of family context and person input variables on learning experiences, self-efficacy, outcome expectancies, interests, and goals by applying causal modeling techniques. Their data were collected from 791 undergraduate students enrolled in psychology classes at two universities. Because of the complex and multidimensional nature of family context and the need to incorporate math and science domain-related indices, students were given instrument packets comprised of a combination of researcher-constructed and existing instruments. By the results of the study, parental encouragement was found to have significant effects on learning experiences and outcome expectancies. On the other hand, learning experiences were found to directly influence self-efficacy and outcome expectancies, while influence of self-efficacy and outcome expectancies on interests and goals are direct and strong.

Yang (2003) investigated the dimensionality of socioeconomic status and its impact on mathematics and science performance at student and school levels by applying two-level structural equation modeling techniques. Data were drawn from 13-year-olds of 17 countries in the Third International Mathematics and Science Study making a sample of 123,031 students and 3,148 schools. For a majority of the countries in that study, student-level economic and cultural dimensions were identified. However, the cultural dimension had the utmost influence on mathematics and science achievement of students. On the other hand, only a general economic dimension was found at school level in most countries. Moreover, the results indicated that dimensions of socioeconomic status had different effects on mathematics and science achievement at individual and school levels.

Özdemir (2003) examined the factors that are related to Turkish students' science achievement in TIMSS-R by using structural equation modeling techniques. Science scores were modeled with respect to instructional practices, affective characteristics, and socioeconomic status of 7,841 students. Strongest positive relation was between socioeconomic status and science achievement. Also, perception of success in science and science achievement were highly related. However, science achievement had a negative relation with student-centered activities whereas teacher-centered activities had a positive effect on science scores in TIMSS-R for Turkish students.

İş Güzel and Berberoğlu (2005) investigated the factors affecting reading and mathematical literacy skills of 15-year-old students across cultures using the data of Brazil, Japan, and Norway from PISA 2000. They studied reading literacy, attitudes towards reading, student-teacher relations, communication with parents, classroom climate, use of technology, and attitudes towards mathematics with structural equation modeling. For all three countries, the researchers found that, reading literacy had the highest impact on mathematical literacy skills; mathematical literacy had a strong relation to attitudes toward mathematics; attitudes toward reading was negatively related to mathematical literacy but positively related to reading literacy; and communication with parents had a positive relation with reading literacy skills. A disciplined classroom environment promoted success for Japan in PISA tests; however, for Brazil, the situation was the inverse. On the other hand, use of technology had a strong impact on reading skills of Brazilian students, but negative effects were found for others.

Farmer, Wardrop, and Rotella (1999) tried to identify factors that differentiated women in a science career from women in other career fields and to compare those factors for men and women, using longitudinal data of 459 men and women who were at ninth or twelfth grades in 1980. Both 1980 and 1990 questionnaires assessed math and science attitudes and the number of math and science courses taken in high school and college. Latter questionnaire repeated some items from the previous one. Separate models were developed for men and women using structural equation modeling techniques. Models contained demographic, cognitive, environmental, and behavioral factors. Findings from the study suggested that for women who had aspirations for a prestigious career when they were in high school and as young adults, a science career was more likely. That is, valuing math and science for career convenience facilitates pursuit of a science career for women. Moreover, both women and men in science careers were intrinsically motivated to take more science courses than required, but for men, science Grade Point Average (GPA) had a stronger influence on pursuing a science career than taking elective science courses.

2.6 The Present Study

In this chapter, theoretical framework of the study was explained and related studies including the factors that will be analyzed in this study were reviewed. These findings indicated a general overview of the factors having impact on science performance of students and also suggested the need for further studies. PISA project is most fruitful with its comprehensive database to investigate influences of student level factors, school level factors, and attitudes towards science on scientific literacy.

PISA 2006 database is suitable to test different models explaining the associations between student level factors, school level factors and the science performance across different cultural settings. Therefore the present study might provide many results that could be beneficial to educators and policy makers so as to enhance students' scientific literacy skills in Turkey, and also beneficial to comparing the results of Turkey, Canada, and Sweden regarding cultural diversities.

CHAPTER 3

METHODOLOGY

3.1 Population and Sample

To establish comparability between results, an international survey should guarantee a comparability of target populations. However, regarding the characteristics of educational systems of each country, there exist substantial differences between countries. Differences in nature of pre-primary education, age to start formal schooling, institutional structures and such make it almost impossible to establish a basis for comparability of grade levels (OECD, 2007). A disadvantage of defining grade level as the basis for target population is that students at grade levels may result in selection of different ages and maturity levels between education systems. Regarding this, PISA prefers defining its target population with reference to age group. Therefore, PISA assesses the knowledge and skills of individuals born within a reference time period although they may have diverse experiences of education at school or outside the school.

The desired PISA target population in each country is defined as “15-year-old students attending educational institutions located within the country, in grades 7 and higher”. This definition directly depends on the testing dates. By the international requirement of PISA, unless otherwise agreed, the assessment had to be conducted during a 42-day period between defined dates 1 March and 31 August 2006. However, international target population was slightly adapted to better fit the age structure of most of the Northern Hemisphere countries (OECD, 2009). Consequently, the target population of PISA by its own definition is “all students aged from 15 years and 3 completed months to 16 years and 2 completed months at the beginning of the testing window, attending educational institutions located within the country, and in grade 7 or higher" (OECD, 2007; 2009). This target population excludes students in grades 6 or lower, and residents attending schools in a foreign country, but it includes foreign students attending schools in the country of assessment.

National and international surveys usually collect data from a sample rather than the whole population. Drawing a sample can be done in several ways depending on the population characteristics and the survey research questions with the aim of avoiding bias in the selection procedure. A sample is only useful to the extent that it represents and can estimate some characteristics of the whole population.

Selecting members of a population by simple random sampling is the most straightforward procedure. However, it is rarely used in educational surveys because it is costly, not practical, and statistically inconvenient. Educational surveys try to understand the statistical variability of the student's outcome measure by school or class level variables and with simple random sampling it would be almost impossible to link student variables to school, class, or teacher variables. Therefore, educational surveys usually draw student samples in two steps. First, a sample of schools is selected from a complete list of schools containing the student population of interest. Then, a simple random sample of students or classes is drawn from within the selected schools. Because PISA is an educational survey, sampling should be done carefully, rigorously, and in a standardized manner.

The sampling plan of PISA has six main components (PISA Project Consortium, 2005):

i. Population definition of the students to be surveyed.
ii. An inventory of schools in which students suitable to definition are enrolled, or likely to be.
iii. Sampling of schools from the inventory.

iv. An inventory of students within each sampled school.
v. Sampling of students from each sampled school.
vi. Documentation of sampling process and indicators of effectiveness of sampling and survey operations.

3.1.1 Sampling Procedure

The sampling design used for PISA is a two-stage stratified sampling. The first is sampling of individual schools in which appropriate students could be enrolled. The second stage was the sampling of students within sampled schools.

During first-stage of sampling, systematic selection of school with respect to probabilities proportional to size takes place. Here, size is not the actual number of students enrolled, but is the estimated number of 15-year-old students at that particular school. In each country, at least 150 schools were selected, though for national analyses a larger sample was required. If the number of participating schools was fewer than 150 in a country, all schools were selected. As schools were sampled, two replacement schools were also identified from the same sampling frame, in case the selected school would choose not to participate. Minimum response rate required for initially selected schools was 85%. When the initial response rate of schools was within the range of 65% and 85%, use of a replacement school helped meeting the initial response rate.

The second stage of sampling was the sampling of students within sampled schools. Once schools were selected to be in the sample, a list of 15-year-old students was prepared in each sampled school. For each country a target cluster size was set and this value is typically 35. Alternative values could be used among countries with agreement. Equal probability selection of as many students as target cluster size was made and all students were selected if list size is smaller than the cluster size. A minimum participation rate of 80% of students within schools was required. This rate had to be met at the national level, not necessarily by each participating school. Schools in which between 25% and 50% of students were participating were not considered as participating schools, nevertheless data from those schools were included in the database and used for various estimations. Data from schools where student participation did not reach at least 25% were excluded from the database.

3.1.2 Subjects of the Study

Subjects of the present study consist of three samples; Canadian sample, Swedish sample, and Turkish sample. All three samples were extracted from PISA 2006 databases and each includes whole students and schools in the database from the country mentioned. Canadian sample is 22,646 students from 896 schools; Swedish sample is 4,443 students from 197 schools; and Turkish sample is 4,942 students from 160 schools. The total number of students drawn from PISA 2006 student database for the current study is 32,031 and the total number of schools drawn from PISA 2006 school database is 1,253.

OECD-member and non-member participants existent in the whole PISA 2006 database and numbers of students from each country are given in Appendix A.

3.2 Instruments

3.2.1 Main Assessment

In PISA 2006 three subject domains were tested, with science as the major domain for the first time in a PISA administration and reading and mathematics as minor domains. Main assessment in PISA 2006 consists of paper and pencil tests (OECD, 2006, p.13). Pencils, erasers, rulers, and in countries where they were routinely used in the classroom; calculators, were provided. No test items required a calculator, but some mathematics items involved solution steps for which the use of a calculator could be of assistance to some students (OECD, 2009, p.28).

The assessment instruments in PISA 2006 were developed around units of assessment. Each assessment unit is composed of a common stimulus and PISA items grouped around that stimulus. Often in combination, different types of stimulus are used such as text passages, tables, graphs and diagrams. A maximum of four items assessing scientific competencies and knowledge were allocated in a unit and for PISA 2006 about 60% of science units contained one or two items measuring attitudes towards science. Each attitude item involved two or three statements of interest in science or support for science and a four-point scale indicates levels of agreement on interest or support (OECD, 2009, p.28).

A total of 108 cognitive items were arranged in 37 science units, and 31 attitudinal items were embedded in those units. The whole of science units in PISA 2006 takes approximately 210 minutes of testing time. The reading assessment consisted of the same 8 units and 31 items as 2003, representing approximately 60 minutes of testing time, and the mathematics assessment consisted of 31 units and 48 items representing approximately 120 minutes of testing time. The mathematics items were selected from the 167 items used in 2003. 22 of 108 cognitive items of science were from the 2003 test and remaining 86 items were selected from a pool of 222 newly-developed items that had been field tested in all countries in 2005. There was no new item development for reading and mathematics.

The PISA 2006 test units are arranged in 13 clusters each designed to occupy 30 minutes of test time. Of those thirteen clusters, seven are science clusters, two are reading clusters, and four are mathematics clusters. The eight link units of science from 2003 test were distributed across the seven science clusters, in first or second position. The fully-linked design is a “balanced incomplete block design”. Each cluster appears in each of the four possible positions within a booklet once and so each test item appeared in four of the test booklets. Moreover, each pair of clusters appears in one and only one booklet.

Each sampled student undertook two hours of testing with randomly assigned one of the thirteen booklets. Students were allowed to take a short break after one hour. The directions to students emphasized that there were no correct answers to the attitudinal questions, and that they would not count in their test scores, but that it was important to answer them truthfully.

3.2.2 Context Questionnaires

To gather contextual information on students, PISA asks students and the principals of their schools to respond to background questionnaires. These questionnaires are central to the analysis of results in terms of a range of student and school characteristics. Two additional questionnaires are offered as international options but they are not in scope of the present study:

- A computer familiarity questionnaire
- A parent questionnaire

3.2.2.1 Student Questionnaire

The student questionnaire was administered after the literacy assessment and it took students about 30 minutes to complete the instrument. The core questions on home background were similar to those used in PISA 2003, however, for some questions the wording was modified to improve the quality of the data collection based on experiences in previous surveys (OECD, 2009, p.58).

The questionnaire covered the following aspects:

- Student characteristics: Grade, study programme, age and gender

- Family background: Occupation of parents, education of parents, home possessions, number of books at home, country of birth for student and parents, language spoken at home
- Students' views on science: Enjoyment of science, confidence in solving science tasks, general and personal value of science, participation in science-related activities, sources of information on science and general interest in learning science
- Students; views on the environment: Awareness of environmental issues, source of information on the environment, perception of the impact of environmental issues, optimism about environmental issues and sense of responsibility for sustainable development
- Students' views of science-related careers: Usefulness of schooling as preparation for the science labor market, information about science-related careers, future-oriented motivations for science and expected occupation at age 30
- Students' reports on learning time: Mode and duration of students' learning time in different subject areas and duration of students' out-of-school lessons
- Students' views on teaching and learning of science: Science course taking in current and previous year, nature of science teaching at school, future-oriented motivations to learn science, importance of doing well in subject areas and academic self-concept in science

3.2.2.2 School Questionnaire

The school questionnaire was administered to the school principal and took about 20 minutes to be completed. It covered a variety of school-related aspects:

- Structure and organization of the school

- Staffing and management
- School's resources
- School admission practices and policies
- Teaching of science
- Activities to facilitate science learning
- Environmental issues
- Career guidance for students

3.3 Validity and Reliability

3.3.1 Content-Related Evidence for Validity

In PISA 2006's main assessment, 108 cognitive items of science were used. Twenty-two of those items were from the PISA 2003's science test and 86 new items were selected from 222 items developed and field tested in all countries in 2005. There was no new item development for reading and mathematics.

Regarding the experience gained from earlier PISA assessments, to achieve conceptually rigorous material that has the highest possible levels of cross-cultural and cross-national diversity, PISA consortium made use of diverse centers of test development centers and experts from all over the world. In view of that, to prepare new science items for PISA 2006 the consortium expanded its number of test development centers over the number used for PISA 2003. Test development teams were established in five culturally-diverse and well-known institutions: the Australian Council for Educational Research (ACER, Australia), the National Institute for Educational Measurement (CITO, the Netherlands), the Department of Teacher Education and School Development, University of Oslo (ILS, Norway), the Leibniz Institute for Science Education, University of Kiel

(IPN, Germany), and the National Institute for Educational Policy Research, (NIER, Japan) (OECD, 2009, p.29).

Cognitive item development was guided by a comprehensive set of guidelines prepared at the start of the project and approved by the first meeting of the PISA 2006 Science Expert Group (SEG). The guidelines included an overview of the development process and timelines, a specification of item requirements, including the importance of framework fit, and a discussion of issues affecting item difficulty. A number of sample items were also provided. These guidelines were expected to be followed by item developers at each of the five test development centers (OECD, 2009, p.31).

In the first phase of item development, initially, test developers prepared units in the local language in a standard format, including stimulus, one or more items, and a proposed coding guide for each item. Then items were exposed to a series of cognitive laboratory activities. Each unit underwent extensive scrutiny at a meeting of members of the relevant test development team. Item writers vigorously analyze all aspects of the items from the point of view of a student, and from the point of view of a coder. Items were revised, often extensively, following item paneling. When substantial revisions were required, items went back to the paneling stage for further consideration. Next, cognitive interviews were made with students to find out the cognitive processes typically utilized as students attempted to answer the items. This stage is particularly useful in clarifying wording of questions, and informing on likely student responses that are used in refining the response coding guides. As the final step in the first phase, sets of units were piloted with several classes of 15-year-olds in schools in the country in which they were developed. As well as providing statistical data on item parameters, including the relative difficulty of items, this enabled real student responses derived under formal test conditions to be obtained, thereby enabling more detailed development of coding guides. Pilot test data were used to inform further revision of items where necessary or sometimes to discard items altogether.

At the beginning of the second phase of item development, each unit was reviewed by at least one team that was not responsible for its primary development. International item paneling brings further improvements to the items and provide feedback on operation of items in different cultures and national contexts. Then, test booklets were prepared for international pilot study, using a number of units developed at different test development centers and trialed with students in whole classes at several different schools.

An international comparative study must be culturally and contextually diverse, that is to say it should ensure an appropriate and valid content. The diversity of the PISA consortium and extensive consultation between expert groups from various cultures provides evidence for content validity. On the other hand, all members of PISA consortium with technical advisory group, expert groups, test development centers, national project managers, and consultants are all knowledgeable experts. This approach of PISA for employing experts also helps validation of the content.

3.3.2 Construct-Related Evidence for Validity

The PISA 2006 context questionnaires included numerous items on student characteristics, student family background, student perceptions, school characteristics and perceptions of school principals. Some of the items were suitable to use as single items in analyses but most questionnaire items were combined to measure latent constructs that cannot be observed directly. Scaling procedures and some transformations were needed to build meaningful indices with these items (OECD, 2009, p.304).

Simple questionnaire indices were constructed through the arithmetical transformation or recoding of one or more items whereas scale indices were constructed through the scaling of items. Typically, scale scores for these indices are estimates of latent traits derived through IRT scaling of dichotomous or Likert-type items. There are different methodological approaches for validating questionnaire constructs, each with their advantages, limitations and problems. Cross-country validity of these constructs is of particular importance as measures derived from questionnaires are often used to explain differences in student performance within and across countries and are, thus, potential sources of policy-relevant information about ways of improving educational systems (OECD, 2009, p.311).

Cross-country validity of the constructs not only requires a well-monitored process of translation into different languages but also makes assumptions about measuring similar characteristics, attitudes and perceptions in different national and cultural contexts. Psychometric techniques can be used to analyze the extent to which constructs have consistent dimensionality and consistent construct validity across participating countries (OECD, 2009, p.312).

Structural Equation Modeling (SEM) was used to confirm theoretically expected dimensions and, if necessary, to re-specify the dimensional structure. Model fit was assessed using the root-mean square error of approximation (RMSEA), the root-mean square residual (RMR), the comparative fit index (CFI) and the non-normed fit index (NNFI). RMSEA values over 0.10 are usually interpreted as a sign of unacceptable model fit whereas values below 0.05 indicate a close model fit. RMR values should be less than 0.05. Both CFI and NNFI are bound between 0 and 1 and values between 0.90 and 0.95 indicate an acceptable model fit, with values greater than 0.95 indicating a close model fit (OECD, 2009, p.312).

Using Confirmatory Factor Analysis (CFA) requires a theoretical model of item dimensionality. Confirmatory factor analyses of student data were based on the international calibration sample in order to have comparable sample sizes across OECD countries. For the comparative analysis of item dimensionality the use of random OECD sub-samples was considered as appropriate. CFA were carried out only for the student questionnaire data. Analyses of data show that depending on the structure of educational systems in some countries a considerable amount of variation is found between schools. For most of the student-level indices the average proportion of between-school variance is below 10%. However, for some indices there is a considerable variance between schools.

Models were estimated both for international pooled sample and country samples. Results for similar indices showed a high degree of consistency across countries. Comparing within-country results for principal component analyses indicated that loadings of factor generally followed similar patterns across countries (OECD, 2009, p.315-348).

3.3.3 Evidence for Validity of Embedded Attitudinal Scales

For the main study in PISA 2006, attitudinal items were embedded within units of the science test in order to obtain measures of two attitudinal dimensions: interest in science and support for scientific inquiry. To ensure the cross-national validity of attitudinal scales, development process consists of four steps. First, the construct should have well-established theoretical underpinnings. Second, there must be wide agreement that the items that are used in PISA are reflective of the underlying conceptual definition of the domain. Third, psychometric analyses are undertaken to ensure that the sets of items that are deemed to be reflective of the underlying construct can be brought together in a coherent way to provide indicators of the underlying construct. Fourth, the constructed scales are reviewed for the extent to which relations with other variables make conceptual sense.

Some of the procedures that PISA follows to achieve attitudinal validity include (OECD, 2009, p.352):

- The use of skilled professional test development teams from a variety of participating countries
- Review of the items as they are prepared by experts who have been directly involved
- Opportunities for review and evaluation of the drafted items by participating countries on multiple occasions
- A detailed set of translation and translation verification protocols that are aimed at ensuring the conceptual and psychometric equivalence of the items across languages and cultures
- A range of small, medium and large trial testing activities where students are asked to respond to the item and to reflect upon the meaning of the items to them

The results of a two dimensional confirmatory factor showed that interest in science items generally loaded on one dimension whereas items selected for main study for support for inquiry domain items loaded on one factor. Estimated latent correlation between the two dimensions was 0.594 and RMSEA measure of model fit was 0.025, which was considered quite acceptable (OECD, 2009, p.353). These findings indicate well established dimensions and provide evidence for validity of attitudinal scale.

An alternative approach to assessing item dimensionality is to assess the fit of the data to a multi-dimensional IRT model. A five-dimensional model (reading, mathematics, science, interest in science, and support for scientific inquiry) was fit to the data. The fit statistics showed an acceptable fit to the multi-dimensional item response model and the fit mean squares were close to normally distributed. This distribution, too, supports the validity of attitudinal scales.

The basic psychometric characteristics of the embedded attitudinal scales appear to be sound. Existence of two factors is confirmed and both the fit to the scaling model and the reliabilities of scales appear to be adequate. The review of differential item functioning (DIF) with respect to country and gender shows that the embedded attitude items have fewer instances of DIF (by country and gender) than do the PISA cognitive items (OECD, 2009, p.357).

The embedded items behave in expected and predictable ways with the other PISA variables. Principal component analysis supports that they are distinct dimensions that correlate appropriately with parallel scales that were included in the context questionnaires (OECD, 2009, p.364).

3.3.4 Reliability

At the international level, reliability of the test is calculated for each of the five overall scales (mathematical literacy, reading literacy, combined scientific literacy and the attitude scales interest and support) before conditioning and based upon five separate scalings, using plausible values and using weighted likelihood estimates (WLEs). The results are displayed in Table 3.1 as follows:

Table 3.1 Reliabilities of Each of the Overall Scales When Scaled Separately

Domain	Reliability (PV)	Reliability (WLE)
Mathematics Literacy	0.613	0.784
Reading Literacy	0.429	0.780
Science Literacy	0.856	0.832
Interest in science	0.886	0.867
Support for scientific inquiry	0.725	0.705

3.4 Procedures

3.4.1 Design of the Study

The present study is a quantitative research with non-experimental study. Regarding the structure and properties of PISA, which is the supplier of the data of the present study, and the method used in the present study; namely multilevel modeling, this study can be considered as a cross-sectional correlational research. PISA assesses mathematical, scientific and reading literacies of 15-year-old students, or as PISA's official saying; students nearing the end of their compulsory education. Thus, PISA tries to assess all 15-year-old students worldwide, or the member and partner countries subset in 2006. Because of this definition of the population, PISA is a cross-sectional study and data of the present study is cross-sectional data. Moreover, hierarchical linear modeling techniques are used and the relationship of students' background characteristics and school characteristics with students' performance in science is investigated with no causal concern. Therefore, the present study is a correlational study.

3.4.2 Research Questions

The present study aims to determine:

1. Which student level factors have significant effect on students' scientific literacy skills in PISA 2006 across Canada, Sweden, and Turkey?
2. Which school level factors have significant effect on students' scientific literacy skills in PISA 2006 across Canada, Sweden, and Turkey?

3. What proportions of variances in scientific literacy scores in PISA 2006 is explained by school-related factors across Canadian, Swedish, and Turkish students?

Research questions of the present study are:

1. Are there differences in the students' scientific literacy skills among schools in Canada?
2. Which school characteristics are associated with the differences in students' scientific literacy skills in Canada?
3. Which student characteristics are associated with the differences in students' scientific literacy skills in Canada?
4. Which cross-level interactions of school characteristics and student characteristics affect the students' scientific literacy skills in Canada?
5. Are there differences in the students' scientific literacy skills among schools in Sweden?
6. Which school characteristics are associated with the differences in students' scientific literacy skills in Sweden?
7. Which student characteristics are associated with the differences in students' scientific literacy skills in Sweden?
8. Which cross-level interactions of school characteristics and student characteristics affect the students' scientific literacy skills in Sweden?
9. Are there differences in the students' scientific literacy skills among schools in Turkey?
10. Which school characteristics are associated with the differences in students' scientific literacy skills in Turkey?
11. Which student characteristics are associated with the differences in students' scientific literacy skills in Turkey?
12. Which cross-level interactions of school characteristics and student characteristics affect the students' scientific literacy skills in Turkey?

3.4.3 Threats to Internal Validity

Because PISA 2006 is an assessment study and there is no treatment, and also one-time testing not more than three hours occurs, history, maturation and subject attrition are not threats to the present study.

PISA uses main tests and context questionnaires only one time and tests are not public. Thus, participants' performance may not differ because they are familiar with the measure. Hence, testing could not be a threat to the present study.

Although PISA follows an age-based selection rule for its sample, all the characteristics of subjects cannot be controlled and initial differences may interfere with the differences observed. Also, oversampling of some strata for national purposes of participating countries in PISA could be observed. Therefore, selection could be a threat to the present study.

The instrumentation threat is caused by inconsistencies with the testing instrument i.e., interviewer, grader, or the test itself. However, the tests of PISA 2006 do not change during a particular study. On the other hand, for the test administrators and graders, PISA consortium provides clear and comprehensive instructions, guidelines, and grading manuals and all the procedures are carried out by trained people. Therefore, instrumentation could not be a threat to the present study.

Finally, the chance factors of extreme scorers remain unaccounted for and there is no intervention in PISA 2006 due to its nature. Thus, regression could not be an internal threat to the present study.

3.5 Data Collection

PISA consortium assigned each country a native expert called National Project Manager (NPM) to implement PISA study according to the procedures prepared by the consortium. A base working location for each national project group, called national center, was established and several assistants were assigned to National Project Manager. For the school level operations the NPM coordinated activities with school level staff, called school coordinators. Trained test administrators administered the PISA assessment in schools (OECD, 2009, p.106). NPMs used the detailed instructions in the School Sampling Preparation Manual to document their school sampling plan and to prepare their school sampling frame. The national target population was defined, school and student level exclusions were identified, and aspects such as the extent of small schools and the homogeneity of students within schools were considered in the preparation of the school sampling plan. Following the selection of the school sample by the consortium, the list of sampled schools was returned to national centers. NPMs then contacted these schools and requested a list of all PISA-eligible students from each school. This was provided on the student listing form, and was used by NPMs to select the student sample. NPMs were required in most cases to select the student sample using KeyQuest, the PISA student sampling and data entry software prepared by the consortium. KeyQuest generated the list of sampled students for each school, known as the "Student Tracking Form" (OECD, 2009, p.109).

National centers establish database of schools before testing began to record the delivery of materials to and from schools, inventories of materials sent and returned, and to monitor the progress of the materials throughout the booklet processing after the testing. Upon receipt of materials back from schools, the counts of completed and unused booklets are checked against the participation status information recorded on the student tracking form by the test administrators.

Test administrators record all key test session information using a test session report. This report provides detailed data on test administration, including session date and timing, the position of the test administrator, conduct of the students, and testing environment.

PISA quality monitors (PQMs) are individuals employed by the consortium and located in participating countries. They visit a sample of schools to record the implementation of the documented field operations in the main study. Typically, two PQMs were engaged for each country, each to visit seven or eight schools. The majority of school visits were unannounced.

The consortium prepared coding guides with comprehensive criteria for coding and provided them to NPMs for each domain of science, reading and mathematics. Those coding guides included numerous examples of acceptable and unacceptable responses. Representatives from each national center attended two international training sessions about coding just before field trial and before the main study. A full-time overall supervisor of coding process who was familiar with the design, reliability-checking procedures, and schedules of coding and also familiar with the content of the tests and coding guides handled and oversaw the complex activities of coding. To prevent disclosure of PISA material, prospective coders were required to sign a confidentiality form before acquiring any copies of materials.

The consortium provided participating countries with the data entry software KeyQuest, which contained the database structures for all of the booklets, questionnaires and tracking forms used in the main survey. Variables could be added or deleted as needed for national options and approved adaptations to response categories could be made. Student response data were entered directly from the test booklets and questionnaires. The software performed validation checks as data were entered (OECD, 2009, p.120).

3.6 Variables of the Study

The PISA 2006 context questionnaires included items on student characteristics, student family background, student perceptions, school characteristics and perceptions of school principals. Optionally, in 16 countries parent questionnaires were administered to the parents of the tested students. However, because only Turkey among the three countries in scope administered parent questionnaire, information from parent questionnaires was not used in this study. Some of the items were designed to be used in analyses as single items. On the other hand, most questionnaire items were designed to be combined in some way so as to measure latent constructs that cannot be observed directly. Then, student, school and parent questionnaire indices were constructed and validated. For the cause of this study, besides some simple indices which were constructed through the arithmetical transformation or recoding of one or more items, student-level and school-level indices constructed by PISA consortium were included as variables of the present study.

Those indices which involve multiple responses were scaled using a weighted maximum likelihood estimate (Warm, 1985) using one-parameter item response model. Except the ones about occupational status of parents, educational level of parents, immigration background and socioeconomic status, all student-level indices used in the present study are standardized WLE scales with mean of the OECD student population is zero and the standard deviation is one. During the construction of the indices, countries were given equal weight for standardization process (OECD, 2007, p.332). On the other hand, school-level indices about activities for learning environmental topics, activities to promote learning of science, quality of educational resources, and teacher shortage are of standardized WLE scales. However, the outcome variable; scientific literacy score is not actual score but is an expected score from distribution of scores that could be reasonably assigned to each individual.

3.6.1 Student-Level Variables

3.6.1.1 Highest Occupational Status of Parents

Students were asked open-ended questions to determine both the father's and mother's occupation. Highest occupational status of parents (hisei) corresponds to the higher PISA's international socio-economic index of occupational status (ISEI) score of either parent or to the only available parent's ISEI score.

3.6.1.2 Highest Educational Level of Parents

The PISA index of highest educational level of parents (hisced) corresponds to the higher International Standard Classification of Education (ISCED) level of either parent. ISCED was designed by UNESCO to serve as an instrument suitable for assembling, compiling and presenting statistics of education both within individual countries and internationally.

3.6.1.3 Immigration Background

Information on the country of birth of the students and their parents was collected. The PISA index on immigrant background (immig) is derived from this information, and has the following categories: (1) native students (those students who had at least one parent born in the country), (2) first-generation students (those students born outside the country of assessment and whose parents were also born in another country), and (3) second generation' students (those born in the country of assessment but whose parent(s) were born in another country).

3.6.1.4 Economic, Social and Cultural Status

The PISA index of economic, social and cultural status (escs) was derived from the highest international socio-economic index of occupational status of father or mother, index of highest educational level of parents, and home possessions. Education, occupational status, and wealth are regarded as determinants of socio-economic status. As there was no straight way to define parental income except for the few countries which took parent questionnaire, amount of household items was taken as the indicator of wealth. The scale of the index is standardized to have an OECD mean of zero and a standard deviation of one.

3.6.1.5 Availability of Household Possessions Indicating Family Wealth

The index of availability of household possessions indicating family wealth (wealth) was derived from three sets of items: i) whether students had a room of their own, access to internet, a dishwasher and a DVD or VCR player; ii) number of items they had at home among cellular phones, televisions, computers and cars; and iii) three country-specific items thought to indicate wealth defined by each country.

3.6.1.6 Home Possessions

The PISA index of home possessions (homepos) is a summary index of all household items and also included the variable indicating the number of books at home. It is obtained by asking students whether they had at their home: a desk to study at, a room of their own, a quiet place to study, a computer they can use for school, an educational software, a link to the Internet, their own calculator, classic literature, books of poetry, works of art (e.g. paintings), books to help with their school work, a dictionary, a dishwasher, a DVD player or VCR, the number of cellular phones, televisions, computers, cars and books at home, and three other country-specific items.

3.6.1.7 Home Educational Resources

The index of home educational resources (hedres) was derived from students' reports on availability of items in their home: a desk to study, a quiet place to study, computer used for school work, educational software, calculator, books to help with school work, and dictionary. During scaling, national item parameters were estimated for each country instead of using parameters estimated for the combined OECD sample.

3.6.1.8 Cultural Possessions at Home

The index of cultural possessions at home (cultposs) was derived from students' reports on the availability of classic literature, books of poetry and works of art in their home. These items were exemplified so that students recognize cultural possessions. National item parameters were estimated for each country.

3.6.1.9 General Interest in Science

The index of general interest in science (intscie) indicates the student's level of interest in learning the topics physics, chemistry, biology of plants, human biology, astronomy, geology, designing science experiments, and requirements for scientific explanations. A four-point scale was used and items were recoded so that positive values indicate higher interest.

3.6.1.10 Enjoyment of Science

The index of enjoyment of science (joyscie) indicates the agreement level of students' on statements of having fun in learning science topics, liking reading about science, being happy with doing science problems, enjoyment of gaining knowledge in science, and interest in learning science. A four-point scale of agreement was used and items were recoded so that positive values indicate higher levels of enjoyment of science.

3.6.1.11 Instrumental Motivation in Science

Instrumental motivation in science index of PISA (instscie) indicates the agreement level of students' on statements of worthiness of learning school science for it will be useful in work life later on, importance of learning school science subjects because of the need to study later on, usefulness of studying school science, studying school science is helping to improve career prospects, and learning school science subjects helps to get a job. A four-point scale of agreement was used and items were recoded so that positive values indicate higher levels of motivation.

3.6.1.12 Future-Oriented Motivation to Learn Science

The index of future-oriented motivation to learn science (sciefut) indicates the agreement level of students' on statements of desire to work in career involving science, desire to study science after secondary school, desire to spend a life doing advanced science, and desire work on science projects as an adult. A four-point scale of agreement was used and items were recoded so that positive values indicate higher levels of motivation to use science in the future.

3.6.1.13 Self-Efficacy in Science

The index of self-efficacy in science (scieeff) was developed from the beliefs of students' in their abilities to perform the tasks: recognizing science question about a health issue in a newspaper report; explaining the cause of occurrence of earthquakes more in some areas than in others; describing the role of antibiotics in disease treatment; identifying the science question for garbage disposal; predicting the effect of changes in environment on certain species; interpreting the scientific information on the labels of food; and distinguishing the better one from other explanations about acid rain formation. Response categories were that the student could do it easily, with some effort, with great effort or that student could not do. Items were recoded so that positive values point to higher degrees of self-efficacy in science.

3.6.1.14 Self-Concept in Science

The PISA index of self-concept in science (scscie) indicates the agreement level of students' on statements that: learning advanced school science is easy; they can give good answers to science questions; they can learn school science topics quickly; school science topics are easy; they can understand well the science concepts when taught; and they can understand new ideas easily in school science. A four-point scale of agreement was used and items were recoded so that positive values indicate a positive self-concept in science.

3.6.1.15 General Value of Science

The index of general value of science (genscie) indicates the agreement level of students' on statements: advancements in science and technology improve living conditions of people; science is important for understanding the natural world; advancements in science and technology improve economy; science is valuable to society; and advancements in science and technology bring benefits to society. A four-point scale of agreement was used and items were recoded so that positive values indicate higher levels of general value rated to science.

3.6.1.16 Personal Value of Science

The index of personal value of science (perscie) indicates the agreement level of students' on statements: science concepts help to realize relations to other people; they will use science in several ways in adult life; science is relevant to them; science helps to understand things around; and there will be many opportunities to use after school life. A four-point scale of agreement was used and responses to items were recoded so that positive values indicate positive perceptions of personal value of science.

3.6.1.17 Science Activities

The index of students' science-related activities (scieact) was derived from the frequency of student's activities of watching television programs about science, obtaining books about science topics, visiting web sites related to science, listening to radio for advances in science, reading magazines and newspapers articles about science, and attending a science club. Positive values on this index indicate higher frequencies of science activities taken by the student.

3.6.1.18 Awareness of Environmental Issues

The index of students' awareness of environmental issues (envaware) shows beliefs of students about their own level of information on environmental issues: increase of greenhouse gases in the atmosphere; use of genetically modified organisms; acid rain, nuclear waste; and consequences of clearing forests for other land use. A four-point scale of awareness was used and items were recoded so that higher values indicate more awareness of environmental issues.

3.6.1.19 Level of Concern for Environmental Issues

The index of students' level of concern for environmental issues (envperc) was derived from students' level of concern about the environmental issues: air pollution; energy shortages; extinction of plants and animals; clearing of forests for other land use; water shortages; and nuclear waste. A four-point scale of concern was used and items were recoded so that higher values on the index indicate more concern about environmental issues.

3.6.1.20 Optimism Regarding Environmental Issues

The index of students' optimism regarding environmental issues (envopt) was derived from students' optimism concerning the development over the next 20 years of the problems associated with the environmental issues: air pollution; energy shortages; extinction of plants and animals; clearing of forests for other land use; water shortages; and nuclear waste. A three-point scale with response categories "improve", "stay about the same", and "get worse" was used. Items were recoded so that higher values indicate higher levels of optimism regarding environmental issues.

3.6.1.21 Responsibility for Sustainable Development

The index of students' responsibility for sustainable development (respdev) was derived from students' level of agreement with importance of checking the exhaust emissions of cars regularly as a condition of use; with discomfort about energy waste by unnecessary electrical appliances; with favoring laws that regulate factory emissions despite the increase in prices of products because of that; with minimizing the use of plastic packaging to reduce waste; with forcing industries to prove that dangerous waste disposed of safely; with protecting the habitats of endangered species by law; and with favoring production of electricity from renewable sources even if that increases the production cost. A four-point scale of agreement was used and items were recoded so that higher values on the scale indicate more responsibility for sustainable development.

3.6.1.22 School Preparation for Science-Related Careers

PISA index of school preparation for science-related careers (carprep) displays agreement level of students with the statements: science subjects available at school equip with knowledge and skills for a science-related career or for different careers; the subjects studied provide them with knowledge and skills for science-related careers; and teachers provide them with the skills and knowledge needed for science-related careers. A four-point scale of agreement was used and items were recoded so that higher values indicate higher levels of agreement with school's utility in preparing students for science careers.

3.6.1.23 Student Information on Science-Related Careers

The index of student information on science-related careers (carinfo) was derived from students' levels of information about science-related careers available in the job market, places to get information about them, what to do if a science career is desired, and employers of science-related jobs. A four-point scale indicating information levels was used. Items were recoded so that higher values indicate that students are more informed about science-related careers.

3.6.1.24 Interaction in Science Teaching and Learning

The index of interaction in science teaching and learning (scintact) indicates students' views about the frequencies of opportunities to explain own ideas; involving student opinions about the topics in lessons; class debate or discussion; and student discussions about topics and activities, during science classes at school. A four-point scale of frequency was used and items were recoded so that positive values indicate higher frequencies of interactive science teaching.

3.6.1.25 Hands-On Activities in Science Teaching and Learning

The index of hands-on activities in science teaching and learning (schands) was developed from students' responses about the frequencies of and time spent for hands-on activities such as doing practical experiments in laboratory; designing school science experiments applicable in the laboratory; drawing conclusions from an experiment; and doing experiments by following the instructions of teacher during learning school science topics. A four-point scale of frequency was used and items were recoded so that positive values indicate higher frequencies of science teaching with hands-on activities.

3.6.1.26 Student Investigations in Science Teaching and Learning

The index of student investigations in science teaching and learning (scinvest) indicates the frequencies of student activities: designing own experiments; choosing own investigations; and investigating and testing own ideas, when learning school science topics at school. A four-point scale of frequency was used and items were recoded so that higher values indicate higher frequencies of student investigations in science teaching.

3.6.1.27 Focus on Model or Applications in Science Teaching and Learning

The index of focus on model or applications in science teaching and learning (scapply) indicates students' views about the frequency of activities where teacher explains application of a school science idea to different phenomena; uses science make students understand the outside world; clearly explains the relevance of science concepts to human life; and uses examples of technological application to show the relevance to society, during teaching science topics at school. A four-point scale of frequency was used and items were recoded so that positive values indicate more teaching activities with focus on model or applications.

3.6.1.28 ICT Internet/Entertainment Use

The index of ICT Internet/entertainment use (intuse) indicates the frequencies of students' computers use to: browse the Internet for information; play games; use the Internet to collaborate with a group or team; download software; download music; and communicate with others. A five-point scale of frequency was used and items were recoded so that positive values indicate higher frequencies of ICT use.

3.6.1.29 ICT Program/Software Use

The index of ICT program/software use (prguse) displays frequencies of students' computers use to: write documents; use spreadsheets; draw, paint or form graphics; use educational software; and write computer programs. A five-point scale of frequency was used and items were recoded so that positive values indicate higher frequencies of ICT use.

3.6.1.30 Self-Confidence in ICT Internet Tasks

The index of self-confidence in ICT Internet tasks (intconf) indicates students' confidences to perform computer tasks such as making online chat; searching the Internet for information; downloading files, getting computer programs; attaching files to e-mail messages; downloading music; and writing and sending e-mails. A four-point scale with response categories "I can do this very well by myself", "I can do this with help", "I know what this means but I cannot do this", and "I don't know what this means" was used. Items were recoded so that positive scores indicate high self-confidence.

3.6.1.31 Self-Confidence in ICT High-Level Tasks

The index of self-confidence in ICT high-level tasks (highconf) indicates students' confidences to perform high-level computer tasks such as using software to detect and clean computer viruses; editing digital photographs or other graphic images; creating a database; using a word processor; using spreadsheet to plot a graph; creating presentations; creating multimedia presentations; and constructing web pages. A four-point scale with response categories "I can do this very well by myself", "I can do this with help", "I know what this means but I cannot do this", and "I don't know what this means" was used. Items were recoded so that positive scores indicate high self-confidence.

3.6.2 School-Level Variables

3.6.2.1 School Size

The index of school size (schsize) indicates the total number of students at school that is reported by the school principal while filling the school questionnaire. This number is obtained by summing the numbers of females and males reported.

3.6.2.2 Class Size

The class size (clsize) indicates the average size of classes of grade corresponding to 15-year-olds at that school.

3.6.2.3 School Type

The index of school type (schltype) indicates the classification of the school as either public or private according to whether a private entity or a public agency has the ultimate power to make decisions concerning its affairs. The index has three categories: i) public schools; ii) “government-dependent” private schools managed by a non-governmental organization; and iii) “government-independent” private schools.

3.6.2.4 Proportion of Females Enrolled at School

The index of proportion of females enrolled at school (pcgirls) indicates the proportion of females to the total number of students at a school reported by the school administrator.

3.6.2.5 Academic Selectivity

The index of academic selectivity (selsch) was constructed from school administrators' responses about criteria of admittance to school, based on a scale of response categories "not considered", "considered", "high priority" or "pre-requisite". Those criteria are: residence in a particular area; academic record of the students; placement tests; recommendation of the previous schools; parents' will about the philosophy of the school on instruction; having special instruction programmes; and involvement of family members at school affairs.

3.6.2.6 Ability Grouping

The index of ability grouping (abgroup) was developed according to the responses of principals about within-school ability grouping of students into different classes or within-class ability grouping of students and whether this is the case for all students, for some students or for none. The categories were "no ability grouping", "grouping for some subjects", and "grouping for all".

3.6.2.7 Resource Autonomy

The index of resource autonomy (respres) was developed according to the responses of principals about number of decisions related to school resources on school's responsibility. Those responsibilities are hiring teachers or firing them, defining teachers' salaries and salary increases, school budgets and budget allocations within the school.

3.6.2.8 Curricular Autonomy

The index of curricular autonomy (respcurr) was developed according to the responses of principals about number of decisions related to curriculum on school's responsibility. Those responsibilities are defining student assessment policies, deciding textbooks to use, determining content of courses, and deciding courses to be offered.

3.6.2.9 School's Educational Resources

School's educational resources index (scmatedu) was derived from items about perceptions of school principals on factors hindering instruction at their schools. Those hindrances are shortage or inadequacy of science laboratory equipment, instructional materials, library materials, computers used for instruction, inadequacy of Internet connectivity, computer software used for instruction, and audio-visual resources. A four-point scale with categories of response "not at all", "very little", "to some extent", and "a lot" was used. All items were recoded so that higher values indicate less hindrance due to lack, shortage, and inadequacy of these resources.

3.6.2.10 Teacher Shortage

PISA's teacher shortage index (tcshort) was derived from perceptions of school principals on teacher factors hindering instruction at school. Actually, it does not necessarily indicate physical absence of teachers, however, PISA chooses to use this name for the index. Those factors are qualifications of science teachers, qualifications of mathematics teachers; qualifications of language teachers; and qualifications of teachers of other subjects. A four-point scale with the response categories "not at all", "very little", "to some extent", and "a lot" was used. The scale was not inverted thus positive values indicate school principals' perceptions of higher teacher shortage at a school.

3.6.2.11 School Activities to Promote the Learning of Science

The index of school activities to promote the learning of science (sciprom) was derived from responses of school administrators about involvement of the school in activities to promote engagement with science among students in grades corresponding to 15-year-olds. Those activities are science fairs, extracurricular science projects, science competitions; excursions related to science activities and field trips. Higher values on this index point to higher levels of these school activities.

3.6.2.12 School Activities for Learning of Environmental Topics

The index of school activities for learning of environmental topics (envlearn) was derived from school principals' responses indicating whether their school organizes activities for students in grades corresponding to 15-year-olds so that they learn about environmental issues. Those activities are among outdoor education, museums trips, visits to science and technology centers, extracurricular projects regarding environmental issues, lectures and seminars about environment. Positive values indicate higher levels of school activities.

3.6.2.13 Student/Teacher Ratio

Student/teacher ratio (stratio) was computed from principals' reports about the total number of students enrolled at school and total number of teachers employed.

3.6.2.14 Ratio of Computers to School Size

Ratio of computers to school size (ratcomp) was computed from the total number of computers available at school and total number of students enrolled at school.

3.6.2.15 Ratio of Computers for Instruction to School Size

Ratio of computers for instruction to school size (iratcomp) was computed from responses to items about the number of computers used for instructional activities at school and total number of students enrolled at school.

3.6.2.16 Proportion of Computers Connected to Web

Proportion of computers connected to web (compweb) was computed from the number of computers which has access to Internet and the total number of computers at school.

3.7 Data Analysis

As preliminary studies of data, frequency distributions analysis, descriptive analysis, missing value analysis, and treatment of data were done by SPSS (Statistics Package for Social Sciences) computer program version 11.0.1. By frequency distributions analysis, patterns of responses were examined. Analysis of descriptives described the distributions and characteristics of data. Treatment part was the analysis of missing data; deletion of variables either not used in the present study or that have missing values at a proportion exceeding the accepted limits; and recoding of some variables and indexes to a meaningful scale.

Variables with missing values above 10% were not used in this research and thus deleted individually within subsamples.

As the main analysis of the present study, data for the selected countries were analyzed and scientific literacy scores of students were modeled with two level multilevel models using HLM (Hierarchical Linear Models) computer program version 6.08 build 6.08.29257.1. Significance level was accepted as $p=0.05$ for testing the hypotheses throughout the present study.

3.7.1 Hierarchical Linear Modeling

In social research and other fields, structure of research data is mostly hierarchical and each grouping in hierarchical structure may have properties that influence the study. Individuals in the same group are likely to exhibit more similar characteristics than those in other groups. Variables may describe individuals but each higher-level unit may also have its own variables. Repeated measurement data are also hierarchical in the sense that multiple measurements are nested within the individual. Hierarchical linear and nonlinear models, also called multilevel models, are appropriate for the study of relationships at any level of the hierarchy with a single analysis, taking the variability associated with each level into account. Hierarchical linear models (HLM) not only estimates model coefficients at each level, but it also predicts the random effects associated with each sampling unit at every level, which is a means to cope with the inadequacy of traditional statistical techniques for modeling hierarchy.

With hierarchical linear models, each of the levels in the structure represented by the data is formally represented by its own submodel. These submodels convey relationships among variables within a given level, and specify how variables at a level influence relations at another level (Raudenbush & Bryk, 2002, p.7). Hierarchical linear models have use in achieving three general research purposes: improved estimation of effects within individual units;

modeling cross-level effects; and partitioning of variance and covariance components among levels.

For the present study, a two-level model and submodels were built for each country. Level-1 units; students were nested in the level-2 units; schools. The submodels run were one-way ANOVA model with random effects, regression model with means-as-outcomes, random-coefficients regression model, and intercepts- and slopes-as-outcomes model.

3.7.1.1 One-Way ANOVA Model with Random Effects

The simplest possible hierarchical linear model is equivalent to a one-way ANOVA with random effects. A basic regression equation for an outcome is such:

$$Y_i = \beta_0 + \beta_1 X_i + r_i$$

where

Y_i is the outcome variable or dependent variable,
β_0 is the expected outcome for observation *i,* or the intercept,
β_1 is the expected change in Y_i associated with a unit increase in X_i
X_i is an independent variable
r_i is the error term representing the unique effect associated with observation *i*

In the case of one-way ANOVA model, coefficients in level-1 model are all set to zero and only the intercept is left. Therefore, the model is also known as Intercept-Only model.

$$Y_{ij} = \beta_{0j} + r_{ij}$$

where level-1 error, r_{ij} is assumed normally distributed with a mean of zero and a constant level-1 variance, σ^2.

The level-2 model for one-way ANOVA with random effects is:

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

where γ_{00} is the grand-mean outcome in the population, and u_{0j} is the random effect associated with unit j. u_{0j} is assumed to have a mean of zero and variance τ_{00}. Substituting level-2 equation into level-1 equation yields the combined model:

$$Y_{ij} = \gamma_{00} + u_{0j} + r_{ij}$$

which is, in fact, the one-way ANOVA model with grand mean, γ_{00}; with a group effect, u_{0j} ; and with a person effect, r_{ij} . The model is a random-effects model because the group effects are taken as random. Here, the variance of the outcome is

$$\mathrm{Var}(Y_{ij}) = \mathrm{Var}(u_{0j} + r_{ij}) = \tau_{00} + \sigma^2$$

Estimating the one-way ANOVA model is often useful as a preliminary step in HLM because it estimates a point estimate and confidence interval for the grand mean, γ_{00}. More importantly, it provides information about the outcome variability at each of the two levels. The σ^2 parameter represents the within-group variability, and τ_{00} represents the between-group variability. Because no predictor variables are specified at level-1 and level-2 equations of the model, the hierarchical model is called fully unconditional model (Raudenbush & Bryk, 2002, p.24).

A useful parameter associated with one-way ANOVA with random effects model is the intraclass correlation coefficient and measures the proportion of the variance in the outcome that is between level-2 units:

$$\rho = \tau_{00}/(\tau_{00} + \sigma^2)$$

3.7.1.2 Means-as-Outcomes Regression Model

Means-as-outcomes regression model demonstrates whether means from each of many groups as an outcome to be predicted by group characteristics. Level-1 equation for this submodel is the same as unconditional model:

$$Y_{ij} = \beta_{0j} + r_{ij}.$$

and level-2 model is:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}W_j + u_{0j}$$

where only one level-2 predictor, W_j is assumed. The combined model is then:

$$Y_{ij} = \gamma_{00} + \gamma_{10}W_j + u_{0j} + r_{ij}$$

Here, u_{0j} had been the deviation of unit j's mean from the grand mean, it now represents the residual

$$u_{0j} = \beta_{0j} - \gamma_{00} - \gamma_{01}W_j.$$

Similarly, the variance in u_{0j}, τ_{00} is now the residual or conditional variance in β_{0j} after controlling for W_j (Raudenbush & Bryk, 2002, p.25).

3.7.1.3 Random-Coefficients Regression Model

One-way ANOVA and means-as-outcomes submodels were random-intercept models and only the level-1 intercept coefficient, β_{0j}, was viewed as random. Regression slopes did not exist in those models. However, a major class of applications of hierarchical linear models conceives level-1 slopes as varying randomly over the population of level-2 units. Random-coefficients regression model is the simplest case of this type. In these models, both the level-1 intercept and one or more level-1 slopes vary randomly, but no attempt is made to predict this variation (Raudenbush & Bryk, 2002, p.26).

Level-1 equation in a simple model is:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \overline{X}_{.j}) + r_{ij}$$

and the level-2 model becomes

$$\beta_{0j} = \gamma_{00} + u_{0j}$$

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

where

γ_{00} is the average intercept across the level-2 units;
γ_{10} is the average regression slope across the level-2 units;
u_{0j} is the unique increment to the intercept associated with level-2 unit *j*;
u_{1j} is the unique increment to the slope associated with level-2 unit *j*.

The dispersion of the level-2 random effects as a variance-covariance matrix is represented as:

$$\mathrm{Var}\begin{bmatrix} u_{0j} \\ u_{1j} \end{bmatrix} = \begin{bmatrix} \tau_{00} & \tau_{01} \\ \tau_{10} & \tau_{11} \end{bmatrix} = T$$

where

$\mathrm{Var}(u_{0j}) = \tau_{00}$ = unconditional variance in the level-1 intercepts;

$\mathrm{Var}(u_{1j}) = \tau_{11}$ = unconditional variance in the level-1 slopes; and

$\mathrm{Cov}(u_{0j}, u_{1j}) = \tau_{01}$ = unconditional covariance between the level-1 intercepts and slopes.

These components are referred to as unconditional variance-covariance components because no level-2 predictors are included in level-2 equations. The combined model is:

$$Y_{ij} = \gamma_{00} + \gamma_{10}(X_{ij} - \overline{X}_{.j}) + u_{0j} + u_{1j}(X_{ij} - \overline{X}_{.j}) + r_{ij}$$

This combined model implies that the outcome Y_{ij} is a function of the average regression equation, $\gamma_{00} + \gamma_{10}(X_{ij} - \overline{X}_{.j})$ plus a random error having three components: u_{0j}, the random effect of unit j on the mean; $u_{1j}(X_{ij} - \overline{X}_{.j})$, where u_{1j} is the random effect of unit j on the slope β_{1j}; and the level-1 error, r_{ij}.

3.7.1.4 Intercepts- and Slopes-as-Outcomes Model

The random coefficients regression model allows estimating the variability in the regression coefficients (both intercepts and slopes) across the level-2 units. The next step is to model this variability. Questions on the interactions of effects of level-1 and level-2 predictors may be addressed by employing the full model. This model may be expanded to incorporate the effects of multiple level-1 predictors and of multiple level-2 predictors (Raudenbush & Bryk, 2002, p. 27-28). In general, hierarchical linear models may involve multiple level-1 predictors where any combination of random, nonrandomly varying, and fixed slopes can be specified.

3.7.2 Centering the Variables

In all quantitative research, variables under study should have precise meaning so that statistical results can be related to the theoretical concerns of the research. In the case of hierarchical linear models, the intercept and slopes in the level-1 model become outcome variables at level-2. It is important that the meaning of these outcome variables be clearly understood. If a value for a variable does not make a sense, the researcher may prefer to transform the variable, or "choose a location" without changing the scaling concerns. A proper choice of location will help ensure numerical stability in estimating hierarchical linear models. The numerical stability of the estimation is not affected by the location for the level-2 predictors, but a suitable choice will ease interpretation of results.

In general, reasonable choices of location depend on the purposes of the research. No superior rule exists to wrap up all cases. The important consideration must be to carefully analyze these options and to bear in mind the location of predictors to interpret the results. The centering type for level-2 variables is not as critical as for the level-1 predictors. Numerical instability is less likely. Nevertheless, it is often convenient to center all of the level-2 predictors around their corresponding grand means (Raudenbush & Bryk, 2002, p. 31-35).

The main advantages of centering are:

- Obtaining estimates of intercepts and other effects that are easier to interpret
- Removing high correlations between the random intercept and slopes, and high correlations between level-1 and level-2 variables and cross-level interactions

For the present study, both level-1 and level-2 variables are centered around their respective grand means. Centering level-1 variables around their group means yields high reliability but grand mean centering also yields close values of reliability besides it tends to explain higher proportions of between group variances.

3.7.3 Random versus Fixed Effects

For modeling β_{ij}, there are three options. First is that the effect of level-1 predictor is constrained to be invariant across level-2 units:

$$\beta_{ij} = \gamma_{i0}$$

where $\gamma_{i0}$is the common effect of level-1 variable in every level-2 unit. That is, the effect of β_{ij} is fixed across level-2 units. Second option is that the slope β_{ij} is modeled as a function of an average value, γ_{ij}, plus a random effect associated with each level-2 unit:

$$\beta_{ij} = \gamma_{i0} + u_{ij}$$

Here, β_{ij} is random. If predictors for β_{ij} are specified, then the slopes-as-outcomes model is formulated. Such as:

$$\beta_{ij} = \gamma_{i0} + \gamma_{i1} W_j + u_{ij}$$

By this model, part of the variation of the slope β_{ij} can be predicted by W_j, but a random component, u_{ij}, remains unexplained. However, once the effect of W_j is taken into account, the residual variation is negligible. Then a model constraining that residual variation to be null would be sensible:

$$\beta_{ij} = \gamma_{i0} + \gamma_{i1} W_j$$

In this third case, the slope is nonrandomly varying because it varies strictly as a function of the predictor W_j.

The decision to take an effect as fixed or random may bring about problems. When the effect is taken as fixed while it is random in fact, the estimates will be biased.

In the present study, all effects are accepted as random in the first place. While building the two-level hierarchical linear model, fixed effects tables and variance components tables were examined from the output file in each step of the building process to decide the type of the effect. Error terms for all slopes were turned on, assuming all the slopes had a fixed and a random effect component. After building the model for fixed effects components, random effects components were examined first as a whole. If problem occurred, in other words, there were some insignificant results for random effect components, all the error terms representing the random effects were turned off and a step-by-step building strategy was employed to detect randomly varying slopes. Beginning from the intercept and moving in order and adding the error term for each slope equation one by one, variables with fixed and random effects were detected. The rationale for following this procedure is to gain the whole set of randomly varying variables because one can also gain significant results in models with incomplete sets or varying subsets of actual variables having random effects. That is, the results would be extremely biased.

A very similar procedure for building the model and deciding which variables to add was used. Again a step-by-step inclusion method was followed during the building process of both level-1 and level-2 equations.

3.7.4 Outcome Variable

The outcome variable modeled in the present study is the overall scientific literacy scores of students. However, not all students responded to all of the science items. Thus, students' scientific literacy scores are not actually observed scores, but are imputed ones. PISA and many other international surveys report student performance through plausible values. Usually, five plausible values are allocated to each student on each performance scale. Working with one plausible value instead of five will not bias the estimation of population parameters but the imputation error displaying the influence of test unreliability for the parameter estimation will not be estimated too. Statistical analyses should be performed independently on each of these five plausible values and results should be aggregated to obtain the final estimates of the statistics and their respective standard errors.

Plausible values are not actual test scores and should not be treated as such. They are multiple imputations of the unobservable latent achievement for each student, and are random numbers drawn from the marginal posterior distribution. As such, plausible values contain random error variance components and are not suitable to use as scores for individuals. They are better suited to describe the performance of the population as a set.

The PISA student file contains 40 plausible values, five for each of the eight PISA 2006 scales. PV1MATH to PV5MATH are for mathematical literacy; PV1SCIE to PV5SCIE for scientific literacy, PV1READ to PV5READ for reading literacy, PV1INTR to PV5INTR for interest in science and PV1SUPP to PV5SUPP for support for scientific inquiry. For the three scientific literacy scales, explaining phenomena scientifically, identifying scientific issues, using scientific evidence, the plausible values variables are PV1EPS to PV5EPS, PV1ISI to PV5ISI, and PV1USE to PV5USE, respectively (OECD, 2009, p.156).

Regarding this nature of the PISA database and definition of plausible values, the outcome variable in the current study is chosen as the overall scientific literacy scores of students, which is represented by the five plausible values in the PISA 2006 database; PV1SCIE to PV5SCIE. Because plausible values cannot be treated as actual scores, those five plausible values are included in the multivariate data matrices of the HLM 6.08 computer program and the program is instructed to make the analysis with plausible values. HLM computer program takes the plausible values into account in generating the estimates. For each HLM model, the program runs with each of the specified plausible values internally, and produces their average value and the correct standard errors. The user seems to be producing one estimate, but HLM estimates as many as the number of specified plausible values are produced and their average and measurement error calculated correctly, thus ensuring an accurate treatment of plausible value data.

3.7.5 Hypotheses Testing

Hypotheses may be formulated and tested about the fixed effects, random level-1 coefficients, and the variance-covariance parameter. The tests could be single-parameter or multiparameter tests.

For single-parameter test of fixed effects, typical null hypothesis is:

$$\mathrm{H_0}: \gamma_{\mathrm{qs}} = 0$$

which implies that the effect of a level-2 predictor, W_{sj}, on a particular level-2 parameter, β_{qj}, is null.

There are several types of multiparameter hypothesis tests one can do with HLM. One approach to multiparameter test for fixed effects is to test if a set of level-2 predictors adds significant explained variance to an outcome. Multiparameter tests regarding γ are useful for omnibus tests of the relationship between categorical level-2 predictor and a β_{qj} parameter; contrasts between categories of a level-2 predictor; examining whether a level-2 characteristics interacts with any of several level-1 predictors; and examining whether some subset of level-2 predictors is needed in a particular β_{qj} model.

For single-parameter test of random level-1 coefficients, typical null hypothesis is:

$$\mathrm{H_0}: \beta_{\mathrm{qs}} = 0$$

In nearly all applications of hierarchical analysis, it is the common need to decide whether level-1 coefficients should be specified as fixed, random, or nonrandomly varying. For single-parameter test for variance and covariance components to decide whether random variation exists, typical null hypothesis is:

$$H_0: \tau_{qq} = 0$$

3.7.6 Weighting of Data

The weighting process does not make the analysis procedures more complex however guarantees that population estimates will be unbiased. Analyses should therefore always be weighted, at any stage of the process. The bias of unweighted estimates could be large. Weights are associated to each student and to each school in PISA because, students and schools in a particular country did not have the same probability of selection due to stratified sampling; differential participation rates of schools or students required adjustments; and some explicit strata were over-sampled for national reporting purposes.

All statistical analyses or procedures concerning the PISA data should be weighted. Unweighted analyses will provide biased population parameter estimates. PISA database files contain weight variables. The weight included in the database is multiplied by a ratio of the number of observations to the sum of the weights. In other words, the weights should be multiplied by the total number of students and divided by the weighted total number of students and the same for schools. This linear transformation ensures that the sum of the weights is equal to the number of observations. Therefore, for the present study, in both the data preparation step and modeling, data are weighted.

During the preparation of data for analysis, weight option of SPSS program was on and the weighting variable for students was the final student weight (W_FSTUWT) and for schools was final school weight (W_FSCHWT). The variables mentioned were also included in the multivariate data matrices

(MDMs) and stated to HLM program as level-1 and level-2 weight variables for modeling process.

3.7.7 Assumptions of Two-Level Hierarchical Linear Model

Model misspecification occurs when some component included in the error term is associated with one or more of the predictors in the model. In HLM, specification assumptions apply at each level, and misspecification at one level may affect results at other levels. In a hierarchical model, normality assumptions of error forms are made for both level-1 and level-2. Although a violation of these assumptions does not bias the level-2 coefficient estimates, it can adversely affect their estimated standard errors and inferential statistics.

Key assumptions of the two-level HLM of the present study were:

1. Each r_{ij} is independent and normally distribute with a mean of 0 and variance σ^2 for every level-1 unit *i* within each level-2 unit
2. The level-1 predictors, X_{qij}, are independent of r_{ij}
3. The vectors of random errors at level-2 are multivariate normal, each with a mean of zero, some variance, τ_{qq}, and covariance among the random elements
4. The set of level-2 predictors are independent of every u_{qj}
5. The errors at level-1 and level-2 are independent
6. The predictors at each level are not correlated with the random effects at the other level

3.8 Missing Data

To prepare multivariate data matrices for hierarchical linear modeling, SPSS data files were created with data extracted from the PISA 2006 database.

Two files were created for each country; one student file as level-1 data and one school file as level-2 file. Originally, besides others, student files contained the 31 student background variables, five plausible values, and weight variables, whereas school files contained the 16 school variables and the weight variable. Cut-point for deletion of a variable is defined as 10% prior to handling of missing values. By missing value analysis, variables for deletion from data files were detected. None of the data files of Sweden and Turkey had missing values greater than or equal to ten percent, but both the student and school files of Canada were problematic. In student file of Canada, the variables *Instrumental Motivation in Science, Self Concept in Science, Interaction in Science Teaching and Learning, Hands-On Activities in Science Teaching and Learning, Student Investigations in Science Teaching and Learning, and Focus on Model or Applications in Science Teaching and Learning* had missing values over 10%. Therefore, these variables were deleted from the student file of Canada and were not used in the modeling of Canadian students' scientific literacy. On the other hand, variables *Student/Teacher Ratio, Ratio of Computers to School Size, Ratio of Computers for Instruction to School Size, and Proportion of Computers Connected to Web* had missing values over 10%, and these variables were deleted from the school file of Canada.

HLM program assumes no missing values in level-2 files, thus missing values in the school data files of the three countries were replaced by their respective grand means.

For level-1 data files, a more complicated procedure was followed. To reduce bias, missing values were replaced by weighted aggregated data instead of grand means. With the weighting option, school means for each variable were calculated and missing values were replaced with those respective weighted means. The flaw with this method is that there could be insufficient data for a group to aggregate data. Total nonresponse for a particular item across the whole school is an example. However, after handling missing values by this method, variables remained with missing values had a proportion of missing less than one in ten thousandth for Canada, less than one in four thousandth for Sweden, and less than 0.2% for Turkey. These values are much suitable for listwise deletion from level-1 data in HLM analysis.

CHAPTER 4

RESULTS

4.1 Preliminary Analyses

All variables included in the present study were pre-analyzed for frequency distributions, descriptive measures and missing values. For the countries of interest, the goals of these pre-analyses are to define general characteristics of responses, to describe distributions, and to handle missing responses appropriately.

4.1.1 Descriptive Statistics for Participants

For the present study, all participants from Canada, Sweden, and Turkey were extracted from the international dataset of PISA. Student data were taken from student database and used as level-1 data; whereas school administrators'/principals' data were taken from school database and used as level-2 data during HLM analyses.

There were a total of 32,031 students from 1,253 schools of the three countries in the sample of the present study. This sample comprised the 8.0 % of total 398,750 students who participated in PISA 2006 from 57 countries.

Table 4.1 Frequency Distribution of Students in the Present Study

Country	No of Students	Percent of Sample	Percent of Total
Canada	22,646	70.7	5.7
Sweden	4,443	13.9	1.1
Turkey	4,942	15.4	1.2
Total:	32,031	100.0	8.0

Total number of schools in the sample of this study was 1,253, which was 8.7 % of total 14,365 schools who participated in PISA 2006 from 57 countries.

Table 4.2 Frequency Distribution of Schools in the Present Study

Country	No of Schools	Percent of Sample	Percent of Total
Canada	896	71.5	6.2
Sweden	197	15.7	1.4
Turkey	160	12.8	1.1
Total:	1,253	100.0	8.7

4.1.2 Descriptive Statistics for Variables

For the present study, initial 31 student-level factors were included in all three level-1 files and 16 school-level factors were included in all three level-2 files. However, initial proportion of missing values was the criteria and some of those variables were excluded from the study country-wise. There was no problematic variable in student-level files of Sweden and Turkey. However, for Canada, five student factors had more than 10% missing values and therefore excluded from the corresponding file. Therefore, 25 variables were included in Canada analyses, and 31 variables were included in both Sweden and Turkey analyses as student level factors.

On the other hand, Canada also had missing value problem in school variables. Four of the sixteen variables exceeded the initial missing value proportion and thus excluded from analyses for Canadian data. As a result, 12 variables were included in Canadian analyses, and 16 variables were included in both Swedish and Turkish analyses as school level factors.

4.2 Hierarchical Linear Modeling Analyses

Four models were built for each of Canada, Sweden, and Turkey to investigate the effects of student-level and school-level factors on students' scientific literacy skills.

4.2.1 HLM Analyses for Canada

4.2.1.1 One-Way ANOVA Model for Canada

The first submodel, one-way Analysis of Variance with random effects model provides information whether there exist differences in Canadian students' scientific literacy scores among schools. That is to say, first question of this research is resolved.

The student-level equation of this submodel is:

$$(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) = \beta_0 + r$$

and school-level equation is:

$$\beta_0 = \gamma_{00} + u_0$$

where

Y : scientific literacy score computed with imputation of five plausible values

β_0 : the intercept or the expected literacy score

r : student-level error

γ_{00} : the grand mean scientific literacy score of schools

u_0 : the random effect associated with school

The final estimation of fixed effects computed for one-way ANOVA with random effects for Canada is given in the table below.

Table 4.3 Final Estimation of Fixed Effects in One-Way ANOVA Model of Canada

Fixed Effect	Coefficient	Standard Error	t-ratio	Approximate d.f.	p-value
For intercept, overall literacy mean, γ_{00}	518.014	3.498	148.094	894	0.000

The model estimates indicate that there were significant differences in students' scientific literacy skills among schools in Canada. The estimated grand mean of scientific literacy scores was 518.01 with a standard error of 3.50 with a 95% confidence interval:

$$\text{C.I.} = 518.01 \pm 1.96 \times 3.50 \ = (511.15, 524.87)$$

Final estimation of variance components given in Table 4.4 presents the school-level effect and student-level effect.

Table 4.4 Final Estimation of Variance Components in One-Way ANOVA Model of Canada

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
School mean, u_0	50.626	2563.026	894	8319.019	0.000
Student effect, r	85.326	7280.494			

Test statistic displayed in the table indicates a significant variation among schools in students' scientific literacy (χ^2 = 8,319.019, d.f. = 894, p < 0.050). The intraclass correlation coefficient was calculated as:

$$\rho = \tau_{00}/(\tau_{00} + \sigma^2) = \frac{2563.026}{2563.026 + 7280.494} = 0.260$$

This coefficient tells us that 26% of the variance in scientific literacy scores of Canadian students was between-schools and 74% was at the individual level. The results suggest that school factors might account for the differences in the scientific literacy skills of Canadian students.

HLM computes an overall or average reliability for the least squares estimates of each level-1 coefficient across the set of level-2 units. The overall estimate of reliability is the average of school reliabilities. From the output file of HLM, reliability of random level-1 coefficient β_0 was found to be 0.90, which shows that the sample mean was much likely to be a reliable indicator of true school mean.

4.2.1.2 Means-as-Outcomes Regression Model for Canada

The second submodel, means-as-outcomes regression model provides information about which school-level factors are associated with the differences in Canadian students' scientific literacy scores, and the second question of this research is resolved.

Level-1 equation of the submodel is the same as one-way ANOVA submodel:

$$(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) = \beta_0 + r$$

On the other hand, school-level factors were cast into the level-2 equation. The building strategy was to insert one school-level variable into equation at a time, run the model built, check significance of effects of the variable, keep if appropriate or discard from the equation otherwise, and proceed with the next variable. School data file contained 12 school factor variables and the aforementioned procedure was run with them. Table 4.5 displays the resultant factors that were associated with students' scientific literacy skills.

Table 4.5 Final Estimation of Fixed Effects in Means-as-Outcomes Regression Model of Canada

Fixed Effect	Coefficient	Standard Error	t-ratio	Approximate d.f.	p-value
For intercept, β_0					
Overall Literacy Mean, γ_{00}	518.368	3.209	161.526	892	0.000
School Type, γ_{01}	-28.170	5.392	-5.224	892	0.000
School Size, γ_{02}	0.032	0.006	5.169	892	0.000

The school-level equation becomes:

$$\beta_0 = \gamma_{00} + \gamma_{01} * \boldsymbol{SCHLTYPE} + \gamma_{02} * \boldsymbol{SCHSIZE} + u_0$$

where bold and italic style of the font for a variable name indicates grand mean centering and

β_0 : the school mean on scientific literacy

γ_{00}: the average of the school means on scientific literacy

γ_{01}: the differentiating effect of type of school on school mean of scientific literacy

γ_{02}: the differentiating effect of school size on school mean of scientific literacy

u_0 : the residual

The combined model is:

$$(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) = \gamma_{00} + \gamma_{01} * \boldsymbol{SCHLTYPE} + \gamma_{02} * \boldsymbol{SCHSIZE} + u_0 + r$$

The results indicate a significant association between overall scientific literacy scores of schools (γ_{00}= 518.368, s.e.= 3.209, p<0.050) and type of the school (γ_{01}= -28.170, s.e.= 5.392, p<0.050); size of the school (γ_{02}= 0.032, s.e.= 0.006, p<0.050) for Canada. These two school-level factors were nominated for final full model. There was a negative association between scientific literacy scores and school type, whereas the association was positive for school size. As can be seen, Canadian schools with lower values in School Type variable tend to have higher scientific literacy scores. That is, regarding the sign of this relation, public schools managed by a public education authority, tend to have higher scientific literacy skills of students than government-dependent private schools; and government-dependent private schools tend to have higher scientific literacy skills of students than government-independent private schools.

Final estimation of variance components for means-as-outcomes model of Canada is given in Table 4.6.

Table 4.6 Final Estimation of Variance Components in Means-as-Outcomes Regression Model of Canada

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	47.034	2212.146	892	7276.353	0.000
Student effect, r	85.340	7282.853			

The residual variance between schools in means-as-outcomes model (τ_{00}=2,212.146) was smaller than that in one-way ANOVA model (τ_{00}=2,563.026). The variance explained or proportion reduction in variance based on the level-2 predictors School Type (SCHLTYPE) and School Size (SCHSIZE):

$$\frac{\tau_{00}(One-way\ ANOVA) - \tau_{00}(Means-as-outcomes)}{\tau_{00}(One-way\ ANOVA)}$$

$$= \frac{2563.026 - 2212.146}{2563.026} = 0.137$$

So, school type and school size explains 14% of between-school variance in students' scientific literacy scores in Canada. However, the 12 school factors did not account for all the variation in the intercepts.

4.2.1.3 Random-Coefficients Regression Model for Canada

The third submodel, random-coefficients regression model provides information about which student-level factors are associated with the differences in Canadian students' scientific literacy scores, and the third question of this research is resolved.

To build the model, a one-by-one check of the student-level variables was performed first. Those detected variables which have significant effect on literacy scores were entered to the level-1 equation as a whole and the model estimates were checked. During this check, the factors that turned out to be not significant when employed together were discarded. Those insignificances indicate covariance between factors.

Finally level-1 equation of the model becomes:

$$
\begin{aligned}
(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) \\
&= \beta_0 + \beta_1 * \boldsymbol{HISEI} + \beta_2 * \boldsymbol{HISCED} + \beta_3 * \boldsymbol{IMMIG} + \beta_4 \\
&* \boldsymbol{CARINFO} + \beta_5 * \boldsymbol{CARPREP} + \beta_6 * \boldsymbol{CULTPOSS} + \beta_7 \\
&* \boldsymbol{ENVAWARE} + \beta_8 * \boldsymbol{ENVOPT} + \beta_9 * \boldsymbol{ENVPERC} + \beta_{10} \\
&* \boldsymbol{GENSCIE} + \beta_{11} * \boldsymbol{HEDRES} + \beta_{12} * \boldsymbol{HOMEPOS} + \beta_{13} \\
&* \boldsymbol{INTSCIE} + \beta_{14} * \boldsymbol{JOYSCIE} + \beta_{15} * \boldsymbol{PERSCIE} + \beta_{16} \\
&* \boldsymbol{RESPDEV} + \beta_{17} * \boldsymbol{SCIEEFF} + \beta_{18} * \boldsymbol{SCIEFUT} + \beta_{19} \\
&* \boldsymbol{WEALTH} + \beta_{20} * \boldsymbol{HIGHCONF} + \beta_{21} * \boldsymbol{INTCONF} + \beta_{22} \\
&* \boldsymbol{INTUSE} + \beta_{23} * \boldsymbol{PRGUSE} + r
\end{aligned}
$$

where

β_0 : mean on scientific literacy

β_1 : the differentiating effect of Highest Occupational Status of Parents

β_2 : the differentiating effect of Highest Educational Level of Parents

β_3 : the differentiating effect of Immigration Background

β_4 : the differentiating effect of Student Information on Science-Related Careers

β_5 : the differentiating effect of School Preparation for Science-Related Careers

β_6 : the differentiating effect of Cultural Possessions at Home

β_7 : the differentiating effect of Awareness of Environmental Issues

β_8 : the differentiating effect of Optimism Regarding Environmental Issues

β_9 : the differentiating effect of Level of Concern for Environmental Issues

β_{10} : the differentiating effect of General Value of Science

β_{11} : the differentiating effect of Home Educational Resources

β_{12} : the differentiating effect of Home Possessions

β_{13} : the differentiating effect of General Interest in Science

β_{14} : the differentiating effect of Enjoyment of Science

β_{15} : the differentiating effect of Personal Value of Science

β_{16} : the differentiating effect of Responsibility for Sustainable Development

β_{17} : the differentiating effect of Self-Efficacy in Science

β_{18} : the differentiating effect of Future-Oriented Motivation to Learn Science

β_{19} : the differentiating effect of Availability of Household Possessions Indicating Family Wealth

β_{20} : the differentiating effect of Self-Confidence in ICT High-Level Tasks

β_{21} : the differentiating effect of Self-Confidence in ICT Internet Tasks

β_{22} : the differentiating effect of ICT Internet/Entertainment Use

β_{23} : the differentiating effect of ICT Program/Software Use

and level-2 equations become:

$$\beta_0 = \gamma_{00} + u_0$$
$$\beta_1 = \gamma_{10} + u_1$$
$$\beta_2 = \gamma_{20} + u_2$$
$$\beta_3 = \gamma_{30} + u_3$$
$$\beta_4 = \gamma_{40}$$
$$\beta_5 = \gamma_{50}$$
$$\beta_6 = \gamma_{60} + u_6$$
$$\beta_7 = \gamma_{70}$$
$$\beta_8 = \gamma_{80}$$
$$\beta_9 = \gamma_{90}$$
$$\beta_{10} = \gamma_{100}$$
$$\beta_{11} = \gamma_{110} + u_{11}$$

$\beta_{12} = \gamma_{120}$

$\beta_{13} = \gamma_{130} + u_{13}$

$\beta_{14} = \gamma_{140}$

$\beta_{15} = \gamma_{150} + u_{15}$

$\beta_{16} = \gamma_{160}$

$\beta_{17} = \gamma_{170}$

$\beta_{18} = \gamma_{180}$

$\beta_{19} = \gamma_{190} + u_{19}$

$\beta_{20} = \gamma_{200} + u_{20}$

$\beta_{21} = \gamma_{210}$

$\beta_{22} = \gamma_{220} + u_{22}$

$\beta_{23} = \gamma_{230} + u_{23}$

There were some problems in slopes when all were assumed randomly varying. Therefore, to define fixed slopes and random slopes, all error terms or in other words the random components for level-2 equations were turned off and step-by-step inclusion off random components to fixed effects was done. Significance of error terms were checked from variance components tables from the output of HLM program. After determining the randomly varying slopes, variances of errors were examined. A high τ_{qq} correlation indicates that the same error variation across school-level units is being carried and a reduction may be brought by fixing one of the variables. From the tau correlations table, high correlations for slopes of *Home Possessions* and *Self-Confidence in ICT Internet Tasks* were observed. The PISA index of home possessions is a summary index of all household items and also included the items used to construct the variables *Cultural Possessions at Home, Home Educational Resources,* and *Availability of Household Possessions Indicating Family Wealth*. When scatter plots for these variables were observed, high individual correlations between *Home Possessions* and *Cultural Possessions at Home, Home Educational Resources,* and *Availability* *of Household Possessions Indicating Family Wealth* were seen, but none between latter ones were observed. Therefore, slope for *Home Possessions* was excluded from the set of randomly varying slopes. Similar relation between *Self-Confidence in ICT High-Level Tasks* and *Self-Confidence in ICT Internet Tasks* was also observed.

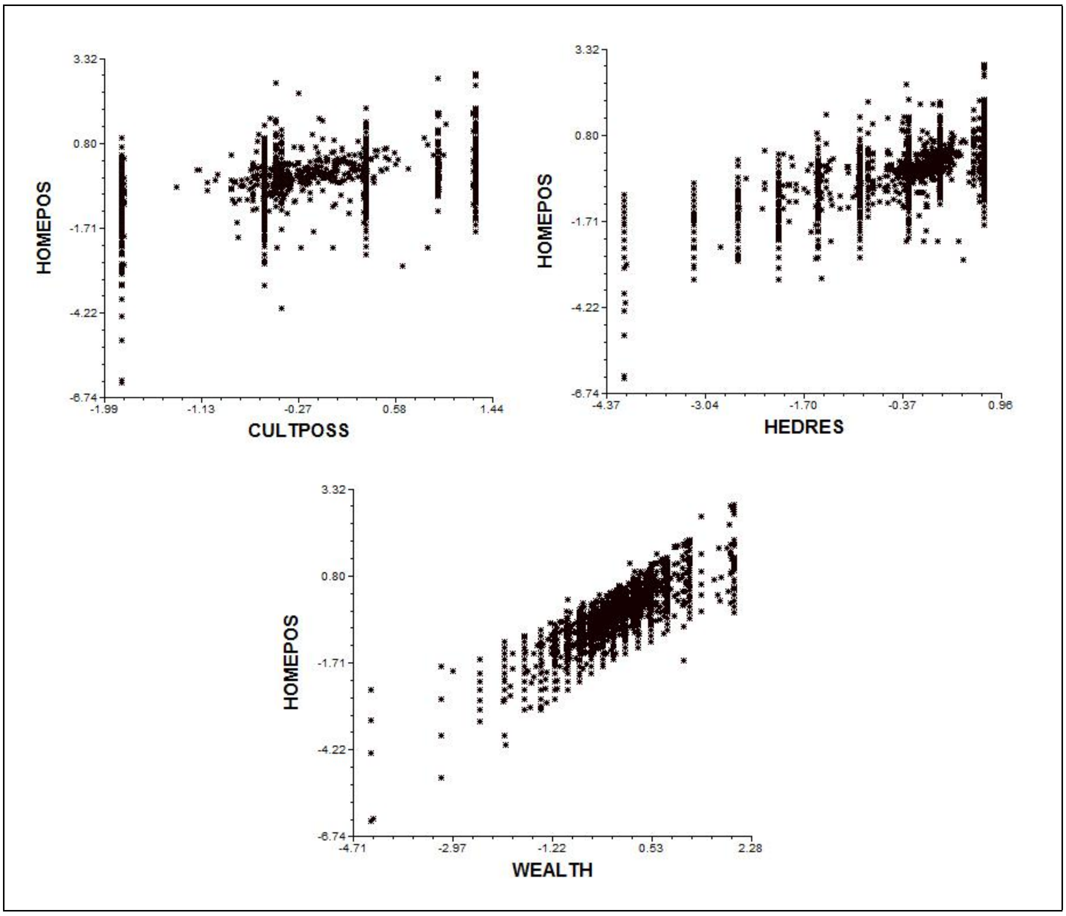


Figure 4.1 Scatter Plots of Home Possessions with Cultural Possessions at Home, Home Educational Resources, and Possessions Indicating Wealth across Schools of Canada

At the end of the process, slopes of *Student Information on Science-Related Careers, School Preparation for Science-Related Careers, Awareness of Environmental Issues, Optimism Regarding Environmental Issues, Level of Concern for Environmental Issues, General Value of Science, Home Possessions,* *Enjoyment of Science, Responsibility for Sustainable Development, Self-Efficacy in Science, Future-Oriented Motivation to Learn Science,* and *Self-Confidence in ICT Internet Tasks* were taken as fixed across schools and the slopes of remaining ones were random.

Final estimation of fixed effects in random-coefficient regression model of Canada is given in Table 4.7 and student-level factors that are associated with the differences in Canadian students' scientific literacy scores were defined accordingly.

Table 4.7 Final Estimation of Fixed Effects in Random-Coefficients Regression Model of Canada

Fixed Effect	Coefficient	s.e.	t-ratio	p-value
For intercept, β_0 Overall Literacy Mean, γ_{00}	522.173	2.687	194.345	0.000
For HISEI slope, β_1, intercept, γ_{10}	0.678	0.072	9.403	0.000
For HISCED slope, β_2, intercept, γ_{20}	2.287	0.865	2.644	0.009
For IMMIG slope, β_3, intercept, γ_{30}	-12.781	2.291	-5.580	0.000
For CARINFO slope, β_4, intercept, γ_{40}	-10.730	0.927	-11.578	0.000
For CARPREP slope, β_5, intercept, γ_{50}	5.401	1.206	4.477	0.000
For CULTPOSS slope, β_6, intercept, γ_{60}	-5.289	1.387	-3.812	0.000
For ENVAWARE slope, β_7, intercept, γ_{70}	12.043	1.292	9.325	0.000
For ENVOPT slope, β_8, intercept, γ_{80}	-3.280	0.861	-3.811	0.000
For ENVPERC slope, β_9, intercept, γ_{90}	-3.690	1.028	-3.590	0.001
For GENSCIE slope, β_{10}, intercept, γ_{100}	4.528	1.064	4.256	0.000
For HEDRES slope, β_{11}, intercept, γ_{110}	-4.430	1.532	-2.892	0.005
For HOMEPOS slope, β_{12}, intercept, γ_{120}	25.325	4.468	5.668	0.000
For INTSCIE slope, β_{13}, intercept, γ_{130}	-4.504	1.511	-2.980	0.003

Table 4.7 Continued

Fixed Effect	Coefficient	s.e.	t-ratio	p-value
For JOYSCIE slope, β_{14}, intercept, γ_{140}	13.572	1.620	8.376	0.000
For PERSCIE slope, β_{15}, intercept, γ_{150}	-4.111	1.446	-2.843	0.007
For RESPDEV slope, β_{16}, intercept, γ_{160}	7.558	1.030	7.339	0.000
For SCIEEFF slope, β_{17}, intercept, γ_{170}	22.731	1.098	20.695	0.000
For SCIEFUT slope, β_{18}, intercept, γ_{180}	10.606	1.220	8.696	0.000
For WEALTH slope, β_{19}, intercept, γ_{190}	-27.093	4.504	-6.015	0.000
For HIGHCONF slope, β_{20}, intercept, γ_{200}	-4.944	1.232	-4.013	0.000
For INTCONF slope, β_{21}, intercept, γ_{210}	17.581	1.866	9.421	0.000
For INTUSE slope, β_{22}, intercept, γ_{220}	-3.903	1.142	-3.417	0.001
For PRGUSE slope, β_{23}, intercept, γ_{230}	-13.304	1.339	-9.938	0.000

Coefficients of student factors had both positive and negative values depending on the structure of the indices formed by PISA; and together with the sign of the coefficient, indices with high values reveal important characteristics. Because these variables except *Highest Occupational Status of Parents, Highest Educational Level of Parents* and *Immigration Background* are standardized scores, the magnitudes of their coefficients can be regarded as comparable. Among the non-likelihood estimate ones, coefficient of immigration background variable was -12.781 and this negative relation shows that native students tend to have higher scientific literacy scores than first- and second-generation immigrants to Canada. However, *Highest Occupational Status of Parents* and *Highest Educational Level of Parents* were positively related to Canadian students' scientific literacy skills.

On the other hand, the high value -27.093 indicated that the more household possessions indicating family wealth, the lower the scientific literacy skills. However, home possessions, which is derived from some of the items constituting wealth indicator index and from the availability of books at home, had a strong positive association with literacy scores (γ_{120}= 25.325, p< 0.050). Other factors which had strong negative associations with scientific literacy scores are Student Information on Science-Related Careers (γ_{40}= -10.730, p< 0.050), and ICT Program/Software Use (γ_{230}= -13.304, p< 0.050). Besides, the associations of enjoyment of science (γ_{140}= 13.572, p< 0.050), awareness of environmental issues (γ_{70}= 12.043, p< 0.050), and self-efficacy in science (γ_{170}= 22.731, p<0.050) with scientific literacy scores of Canadian students were strong and positive.

Table 4.8 Final Estimation of Variance Components in Random-Coefficients Regression Model of Canada

Random Effect	Standard Deviation	Variance Component	χ^2	p-value
Schools mean, u_0	37.116	1377.614	1326.892	0.000
For HISEI slope, u_1	0.316	0.100	555.674	0.012
For HISCED slope, u_2	4.381	19.195	610.714	0.000
For IMMIG slope, u_3	14.249	203.054	657.333	0.000
For CULTPOSS slope, u_6	5.502	30.276	675.613	0.000
For HEDRES slope, u_{11}	6.481	42.005	592.620	0.001
For INTSCIE slope, u_{13}	10.219	104.424	611.105	0.000
For PERSCIE slope, u_{15}	5.784	33.457	588.902	0.001
For WEALTH slope, u_{19}	11.982	143.561	598.571	0.000
For HIGHCONF slope, u_{20}	5.239	27.445	600.542	0.000
For INTUSE slope, u_{22}	6.640	44.085	657.900	0.000
For PRGUSE slope, u_{23}	6.233	38.845	682.501	0.000
Student effect, *r*	66.505	4422.940		

Estimations of variance components are given in Table 4.8, where student factors with random components are presented with variable names. Those factors with significant error terms indicate that the slopes of the variables were much steeper for some schools.

The variance explained with the inclusion of those factors in the model can be examined by comparing the variances of one-way ANOVA submodel with the present submodel. This is done by calculating a proportion index for the reduction with student level variances of the submodels:

$$\rho = \frac{\sigma^2(One - Way\ ANOVA) - \sigma^2(Random\ Coef.Regr.)}{\sigma^2(One - Way\ ANOVA)}$$

$$= \frac{7280.494 - 4422.940}{7280.494} = 0.392$$

Including the student-level factors; *Highest Occupational Status of Parents, Highest Educational Level of Parents, Immigration Background, Student Information on Science-Related Careers, School Preparation for Science-Related Careers, Cultural Possessions at Home, Awareness of Environmental Issues, Optimism Regarding Environmental Issues, Level of Concern for Environmental Issues, General Value of Science, Home Educational Resources, Home Possessions, General Interest in Science, Enjoyment of Science, Personal Value of Science, Responsibility for Sustainable Development, Self-Efficacy in Science, Future-Oriented Motivation to Learn Science, Availability of Household Possessions Indicating Family Wealth, Self-Confidence in ICT High-Level Tasks, Self-Confidence in ICT Internet Tasks, Internet/Entertainment Use,* and *ICT Program/Software Use*, as predictors of scientific literacy scores, within school variance was reduced. Therefore, these factors accounted for the 39% of variance in Canadian students' scientific literacy scores.

4.2.1.4 Intercepts- and Slopes-as-Outcomes Model for Canada

The fourth submodel, intercepts- and slopes-as-outcomes model provides information about which school-level factors influence students' characteristics that affect scientific literacy scores in Canada, and the fourth question of this research is resolved. The cross-level interactions of school factors with student factors are investigated.

This submodel incorporates the three previous submodels; one-way ANOVA with random effects, means-as-outcomes regression, and random-coefficients regression. Additionally, this fourth submodel examines varying slopes that can be accounted with randomly varying school effects. That is, it models the variability in intercepts and slopes across schools.

Analysis for this model begins from the random-coefficients regression stage and includes school-level factors defined in means-as-outcomes model as associated with the differences in Canadian students' scientific literacy scores. Sequentially, for each of the level-2 equations of random-coefficients model having random effects, school variables School Type and School Size were entered and checked for significant effects on corresponding slopes. School variables with significant effect on slopes were kept and others were discarded. Then, model checking continued with the next slope.

Finally, level-1 equation of the full model becomes:

$$(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) = \beta_0 + \beta_1 * \boldsymbol{HISEI} + \beta_2 * \boldsymbol{HISCED} + \beta_3 * \boldsymbol{IMMIG} + \beta_4 * \boldsymbol{CARINFO} + \beta_5 * \boldsymbol{CARPREP} + \beta_6 * \boldsymbol{CULTPOSS} + \beta_7 * \boldsymbol{ENVAWARE} + \beta_8 * \boldsymbol{ENVOPT} + \beta_9 * \boldsymbol{ENVPERC} + \beta_{10} * \boldsymbol{GENSCIE} + \beta_{11} * \boldsymbol{HEDRES} + \beta_{12} * \boldsymbol{HOMEPOS} + \beta_{13} * \boldsymbol{INTSCIE} + \beta_{14} * \boldsymbol{JOYSCIE} + \beta_{15} * \boldsymbol{PERSCIE} + \beta_{16} * \boldsymbol{RESPDEV} + \beta_{17} * \boldsymbol{SCIEEFF} + \beta_{18} * \boldsymbol{SCIEFUT} + \beta_{19} * \boldsymbol{WEALTH} + \beta_{20} * \boldsymbol{HIGHCONF} + \beta_{21} * \boldsymbol{INTCONF} + \beta_{22} * \boldsymbol{INTUSE} + \beta_{23} * \boldsymbol{PRGUSE} + r$$

and for level-2, the final equations become:

$$\beta_0 = \gamma_{00} + \gamma_{01} * \boldsymbol{SCHLTYPE} + \gamma_{02} * \boldsymbol{SCHSIZE} + u_0$$
$$\beta_1 = \gamma_{10} + u_1$$
$$\beta_2 = \gamma_{20} + \gamma_{21} * \boldsymbol{SCHLTYPE} + \gamma_{22} * \boldsymbol{SCHSIZE} + u_2$$
$$\beta_3 = \gamma_{30} + u_3$$
$$\beta_4 = \gamma_{40}$$
$$\beta_5 = \gamma_{50}$$
$$\beta_6 = \gamma_{60} + u_6$$
$$\beta_7 = \gamma_{70}$$
$$\beta_8 = \gamma_{80}$$
$$\beta_9 = \gamma_{90}$$
$$\beta_{10} = \gamma_{100}$$
$$\beta_{11} = \gamma_{110} + u_{11}$$
$$\beta_{12} = \gamma_{120}$$
$$\beta_{13} = \gamma_{130} + u_{13}$$
$$\beta_{14} = \gamma_{140}$$
$$\beta_{15} = \gamma_{150} + u_{15}$$
$$\beta_{16} = \gamma_{160}$$
$$\beta_{17} = \gamma_{170}$$
$$\beta_{18} = \gamma_{180}$$
$$\beta_{19} = \gamma_{190} + u_{19}$$
$$\beta_{20} = \gamma_{200} + u_{20}$$
$$\beta_{21} = \gamma_{210}$$
$$\beta_{22} = \gamma_{220} + \gamma_{221} * \boldsymbol{SCHSIZE} + u_{22}$$
$$\beta_{23} = \gamma_{230} + u_{23}$$

The final estimations of fixed effects for intercepts- and slopes-as-outcomes model of Canada is presented in Table 4.9.

Table 4.9 Final Estimation of Fixed Effects in Intercepts- and Slopes-as-Outcomes Model of Canada

Fixed Effect	Coefficient	s.e.	p-value
For Intercept, β_0			
Intercept, γ_{00}	523.169	2.545	0.000
SCHLTYPE, γ_{01}	-19.309	4.043	0.000
SCHSIZE, γ_{02}	0.019	0.005	0.000
For HISEI slope, β_1, intercept, γ_{10}	0.661	0.070	0.000
For HISCED slope, β_2			
Intercept, γ_{20}	1.901	0.822	0.021
SCHLTYPE, γ_{21}	9.163	3.082	0.005
SCHSIZE, γ_{22}	-0.004	0.002	0.018
For IMMIG slope, β_3, intercept, γ_{30}	-14.017	2.182	0.000
For CARINFO slope, β_4, intercept, γ_{40}	-10.681	0.917	0.000
For CARPREP slope, β_5, intercept, γ_{50}	5.295	1.214	0.000
For CULTPOSS slope, β_6, intercept, γ_{60}	-5.248	1.362	0.000
For ENVAWARE slope, β_7, intercept, γ_{70}	11.962	1.268	0.000
For ENVOPT slope, β_8, intercept, γ_{80}	-3.249	0.860	0.000
For ENVPERC slope, β_9, intercept, γ_{90}	-3.741	1.014	0.001
For GENSCIE slope, β_{10}, intercept, γ_{100}	4.276	1.076	0.000
For HEDRES slope, β_{11}, intercept, γ_{110}	-4.604	1.541	0.004
For HOMEPOS slope, β_{12}, intercept, γ_{120}	25.151	4.455	0.000
For INTSCIE slope, β_{13}, intercept, γ_{130}	-4.392	1.504	0.004
For JOYSCIE slope, β_{14}, intercept, γ_{140}	13.276	1.620	0.000
For PERSCIE slope, β_{15}, intercept, γ_{150}	-3.675	1.461	0.015
For RESPDEV slope, β_{16}, intercept, γ_{160}	7.478	1.034	0.000
For SCIEEFF slope, β_{17}, intercept, γ_{170}	22.749	1.095	0.000

Table 4.9 Continued

Fixed Effect	Coefficient	s.e.	p-value
For SCIEFUT slope, β_{18}, intercept, γ_{180}	10.649	1.213	0.000
For WEALTH slope, β_{19}, intercept, γ_{190}	-27.171	4.502	0.000
For HIGHCONF slope, β_{20}, intercept, γ_{200}	-4.938	1.228	0.000
For INTCONF slope, β_{21}, intercept, γ_{210}	17.595	1.844	0.000
For INTUSE slope, β_{22}			
Intercept, γ_{220}	-3.645	1.109	0.002
SCHSIZE, γ_{221}	0.004	0.002	0.032
For PRGUSE slope, β_{23}, intercept, γ_{230}	-13.064	1.334	0.000

The results indicated that mean scientific literacy score across schools is significantly associated with students' scientific literacy scores. School type was strongly and negatively related to scientific literacy of students, again in favor of public schools for high literacy scores (γ_{01}= -19.309, p< 0.050). On the other hand, school size was positively related to literacy scores of students (γ_{02}= 0.019, p< 0.050).

Another set of cross-level interactions occurred between school type, school size and student variable Highest Educational Level of Parents. There was a strong positive relation between the type school that student is enrolled and the slope of highest education level of student's parents. That is to say, students in government-independent private schools tend to have parents with higher levels of education than those in government-dependent private schools; and students in government-dependent private schools tend to have parents with higher levels of education than those in public schools.

Lastly, Internet/entertainment use had a negative effect on students' scientific literacy scores (γ_{220}= -3.645, p< 0.050), however, the slope of internet/entertainment use was positively related to school size (γ_{221}= 0.004, p< 0.050). Therefore, more crowded schools tend to have less negative slopes for

Internet/entertainment use, and thus have higher scientific literacy scores for students.

Estimations of variance components for intercepts- and slopes-as-outcomes submodel, or the final full model are given in Table 4.10.

Table 4.10 Final Estimation of Variance Components in Intercepts- and Slopes-as-Outcomes Model of Canada

Random Effect	Standard Deviation	Variance Component	p-value
Schools mean, u_0	35.746	1277.763	0.000
For HISEI slope, u_1	0.300	0.090	0.012
For HISCED slope, u_2	3.982	15.857	0.001
For IMMIG slope, u_3	14.144	200.068	0.000
For CULTPOSS slope, u_6	5.400	29.164	0.000
For HEDRES slope, u_{11}	6.562	43.061	0.001
For INTSCIE slope, u_{13}	10.210	104.243	0.000
For PERSCIE slope, u_{15}	5.646	31.879	0.001
For WEALTH slope, u_{19}	11.884	141.231	0.000
For HIGHCONF slope, u_{20}	5.264	27.707	0.000
For INTUSE slope, u_{22}	6.458	41.701	0.000
For PRGUSE slope, u_{23}	6.150	37.827	0.000
Student effect, *r*	66.482	4419.798	

The proportion reduction in variance based on the level-2 predictors *School Type* and *School Size* for the intercept is:

$$\frac{\tau_{qq}(Random\ Coef.Regr.) - \tau_{qq}(Intercepts - and\ Slopes -)}{\tau_{qq}(Random\ Coef.Regr.)}$$

$$= \frac{1377.614 - 1277.763}{1377.614} = 0.072$$

So, *School Type* and *School Size* explained 7% of the variance in scientific literacy scores of Canadian students when controlled for other predictors in the model. Likewise, the proportion reduction in variance of *Highest Educational Level of Parents* slope, β_2 was:

$$\frac{19.195 - 15.857}{19.195} = 0.174$$

and the proportion reduction in variance of *Internet/Entertainment Use* slope, β_{22} was:

$$\frac{44.085 - 41.701}{44.085} = 0.054$$

Proportion reduction in variance for intercepts- and slopes-as-outcomes submodel (ρ = 0.072) was lower than that of means-as-outcomes submodel (ρ = 0.137) and that is a result of the differences in the samples of the two models. However, reductions of 17% in variance of *Highest Educational Level of Parents* slope and 5% in variance of *Internet/Entertainment Use* slope were accounted for by cross-level interactions with school variables *School Type* and *School Size*.

4.2.2 HLM Analyses for Sweden

4.2.2.1 One-Way ANOVA Model for Sweden

The first submodel for Sweden, one-way Analysis of Variance with random effects model provides information whether there exist differences in Swedish students' scientific literacy scores among schools and resolves the fifth question of this research.

The student-level equation of this submodel is:

$$(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) = \beta_0 + r$$

and school-level equation is:

$$\beta_0 = \gamma_{00} + u_0$$

The final estimation of fixed effects computed for one-way ANOVA with random effects for Sweden is given in Table 4.11.

Table 4.11 Final Estimation of Fixed Effects in One-Way ANOVA Model of Sweden

Fixed Effect	Coefficient	Standard Error	t-ratio	Approximate d.f.	p-value
For intercept, overall literacy mean, γ_{00}	505.377	6.140	81.989	194	0.000

The model estimates indicate that there were significant differences in students' scientific literacy skills among schools in Sweden. The estimated grand mean of scientific literacy scores was 505.38 with a standard error of 6.14 and thus a 95% confidence interval is:

$$\text{C.I.} = 505.38 \pm 1.96 \times 6.14 \ = (493.35, 517.41)$$

Final estimation of variance components given in Table 4.12 presents the school-level effect and student-level effect for Sweden.

Table 4.12 Final Estimation of Variance Components in One-Way ANOVA Model of Sweden

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
School mean, u_0	38.370	1472.287	194	966.756	0.000
Student effect, r	88.444	7822.430			

There was a significant variation among schools in students' scientific literacy scores for Sweden (χ^2 = 966.756, d.f. = 194, p < 0.050). The intraclass correlation coefficient was calculated as:

$$\rho = \tau_{00}/(\tau_{00} + \sigma^2) = \frac{1472.287}{1472.287 + 7822.430} = 0.158$$

This intraclass correlation coefficient indicates that 16% of the variance in scientific literacy scores of Swedish students was between-schools and 84% was at the individual level. The results suggest that school-level factors might account for the differences in the scientific literacy skills of Swedish students.

The overall estimate of reliability is the average of school reliabilities. For this submodel, reliability of random level-1 coefficient β_0 was found to be 0.81, and it indicates that the sample mean was much likely to be a reliable indicator of true school mean.

4.2.2.2 Means-as-Outcomes Regression Model for Sweden

The second submodel for Sweden, means-as-outcomes regression model provides information about which school-level factors are associated with the differences in Swedish students' scientific literacy scores, and the sixth question of this research is resolved.

Level-1 equation of the submodel is the same as one-way ANOVA submodel:

$$(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) = \beta_0 + r$$

There was no severe missing data problem and thus after handling missing data, the school data file of Sweden contained all original 16 school factor variables. The building procedure for school-level equation was run with them. Table 4.13 displays the factors that were associated with Swedish students' scientific literacy skills.

Table 4.13 Final Estimation of Fixed Effects in Means-as-Outcomes Regression Model of Sweden

Fixed Effect	Coefficient	Standard Error	t-ratio	Approximate d.f.	p-value
For intercept, β_0					
Overall Literacy Mean, γ_{00}	505.555	3.035	166.563	190	0.000
Proportion of Females Enrolled at School, γ_{01}	141.031	70.548	1.999	162	0.047
School Type, γ_{02}	-46.868	11.512	-4.071	118	0.000
School Size, γ_{03}	0.064	0.015	4.203	190	0.000
Student/Teacher Ratio, γ_{04}	4.443	1.837	2.418	190	0.017

The school-level equation becomes:

$$\beta_0 = \gamma_{00} + \gamma_{01} * \boldsymbol{PCGIRLS} + \gamma_{02} * \boldsymbol{SCHLTYPE} + \gamma_{03} * \boldsymbol{SCHSIZE} + \gamma_{04} * \boldsymbol{STRATIO} + u_0$$

where bold and italic style of the font for a variable name indicates grand mean centering and

β_0 : the school mean on scientific literacy

γ_{00}: the average of the school means on scientific literacy

γ_{01}: the differentiating effect of proportion of females enrolled at school on school mean of scientific literacy

γ_{02}: the differentiating effect of type of school on school mean of scientific literacy

γ_{03}: the differentiating effect of school size on school mean of scientific literacy

γ_{04}: the differentiating effect of student/teacher ratio on school mean of scientific literacy

u_0 : the residual

The combined model is:

$$(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) = \gamma_{00} + \gamma_{01} * \boldsymbol{PCGIRLS} + \gamma_{02} * \boldsymbol{SCHLTYPE} + \gamma_{03} * \boldsymbol{SCHSIZE} + \gamma_{04} * \boldsymbol{STRATIO} + u_0 + r$$

The results indicated that there exist significant associations between students' scientific literacy scores and overall scientific literacy scores of schools (γ_{00}= 505.555, s.e.= 3.035, p<0.050); proportion of females enrolled at school (γ_{01}= 141.031, s.e.= 70.549, p<0.050); type of the school (γ_{02}= -46.868, s.e.= 11.512, p<0.050); size of the school (γ_{03}= 0.064, s.e.= 0.015, p<0.050); student/teacher ratio (γ_{04}= 4.443, s.e.= 1.837, p<0.050) for Sweden. These four school-level factors were nominated for final full model. There was a negative association between scientific literacy scores and school type, whereas the associations were positive with proportion of females, student/teacher ratio, and school size. As a result, Swedish schools with lower values in School Type variable tended to have higher scientific literacy scores. That is, public schools managed by a public education authority, tend to have higher scientific literacy skills of students than government-dependent private schools; and government-dependent private schools tend to have higher scientific literacy skills of students than government-independent private schools. On the other hand, higher proportions of females at school, larger schools, and higher student ratios over teachers were all associated with better scientific literacy skills of students in Sweden.

Final estimation of variance components for means-as-outcomes model of Sweden is given in Table 4.14.

Table 4.14 Final Estimation of Variance Components in Means-as-Outcomes Regression Model of Sweden

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
School mean, u_0	27.451	753.537	190	720.394	0.000
Student effect, r	88.294	7795.902			

The residual variance between schools in means-as-outcomes model (τ_{00}=753.537) was smaller than that in one-way ANOVA model (τ_{00}= 1472.287). Proportion reduction in variance based on including level-2 predictors Proportion of Females Enrolled at School (PCGIRLS), School Type (SCHLTYPE), School Size (SCHSIZE), and Student/Teacher Ratio (STRATIO) was:

$$\frac{\tau_{00}(One-way\ ANOVA)-\tau_{00}(Means-as-outcomes)}{\tau_{00}(One-way\ ANOVA)}$$

$$=\frac{1472.287-753.537}{1472.287}=0.488$$

Thus, proportion of females at school, school type, school size, and student ratio over teachers explained 49% of between-school variance in students' scientific literacy scores in Sweden. However, the 16 school factors did not account for all the variation in the intercepts.

4.2.2.3 Random-Coefficients Regression Model for Sweden

The third submodel for Sweden, random-coefficients regression model provides information about which student-level factors are associated with the differences in Swedish students' scientific literacy scores, and the seventh question of this research is resolved.

To build this model, again one-by-one check and addition of student variables was performed. Those variables previously checked for individual significance in effect on students' scientific literacy scores were used and factors that turn out to be not significant when employed together were discarded.

As a result, final level-1 equation for random-coefficients regression submodel of Sweden becomes:

$$\begin{aligned}(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE)\\ &= \beta_0+\beta_1*\boldsymbol{HISCED}+\beta_2*\boldsymbol{IMMIG}+\beta_3*\boldsymbol{CULTPOSS}+\beta_4\\ &*\boldsymbol{ENVAWARE}+\beta_5*\boldsymbol{ENVOPT}+\beta_6*\boldsymbol{GENSCIE}+\beta_7\\ &*\boldsymbol{HOMEPOS}+\beta_8*\boldsymbol{JOYSCIE}+\beta_9*\boldsymbol{PERSCIE}+\beta_{10}\\ &*\boldsymbol{RESPDEV}+\beta_{11}*\boldsymbol{SCIEEFF}+\beta_{12}*\boldsymbol{SCINVEST}+\beta_{13}\\ &*\boldsymbol{SCSCIE}+\beta_{14}*\boldsymbol{HIGHCONF}+\beta_{15}*\boldsymbol{INTCONF}+\beta_{16}\\ &*\boldsymbol{INTUSE}+\beta_{17}*\boldsymbol{ESCS}+r\end{aligned}$$

where

β_0 : mean on scientific literacy

β_1 : the differentiating effect of Highest Educational Level of Parents

β_2 : the differentiating effect of Immigration Background

β_3 : the differentiating effect of Cultural Possessions at Home

β_4 : the differentiating effect of Awareness of Environmental Issues

β_5 : the differentiating effect of Optimism Regarding Environmental Issues

β_6 : the differentiating effect of General Value of Science

β_7 : the differentiating effect of Home Possessions

β_8 : the differentiating effect of Enjoyment of Science

β_9 : the differentiating effect of Personal Value of Science

β_{10} : the differentiating effect of Responsibility for Sustainable Development

β_{11} : the differentiating effect of Self-Efficacy in Science

β_{12} : the differentiating effect of Student Investigations in Science Teaching and Learning

β_{13} : the differentiating effect of Self-Concept in Science

β_{14} : the differentiating effect of Self-Confidence in ICT High-Level Tasks

β_{15} : the differentiating effect of Self-Confidence in ICT Internet Tasks

β_{16} : the differentiating effect of Internet/Entertainment Use

β_{17} : the differentiating effect of Economic, Social and Cultural Status

and level-2 equations become:

$$\beta_0 = \gamma_{00} + u_0$$
$$\beta_1 = \gamma_{10}$$
$$\beta_2 = \gamma_{20}$$
$$\beta_3 = \gamma_{30}$$
$$\beta_4 = \gamma_{40} + u_4$$

$$\beta_5 = \gamma_{50}$$
$$\beta_6 = \gamma_{60}$$
$$\beta_7 = \gamma_{70} + u_7$$
$$\beta_8 = \gamma_{80}$$
$$\beta_9 = \gamma_{90}$$
$$\beta_{10} = \gamma_{100}$$
$$\beta_{11} = \gamma_{110}$$
$$\beta_{12} = \gamma_{120}$$
$$\beta_{13} = \gamma_{130}$$
$$\beta_{14} = \gamma_{140}$$
$$\beta_{15} = \gamma_{150}$$
$$\beta_{16} = \gamma_{160}$$
$$\beta_{17} = \gamma_{170}$$

To define fixed slopes and random slopes, all random components for level-2 equations were turned off and step-by-step inclusion off random components to fixed effects was done. After determining the randomly varying slopes, variances of errors were examined. A high τ_{qq} correlation indicates that the same error variation across school-level units is being carried and a reduction may be brought by fixing one of the variables. However, correlation between the two randomly varying slopes of Awareness of Environmental Issues and Home Possessions was moderate to low (r = -0.37). Therefore, none of the slopes was fixed across schools.

Final estimation of fixed effects in random-coefficient regression model of Sweden is given in Table 4.15 and student-level factors that are associated with the differences in Swedish students' scientific literacy scores were defined accordingly.

Table 4.15 Final Estimation of Fixed Effects in Random-Coefficients Regression Model of Sweden

Fixed Effect	Coefficient	s.e.	p-value
For intercept, β_0 Overall Literacy Mean, γ_{00}	505.279	3.904	0.000
For HISCED slope, β_1, intercept, γ_{10}	-7.201	1.751	0.000
For IMMIG slope, β_2, intercept, γ_{20}	-30.682	3.496	0.000
For CULTPOSS slope, β_3, intercept, γ_{30}	5.614	1.835	0.003
For ENVAWARE slope, β_4, intercept, γ_{40}	19.028	1.589	0.000
For ENVOPT slope, β_5, intercept, γ_{50}	-5.932	1.502	0.001
For GENSCIE slope, β_6, intercept, γ_{60}	12.458	2.310	0.000
For HOMEPOS slope, β_7, intercept, γ_{70}	-12.275	3.255	0.000
For JOYSCIE slope, β_8, intercept, γ_{80}	6.659	2.642	0.013
For PERSCIE slope, β_9, intercept, γ_{90}	-11.968	2.572	0.000
For RESPDEV slope, β_{10}, intercept, γ_{100}	4.782	1.516	0.002
For SCIEEFF slope, β_{11}, intercept, γ_{110}	9.246	2.102	0.000
For SCINVEST slope, β_{12}, intercept, γ_{120}	-19.067	1.528	0.000
For SCSCIE slope, β_{13}, intercept, γ_{130}	22.617	1.875	0.000
For HIGHCONF slope, β_{14}, intercept, γ_{140}	-11.321	1.721	0.000
For INTCONF slope, β_{15}, intercept, γ_{150}	11.213	2.690	0.000
For INTUSE slope, β_{16}, intercept, γ_{160}	-7.620	1.697	0.000
For ESCS slope, β_{17}, intercept, γ_{170}	30.439	3.747	0.000

Coefficient of immigration background variable was -30.682 and this negative relation shows that native students tended to have higher scientific literacy scores than first- and second-generation immigrants to Sweden. Also, *Highest Educational Level of Parents* was negatively related to Swedish students' scientific literacy skills, indicating that those students with higher levels of parental education tended to develop lower levels of scientific literacy skills in Sweden.

Among the WLE scale variables, Home Possessions (γ_{70}= -12.275, p<0.050), Personal Value of Science (γ_{90}= -11.968, p< 0.050), Student Investigations in Science Teaching and Learning (γ_{120}= -19.067, p< 0.050), and Self-Confidence in ICT High-Level Tasks (γ_{140}= -11.321, p< 0.050) were strongly and negatively related to Swedish students' scientific literacy scores.

On the other hand, student-level factors which have strong positive associations with scientific literacy scores were Awareness of Environmental Issues (γ_{40}= 19.028, p< 0.050), General Value of Science (γ_{60}= 12.458, p< 0.050), Self-Efficacy in Science (γ_{110}= 9.246, p< 0.050), Self-Concept in Science (γ_{130}= 22.617, p< 0.050), and Economic, Social and Cultural Status (γ_{170}= 30.439, p< 0.050).

Estimations of variance components are given in Table 4.16, where student factors with random components are presented with variable names. Student factors with significant error terms indicate that those slopes of the variables were much steeper for some schools.

Table 4.16 Final Estimation of Variance Components in Random-Coefficients Regression Model of Sweden

Random Effect	Standard Deviation	Variance Component	p-value
Schools mean, u_0	25.404	645.353	0.000
For ENVAWARE slope, u_4	4.817	23.206	0.009
For HOMEPOS slope, u_7	7.539	56.837	0.004
Student effect, *r*	66.505	4422.952	

The variance explained with the inclusion of those student-level factors in the model was:

$$\rho = \frac{\sigma^2(One - Way\ ANOVA) - \sigma^2(Random\ Coef.Regr.)}{\sigma^2(One - Way\ ANOVA)}$$

$$= \frac{7822.430 - 4422.952}{7822.430} = 0.435$$

Including the student-level factors; *Highest Educational Level of Parents, Immigration Background, Cultural Possessions at Home, Awareness of Environmental Issues, Optimism Regarding Environmental Issues, General Value of Science, Home Possessions, Enjoyment of Science, Personal Value of Science, Responsibility for Sustainable Development, Self-Efficacy in Science, Student Investigations in Science Teaching and learning, Self-Concept in Science, Self-Confidence in ICT High-Level Tasks, Self-Confidence in ICT Internet Tasks, Internet/Entertainment Use,* and *Economic, Social and Cultural Status*, as predictors of scientific literacy scores, within school variance was reduced. Therefore, these factors accounted for the 44% of variance in Swedish students' scientific literacy scores in PISA 2006.

4.2.2.4 Intercepts- and Slopes-as-Outcomes Model for Sweden

The fourth submodel, intercepts- and slopes-as-outcomes model provides information about which school-level factors influence students' characteristics that affect scientific literacy scores in Sweden, and the eighth question of this research is resolved. The cross-level interactions of school factors with student factors that are affecting scientific literacy skills of Swedish students are investigated.

This submodel incorporates the three previous submodels and examines varying slopes that can be accounted with randomly varying school effects. To put it another way, it models the variability in intercepts and slopes of scientific literacy scores of students across schools of Sweden.

Analysis for this model began from the random-coefficients regression stage and includes school-level factors defined in means-as-outcomes model as associated with the differences in Swedish students' scientific literacy scores. Sequentially, for each of the level-2 equations of random-coefficients model having random effects, school variables Proportion of Girls Enrolled at School, School Type, School Size, and Student Ratio over Teachers were entered and checked for significant effects on corresponding slopes. School variables with significant effect on slopes were kept and others were discarded. Then, model checking continued with the next slope.

Finally, level-1 equation of the full model, which is the same as random-coefficients regression model, becomes:

$$
\begin{aligned}
&(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) \\
&\qquad = \beta_0 + \beta_1 * \boldsymbol{HISCED} + \beta_2 * \boldsymbol{IMMIG} + \beta_3 * \boldsymbol{CULTPOSS} + \beta_4 \\
&\qquad * \boldsymbol{ENVAWARE} + \beta_5 * \boldsymbol{ENVOPT} + \beta_6 * \boldsymbol{GENSCIE} + \beta_7 \\
&\qquad * \boldsymbol{HOMEPOS} + \beta_8 * \boldsymbol{JOYSCIE} + \beta_9 * \boldsymbol{PERSCIE} + \beta_{10} \\
&\qquad * \boldsymbol{RESPDEV} + \beta_{11} * \boldsymbol{SCIEEFF} + \beta_{12} * \boldsymbol{SCINVEST} + \beta_{13} \\
&\qquad * \boldsymbol{SCSCIE} + \beta_{14} * \boldsymbol{HIGHCONF} + \beta_{15} * \boldsymbol{INTCONF} + \beta_{16} \\
&\qquad * \boldsymbol{INTUSE} + \beta_{17} * \boldsymbol{ESCS} + r
\end{aligned}
$$

and for level-2, the final equations become

$$\beta_0 = \gamma_{00} + \gamma_{01} * \boldsymbol{SCHLTYPE} + \gamma_{02} * \boldsymbol{SCHSIZE} + u_0$$
$$\beta_1 = \gamma_{10}$$
$$\beta_2 = \gamma_{20}$$
$$\beta_3 = \gamma_{30}$$
$$\beta_4 = \gamma_{40} + u_4$$
$$\beta_5 = \gamma_{50}$$

$\beta_6 = \gamma_{60}$

$\beta_7 = \gamma_{70} + \gamma_{71} * \boldsymbol{PCGIRLS} + u_7$

$\beta_8 = \gamma_{80}$

$\beta_9 = \gamma_{90}$

$\beta_{10} = \gamma_{100}$

$\beta_{11} = \gamma_{110}$

$\beta_{12} = \gamma_{120}$

$\beta_{13} = \gamma_{130}$

$\beta_{14} = \gamma_{140}$

$\beta_{15} = \gamma_{150}$

$\beta_{16} = \gamma_{160}$

$\beta_{17} = \gamma_{170}$

The final estimations of fixed effects for full model of Sweden are presented in Table 4.17.

Table 4.17 Final Estimation of Fixed Effects in Intercepts- and Slopes-as-Outcomes Model of Sweden

Fixed Effect	Coefficient	s.e.	p-value
For Intercept, β_0			
Intercept, γ_{00}	506.306	2.710	0.000
SCHLTYPE, γ_{01}	-29.001	12.240	0.020
SCHSIZE, γ_{02}	0.036	0.015	0.023
For HISCED slope, β_1, Intercept, γ_{10}	-7.116	1.722	0.000
For IMMIG slope, β_2, Intercept, γ_{20}	-31.538	3.479	0.000
For CULTPOSS slope, β_3, Intercept, γ_{30}	5.667	1.763	0.002
For ENVAWARE slope, β_4, Intercept, γ_{40}	19.140	1.632	0.000
For ENVOPT slope, β_5, Intercept, γ_{50}	-6.039	1.448	0.000
For GENSCIE slope, β_6, Intercept, γ_{60}	11.894	2.243	0.000

Table 4.17 Continued

Fixed Effect	Coefficient	s.e.	p-value
For HOMEPOS slope, β_7,			
Intercept, γ_{70}	-11.900	3.094	0.000
PCGIRLS, γ_{71}	-83.419	36.003	0.022
For JOYSCIE slope, β_8, Intercept, γ_{80}	6.741	2.644	0.012
For PERSCIE slope, β_9, Intercept, γ_{90}	-11.409	2.590	0.000
For RESPDEV slope, β_{10}, Intercept, γ_{100}	5.138	1.501	0.001
For SCIEEFF slope, β_{11}, Intercept, γ_{110}	9.189	2.127	0.000
For SCINVEST slope, β_{12}, Intercept, γ_{120}	-19.044	1.468	0.000
For SCSCIE slope, β_{13}, Intercept, γ_{130}	22.308	1.751	0.000
For HIGHCONF slope, β_{14}, Intercept, γ_{140}	-11.066	1.712	0.000
For INTCONF slope, β_{15}, Intercept, γ_{150}	10.818	2.715	0.000
For INTUSE slope, β_{16}, Intercept, γ_{160}	-7.260	1.717	0.000
For ESCS slope, β_{17}, intercept, γ_{170}	29.526	3.720	0.000

Estimations of variance components for intercepts- and slopes-as-outcomes model, or the final full model for Sweden, are given in Table 4.18.

Table 4.18 Final Estimation of Variance Components in Intercepts- and Slopes-as-Outcomes Model of Sweden

Random Effect	Standard Deviation	Variance Component	p-value
Schools mean, u_0	22.163	491.214	0.000
For ENVAWARE slope, u_4	5.063	25.638	0.009
For HOMEPOS slope, u_7	6.202	38.464	0.007
Student effect, *r*	66.451	4415.690	

The proportion reduction in variance based on the level-2 predictors *School Type* and *School Size* for the schools-mean scientific literacy score intercept was:

$$\frac{\tau_{qq}(Random\ Coef.Regr.) - \tau_{qq}(Intercepts - and\ Slopes -)}{\tau_{qq}(Random\ Coef.Regr.)}$$

$$= \frac{645.353 - 491.214}{645.353} = 0.239$$

So, *School Type* and *School Size* explained 24% of the variance in scientific literacy scores of Swedish students when controlled for other predictors in the model. Likewise, the proportion reduction in variance of *Home Possessions* slope, β_7 with interaction of schools' mean proportion of girls when controlled for other predictors in the model was:

$$\frac{56.837 - 38.464}{56.837} = 0.323$$

Proportion reduction in variance for intercepts- and slopes-as-outcomes submodel (ρ = 0.239) was lower than that of means-as-outcomes submodel (ρ = 0.488) and that is a result of the differences in the samples of the two models. However, a reduction of 32% in variance of *Home Possessions* slope was accounted for by cross-level interactions with school variable *Proportion of Girls Enrolled at School*.

4.2.3 HLM Analyses for Turkey

4.2.3.1 One-Way ANOVA Model for Turkey

The first submodel for Turkey, one-way Analysis of Variance with random effects model provides information whether there exist differences in Turkish students' scientific literacy scores among schools and resolves the ninth question of this research.

The level-1 equation of this submodel is:

$$(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) = \beta_0 + r$$

and level-2 equation is:

$$\beta_0 = \gamma_{00} + u_0$$

The final estimation of fixed effects computed for one-way ANOVA with random effects for Turkey is given in Table 4.19.

Table 4.19 Final Estimation of Fixed Effects in One-Way ANOVA Model of Turkey

Fixed Effect	Coefficient	Standard Error	t-ratio	Approximate d.f.	p-value
For intercept, overall literacy mean, γ_{00}	400.227	8.615	46.458	158	0.000

The model estimates indicate that there were significant differences in students' scientific literacy skills among schools in Turkey. The estimated grand mean of scientific literacy scores was 400.23 with a standard error of 8.62; and a 95% confidence interval was:

$$\text{C.I.} = 400.23 \pm 1.96 \times 8.62 = (383.33, 417.13)$$

Final estimation of variance components given in Table 4.20 presents the school-level effect and student-level effect for Turkey.

Table 4.20 Final Estimation of Variance Components in One-Way ANOVA Model of Turkey

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	58.288	3397.479	158	7270.795	0.000
Student effect, r	55.725	3105.256			

There was a significant variation among schools in students' scientific literacy scores for Turkey (χ^2 = 7270.795, d.f. = 158, p < 0.050). The intraclass correlation coefficient was calculated as:

$$\rho = \tau_{00}/(\tau_{00} + \sigma^2) = \frac{3397.479}{3397.479 + 3105.256} = 0.522$$

This intraclass correlation coefficient indicates that 52% of the variance in scientific literacy scores of Turkish students was between-schools and 48% was at the individual level. The results suggest that school-level factors might account for the differences in the scientific literacy skills of Turkish students to a great extent.

The overall reliability of random level-1 coefficient β_0 was found to be 0.98, and it indicates that the sample mean was much likely to be a reliable indicator of true school mean.

4.2.3.2 Means-as-Outcomes Regression Model for Turkey

The second submodel for Turkey, means-as-outcomes regression model provides information about which school-level factors are associated with the differences in Turkish students' scientific literacy scores, and the tenth question of this research is resolved.

Student-level equation of the submodel is:

$$(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) = \beta_0 + r$$

There was no severe missing data problem and thus after handling missing data, the school data file of Turkey contained all original 16 school factor variables of the present study. The building procedure for school-level equation was run with them. Table 4.21 displays the factors that were associated with Turkish students' scientific literacy skills.

Table 4.21 Final Estimation of Fixed Effects in Means-as-Outcomes Regression Model of Turkey

Fixed Effect	Coefficient	Standard Error	t-ratio	Approximate d.f.	p-value
For intercept, β_0					
Overall Literacy Mean, γ_{00}	415.958	5.245	79.312	154	0.000
Academic Selectivity of School, γ_{01}	32.952	5.067	6.503	154	0.000
School Activities to Promote Learning of Science, γ_{02}	14.695	3.842	3.824	154	0.000
School's Educational Resources, γ_{03}	17.330	6.123	2.830	154	0.006
Teacher Shortage, γ_{04}	16.200	5.203	3.114	154	0.003

The school-level equation becomes:

$$\beta_0 = \gamma_{00} + \gamma_{01} * \boldsymbol{SELSCH} + \gamma_{02} * \boldsymbol{SCIPROM} + \gamma_{03} * \boldsymbol{SCMATEDU} + \gamma_{04} * \boldsymbol{TCSHORT} + u_0$$

where bold and italic style of the font for a variable name indicates grand mean centering and

β_0 : the school mean on scientific literacy
γ_{00}: the average of the school means on scientific literacy
γ_{01}: the differentiating effect of academic selectivity of school
γ_{02}: the differentiating effect of school activities to promote learning of science
γ_{03}: the differentiating effect of school's educational resources
γ_{04}: the differentiating effect of teacher shortage
u_0 : the residual

The combined model is:

$$(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) = \gamma_{00} + \gamma_{01} * \boldsymbol{SELSCH} + \gamma_{02} * \boldsymbol{SCIPROM} + \gamma_{03} * \boldsymbol{SCMATEDU} + \gamma_{04} * \boldsymbol{TCSHORT} + u_0 + r$$

The results indicated that there exist significant associations between students' scientific literacy scores and overall scientific literacy scores of schools (γ_{00}= 415.958, s.e.= 5.245, p<0.050); academic selectivity of school (γ_{01}= 32.952, s.e.= 5.067, p<0.050); the extent of school activities to promote learning of science (γ_{02}= 14.695, s.e.= 3.842, p<0.050); adequacy of school's educational resources (γ_{03}= 17.330, s.e.= 6.123, p<0.050); the degree of teacher shortage (γ_{04}= 16.200, s.e.= 5.203, p<0.050) for Turkey. These four school-level factors were nominated for final full model. All associations of these variables with the intercept seemed to be positive. However, the index of teacher shortage was not rescaled to represent less shortage for higher values on that scale. As a result, Turkish schools that are short for teachers tended to have higher scientific literacy scores regarding the positive coefficient for teacher shortage index. On the other hand, students at schools with more activities to promote learning of science such as science clubs, science fairs, science competitions, and science projects tended to develop higher levels of scientific literacy skills in Turkey. Moreover, students from schools applying selective admittance policies got higher scientific literacy scores directly proportional to the degree of the selectivity. Finally, scientific literacy scores of students were higher for those schools having more educational resources lack of which might hinder instruction, such as computers, Internet connectivity, software, library materials, and audio-visual devices.

Final estimation of variance components for means-as-outcomes model of Turkey is given in Table 4.22.

Table 4.22 Final Estimation of Variance Components in Means-as-Outcomes Regression Model of Turkey

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	40.977	1679.121	154	3696.767	0.000
Student effect, r	55.729	3105.696			

The residual variance between schools in means-as-outcomes model (τ_{00}=1679.121) was smaller than that in one-way ANOVA model (τ_{00}= 3397.479). Proportion reduction in variance based on including school-level predictors Academic Selectivity of School (SELSCH), School Activities to Promote

Learning of Science (SCIPROM), School's Educational Resources (SCMATEDU), and Teacher shortage (TCSHORT) in the model was:

$$\frac{\tau_{00}(One - way\ ANOVA) - \tau_{00}(Means - as - outcomes)}{\tau_{00}(One - way\ ANOVA)}$$

$$= \frac{3397.479 - 1679.121}{3397.479} = 0.506$$

Therefore, the degrees of academic selectivity of school for admittance, school activities to promote learning of science, school's educational resources, and teacher shortage explained 51% of between-school variance in students' scientific literacy scores in Turkey. However, the 16 school factors did not account for all the variation in the intercepts.

4.2.3.3 Random-Coefficients Regression Model for Turkey

The third submodel for Turkey, random-coefficients regression model provides information about which student-level factors are associated with the differences in Turkish students' scientific literacy scores, and the eleventh question of this research is resolved.

To build this model, again one-by-one check and addition of student variables was performed. Those variables previously checked for individual significance in effect on students' scientific literacy scores were used and factors that turn out to be not significant when employed together were discarded.

As a result, final level-1 equation for random-coefficients regression submodel of Turkey becomes:

$$
\begin{aligned}
(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) \\
&= \beta_0 + \beta_1 * \boldsymbol{HISEI} + \beta_2 * \boldsymbol{ENVAWARE} + \beta_3 * \boldsymbol{ENVOPT} + \beta_4 \\
&* \boldsymbol{GENSCIE} + \beta_5 * \boldsymbol{HEDRES} + \beta_6 * \boldsymbol{HOMEPOS} + \beta_7 \\
&* \boldsymbol{JOYSCIE} + \beta_8 * \boldsymbol{RESPDEV} + \beta_9 * \boldsymbol{SCHANDS} + \beta_{10} \\
&* \boldsymbol{SCIEEFF} + \beta_{11} * \boldsymbol{HIGHCONF} + \beta_{12} * \boldsymbol{INTCONF} + \beta_{13} \\
&* \boldsymbol{PRGUSE} + r
\end{aligned}
$$

where

β_0 : mean on scientific literacy

β_1 : the differentiating effect of Highest Occupational Status of Parents

β_2 : the differentiating effect of Awareness of Environmental Issues

β_3 : the differentiating effect of Optimism Regarding Environmental Issues

β_4 : the differentiating effect of General Value of Science

β_5 : the differentiating effect of Home Educational Resources

β_6 : the differentiating effect of Home Possessions

β_7 : the differentiating effect of Enjoyment of Science

β_8 : the differentiating effect of Responsibility for Sustainable Development

β_9 : the differentiating effect of Hands-On Activities in Science Teaching and Learning

β_{10} : the differentiating effect of Self-Efficacy in Science

β_{11} : the differentiating effect of Self-Confidence in ICT High-Level Tasks

β_{12} : the differentiating effect of Self-Confidence in ICT Internet Tasks

β_{13} : the differentiating effect of ICT Program/Software Use

and level-2 equations become:

$$\beta_0 = \gamma_{00} + u_0$$
$$\beta_1 = \gamma_{10} + u_1$$
$$\beta_2 = \gamma_{20} + u_2$$
$$\beta_3 = \gamma_{30} + u_3$$

$\beta_4 = \gamma_{40} + u_4$

$\beta_5 = \gamma_{50}$

$\beta_6 = \gamma_{60}$

$\beta_7 = \gamma_{70} + u_7$

$\beta_8 = \gamma_{80}$

$\beta_9 = \gamma_{90}$

$\beta_{10} = \gamma_{100} + u_{10}$

$\beta_{11} = \gamma_{110}$

$\beta_{12} = \gamma_{120}$

$\beta_{13} = \gamma_{130}$

To define fixed slopes and random slopes, all random components for level-2 equations were turned off and step-by-step inclusion off random components to fixed effects was done. After determining the randomly varying slopes, variances of errors were examined. High τ_{qq} correlations were observed between slopes of Home Possessions and Awareness of Environmental Issues, between slopes of Home Possession and Enjoyment of Science. Moreover, high τ_{qq} correlations were observed between slopes of Hands-On Activities in Science Teaching and Learning and Highest Occupational Status of Parents, between slopes of Hands-On Activities in Science Teaching and Awareness of Environmental Issues, and between slopes of Hands-On Activities in Science Teaching and Enjoyment of Science. Therefore, the effects of variables Home Possessions and Hands-On Activities in Science Teaching and Learning were fixed across schools of Turkey, giving the final state of the random coefficients model.

Final estimation of fixed effects in random-coefficient regression model of Turkey is given in Table 4.23 and student-level factors that are associated with the differences in Turkish students' scientific literacy scores were defined accordingly.

Table 4.23 Final Estimation of Fixed Effects in Random-Coefficients Regression Model of Turkey

Fixed Effect	Coefficient	s.e.	p-value
For intercept, β_0 Overall Literacy Mean, γ_{00}	418.761	7.039	0.000
For HISEI slope, β_1, intercept, γ_{10}	0.269	0.113	0.019
For ENVAWARE slope, β_2, intercept, γ_{20}	7.191	2.039	0.001
For ENVOPT slope, β_3, intercept, γ_{30}	-7.217	1.609	0.000
For GENSCIE slope, β_4, intercept, γ_{40}	7.286	1.982	0.001
For HEDRES slope, β_5, intercept, γ_{50}	7.614	1.649	0.000
For HOMEPOS slope, β_6, intercept, γ_{60}	-4.398	2.104	0.039
For JOYSCIE slope, β_7, intercept, γ_{70}	6.525	2.141	0.003
For RESPDEV slope, β_8, intercept, γ_{80}	4.070	1.388	0.005
For SCHANDS slope, β_9, intercept, γ_{90}	-7.032	1.655	0.000
For SCIEEFF slope, β_{10}, intercept, γ_{100}	5.853	1.818	0.002
For HIGHCONF slope, β_{11}, intercept, γ_{110}	-6.743	2.183	0.003
For INTCONF slope, β_{12}, intercept, γ_{120}	14.110	1.807	0.000
For PRGUSE slope, β_{13}, intercept, γ_{130}	-7.529	1.338	0.000

The positive coefficient of Highest Occupational Status of Parents slope (γ_{10}= 0.269, p< 0.050) indicates that students whose parents stand higher on International Socio-Economic Index of Occupational Status (ISEI) tended to have higher scientific literacy scores. The index captures the attributes of occupations that convert parents' education into income. That is to say, students from families of higher income related to higher education of parents developed better scientific literacy skills in Turkey.

Among the WLE scale variables, Awareness of Environmental Issues, General Value of Science, Home Educational Resources , Enjoyment of Science, Responsibility for Sustainable Development, Self-Efficacy in Science, and Self-

Confidence in ICT Internet Tasks were positively related to Turkish students' scientific literacy scores. On the other hand, Optimism Regarding Environmental Issues, Home Possessions, Hands-On Activities in Science Teaching and Learning, Self-Confidence in ICT High-Level Tasks, and ICT Program/Software Use were negatively related to students' scientific literacy scores.

Estimations of variance components for random-coefficients model of Turkey are given in Table 4.24, where student factors with random components are presented with variable names. Student factors with significant error terms indicate that those slopes of the variables were much steeper for some schools.

Table 4.24 Final Estimation of Variance Components in Random-Coefficients Regression Model of Turkey

Random Effect	Standard Deviation	Variance Component	p-value
Schools mean, u_0	42.608	1815.458	0.000
For HISEI slope, u_1	0.198	0.039	0.000
For ENVAWARE slope, u_2	5.215	27.196	0.000
For ENVOPT slope, u_3	3.813	14.539	0.001
For GENSCIE slope, u_4	4.723	22.310	0.004
For JOYSCIE slope, u_7	5.039	25.389	0.000
For SCIEEFF slope, u_{10}	5.500	30.253	0.000
Student effect, *r*	46.149	2129.773	

The variance explained with the inclusion of those student-level factors in the model was:

$$\rho = \frac{\sigma^2(One - Way\ ANOVA) - \sigma^2(Random\ Coef.Regr.)}{\sigma^2(One - Way\ ANOVA)}$$

$$= \frac{3105.256 - 2129.773}{3105.256} = 0.314$$

Therefore, including the student-level factors; *Highest Occupational Status of Parents, Awareness of Environmental Issues, Optimism Regarding Environmental Issues, General Value of Science, Home Educational Resources, Home Possessions, Enjoyment of Science, Responsibility for Sustainable Development, Hands-On Activities in Science Teaching, Self-Efficacy in Science, Self-Confidence in ICT High-Level Tasks, Self-Confidence in ICT Internet Tasks,* and *ICT Program/Software Use* as predictors of scientific literacy scores in the model, within school variance was reduced. Therefore, these factors accounted for the 31% of variance in Turkish students' scientific literacy scores in PISA 2006.

4.2.3.4 Intercepts- and Slopes-as-Outcomes Model for Turkey

The fourth submodel for Turkey, intercepts- and slopes-as-outcomes model provides information about which school-level factors influence students' characteristics that affect scientific literacy scores in Turkey, and the twelfth question of this research is resolved. The cross-level interactions of school factors with student factors that are affecting scientific literacy skills of Turkish students are investigated.

This model incorporates the three previous submodels of Turkey and examines varying slopes that can be accounted with randomly varying school effects. It models the variability in intercepts and slopes of scientific literacy scores of students across schools of Turkey.

Analysis for this model began from the random-coefficients regression stage and included school-level factors defined in means-as-outcomes model as associated with the differences in Turkish students' scientific literacy scores. Sequentially, for each of the level-2 equations of random-coefficients model having random effects, school variables Academic Selectivity of School, School

Activities to Promote Learning of Science, School's Educational Resources, and Teacher Shortage were entered and checked for significant effects on corresponding slopes. School variables with significant effect on slopes were kept and others were discarded. Then, model checking continued with the next slope.

Finally, level-1 equation of the full model of Turkey, which is the same as random-coefficients regression model, becomes:

$$\begin{aligned}(Y|PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE) \\ = \beta_0 + \beta_1 * \boldsymbol{HISEI} + \beta_2 * \boldsymbol{ENVAWARE} + \beta_3 * \boldsymbol{ENVOPT} + \beta_4 \\ * \boldsymbol{GENSCIE} + \beta_5 * \boldsymbol{HEDRES} + \beta_6 * \boldsymbol{HOMEPOS} + \beta_7 \\ * \boldsymbol{JOYSCIE} + \beta_8 * \boldsymbol{RESPDEV} + \beta_9 * \boldsymbol{SCHANDS} + \beta_{10} \\ * \boldsymbol{SCIEEFF} + \beta_{11} * \boldsymbol{HIGHCONF} + \beta_{12} * \boldsymbol{INTCONF} + \beta_{13} \\ * \boldsymbol{PRGUSE} + r\end{aligned}$$

and for level-2, the final equations become

$$\beta_0 = \gamma_{00} + \gamma_{01} * \boldsymbol{SELSCH} + \gamma_{02} * \boldsymbol{SCIPROM} + \gamma_{03} * \boldsymbol{TCSHORT} + u_0$$
$$\beta_1 = \gamma_{10} + u_1$$
$$\beta_2 = \gamma_{20} + u_2$$
$$\beta_3 = \gamma_{30} + u_3$$
$$\beta_4 = \gamma_{40} + u_4$$
$$\beta_5 = \gamma_{50}$$
$$\beta_6 = \gamma_{60}$$
$$\beta_7 = \gamma_{70} + u_7$$
$$\beta_8 = \gamma_{80}$$
$$\beta_9 = \gamma_{90}$$
$$\beta_{10} = \gamma_{100} + u_{10}$$
$$\beta_{11} = \gamma_{110}$$
$$\beta_{12} = \gamma_{120}$$
$$\beta_{13} = \gamma_{130}$$

The final estimations of fixed effects for full model of Turkey are presented in Table 4.25.

Table 4.25 Final Estimation of Fixed Effects in Intercepts- and Slopes-as-Outcomes Model of Turkey

Fixed Effect	Coefficient	s.e.	p-value
For Intercept, β_0			
Intercept, γ_{00}	425.540	5.088	0.000
SELSCH, γ_{01}	18.512	3.008	0.000
SCIPROM, γ_{02}	14.388	3.512	0.000
TCSHORT, γ_{03}	9.324	3.623	0.011
For HISEI slope, β_1, Intercept, γ_{10}	0.270	0.110	0.015
For ENVAWARE slope, β_2, Intercept, γ_{20}	7.191	2.086	0.002
For ENVOPT slope, β_3, Intercept, γ_{30}	-7.305	1.579	0.000
For GENSCIE slope, β_4, Intercept, γ_{40}	7.224	1.965	0.001
For HEDRES slope, β_5, Intercept, γ_{50}	7.526	1.662	0.000
For HOMEPOS slope, β_6, Intercept, γ_{60}	-4.244	2.070	0.043
For JOYSCIE slope, β_7, Intercept, γ_{70}	6.452	2.119	0.003
For RESPDEV slope, β_8, Intercept, γ_{80}	4.167	1.393	0.005
For SCHANDS slope, β_9, Intercept, γ_{90}	-7.232	1.612	0.000
For SCIEEFF slope, β_{10}, Intercept, γ_{100}	5.869	1.854	0.002
For HIGHCONF slope, β_{11}, Intercept, γ_{110}	-6.616	2.196	0.004
For INTCONF slope, β_{12}, Intercept, γ_{120}	14.165	1.766	0.000
For PRGUSE slope, β_{13}, Intercept, γ_{130}	-7.552	1.300	0.000

Estimations of variance components for intercepts- and slopes-as-outcomes model, or the final full model for Turkey, are given in Table 4.26 with variable names.

Table 4.26 Final Estimation of Variance Components in Intercepts- and Slopes-as-Outcomes Model of Turkey

Random Effect	Standard Deviation	Variance Component	p-value
Schools mean, u_0	31.845	1014.115	0.000
For HISEI slope, u_1	0.183	0.034	0.000
For ENVAWARE slope, u_2	5.250	27.564	0.000
For ENVOPT slope, u_3	3.789	14.355	0.001
For GENSCIE slope, u_4	4.614	21.288	0.004
For JOYSCIE slope, u_7	5.173	26.762	0.000
For SCIEEFF slope, u_{10}	5.526	30.541	0.000
Student effect, *r*	46.172	2131.851	

The proportion reduction in variance based on the level-2 predictors: *Academic Selectivity of School, School Activities to Promote Learning of Science* and *Teacher Shortage* for the schools mean scientific literacy score intercept was:

$$\frac{\tau_{qq}(Random\ Coef.Regr.) - \tau_{qq}(Intercepts - and\ Slopes -)}{\tau_{qq}(Random\ Coef.Regr.)}$$

$$= \frac{1815.458 - 1014.115}{1815.458} = 0.442$$

Therefore, *Academic Selectivity of School, School Activities to Promote Learning of Science* and *Teacher Shortage* explained 44% of the variance in scientific literacy scores of Turkish students when controlled for other predictors in the model.

Proportion reduction in variance for intercepts- and slopes-as-outcomes submodel ($\rho = 0.442$) was lower than that of means-as-outcomes submodel ($\rho = 0.506$) and that is a result of the differences in the samples of the two models. Nevertheless, a reduction of 44% in variance of intercept of students' scientific literacy scores was accounted for by cross-level interactions with school variables *Academic Selectivity of School, School Activities to Promote Learning of Science* and *Teacher Shortage* in Turkey. But, due to the differences between the samples of intercepts- and slopes-as-outcomes model and means-as-outcomes model, the school factor *School's Educational Resources* did not interact with the intercept.

CHAPTER 5

DISCUSSION, CONCLUSION, AND IMPLICATIONS

5.1 Summary of the Results

The purpose of this study was to investigate the student- and school-level factors influencing students' scientific literacy skills in PISA 2006; and cross-level interactions of those factors. Data of three countries were used for the present study: Canada, Sweden, and Turkey.

Results for Canada indicate that 26% of the variance in scientific literacy scores of Canadian students was between-schools and 74% was at the individual level. From this existence of between-schools variation, one can conclude that school-level factors might account for variations in individual scientific literacy scores of students in Canada. The school-level factors type of school; whether it is public, government-controlled private, or government-independent, and the size of the school were found to affect individual scientific literacy skills of Canadian students. On the other hand, mean school scientific literacy scores were also significantly associated with individual scientific literacy scores of students. As a result of considering school type and school size as school-level predictors, explanation of some 14% of between-school variance in students' scientific literacy scores in Canada was feasible.

As for student-level factors, *Highest Occupational Status of Parents, Highest Educational Level of Parents, Immigration Background, Student Information on Science-Related Careers, School Preparation for Science-Related Careers, Cultural Possessions at Home, Awareness of Environmental Issues, Optimism Regarding Environmental Issues, Level of Concern for Environmental Issues, General Value of Science, Home Educational Resources, Home*

Possessions, General Interest in Science, Enjoyment of Science, Personal Value of Science, Responsibility for Sustainable Development, Self-Efficacy in Science, Future-Oriented Motivation to Learn Science, Availability of Household Possessions Indicating Family Wealth, Self-Confidence in ICT High-Level Tasks, Self-Confidence in ICT Internet Tasks, Internet/Entertainment Use, and *ICT Program/Software Use* accounted for the 39% of variation in scientific literacy scores of Canadian students.

Cross-level interactions occurred between school type, school size and student variable Highest Educational Level of Parents in Canada. Also, students' Internet and entertainment use of computers was affected by the size of the school. Of these interactions, the one between school type, school size and student's education level of parents explained the 17% of variation in slope of student-level variable *Highest Educational Level of Parents*. Likewise, proportion reduction in variance of *Internet/Entertainment Use* slope due to its interaction with school size was 5%.

Results for Sweden suggest that between-school variations exist among schools of Sweden. The variance in scientific literacy scores of Swedish students in PISA 2006 was 16% between-schools and 84% depended on individual differences. On the other hand, school characteristics detected to impact scientific literacy skills of students in Sweden were the proportion of females enrolled at school, type of the school, size of the school, and student/teacher ratio. There was a negative association of school type with scientific literacy scores. Therefore, we can conclude that public schools in Sweden tend to have higher scientific literacy skills of students than government-dependent private schools; and government-dependent private schools tend to have higher scientific literacy skills of students than government-independent private schools. Besides, higher proportions of females at school, increasing size of school, and higher student/teacher ratios were all associated with better scientific literacy skills of students in Sweden. These school characteristics explained 49% of between-school variations in students' scientific literacy scores for Sweden.

Student factors to impact scientific literacy scores in Sweden were found to be *Highest Educational Level of Parents, Immigration Background, Cultural Possessions at Home, Awareness of Environmental Issues, Optimism Regarding Environmental Issues, General Value of Science, Home Possessions, Enjoyment of Science, Personal Value of Science, Responsibility for Sustainable Development, Self-Efficacy in Science, Student Investigations in Science Teaching and learning, Self-Concept in Science, Self-Confidence in ICT High-Level Tasks, Self-Confidence in ICT Internet Tasks, Internet/Entertainment Use,* and *Economic, Social and Cultural Status*. These factors accounted for the 44% of variance in Swedish students' scientific literacy scores in PISA 2006.

In Sweden, cross-level interactions occurred between the intercept for students' scientific literacy scores and school's type and size. Moreover, proportion of girls enrolled at school seemed to be negatively associated with home possessions of students. That is, *School Type* and *School Size* explained 24% of the variance in scientific literacy scores of Swedish students when controlled for other predictors in the model and proportion of girls explained 32% of random variation in the slope of *Home Possessions* when controlled for other predictors in the model.

Results for Turkey indicated that between-school variations exist among schools of Turkey for scientific literacy skills. Estimates suggested that 52% of the variance in scientific literacy scores of Turkish students was between-schools and 48% was among students. Significant relations of students' scientific literacy scores with overall scientific literacy scores of schools, academic selectivity of school, the extent of school activities to promote learning of science, adequacy of school's educational resources, and the degree of teacher shortage existed in Turkey. All associations of these school factors with students' scientific literacy scores were positive but the index of teacher shortage had not been rescaled to represent less shortage of teachers for higher values on that scale. Nonetheless, these factors explained 51% of between-school variance in students' scientific literacy scores in Turkey.

Student factors affecting scientific literacy scores of Turkish students in PISA 2006 were found to be *Highest Occupational Status of Parents, Awareness of Environmental Issues, General Value of Science, Home Educational Resources , Enjoyment of Science, Responsibility for Sustainable Development, Self-Efficacy in Science, and Self-Confidence in ICT Internet Tasks* were positively related to Turkish students' scientific literacy scores. On the other hand, *Optimism Regarding Environmental Issues, Home Possessions, Hands-On Activities in Science Teaching and Learning, Self-Confidence in ICT High-Level Tasks*, and *ICT Program/Software Use* were negatively related to Turkish students' scientific literacy scores. These factors accounted for the 31% of variance in Turkish students' scientific literacy scores in PISA 2006.

In Turkey, significant cross-level interactions of the students' mean on scientific literacy with school characteristics: academic selectivity, activities to promote learning of science and teacher shortage were observed, but not with school's educational resources. Resultantly, school factors *Academic Selectivity of School, School Activities to Promote Learning of Science* and *Teacher Shortage* explain 44% of the variance in scientific literacy scores of Turkish students when controlled for other characteristics of students.

5.2 Discussion of the Results

5.2.1 Student-Level Factors

Student characteristics that exhibit significant impact on students' scientific literacy scores in PISA 2006 are discussed in this section. The factors are investigated first in a country based classification.

5.2.1.1 Student-Level Factors for Canada

Canada was the country with highest number of student-level factors that have significant effects on students' scientific literacy scores in the present study. Same building procedure was followed for each of the countries and 23 out of 31 student-level variables of interest were found to be effective in the model. Of those variables, 11 had randomly varying characteristics across schools of Canada, which is also the highest number among the three countries. This may be due to the fact that Canadian sample is approximately 4 times greater than that of Sweden and that of Turkey. However, for alpha 0.05, anticipated effect size 0.8 and desired statistical power 0.8, minimum required sample size is calculated as 52 for a two-tailed test and 42 for one-tailed test in t-ratio. The sample sizes of the three countries are far beyond this minimum. Therefore, t-test might have produced valid estimations of parameters.

Slopes of *Highest Educational Level of Parents* and *Highest Occupational Status of Parents* were significantly related to scientific literacy for Canadian students. The association of higher parental education with occupational status of parents and with higher educational level of children has been mentioned numerous times (McNeal, 1999; Dearing et al., 2001; Dubow et al., 2009). Moreover, the index of highest occupational status of parents of PISA incorporates parents' education convertible to occupation or, in other words, income. This is an expected situation for Canada because in such developed countries, higher education levels are associated with more qualified workforce and the effect of high levels of parental education and family economic status on children becomes the incentive for careers similar to their parents.

Education, occupational status, income, and wealth are strong determinants of socio-economic status. Though the PISA index of economic, social and cultural status does not seem to be associated with students' scientific literacy skills, its subdimensions are all related to scientific literacy of Canadian students. That is, index of economic, social and cultural status captures occupational status of parents, highest educational levels of parents and the index of home possessions, and index of home possessions captures the three subdimensions: *Cultural Possessions at Home, Home Educational Resources, Availability of Household Possessions Indicating Family Wealth*. However, all these subdimensions of *Home Possessions* are negatively associated with scientific literacy scores where *Home Possessions* is positively related. This might be the result of the fact that *Cultural Possessions at Home, Home Educational Resources, Availability of Household Possessions Indicating Family Wealth* constitute but they are not the full extent to *Home Possessions*, because it also includes some other country-specific items besides the ones taken into consideration while constructing these dimensions. Moreover, PISA index of economic, cultural and social status cannot be claimed to fully represent socioeconomic status of an individual.

Canada is a country with high immigrant population. Therefore, immigration status of a student is expected to significantly affect education and skills for life. Indeed, student factor *Immigration Background* was found to have significant effect on students' scientific literacy skills in Canada. That association is strong and negative; indicating that immigrant students perform significantly lower in scientific literacy scores than native students. This finding is consistent with the findings of previous research about immigration background (UNICEF-Innocenti Research Centre, 2009; Schleicher, 2006).

Both student factors about science-related careers; *School Preparation for Science-Related Careers* and *Student Information on Science-Related Careers* were found to impact students' scientific literacy skills for Canada. However, students' perception of school preparation was positively associated whereas their perception of being self-informed on science-related careers is negatively related with their scientific literacy scores. Ironically, students thinking that they are informed on availability of and processes of acquiring science-related careers tend to develop less scientific literacy skills in Canada. This result is consistent with the recent tendency of decreasing science studies among young people (OECD, 2007, p.16; Riley & Torrance, 2003; Osborne et al., 2003; Jenkins, 1994). On the other hand, students from schools preparing for or informing them on science-related careers seem to develop better scientific literacy skills. This might be due to developing an awareness of benefits of science-related careers though acquiring such a career is perceived as burdensome for numerous children.

Incorporating environmental issues into students' scientific literacy skills, the factors *Awareness of Environmental Issues* and *Responsibility for Sustainable Development* are strongly and positively related with scientific literacy skills whereas *Optimism Regarding Environmental Issues* and *Level of Concern for Environmental Issues* are related weakly and negatively with scientific literacy skills for Canadian students. Environmental literacy is a recently evolving term mainly originated from scientific literacy (Roth, 1992), and it overlaps with scientific literacy. Developing literacy skills for environmental issues requires awareness and participation (Daudi, 2008). Therefore, high positive associations of environmental awareness and responsibility for sustainable development with scientific literacy skills of Canadian students are reasonable. On the other hand, lower scientific literacy scores of students with more optimism and more concern regarding environmental issues could be a result of inadequate understanding of younglings on the fact that scientifically literate individuals will help reducing and/or solving environmental problems.

Canadian students' perception of general value of science is positively related to their scientific literacy skills whereas their perception of personal value of science is negatively related. The magnitude of *General Value of Science* coefficient is slightly higher than that of *Personal Value of Science*. Students seem to accept the value of science and improvements in human life based on science and technology. The effect of this factor on scientific literacy skills is weak nonetheless positive among Canadian students. Still, this relation is consistent with the finding that students who perceive science as being useful and consider science learning important are likely to have higher aspirations of learning science (Singh et al., 2005). On the other hand, why students allocating higher levels of personal value to science tend to have lower scientific literacy scores remains unresolved.

Students' self-efficacies in science play an important role in developing scientific literacy skills in Canada. The factor *Self-Efficacy in Science* has a positive and very high coefficient for scientific literacy scores. A bulk of research suggests that one's beliefs about competence, efficacy, and confidence are much related to achievement (Andrew, 1998; Eccles et al., 1998; Jinks & Morgan, 1999; Kupermintz, 2002; Singh et al., 2002). Accordingly, the relation found for Canadian students in the present study about self-efficacy in science and scientific literacy skills supports this notion because the tasks used for detecting efficacy beliefs to construct the index of self-efficacy in science require high levels of scientific thinking and reasoning. Furthermore, the high coefficient emphasizes the importance of the effect of self-efficacy in science.

Motivational factors affecting scientific literacy skills of Canadian students were found to be *General Interest in Science, Enjoyment of Science*, and *Future-Oriented Motivation to Learn Science*. The effect of *General Interest in Science* is negative but relatively weak, where *Enjoyment of Science* and *Future-Oriented Motivation to Learn Science* are highly positively associated. Motivation is frequently cited to have significant effect on success in science-related tasks (Eccles et al., 1998; DeBacker & Nelson, 2000; Kupermintz, 2002; Singh et al., 2002; Singh et al. 2005). As DeBacker and Nelson (2000) stated, learning science is influenced by personal enjoyment or satisfaction in science domain, student's evaluation of usefulness of science in future endeavor, and importance of accomplishments in science domain. Thus, developing higher levels of scientific literacy skills while enjoying science and regarding science as useful for future lives is evident for Canadian students. However, general interest in science measured in PISA by a wide range of subjects and methods seems to hinder acquiring higher scientific literacy scores in Canada. Expecting children to pay interest to a wide-ranging of domains concurrently might be resulting with an inverse effect in the case of developing scientific literacy skills.

Among the information and communication technology factors, self-confidence in Internet tasks is positively and strongly related to scientific literacy skills of Canadian students. However, self-confidence in high-level tasks and Internet/entertainment use are moderately and negatively related. Moreover, program and software use is strongly negatively related to scientific literacy skills. The variables *ICT Internet/Entertainment Use* and *ICT Program/Software Use* are about the frequencies of computer use. Actions of computer use are automated processes facilitating reach to information and solving problems. Furthermore, students those are confident in high-level ICT tasks tend to use computer more intensely. This facility might be obstructing students' use of existing skills and their potentials. As for the positive effect of self-confidence in Internet tasks, the tasks asked to construct the PISA index are basics of internet use for an average user. Student performance is related to computer familiarity as defined by amount and type of computer use but ease of task completion, use of more functionality, and the ability to navigate an unfamiliar interface and application are also related to performance (Lennon et al., 2003; Balım et al., 2009). Therefore, typical use of internet for an average user might be facilitating the resolution of a science-related task but the other ICT factors considered in this research and this one are neither definite means nor satisfactory conditions to develop scientific literacy skills.

Among the student factors of Canada, the slopes of the factors *Highest Occupational Status of Parents, Highest Educational Level of Parents, Immigration Background, Cultural Possessions at Home, Home Educational Resources, General Interest in Science, Personal Value of Science, Availability of Household Possessions Indicating Family Wealth, Self-Confidence in ICT High-*

Level Tasks, ICT Internet/Entertainment Use, and *ICT Program/Software Use* were randomly varying, indicating variability across schools of Canada. The present study aims to investigate the factors affecting scientific literacy skills of students, but not why for some schools the effects are higher and not why for some schools the effects are lower. Those are out of the scope of this study and require further in-depth research.

5.2.1.2 Student-Level Factors for Sweden

For Sweden, 17 out of 31 student-level variables of interest were found to be effective in the model. Of those variables, 2 had randomly varying characteristics across schools of Sweden.

Among the student background factors, slopes of *Economic, Social and Cultural Status* and *Cultural Possessions at Home* were significantly related to scientific literacy for Swedish students in the model. The effect of socioeconomic status is substantially high whereas cultural possessions had a much lower but positive effect on scientific literacy scores of Swedish students. On the other hand, *Highest Educational Level of Parents* is weakly, *Home Possessions* is moderately, and *Immigrant Background* is very strongly but all are negatively associated with scientific literacy skills of students. Backgrounds of students represented with all these variables are constituents of socioeconomic status and the effect of socioeconomic status on student performance has been elicited many times (Caldas & Bankston, 1997; Okpala et al., 2001; Yang, 2003; Turmo, 2004; Marks et al., 2006; Kalender & Berberoğlu, 2009). The findings for Sweden are consistent with the findings about the existence of and varying types of relations between socioeconomic status and domain-free performances of students. However, the effect of *Immigration Background* on scientific literacy skills of Swedish students is worth mentioning. Sweden is also a country with significant immigrant ratio with respect to its population. Fourteen percent of Swedish population was born outside of country (http://www.sweden.se/eng/Home/Society/Migration/). The finding for the effect of immigration background on scientific literacy skills of students in Sweden is consistent with the findings of previous research about immigration background (UNICEF-Innocenti Research Centre, 2009; Schleicher, 2006). On top, its coefficient is 2.5 times that of Canada.

Motivational factor *Enjoyment of Science* was positively related to scientific literacy skills of Swedish students. The effect was weak to moderate and other motivational factors in this study did not have significant impact. However, DeBacker's and Nelson's (2000) statement is still valid regarding the influence of personal enjoyment and satisfaction in the science domain. Swedish results for motivational factors support this notion.

The two factors of self-beliefs in science; *Self-Efficacy in Science* and *Self-Concept in Science,* were positively associated with scientific literacy skills of students from Sweden. The effect of efficacy was high but in this time effect of self-concept was significant and twice the higher. Again, this finding on impact of self-beliefs in science on scientific literacy is consistent with the previous research (Andrew, 1998; Eccles et al., 1998; Jinks & Morgan, 1999; Singh et al., 2002). However, Swedish students seem to emphasize competence and confidence dimensions while developing scientific literacy skills instead of pure efficacy. Because, the items used for constructing the self-concept in science index are clear statements of high competence and confidence, such as "Learning advanced school science topics would be easy for me", "I can usually give good answers to questions on science topics", "I learn school science topics quickly", and "I can easily understand new ideas in school science".

Swedish students who develop higher levels of scientific literacy skills seem to pay higher levels of value judgments to science in regard for improvements in human lives, understanding natural world, bringing economic and social benefits. However, for the case of personal value paid to science there is a strong negative association with scientific literacy skills of Swedish students, which in fact is about the same magnitude of relation with *General Value of Science*. It can be said that students claiming that they will make use of science as adults; science is relevant to them; science helps them understand the world; and there will be opportunities for them to use science after school lives cannot establish the connection that the means for using science in human life, especially at personal level, are the scientific literacy skills. Thus, students with positive perceptions of personal value of science tend to develop lower levels of scientific literacy skills in Sweden.

For Swedish students, those factors about concern for the environment are found to be related with scientific literacy scores. *Awareness of Environmental Issues* and *Responsibility for Sustainable Development* are positively related with scientific literacy skills whereas *Optimism Regarding Environmental Issues* is negatively related with scientific literacy skills. Of these, the effect of the awareness of environmental issues measures is the strongest, in accordance with the previous research (Roth, 1992; Rennie, 2005; Daudi, 2008). *Responsibility for Sustainable Development* was weakly positively related with scientific literacy skills. However, optimism for future scientific developments' role in resolving environmental problems is negatively associated with developing scientific literacy skills among Swedish students. This type of optimism brings to mind that in Sweden optimist students for environmental changes in a broad sense may not be willing to use science in future or take action in resolving environmental problems using science.

Within science teaching and learning dimension, *Student Investigations in Science Teaching and Learning* factor was highly negatively associated with scientific literacy scores of Swedish students in PISA 2006. Sweden is a country with homogeneous culture and explicitly regulated content and teaching methods (Åström & Karlsson, 2007). Scientific literacy scores of Swedish students in PISA 2006 are fairly high but not different from OECD average. According to average score ranking, Sweden was the fourteenth of 30 OECD countries and the twentieth of 57 total participants. It can be said that, in Sweden, science education seems to be more teacher-centered than student-centered, and those students more exposed to student-centered activities, in particular to student investigations, tend to have lower scientific literacy scores than their peers getting teacher-centered science education.

Among the information and communication technology factors, *Self-Confidence in ICT Internet Tasks* is positively and strongly related to scientific literacy skills of Swedish students. However, *Self-Confidence in ICT High-Level Tasks* and *Internet/Entertainment Use* are highly negatively and moderately negatively related with scientific literacy skills of Swedish students respectively. Again, *Self-Confidence in ICT Internet Tasks* captures typical use of internet for an average user and ease of those tasks captured, familiarity to contemporary computer use, and functionality may be improving the performance (Lennon et al., 2003). On the other hand, hindering effects of use of Internet and confidence in high-level tasks are not clear.

Among the student factors of Sweden, the slopes of the factors *Awareness of Environmental Issues* and *Home Possessions* were randomly varying, indicating variability across schools of Sweden. For some schools, effects of these factors on students' scientific literacy scores are higher than other schools and for some schools the effects are lower. The present study aims to investigate the factors affecting scientific literacy skills of students, but not to explain why.

5.2.1.3 Student-Level Factors for Turkey

For Turkey, 13 out of 31 student factors of interest were found to have significant effect on scientific literacy scores in the model. Of those variables, six had randomly varying characteristics across schools of Turkey.

The student background factors *Highest Occupational Status of Parents* and *Home Educational Resources* were significantly positively related to Turkish students' scientific literacy skills. Apart from those, *Home Possessions* had significant effect but that effect was negative. Association of educational resources at home with scientific literacy scores was moderate but that of home possessions was negative. The effect of parental occupation on student performance is displayed one more time with these results (McNeal, 1999; Dearing et al., 2001; Dubow et al., 2009). However the effect of educational resources at home is moderate and positive and weak negative effect of home possessions could again be a result of country-specific items asked for constructing the index. For the case Turkey, the extra items regarded as home possessions were “air conditioner”, “treadmill”, “home theater”, and “the number of rooms at home with shower facilities”. Inclusion of these uncommon resources in the index might have been reducing the dispersion of scores on the *Home Possessions* index and thus moderating the magnitude of the effect. On the other hand, inverse association of home resources with scientific literacy scores indicates that in Turkey, students more deprived of these resources develop better scientific literacy skills than less deprived. For those students, science education and science-related careers might seem a way to a higher socioeconomic status and thus less deprivation.

Enjoyment of Science was a motivation moderately and positively related to scientific literacy skills of Turkish students. This result is consistent with the notion that enjoyment and satisfaction improves performance in science (DeBacker & Nelson, 2000; Singh et al., 2002; Singh et al., 2005).

Self-efficacy beliefs of Turkish students in science are positively related to their scientific literacy skills. However, the association is weak and far less from that for Canada and half as that of Sweden. Nonetheless, this finding on impact of self-efficacy in science on scientific literacy is consistent with the previous research (Andrew, 1998; Eccles et al., 1998; Jinks & Morgan, 1999; Singh et al., 2002). This relation supports the previous research findings about positive effect of perceptions of higher efficacy beliefs with higher academic performances, but those findings are mostly sparse (Jinks & Morgan, 1999). However, existing knowledge supports the effect of self-efficacy. It is still a highly probable predictor of scientific literacy scores of Turkish students in PISA.

General value granted by Turkish students to science is positively related to Turkish students' scientific literacy scores in PISA 2006. Turkish students with higher levels of scientific literacy skills tend to have higher levels of value beliefs about science and its impact on their lives. The moderate effect of *General Value of Science* on scientific literacy scores indicates this.

For Turkish students, those factors about concern for the environmental issues found to be related with scientific literacy scores are *Awareness of Environmental Issues, Responsibility for Sustainable Development* and *Optimism Regarding Environmental Issues.* Of these, the effects of the *Awareness of Environmental Issues* and *Responsibility for Sustainable Development* were positively related with scientific literacy skills, in accordance with the previous research (Roth, 1992; Rennie, 2005; Daudi, 2008). However, Turkish students with environmental optimism tend to develop less scientific literacy skills. Regarding this dimension, optimism exhibits an inverse disposition with respect to awareness, concern, and responsibility for environmental issues.

As of science teaching and learning, *Hands-On Activities in Science Teaching and Learning* factor was moderately negatively associated with scientific literacy scores of Turkish students in PISA 2006. It can be seen that in Turkey, teaching of science with hands-on activities is inconvenient. The learning methods asked for frequency of use during construction of the index are teacher-centered, low-level, procedural activities. Therefore, using this type of methods to teach science to Turkish students has an adverse effect on their development processes of scientific literacy skills.

The information and communication technology familiarity factor, *Self-Confidence in ICT Internet Tasks* is positively and strongly related to scientific literacy skills of Turkish students. On the other hand, *Self-Confidence in ICT High-Level Tasks* and *ICT Program/Software Use* are moderately negatively related with scientific literacy skills of Turkish students. Again for Turkish case, the confidence in typical use of Internet seems to improve scientific literacy scores but confidence in high-level tasks and frequency of use of computers for writing documents and/or educational purposes are impede development of better scientific literacy skills.

Among the student factors for Turkey, the slopes of the factors *Highest Occupational Status of Parents, Awareness of Environmental Issues, Optimism Regarding Environmental Issues, General Value of Science, Enjoyment of Science,* and *Home Possessions* were randomly varying, indicating variability across schools of Turkey. For some schools, effects of these factors on students' scientific literacy scores are higher than other schools and for some schools the effects are lower. Again, the present study does not attempt to account for the causes of these variations.

5.2.2 School-Level Factors

PISA studies examined school factors selected on the basis of effective teaching and instruction, organizational and managerial characteristics of schools, and resource inputs.

5.2.2.1 School-Level Factors for Canada

Among the school-level factors, school type and school size are found to have significant impact on students' scientific literacy scores in Canada. The effect of school type was strongly negative, indicating that students from public schools of Canada tend to have higher scientific literacy scores than those from private schools. Even within private school side, students from government-dependent private schools tend to have higher scores than those from government-independent private schools. This finding is not consistent with the notion that students from private schools outperform those from public schools (OECD, 2007). However, this could be a result of oversampling of particular groups. Because, 9% of Canadian schools in PISA 2006 database are public schools, 6% are government-dependent private schools, and 85% are government-independent private schools. On the other hand, students from more crowded schools tend to have higher scientific literacy scores.

As for first group of cross-level interactions of school and student level factors in Canada, school type and school size interacts with the intercept of scientific literacy scores of students. That is, school type negatively affects the average scientific literacy score of the school that affects the mean scientific literacy scores of students. Moreover, school size positively affects the average scientific literacy score of the school that affects the mean scientific literacy scores of students. That is to say, students from public schools with higher numbers of students are expected to get higher mean scientific literacy scores from any other combination of these characteristics.

Next group of interactions occur between school factors *School Type* and *School Size* and student factor *Highest Educational Level of Parents*. In this group, school type positively affects the average of highest parental education level of the school that affects the slope of *Highest Educational Level of Parents* positively. On the other hand, school size negatively affects the average of highest parental education level of the school that affects the slope of *Highest Educational Level of Parents* positively. That is, students from government-independent private schools with lowest school size are expected to have the steepest slope for highest parental education, thus higher scientific literacy scores than those students with any other combination of those school factors.

Final set of interactions are between school factor *School Size* and student factor *ICT Internet/Entertainment Use*. In this set, school size positively affects the average Internet/entertainment use of the school that affects the slope of *ICT Internet/Entertainment Use* of students negatively. Consequently, students from more crowded schools are expected to have gentler slopes of frequency of Internet/entertainment use and thus are expected to get higher scientific literacy scores.

5.2.2.2 School-Level Factors for Sweden

Among the school-level factors, *Proportion of Girls Enrolled at School, School Type, School Size*, and *Student/Teacher Ratio* are found to have significant impact on students' scientific literacy scores in Sweden. The effect of school type was strongly negative, indicating that students from public schools of Sweden tend to have higher scientific literacy scores than those from private schools, and within private schools, students from government-dependent private schools tend to have higher scores than those from government-independent private schools. However, in Sweden, 12% of schools are government-dependent private schools and 88% are independent private schools. Therefore, public schools are nonexistent in PISA study. To conclude, schools of Sweden are private institutions and students from government-dependent ones tend to perform better in developing scientific literacy skills than those from independent ones. Furthermore, regarding the positive relation between school size and scientific literacy scores, one can say that students from more crowded schools tend to have better scientific literacy scores in Sweden.

On the other hand, the coefficient of proportion of girls at school is substantially high for Sweden. Though overall differences between girls and boys are small (OECD, 2007), one in hundredth point change in proportion of girls change 1.4 score points in the model. Boy-girls differences in scientific literacy scores range from 15 score points to 45 score points with an average of 26 in OECD countries. The distribution of proportion of girls at schools of Sweden in PISA 2006 is leptokurtic (N=197, μ=0.48, SD=0.08). Lowest proportion is 0.03 and the highest is 0.75. In several countries, although there are no performance differences between males and females, on average nearly twice as many males as females in OECD countries are graduating with science degrees (OECD, 2007, pp. 61-62). Moreover, there is systematic difference in the way males and females relate to science and science curriculum and males might be better in mastering scientific knowledge (OECD, 2007, p. 71,114). However, gender differences cannot be attributed to education systems only and social dynamics should be taken into consideration. Regarding these facts, males could be developing better scientific literacy skills in Sweden and thus schools with higher male populations could be associated with this effect.

Similarly, the association of Student/Teacher Ratio with scientific literacy scores is positive and distribution of the scale scores is again leptokurtic (N=197, μ=12.13, SD=2.58). The distribution of this ratio is found from the data as between 1.67 and 21.90. Though there are inconsistent findings about the effects of pupil/teacher ratio on students' performances, smaller ratios might have positive effects (Akerhielm, 1995; Finn & Achilles, 1999). However, the effect is multi-faceted and no clear evidence exists (Hoxby, 2000; Bandiera et al., 2010). Therefore, in a peculiar way, *Student/Teacher Ratio* is significantly positively related to scientific literacy scores of Swedish students and students from schools of higher student/teacher ratios tend to develop better scientific literacy scores.

The first set of cross-level interactions for Sweden occurs between the type of the school, size of the school, and students' average scientific literacy scores. In this set of interactions, school's average scientific literacy score affects the students' scientific literacy scores. Then, school type negatively affects the mean scientific literacy scores of students. Finally, school size positively affects the mean scientific literacy scores of students. As a result, students from public schools with higher numbers of students are expected to get higher mean scientific literacy scores from any other combination of these characteristics. However, as no school from Sweden attending PISA 2006 suits the definition of public school, it can be concluded that students from more crowded government-dependent schools tend to obtain higher scientific literacy scores than interactions of being government-independent and school size.

Final set of significant cross-level interactions in Sweden are between the proportion of girls and slope of student's home possessions in the model. In this set, *Proportion of Girls Enrolled at School* negatively affects the average *Home Possessions* score across schools that affects the *Home Possessions* slope of the individuals. That is to say, for some schools with fewer girls enrolled, the slope of student's home possessions is steeper and thus the effect of Home Possessions for students from those schools on their scientific literacy scores is higher. Tracking this path on the other direction, it proposes that schools with higher average home possessions of students tend to have lower proportions of girls.

5.2.2.3 School-Level Factors for Turkey

Among the school-level factors, *Academic Selectivity of School, School Activities to Promote Learning of Science, School's Educational Resources,* and *Teacher Shortage* are found to have significant impact on students' scientific literacy scores in Turkey. The effects of all these school factors are positive on Turkish students' scientific literacy skills.

Schools across Turkey use a selective admittance system for students. There exist placement exams for specific programmes and student's earlier academic record is considered in those exams. On the other hand, academic performance is frequently investigated along with residence in a particular area when admittance other than placement tests is the case. This degree of selectivity gives rise to special emphasis on students' performance in domains such as science. As a result, the high positive association of academic selectivity of schools with students' scientific literacy skills is an expected one.

Students from schools with higher levels of activities that promote the learning of science tend to develop better scientific literacy skills. This highly positive relation indicates that active engagement of Turkish students in school activities about science such as in science clubs, science fairs, science competitions, projects and field trips produce better scientific literacy skills for those students. Therefore, schools those promote such activities have higher differentiating effects on students' scientific literacy skills.

A school's educational or material resources are highly positively related to scientific literacy skills of students enrolled at those schools in Turkey. For PISA 2006, there was little concern about the inadequacy of internet connectivity or instructional materials in schools where their administrators reported that instruction was hindered by shortage of these resources. In fact, administrators from some countries including Turkey were concerned more about the supply of laboratory equipment (OECD, 2007, p.256). However, school administrators did not provide objective measures for conditions of those educational materials in the questionnaires and thus they are not more than perceptions of administrators. Therefore, it seems that school principals in Turkey mainly associate laboratory equipment, audio-visual materials, and library materials with their conceptions of important science education materials and the concern for hindrance in instruction due to inadequacy of these resources is high among principals (OECD, 2007, p.257). Consequently, the effect of *School's Educational Resources* where higher values along the index represent less concern for hindrance of instruction by inadequacy of educational materials on scientific literacy skills of students from that school is a plausible relation.

On the other hand, *Teacher Shortage* is found to have positive association with students' scientific literacy scores in Turkey. However, the index is not inverted and thus higher values indicate higher numbers of reports of teacher shortages by school administrators. Schools where vacant science teaching positions exist covered the 7% of students from Turkey (OECD, 2007, p. 254). But, index of teacher shortage is again constructed with perceptions of school principals and represents their views on the extent to which instruction was hindered by lack of qualifications of teachers rather than actual vacant positions for teachers. In fact, in Turkey 30% or more of the school principals reported lack of science teachers even there were no vacancies (OECD, 2007, p. 254). Therefore, it can be said that at schools of Turkey where school administrators are more concerned for quality of science teachers, students tend to develop better scientific literacy skills than those at schools with less concern for science teacher qualifications. This might be via in-service teacher education, reassigning of staff within school or appointing new teachers at schools with teacher quality concerns.

Sole set of significant cross-level interactions observed in Turkey are between the school factors *Academic Selectivity*; *School Activities to Promote Learning of Science*; and *Teacher Shortage*, and intercept of scientific literacy scores of students. In this set of interactions, the degree of school's selectivity for admittance of students very highly affects the students' mean scientific literacy scores. That is, students getting high scientific literacy scores are mainly expected to come from highly selective schools. Then, the degree of school's involvement in science activities promoting the learning of science highly impacts the students' mean scientific literacy scores. In other words, students getting high scientific literacy scores are expected mainly to come from schools more involved with science activities engaging students. Finally, at schools where administrators are more concerned about the qualifications of their teachers, students develop better scientific literacy skills than those from schools where administrators are less concerned about teacher qualifications.

5.3 Conclusions

Representing the hierarchical structure in which students are nested within schools, some school characteristics are associated with student characteristics. School membership affects student characteristics that affect performance of student. Across the countries investigated in the present study, some student factors and some school factors were both found to have significant impact on students' scientific literacy skills. Allocations of these human and physical resources among students and schools help predict students' scientific literacy scores.

For all three countries, there are differences in students' scientific literacy skills among schools. With the help of hierarchical linear modeling, variances of students' scientific literacy scores are partitioned as between-schools and within schools, which is not feasible by ordinary regression methods. For Canada, 26% of variance in scientific literacy scores is between-schools, indicating differences among schools that might be accounted for by the school factors. Similarly, there exist individual differences between schools of Sweden that might be accounted for by membership of the school. A lower though significant proportion; 16% of variability in Swedish students' scientific literacy scores are across schools. However, Turkey is the top with 52% between-school variance, which is higher than the variance accountable by individual characteristics. Therefore, there are substantial differences or large gaps between schools of Turkey in regard for scientific literacy skills of students.

Public schools facilitate development of higher levels of scientific literacy skills for their students than government-controlled private schools and government-controlled private schools facilitate development of higher levels of scientific literacy skills for their students than government-independent private schools in countries with lower impact of socioeconomic background on science performance.

In countries with higher impact of socioeconomic background and with larger performance differences such as Turkey, schools are more selective in admittance of students. Schools those are effective in science performance in such countries employ highly motivated, competent, and scientifically more literate students. Moreover, material resources and quality of teachers are of higher concern for schools in those countries.

Parental education and education-related occupation of family are of high importance for children in Canada and Turkey, thus related to science performances of students in those countries.

Science performances of students are related to socioeconomic background, in many ways to its attributes in Canada, Sweden, and Turkey. However, neither high nor poor performance could be directly associated with those attributes in any of those cultures.

Immigration status is of serious concern for Canada and Sweden because immigrant students tend to develop lower levels of scientific literacy skills than natives. Moreover, first generation immigrants perform lower than second generation immigrants.

Intrinsic factors of motivation, self-beliefs, and value given by students to learn science are highly effective in developing scientific literacy skills across Canada, Sweden, and Turkey, again free from culture.

Students from Canada, Sweden, and Turkey exhibit the same pattern for their environmental perceptions and the relation of those perceptions to scientific literacy. Students that are aware of environmental issues and feel responsibility for sustainable development develop better scientific literacy skills. However, the more awareness they develop and the more responsibility they feel, students become less optimistic about future benefits of scientific and technological developments in resolving environmental problems.

All students with higher levels of scientific literacy skills from Canada, Sweden, and Turkey are confident in typical personal use of Internet. However, they all are less confident in use of high-level tasks of information and communication technology. Moreover, frequencies of use of Internet and computer programs are lower for those students with higher levels of scientific literacy skills in all three countries.

A common feature in the models of three countries is the lack of relation of scientific literacy skills with some dimensions of student factors core to science education. The major expectancy of science education is science performance and science literacy. However, student factors representing science teaching and learning dimension were little or not related to scientific literacy skills of students. Among those factors, the only related ones were about student-centered learning of science and they were negatively related to developing scientific literacy skills in a very limited manner. In Turkey hands-on activities in science teaching; and in Sweden student investigations in science teaching hindered development of scientific literacy skills. On the other hand, interactions in science teaching and demonstrating the applications and relevancy of science to societal life were not at all associated with scientific literacy skills of students. Though there is no causal explanation of this situation in this study, it seems to be a problem of science education with respect to equipping individuals with knowledge and skills for life.

5.3.1 Country-Based Comparison of Models

- *Highest Occupational Status of Parents* is twice the more effective in Canada than in Turkey on students' scientific literacy scores; however, it is not an effective socioeconomic attribute for scientific literacy scores of Swedish students.
- *Home Possessions* has inverse and a high magnitude effect on students' scientific literacy skills in Canada of those in Sweden and

Turkey. That is, effects for Sweden and Turkey are far lower and negative on students' scientific literacy skills.

- The common motivational factor among three countries is enjoyment of science. Across the three cultures, it moderately to highly contributes to development of higher levels of scientific literacy skills.
- In all three countries, self-efficacy in science has positive impact on students' scientific literacy skills. In Canada, the effect is high; on the other hand, they are moderate for Sweden and Turkey. However, Swedish students regard themselves competent rather than just efficient. Therefore, Swedish students' self-concepts in science are highly related with higher levels of scientific literacy skills.
- Across three cultures, students with higher perception levels of value of science develop better scientific literacy skills. However, in Canada and Sweden personal value of science is negatively associated with students' scientific literacy skills, indicating that those children who establish the relevancy of science to their lives tend to develop lower levels of scientific literacy.
- Free from culture, students who become more aware of environmental issues and feel more responsible for sustainable development grow better scientific literacy skills. On the contrary, those are more optimistic about the future role of scientific and technological developments in resolving environmental problems tend to develop lower levels of scientific literacy skills. That is, the more the environmental awareness and responsibility, the less optimistic students become in all these cultures.
- In all three countries, students who feel more confident in using Internet for typical personal purposes develop higher levels of scientific literacy skills. When it comes to actual use, those who use Internet more frequently tend to develop lower levels of scientific literacy scores in Canada and Sweden and those who use computer

packages more frequently tend to develop lower levels of scientific literacy scores in Canada and Turkey.

- Level of confidence in high-level tasks of information and communication technology is negatively related with higher levels of scientific literacy in all countries.
- The way of science teaching and learning is not much related to students' scientific literacy skills however, student investigations and hands-on activities proved negative on scientific literacy levels of Swedish and Turkish students respectively.
- Only in Canada science career-oriented preparation and informing of students have impact on scientific literacy skills of students. School preparation for science careers has positive association whereas student information on availability of science-related careers has negative association with scientific literacy.
- In all three countries, some school characteristics are determined to have impact on individual levels of scientific literacy skills.
- Both in Canada and Sweden, type of the school in regard for being public or private and government-dependent or independent is found to impact development of students' scientific literacy skills.
- Both in Canada and Sweden, size of the school is found to impact development of students' scientific literacy skills.
- In Sweden, quantities regarding school such as size, proportion of girls and student/teacher ratio are found to impact scientific literacy skills of students, whereas in Turkey qualities regarding school such as academic selectivity, quality of science teachers, promotion of learning science are found to impact scientific literacy skills of students.

5.4 Implications

Results from the present study are associative in nature and the study does not investigate causal relations. However, comparing the relations of factors across different cultures, it provides some suggestions for improving science performance and literacy, and for pursuing quality and equity within education systems.

The relationships of science performance with attitudinal dimensions are limited and moderate among students. That is, most students exhibit a broad interest in science no matter they are high performers or not. Therefore, education systems should be restructured to develop more positive attitudes towards science.

Though students with positive feelings about science and about its importance tend to be more scientifically literate, these attitudes are not results of doing science or experiencing. They are rather broad views. Schools should promote learning with higher levels of active engagement of students, in the mean time forcing them to use scientific knowledge and knowledge about science.

General value given to science focuses on societal benefits of science and is positively related with science performance. On the other hand, personal value of science is related with future relevancy and personal benefits of science. Students who take science personally more valuable develop lower levels of scientific literacy skills. Therefore students should be informed by teachers and schools that scientific literacy is a lifelong process in nature and it is not developed when needed according to personal relevancy of science.

Enjoyment of learning science is a motive to develop higher levels of scientific literacy. However, high performers do not necessarily exhibit future-oriented motivation to learn science. Therefore, students with higher aspirations of science-related careers should be motivated to pursue that goal in order to form science personnel because scientifically more literate persons do not necessarily become scientists or pursue science-related careers.

Across all three cultures, science-related activities outside school do not seem to be related with higher levels of scientific literacy. The implication is twofold: engagement in science could be improved with outside-school science activities, and schools should emphasize that science is not just something you do at school.

High performers in science are not to be related to school preparation for science careers and are less informed on science-related careers. Therefore, special programmes should be implemented among schools to inform and prepare students for science related careers to direct high performers to science, at least partially.

Teaching and learning of science is negatively or not at all related to scientific literacy. This suggests a problem for science education either because it does not facilitate development of scientific literacy skills or because it is perceived as a school subject. Better methods and approaches to science teaching should be employed by education systems in order to grow scientifically more literate citizens.

Higher levels of awareness for environmental issues and feeling responsibility for sustainable development are common for students with higher scientific literacy skills. On the other hand, scientific literacy skills are negatively or not at all related to levels of concern for environmental issues. Being aware but not concerned about environmental issues suggests that schools should provide students with more information on the subject and deeper understandings.

Socioeconomic background seems to be strongly associated with scientific literacy skills across countries. Therefore, schools should provide learning opportunities for students irrespective of their socioeconomic backgrounds to ensure more quality and equity in education.

5.5 Recommendations for Further Research

Results of the present study are just associations and do not propose any causal connection. Experimental studies should be conducted to explain the causal connections of student and school factors with scientific literacy skills.

In-depth analyses are required to understand why variability occurs across schools on some factors but not on others. Whether there is connection between variability and low proportions of effective school factors across countries might be resolved by that way.

To assess cultural effects on students' science performances, levels of HLM could be incremented by within-country regions or region of the country. To a considerable extent, location on performance could be understood.

5.6 Limitations of the Study

The present study is limited to student and school data of Canada, Sweden and Turkey in the Programme for International Student Assessment 2006. Therefore, any generalization of the findings of this study to all countries, OECD countries, or non-OECD participants, also to students of all ages, students of all 15-year-olds, or students of any grade is controversial.

Results of the present study are associative in nature and they cannot explain any causal relation. Therefore the information regarding the effects of school and student factors on scientific literacy skills of students are limited to existent-nonexistent and positive-negative bases.

Important amounts of variances remained unexplained by the school and student factors in the models. The variables of the present study are thus limited to inclusive variables of PISA and selection of those variables by the researcher in the first place. Necessary variables to build more accurate models might not have been sought by PISA.

Sampling could be another limitation of this study. The sampling design used for PISA 2006 was a stratified sampling design. First schools were sampled from strata and then students were sampled from those sampled schools. Therefore identification of stratification variables and sampling variance due to stratification could be a limitation of representation of population.

Use of some information collected from school principals or administrators could be a limitation of this study. Many items in the school questionnaire were to be responded according to the perceptions of administrators, lacking precise measures. Therefore, responses of the administrators might be biased and unreliable. On the other hand, administrator might not have accurate data on the measure asked at the short window of testing and prefer giving approximate responses or even omitting.

REFERENCES

About OECD, (n.d.) Retrieved November 11, 2008 from The Organisation for Economic Co-operation and Development Web site: http://www.oecd.org/pages/0,3417,en_36734052_36734103_1_1_1_1_1,00.html

Akerhielm, K. (1995). Does class size matter? *Economics of Education Review*, 14(3), 229-241.

Akyüz, G. (2006). *Teacher and Classroom Characteristics: Their Relationship with Mathematics Achievement in Turkey, European Union Countries and Candidate Countries.* Unpublished doctoral dissertation, Middle East Technical University, Ankara, Turkey.

Anderson, J. O., Lin, H. S., Treagust, D. F., Ross, S. P., & Yore, L. D. (2007). Using large-scale assessment datasets for research in science and mathematics education: Programme for International Student Assessment (PISA). *International Journal of Science and Mathematics Education*, 5(4), 591-614. doi:10.1007/s10763-007-9090-y

Andrew, S. (1998). Self-efficacy as a predictor of academic performance in science. *Journal of Advanced Nursing*, 27(3), 596-603.

Åström, M., & Karlsson, K. (2007). Using hierarchical models to test differences in Swedish results from OECD's PISA 2003: Integrated and subject-specific science education. *Nordina*, 3(2), 121-131.

Balım, A. G., Evrekli, D. İ., & Deniş, H. (2009). Türkiye'nin PISA 2006'daki durumu üzerine bir inceleme: fen bilimleri yeterlilik düzeyinin bilgi ve iletişim teknolojilerinin kullanımına göre değerlendirilmesi. *e-Journal of New World Sciences Academy*, 4(3), 1053-1066, Article No: 1C0079.

Bandiera, O., Larcinese, V., & Rasul, I. (2010). Heterogeneous class size effects: new evidence from a panel of university students. *The Economic Journal*, no. doi:10.1111/j.1468-0297.2010.02364.x

Beaton, A. E., Postlethwaite, T. N., Ross, K. N., Spearritt, D., & Wolf, R. M., (1999). *The benefits and limitations of international educational achievement studies*. Paris: International Institute for Educational Planning/UNESCO.

Becker, B. J. (1990). Coaching for the Scholastic Aptitude Test: further synthesis and appraisal. *Review of Educational Research*, 60(3), 373-417.

Brunner, M., Artelt, C., Krauss, S., & Baumert, J. (2007). Coaching for the PISA test. *Learning and Instruction*, 17(2), 111-122. doi:10.1016/j.learninstruc.2007.01.002

Bybee R.W. (1997). Towards an understanding of scientific literacy. In: W. Gräber & C. Bolte.(Eds.). *Scientific literacy. An international symposium* (p. 37-68). Institut für die Pädagogikder Naturwissenschaften (IPN): Kiel, Germany.

Bybee, R. W. (2008). Scientific Literacy, Environmental Issues, and PISA 2006: The Paul F-Brandwein Lecture. *Journal of Science Education and Technology*, 17(6), 566-585. doi:10.1007/s10956-008-9124-4

Bybee, R. W., & Stage, E. (2005). No Country Left Behind. *Issues in Science and Technology*, Winter 2005, 69-75.

Caldas, S. J., & Bankston, C. (1997). Effect of school population socioeconomic status on individual academic achievement. *Journal of Educational Research*, 90(5), 269-277.

Cooper, P., & McIntyre, D. (1996). *Effective teaching and learning: Teachers' and students' perspectives*. Buckingham: Open University Press.

Çalışkan, M. (2008). *The Impact of School and Student Related Factors on Scientific Literacy Skills in the Programme for International Student Assessment – PISA 2006.* Unpublished doctoral dissertation, Middle East Technical University, Ankara, Turkey.

Daudi, S. S. (2008). Environmental literacy: a system of best-fit for promoting environmental awareness in low literate communities. *Applied Environmental Education and Communication*, 7, 76-82. doi:10.1080/15330150802502155

Dearing, E., McCartney, K,, & Taylor, B. A. (2001). Change in family income-to-needs matters more for children with less. *Child Development*, 72(6), 1779–1793. doi:10.1111/1467-8624.00378

DeBacker, T. K., & Nelson, R. M. (2000). Motivation to learn science: differences related to gender, class type, and ability. *Journal of Educational Research*, 93(4), 245-254.

DeBoer, G. E. (2000). Scientific literacy: Another look at its historical and contemporary meanings and its relationships to science education reform. *Journal of Research in Science Teaching*, 37(6), 582-601.

Dubow, E. F., Boxer, P., & Huesmann, L. R. (2009). Long-term effects of parents' education on children's educational and occupational success: mediation by family interactions, child aggression, and teenage aspirations. *Merrill Palmer Q (Wayne State Univ. Press)*, 55(3), 224–249. doi:10.1353/mpq.0.0030

Eccles, J. S., Wigfield, A., & Schiefele, U. (1998). Motivation to succeed. In N. Eisenberg (Vol. Ed.) & W. Damon (Series Ed.). *Handbook of Child Psychology* (5th ed., Vol. 3, pp. 1017-1095). New York: John Wiley & Sons.

Farmer, H. S., Wardrop, J. L., & Rotella, S. C. (1999). Antecedent factors differentiating women and men in science/nonscience careers. *Psychology of Women Quarterly*, 23(4), 763-780.

Ferry, T. R., Fouad, N. A., & Smith, P.L. (2000). The role of family context in a social cognitive model for career-related choice behavior: a math and science perspective. *Journal of Vocational Behavior*, 57(3), 348-364. doi:10.1006/jvbe.1999.1743

Finn, J. D., & Achilles, C. M. (1999). Tennessee's Class Size Study: findings, implications, misconceptions. *Educational Evaluation and Policy Analysis*, 21(2), 97-109.

Fonseca, J. M. B., & Conboy, J. E. (2006). Secondary student perceptions of factors effecting failure in science in Portugal. *Eurasia Journal of Mathematics, Science and Technology Education*, 2(2), 83-96.

Fredriksson, U., (2001). International student assessment and its implications for schools. *Education Review*, 15(1), 104-109.

Gardner, P. L. (1975). Attitudes to science. *Studies in Science Education*, 2, 1–41.

Gardner, P. L. (1995). Measuring attitudes to science. *Research in Science Education*, 25(3), 283–289.

Goldstein, H. (2004). International comparisons of student attainment: some issues arising from the PISA study. *Assessment in Education*, 11(3), 319-330.

Harlen, W. (2001). The assessment of scientific literacy in the OECD/PISA project. *Studies in Science Education*, 36(1), 79-103.

Hoxby, C. M. (2000). The effects of class size on student achievement: new evidence from population variation. *The Quarterly Journal of Economics*, 115(4), 1239-1285. doi:10.1162/003355300555060

İş Güzel, Ç. (2006). *A Cross-Cultural Comparison of the Impact of Human and Resource Allocations on Students' Mathematical Literacy Skills in the Programme for International Student Assessment (PISA) 2003.* Unpublished doctoral dissertation, Middle East Technical University, Ankara, Turkey.

İş Güzel, Ç., & Berberoğlu, G. (2005). An analysis of the Programme for International Student Assessment 2000 (PISA 2000) mathematical literacy data for Brazilian, Japanese and Norwegian students. *Studies In Educational Evaluation*, 31(4), 283-314. doi:10.1016/j.stueduc.2005.11.006

Jenkins, E. W. (1994). Public understanding of science and science education for action. *Journal of Curriculum Studies*, 26(6), 601-611. doi:10.1080/0022027940260602

Jinks, J., & Morgan, V. (1999). Children's perceived academic self-efficacy: an inventory scale. *Clearing House*, 72(4), 224-230.

Kalender, İ., & Berberoğlu, G. (2009). An assessment of factors related to science achievement of Turkish students. *International Journal of Science Education*, 31(10), 1379-1394. doi:10.1080/09500690801992888

Kirsch, I., de Jong, J., Lafontaine, D., McQueen, J., Mendelovits, J., & Monseur, C., (2002). *Reading for Change: Performance and Engagement across Countries*. Paris: OECD Publications.

Kjærnsli, M., & Lie, S. (2004). PISA and scientific literacy: similarities and differences between the Nordic countries. *Scandinavian Journal of Educational Research*, 48(3), 271-286. doi:10.1080/00313830410001695736

Klopfer, L. E. (1971). Evaluation of learning in science. In B. S. Bloom, J. T. Hastings & G. F. Madaus (Eds.), *Handbook of formative and summative evaluation of student learning* (London: McGraw-Hill).

Kupermintz, H. (2002). Affective and conative factors as aptitude resources in high school science achievement. *Educational Assessment*, 8(2), 123-137.

LaFontaine, D. & Monseur, C. (2006). *Impact of test characteristics on gender equity indicators in the assessment of reading comprehension*. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, California, USA.

Laugksch, R. C. (2000). Scientific literacy: A conceptual overview. *Science Education*, 84(1), 71-94.

Lee, V. E., & Bryk, A. S. (1989). A multilevel model of the social distribution of high school achievement. *Sociology of Education*, 62(3), 172-192.

Lee, V. E., Smith, J. B., & Croninger, R. G. (1997). How high school organization influences the equitable distribution of learning in mathematics and science. *Sociology of Education*, 70(2), 128-150.

Lennon, M., Kirsch, I., von Davier, M., Wagner, M., & Yamamoto, K. (2003). *Feasibility Study for the PISA ICT Literacy Assessment: Report to Network A* (Long Abstract). Retrieved July 29, 2010 from Organisation for Economic Co-operation and Development Programme for International Student Assessment website: http://www.oecd.org/dataoecd/35/13/33699866.pdf

Lie, S., & Linnakylä, P. (2004). Nordic PISA 2000 in a sociocultural perspective. *Scandinavian Journal of Educational Research*, 48(3), 227-230. doi:10.1080/00313830410001695709

Literacy. (2009). In Wikipedia, The Free Encyclopedia. Retrieved July 20, 2009, from http://en.wikipedia.org/w/index.php?title=Literacy&oldid=303153216

Marks, G. N., Cresswell, J., & Ainley, J. (2006). Explaining socioeconomic inequalities in student achievement: the role of home and school factors. *Educational Research and Evaluation*, 12(2), 105-128. doi:10.1080/13803610600587040

McNeal, R. B. (1999). Parental involvement as social capital: differential effectiveness on science achievement, truancy, and dropping out. *Social Forces,* 78(1), 117-144.

Members and partners, (n.d.) Retrieved November 11, 2008 from The Organisation for Economic Co-operation and Development Web site: http://www.oecd.org/pages /0,3417,en_36734052_36761800_1_1_1_1_1,00.html

Mislevy, R. J. (1995). What can we learn from international assessments? *Educational Evaluation and Policy Analysis*, 17(4), 419-437.

Naumann, J. (2005). TIMSS, PISA, PIRLS and Low Educational Achievement in World Society. *Prospects*, 35(2), 229-248.

Okpala, C. O., Okpala, A. O., & Smith, F. E.(2001).Parental involvement, instructional expenditures, family socioeconomic attributes, and student achievement. *Journal of Educational Research*, 95(2), 110-115.

Organisation for Economic Co-operation and Development, (2001). *Knowledge and Skills for Life: First Results from the OECD Programme for International Student Assesssment (PISA) 2000*. Paris: OECD Publications.

Organisation for Economic Co-operation and Development, (2004). *Learning for Tomorrow's World: First Results from PISA 2003*. Paris: OECD Publications.

Organisation for Economic Co-operation and Development, (2005). *PISA 2003 Technical Report*. Paris: OECD Publications.

Organisation for Economic Co-operation and Development, (2006). *Assessing Scientific, Reading and Mathematical Literacy: A Framework for PISA 2006*. Paris: OECD Publications.

Organisation for Economic Co-operation and Development, (2007). *PISA 2006: Science competencies for tomorrow's world Volume 1: Analysis*. Paris: OECD Publications.

Organisation for Economic Co-operation and Development, (2009). *PISA 2006 Technical Report*. Paris: OECD Publications.

Osborne, J., Simon, S., & Collins, S. (2003). Attitudes towards science: a review of the literature and its implications. *International Journal of Science Education*, 25(9), 1049-1079. doi:10.1080/0950069032000032199

Özdemir, E. (2003). *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.* Unpublished master's thesis, Middle East Technical University, Ankara, Turkey.

Papanastasiou, C. (2002). Achievements in Mathematics and Science in an International Context. *Educational Research and Evaluation*, 8(1), 3-12.

Papanastasiou, E. C., Zembylas, M., & Vrasidas, C. (2003). Can computer use hurt science achievement? The USA results from PISA. *Journal of Science Education and Technology*, 12(3), 325-332.

Papanastasiou, E. C., Zembylas, M., & Vrasidas, C. (2005). An examination of the PISA database to explore the relationship between computer use and science achievement. Educational Research and Evaluation, 11(6), 529-543. doi:10.1080/13803610500254824

Pella, M. O., O'Hearn, G. T., & Gale, C. W. (1966). Referents to scientific literacy. *Journal of Research in Science Teaching*, 4(3), 199-208.

PISA Project Consortium (2005). *School Sampling Preparation Manual: PISA 2006 Main Study. Version One: April 2005.* Paris: OECD Publications.

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

Rennie, L. J. (2005). Science awareness and scientific literacy. *Journal of the Australian Science Teachers Association*, 51(1), 10-14.

Riley, K., & Torrance, H. (2003). Big change question As national policy-makers seek to find solutions to national education issues, do international comparisons such as TIMMS and PISA create a wider understanding, or do they serve to promote the orthodoxies of international agencies? *Journal of Educational Change*, 4(4), 419-425. doi:10.1023/B:JEDU.0000006188.73593.8d

Roth, C. E. (1992). *Environmental Literacy: Its Roots, Evolution, and Directions in the 1990s*. Columbus OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Schleicher, A. (2006). Where immigrant students succeed: a comparative review of performance and engagement in PISA 2003. *Intercultural Education*, 17(5), 507-516. doi:10.1080/14675980601063900

Showalter, V. M. (1974). What is united science education? Part 5. Program objectives and scientific literacy. Prism II, 2(3 4).

Singh, K., Chang, M., & Dika, S. (2005). Affective and motivational factors in engagement and achievement in science. *International Journal of Learning*, 12(6), 207-218.

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

TIMSS & PIRLS International Study Center, (2006a). *PIRLS 2006 Assessment Framework and Specifications*, 2nd Ed. Chestnut Hill, Massachusetts: Boston College Publications.

TIMSS & PIRLS International Study Center, (2006b). *TIMSS Advanced 2008 Assessment Frameworks*. Chestnut Hill, Massachusetts: Boston College Publications.

Topping, K. (2006). PISA/PIRLS data on reading achievement: Transfer into international policy and practice. *The Reading Teacher*, 59(6), 588-590. doi:10.1598/RT.59.6.9

Topping, K., Brozo, W., Dionisio, M. L., Roller, C., & Valtin, R., (2003). *Policy and practice implications of the Program for International Student Assessment (PISA) 2000: Report of the International Reading Association PISA task force.* Retrieved from ERIC database. (ED478184)

Turmo, A. (2004). Scientific literacy and socio-economic background among 15-year-olds - a Nordic perspective. *Scandinavian Journal of Educational Research*, 48(3), 287-305. doi:10.1080/00313830410001695745

UNESCO Education Sector, (2004). *The Plurality of Literacy and its Implications for Policies and Programmes: Position Paper*. Paris: United National Educational, Scientific and Cultural Organization.

UNICEF Innocenti Research Centre (2009). *Children in Immigrant Families in Eight Affluent Countries: Their Family, National and International Context*. Retrieved June 21, 2010 from http://www.unicef-irc.org/publications/pdf/ii_immig_families.pdf

Wang, J., & Staver, J. R. (2001). Examining relationships between factors of science education and student career aspiration. *Journal of Educational Research*, 94(5), 312-319.

Warm, T. A. (1985). *Weighted Maximum Likelihood Estimation of Ability in Item Response Theory with Tests of Finite Length*. CGI-TR-85-08, U.S. Coast Guard Institute, Oklahoma City.

White, P., & Smith, E. (2005). What can PISA tell us about teacher shortages? *European Journal of Education*, 40(1), 93-112.

Yang, Y. (2003). Dimensions of socio-economic status and their relationship to mathematics and science achievement at individual and collective levels. *Scandinavian Journal of Educational Research*, 47(1), 21-41. doi:10.1080/0031383032000033317

Young, D. J., Reynolds, A. J., & Walberg, H. J. (1996). Science achievement and educational productivity: a hierarchical linear model. *Journal of Educational Research*, 89(5), 272-278. doi:10.1080/00220671.1996.9941328

APPENDIX A

PARTICIPATING COUNTRIES IN PISA 2006

Table A.1 PISA 2006 Participating Countries and Number of Participant Students

OECD-Member Participants		**Non-OECD-Member Participants**	
Country	**No of Students in PISA 2006**	**Country**	**No of Students in PISA 2006**
Australia	14,170	Argentina	4,339
Austria	4,927	Azerbaijan	5,184
Belgium	8,857	Brazil	9,295
Canada	22,646	Bulgaria	4,498
Czech Republic	5,932	Chile	5,233
Denmark	4,532	Colombia	4,478
Finland	4,714	Croatia	5,213
France	4,716	Estonia	4,865
Germany	4,891	Hong Kong-China	4,645
Greece	4,873	Indonesia	10,647
Hungary	4,490	Israel	4,584
Iceland	3,789	Jordan	6,509
Ireland	4,585	Kyrgyz Republic	5,904
Italy	21,773	Latvia	4,719
Japan	5,952	Liechtenstein	339
Korea	5,176	Lithuania	4,744
Luxembourg	4,567	Macao-China	4,760
Mexico	30,971	Qatar	6,265
The Netherlands	4,871	Rep. of Montenegro	4,455
New Zealand	4,823	Rep. of Serbia	4,798
Norway	4,692	Romania	5,118
Poland	5,547	Russian Federation	5,799
Portugal	5,109	Slovenia	6,595
Slovak Republic	4,731	Chinese Taipei	8,815
Spain	19,604	Thailand	6,192
Sweden	4,443	Tunisia	4,640
Switzerland	12,192	Uruguay	4,839
Turkey	4,942		
United Kingdom	13,152		
United States	5,611		

APPENDIX B

RANKING OF COUNTRIES ON SCIENCE SCALE OF PISA 2006

Table B.1 Ranking of Participants on Overall Science Scale in PISA 2006

OECD Member	Non-OECD	Science Score	s.e.
Finland		563	2.0
	Hong Kong – China	542	2.5
Canada		534	2.0
	Chinese Taipei	532	3.6
	Estonia	531	2.5
Japan		531	3.4
New Zealand		530	2.7
Australia		527	2.3
Netherlands		525	2.7
	Liechtenstein	522	4.1
Korea		522	3.4
	Slovenia	519	1.1
Germany		516	3.8
United Kingdom		515	2.3
Czech Republic		513	3.5
Switzerland		512	3.2
	Macao – China	511	1.1
Austria		511	3.9
Belgium		510	2.5
Ireland		508	3.2
Hungary		504	2.7
Sweden		503	2.4
Poland		498	2.3
Denmark		496	3.1
France		495	3.4
	Croatia	493	2.4
Iceland		491	1.6
	Latvia	490	3.0
United States		489	4.2
Slovak Republic		488	2.6
Spain		488	2.6
	Lithuania	488	2.8
Norway		487	3.1
Luxembourg		486	1.1
	Russian Federation	479	3.7
Italy		475	2.0
Portugal		474	3.0
Greece		473	3.2
	Israel	454	3.7
	Chile	438	4.3
	Serbia	436	3.0
	Bulgaria	434	6.1

Table B.1 Continued

OECD Member	Non-OECD	Science Score	s.e.
	Uruguay	428	2.7
Turkey		424	3.8
	Jordan	422	2.8
	Thailand	421	2.1
	Romania	418	4.2
	Montenegro	412	1.1
Mexico		410	2.7
	Indonesia	393	5.7
	Argentina	391	6.1
	Brazil	390	2.8
	Colombia	388	3.4
	Tunisia	386	3.0
	Azerbaijan	382	2.8
	Qatar	349	0.9
	Kyrgyzstan	322	2.9

APPENDIX C

CHECKING HLM ASSUMPTIONS OF MODELS

C.1 Assumption Tests for Canada

C.1.1 Homogeneity of Variances Assumption for Canada

Homogeneity of error variances for level-1 model is examined by the distribution of natural logarithm of the residual standard deviation from the final fitted fixed effect model (MDRSVAR). As much as the distribution is normal, there is evidence that level-1 variance varies randomly over level-2 units. However, assuming variances are equal when not true, no bias will arise and standard errors will be robust (Raudenbush & Bryk, 2002, p.263-267).

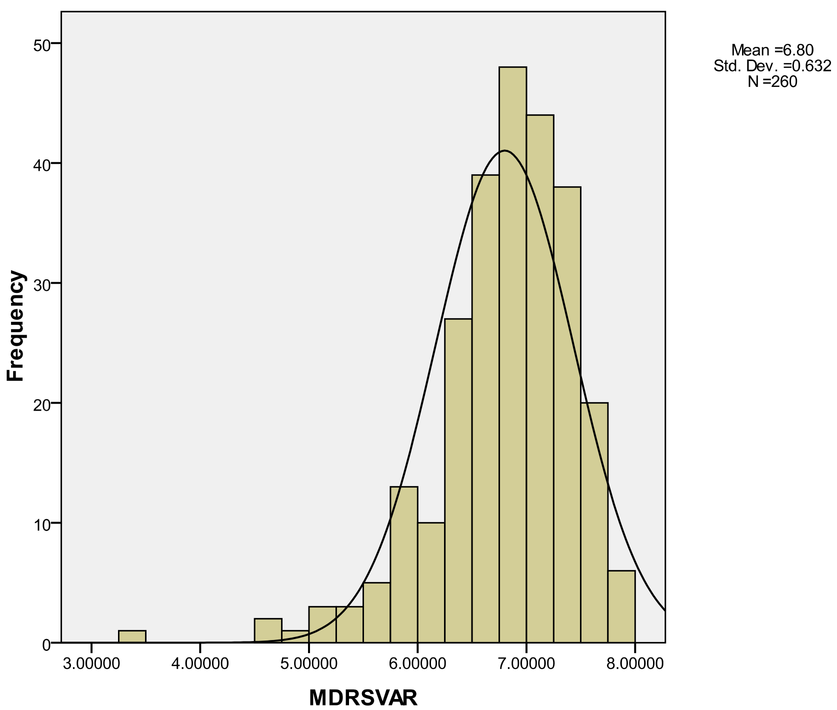


Figure C.1 Histogram of MDRSVAR for Canada

For this distribution, skewness is -1.306 and kurtosis is 3.490. It can be seen that there are a few schools that have smaller within-school variance than others. Therefore this close-to-normal distribution presents evidence for homogeneity of error variances across schools. Moreover, the resemblance of Q-Q plot to a 45 degree line with few outliers at the end also suggests homogeneity of variances.

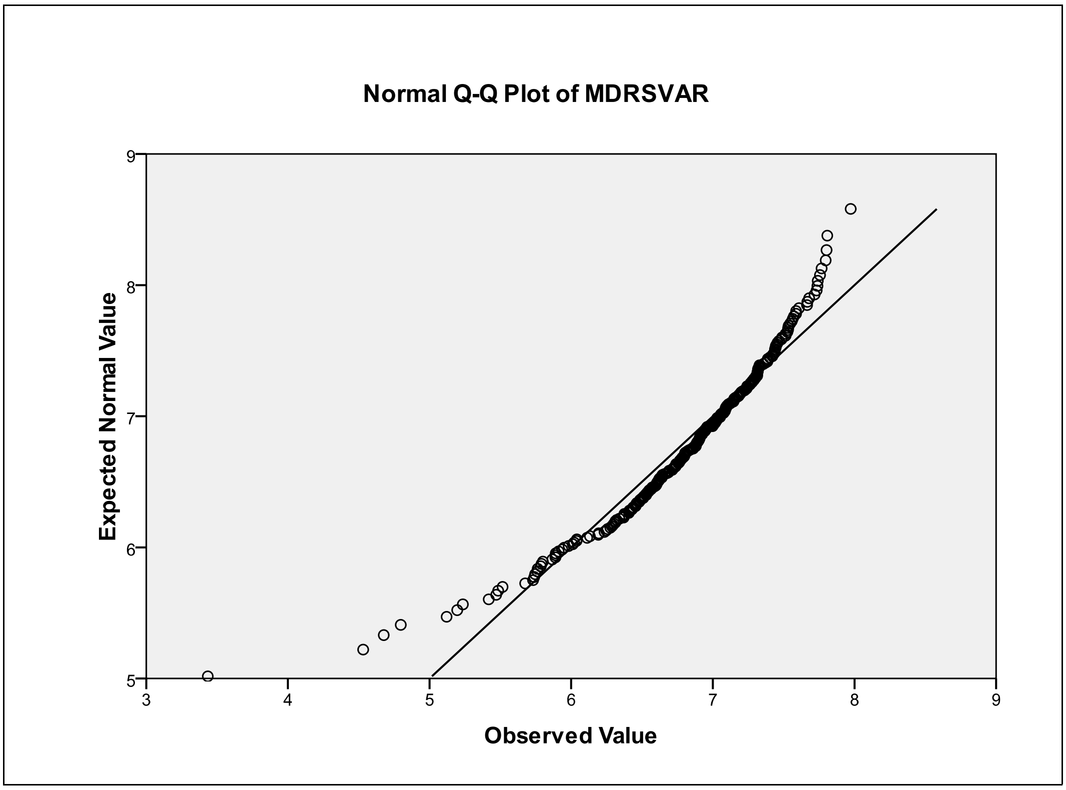


Figure C.2 Q-Q Plot of MDRSVAR for Canada

C.1.2 Normality of Errors Assumption for Canada

Plots of random coefficients empirical Bayes (EB) estimates exhibit high kurtosis. The normality assumption of random coefficients thus fails. However, nonnormality of errors at student level will not bias estimation of school effects, but it will introduce bias into standard errors at both levels (Raudenbush & Bryk, 2002, p.263-267). Nonetheless, little is known about the direction and severity of such effects. That nonnormality might be the result of multicategorical outcomes for observed variables when categories are not clearly ordered, or proportional concerns for categories are not similar.

Table C.1 Distribution of EB Estimates of Random Coefficients for Canada

Estimate	N	Skewness		Kurtosis	
	Statistic	Statistic	Std. Error	Statistic	Std. Error
EBHISEI	895	0.311	0.082	15.441	0.163
EBIMMIG	895	1.940	0.082	28.486	0.163
EBWEALTH	895	-0.048	0.082	6.847	0.163
EBHIGHCO	895	0.064	0.082	11.053	0.163
EBINTUSE	895	0.112	0.082	10.460	0.163
EBPRGUSE	895	-0.335	0.082	10.914	0.163

C.1.3 Normality Assumption of School Residuals for Canada

Under the assumption of normality at level-2, the scatter plot of theoretical values from a chi-square distribution of the number of random factors (CHIPCT) against Mahalanobis distance (MDIST) should represent a 45-degree line. Otherwise, there is evidence against the normality assumption. Mahalanobis distance provides a summary measure of the distance of the empirical Bayes estimates of the fixed effect parameters from their fitted values. For the case of Canada, high deviations from the line are observed, indicating that normality assumption for school residuals is not met.

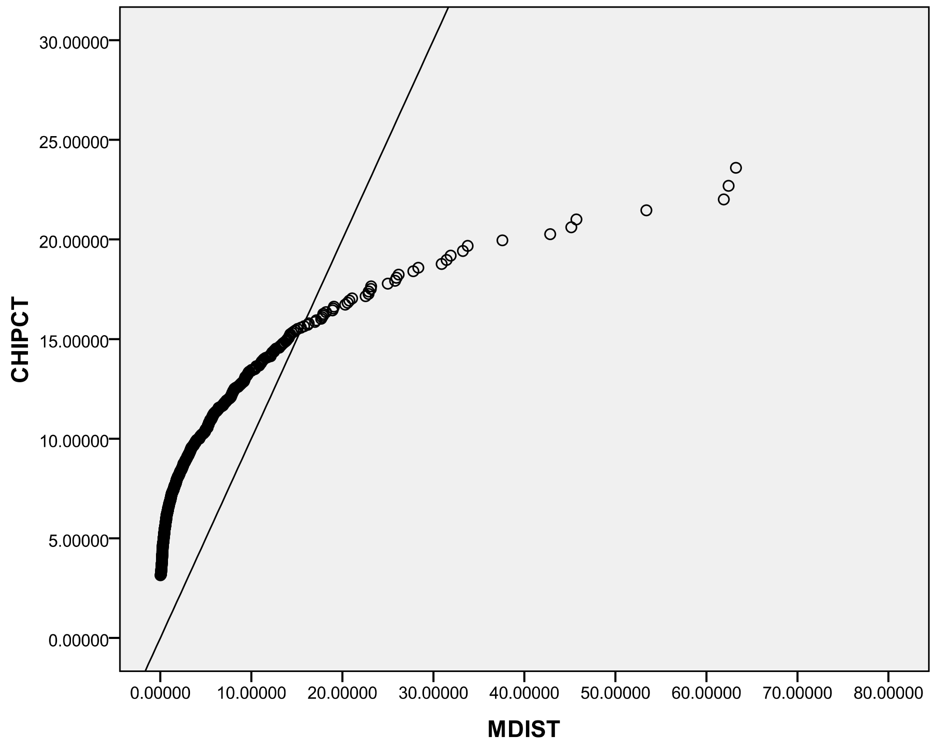


Figure C.3 Scatter Plot of CHIPCT against MDIST for Canada

C.2 Assumption Tests for Sweden

C.2.1 Homogeneity of Variances Assumption for Sweden

Homogeneity of error variances for level-1 is examined by the distribution of natural logarithm of the residual standard deviation from the final fitted fixed effect model (MDRSVAR). For this distribution, skewness is 0.204 and kurtosis is -0.032. Therefore this distribution is normal and presents evidence for homogeneity of error variances across schools of Sweden. Moreover, the resemblance of Q-Q plot of MDRSVAR to a 45 degree line also suggests homogeneity of variances.

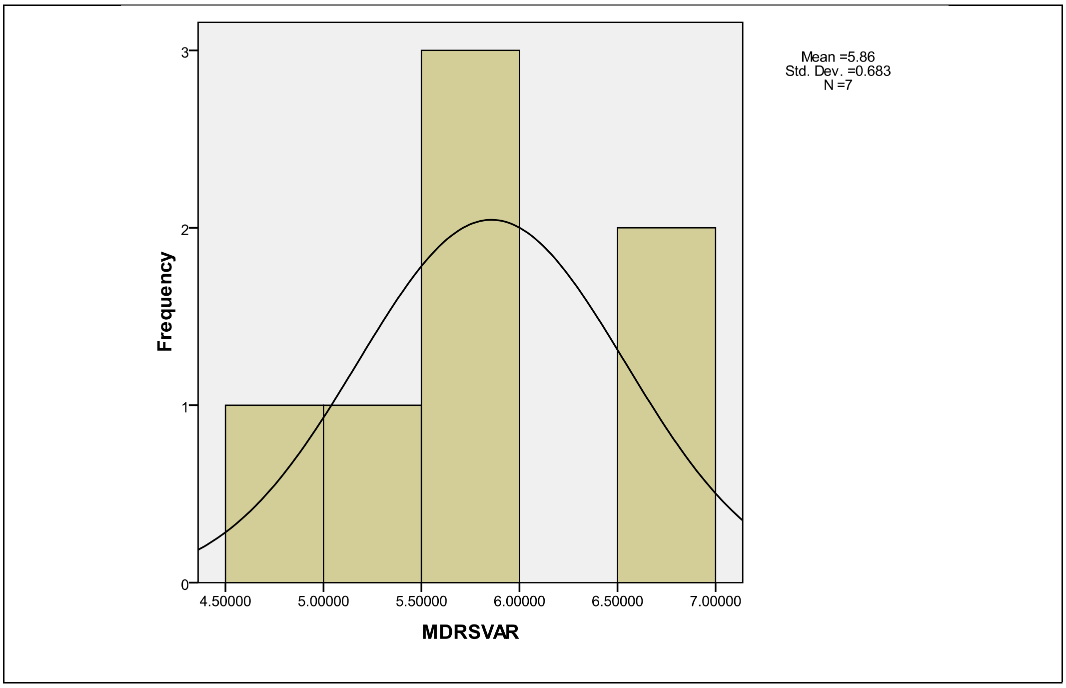


Figure C.4 Histogram of MDRSVAR for Sweden

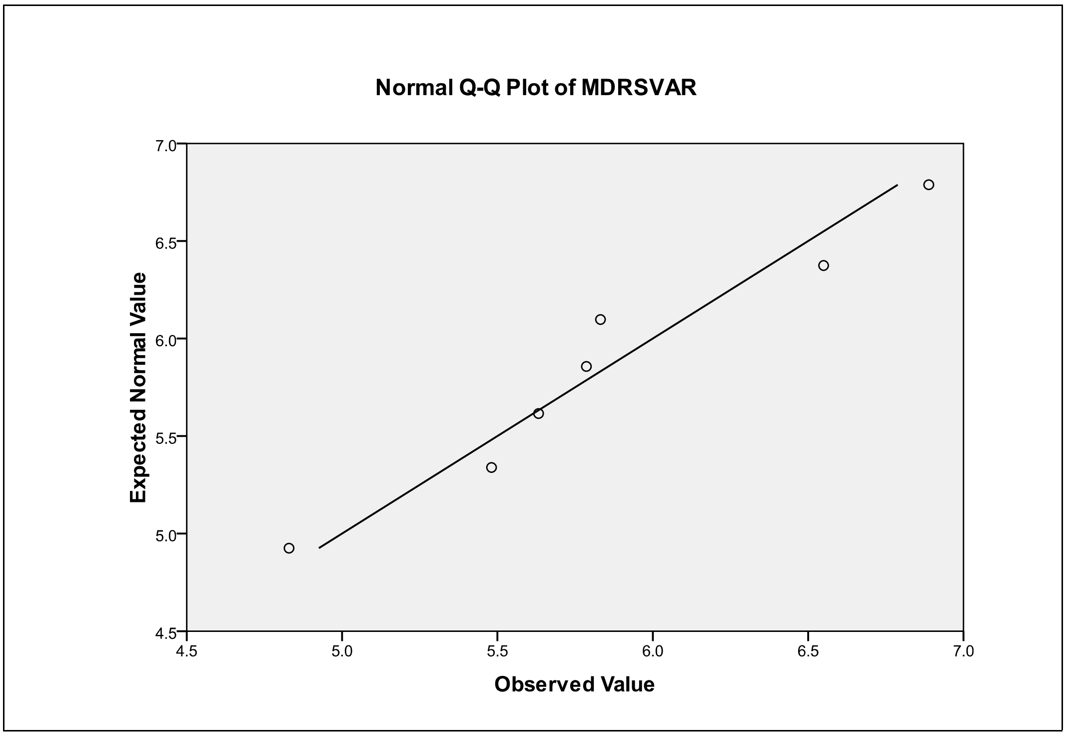


Figure C.5 Q-Q Plot of MDRSVAR for Sweden

C.2.2 Normality of Errors Assumption for Sweden

Plots of random coefficients empirical Bayes estimates exhibit normal distributions. Therefore, there is evidence that errors at level-1 are distributed normally.

Table C.2 Distribution of EB Estimates of Random Coefficients for Sweden

Estimate	N	Skewness		Kurtosis	
	Statistic	Statistic	Std. Error	Statistic	Std. Error
EBENVAWA	195	-0.167	0.174	2.104	0.346
EBHOMEPO	195	0.296	0.174	2.660	0.346

C.2.3 Normality Assumption of School Residuals for Sweden

Under the assumption of normality at level-2, the scatter plot of CHIPCT against MDIST should represent a 45-degree line.

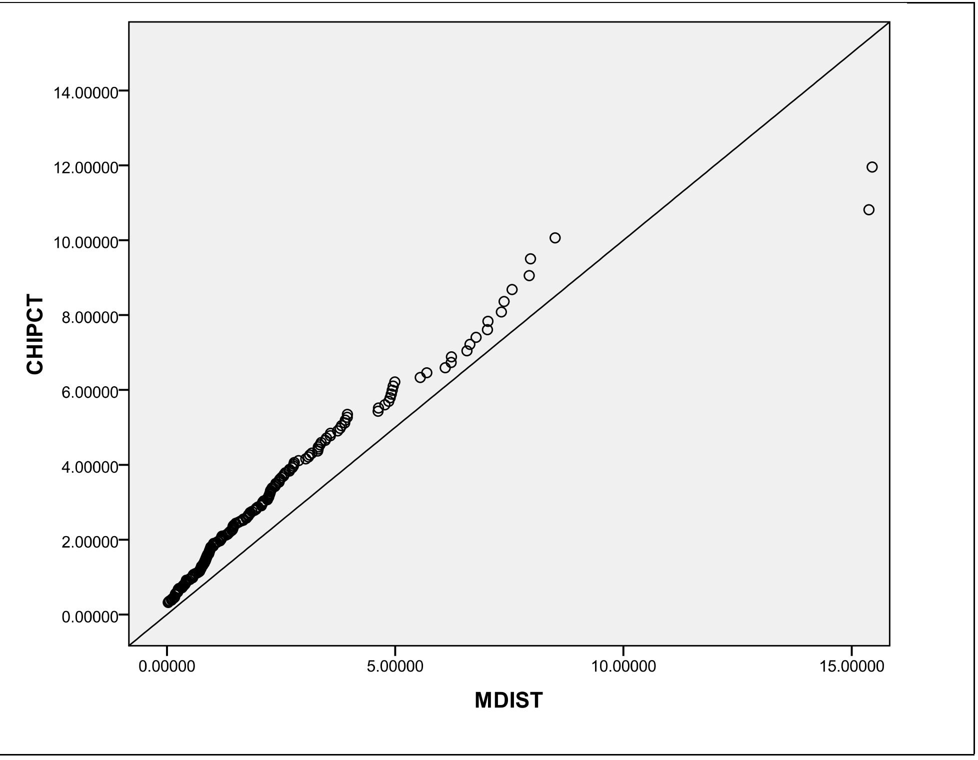


Figure C.6 Scatter Plot of CHIPCT against MDIST for Sweden

The line does not overlap with distribution of points but is highly similar. Moreover, degree of overlap between CHIPCT and MDIST supports this similarity. Therefore there is not enough evidence against the normality of school residuals assumption for Sweden.

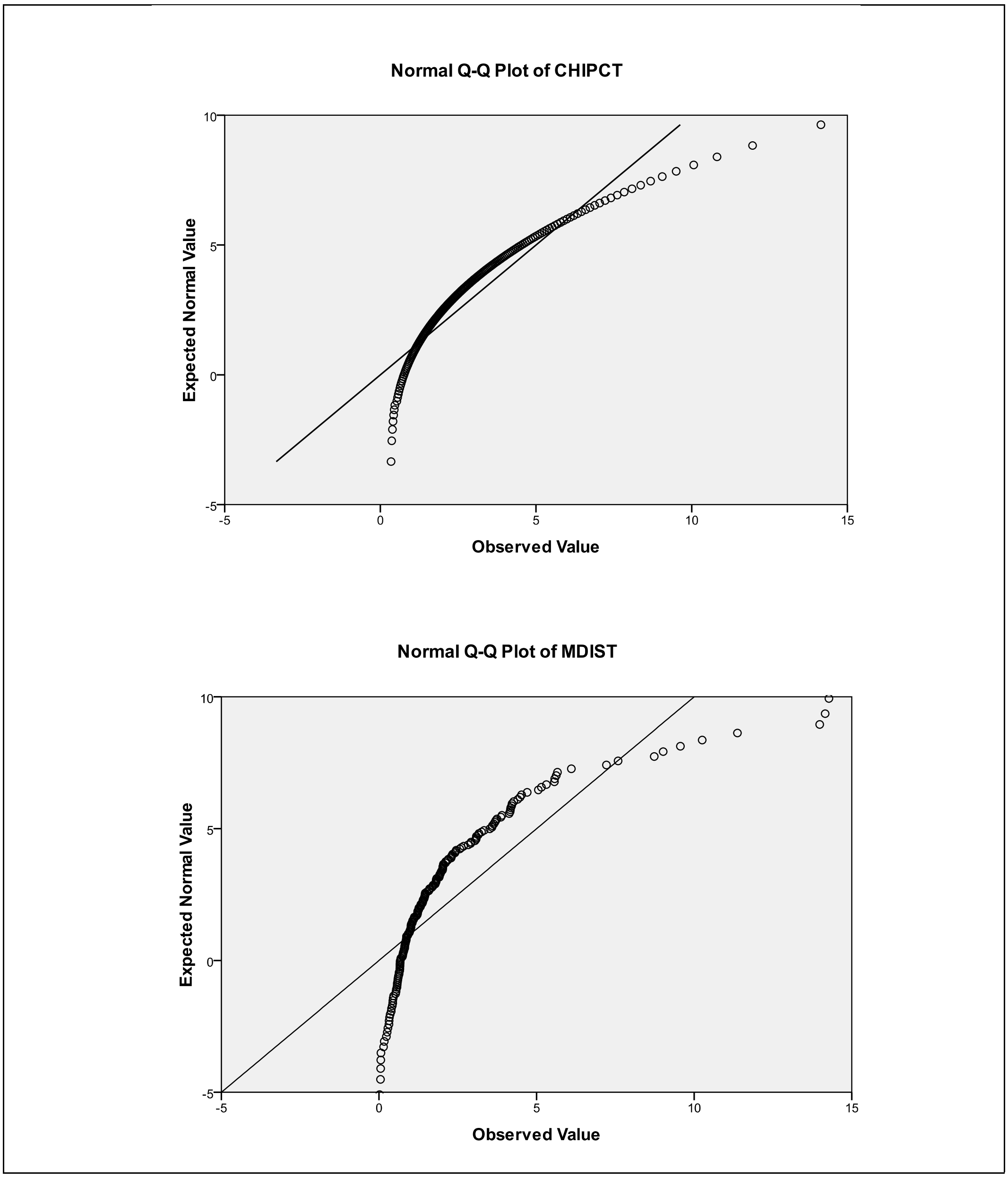


Figure C.7 Q-Q Plots of CHIPCT and MDIST for Sweden

C.3 Assumption Tests for Turkey

C.3.1 Homogeneity of Variances Assumption for Turkey

Homogeneity of error variances for level-1 is examined by the distribution of natural logarithm of the residual standard deviation from the final fitted fixed effect model (MDRSVAR). For this distribution, skewness is -1.421 and kurtosis is 2.699. There are a few schools that have smaller within-school variance than others. Also, there are some outliers in Q-Q plot. However, by this distribution, there is not enough evidence for heterogeneity of error variances across schools of Turkey.

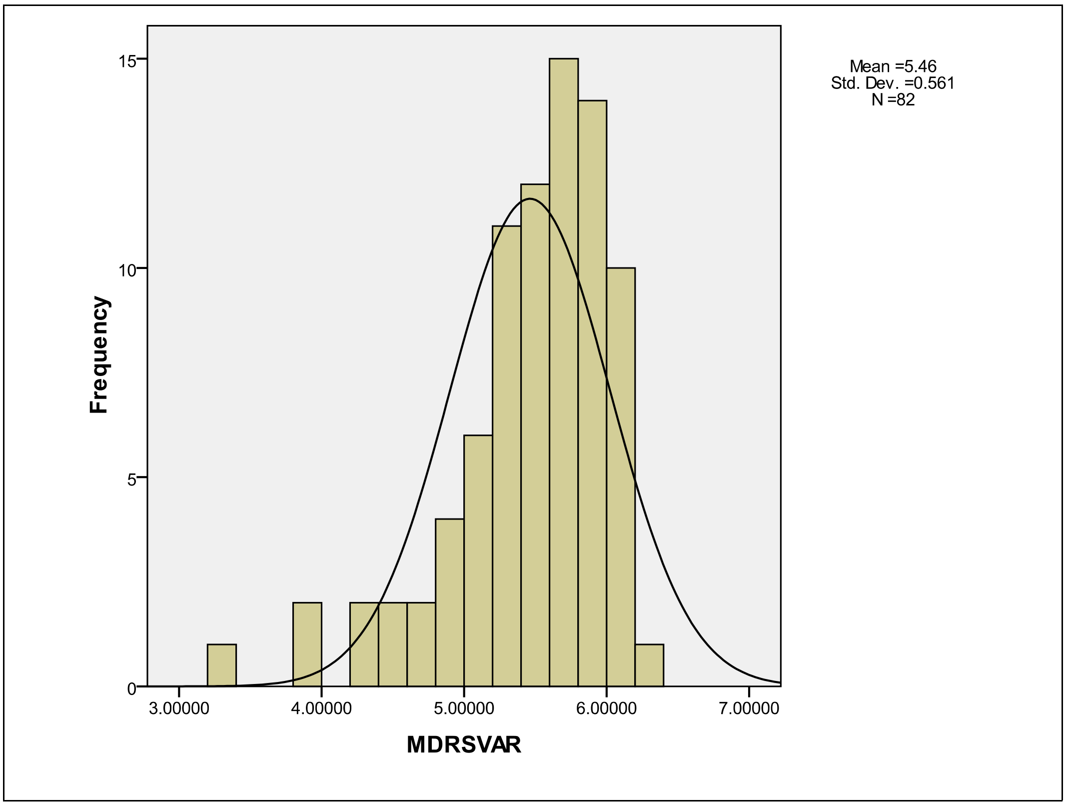


Figure C.8 Histogram of MDRSVAR for Turkey

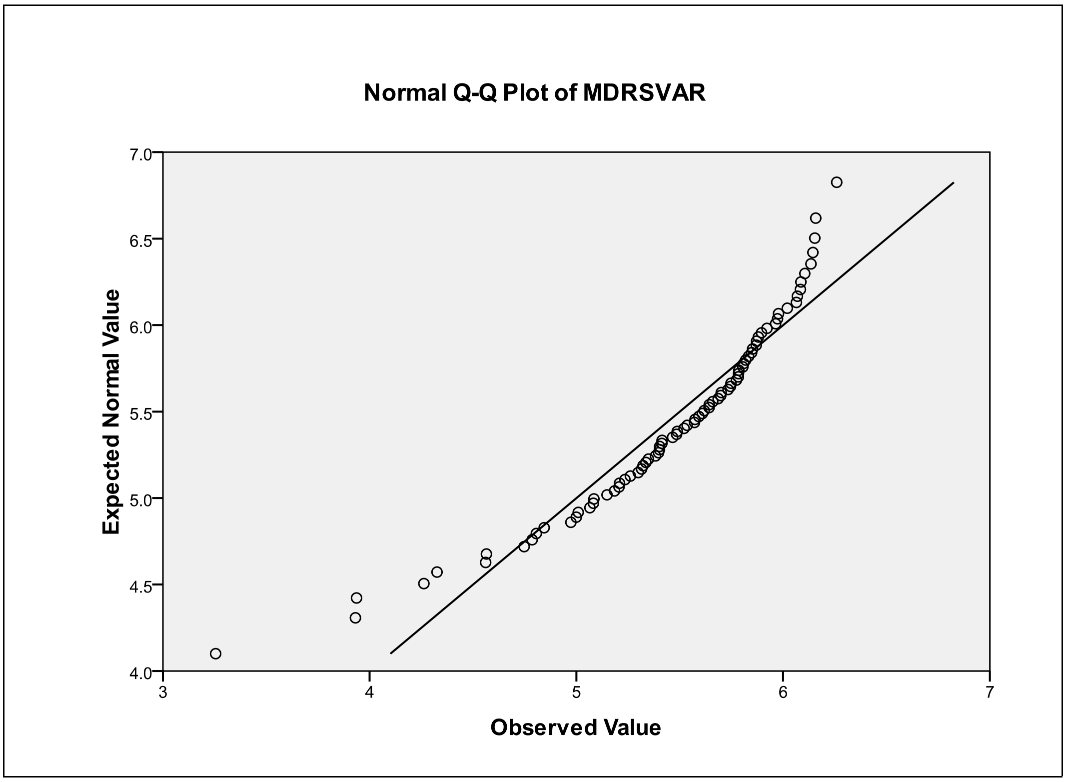


Figure C.9 Q-Q Plot of MDRSVAR for Turkey

C.3.2 Normality of Errors Assumption for Turkey

Plots of random coefficients empirical Bayes estimates exhibit high kurtosis and considerable skewness.

Table C.3 Distribution of EB Estimates of Random Coefficients for Turkey

Estimate	N	Skewness		Kurtosis	
	Statistic	Statistic	Std. Error	Statistic	Std. Error
EBENVAWA	159	-1.265	0.192	5.970	0.383
EBENVOPT	159	-0.404	0.192	9.057	0.383
EBGENSCI	159	-0.965	0.192	6.984	0.383
EBJOYSCI	159	1.358	0.192	9.525	0.383
EBSCIEEF	159	-1.150	0.192	8.234	0.383

The normality assumption of random coefficients thus fails. However, that will not bias estimation of school effects, but it will introduce bias into standard errors at both levels (Raudenbush & Bryk, 2002, p.263-267).

C.3.3 Normality Assumption of School Residuals for Turkey

Under the assumption of normality at level-2, the scatter plot of CHIPCT against Mahalanobis distance should represent a 45-degree line.

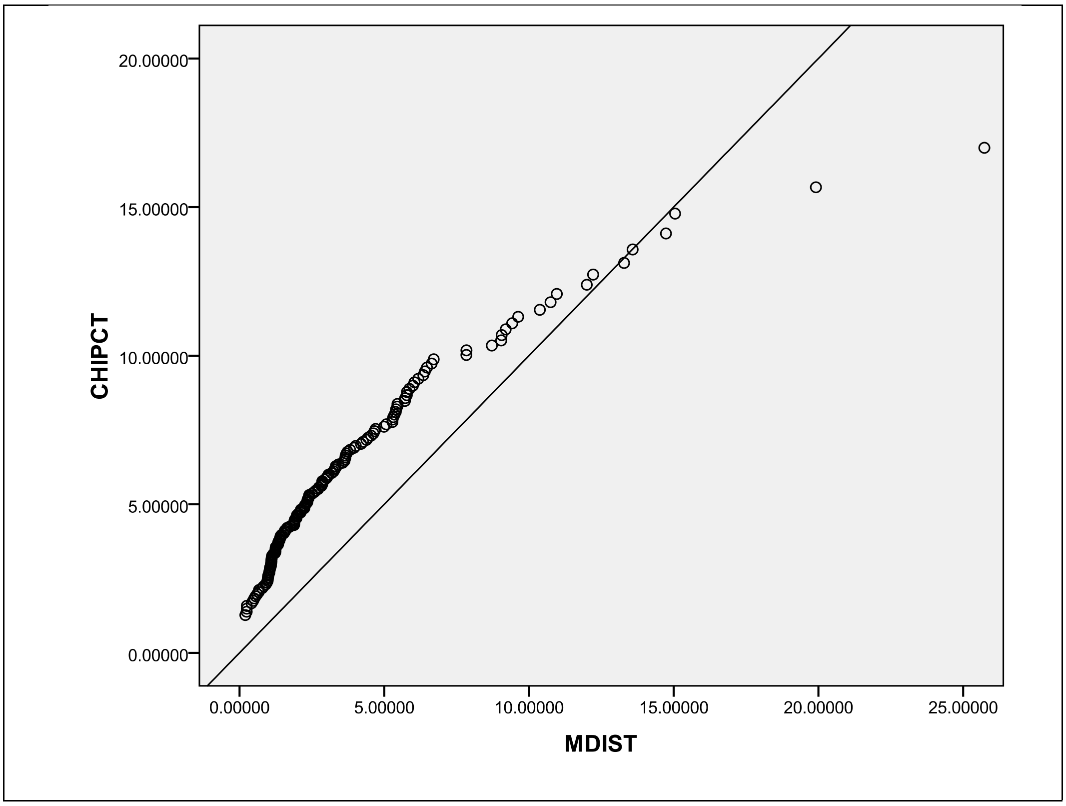


Figure C.10 Scatter Plot of CHIPCT against MDIST for Turkey

The line does not overlap with distribution of points but is moderately similar. Moreover, degree of overlap between CHIPCT and MDIST supports this similarity. Therefore there is not enough evidence against the normality of school residuals assumption for Turkey.

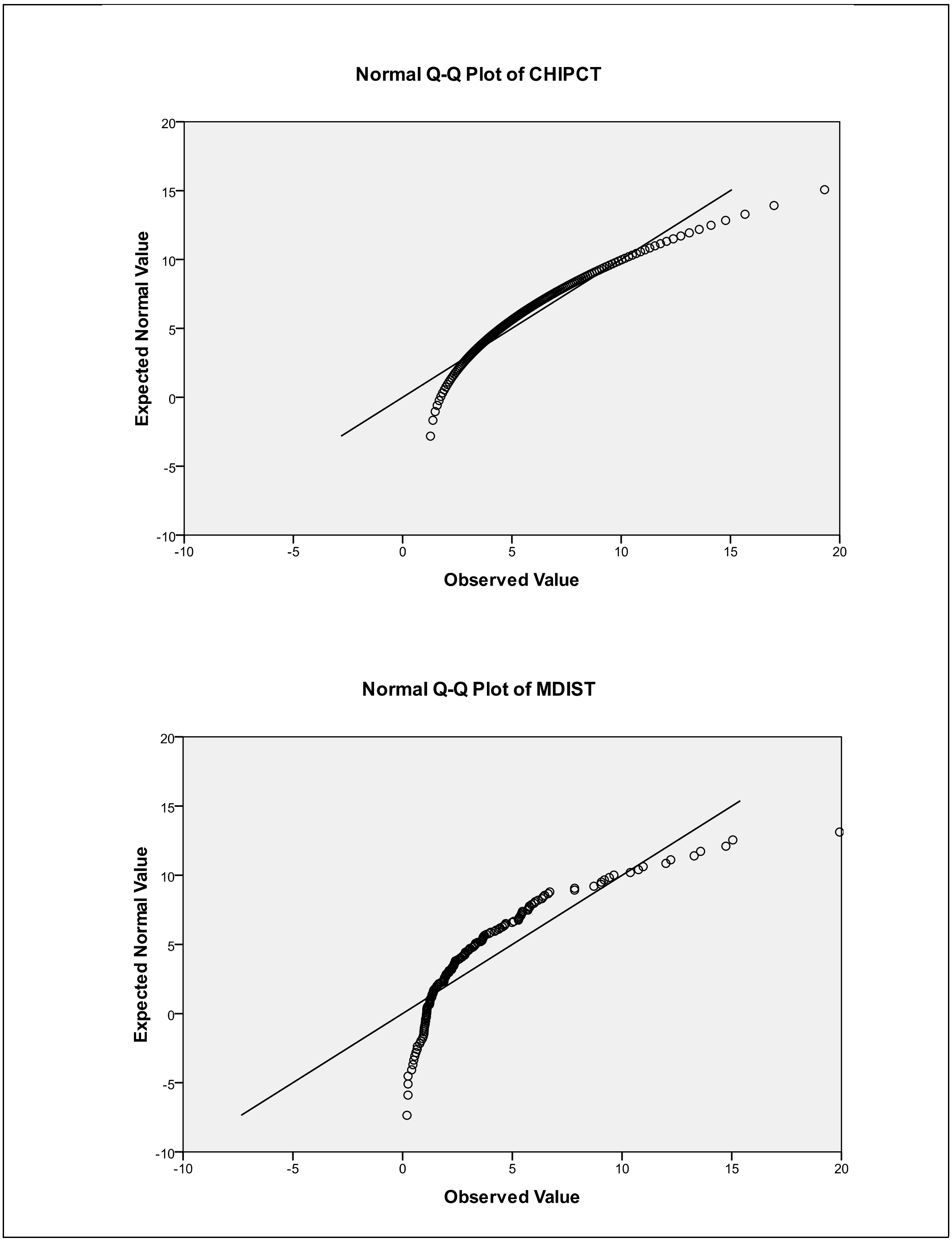


Figure C.11 Q-Q Plots of CHIPCT and MDIST for Turkey

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name	**:**	Çelebi, Özgür
Nationality	**:**	Turkish (T.C.)
Date of Birth	**:**	May 28th, 1977
Phone	**:**	+90 532 300 62 07
e-mail	**:**	ozgurclb@gmail.com ozgurclb@yahoo.com

EDUCATION

Degree	**Institution**	**Year of Graduation**
M.Sc.	METU Dept. of SSME	2004
B.S.	METU Dept. of SSME	2001
High School	Bursa Anatolian Lycée	1994

CERTIFICATES

June 2003: *Cisco Networking Academy Cisco Certified Network Associate Education Programme,* Continuing Education Center, Middle East Technical University, Ankara, Turkey.

June 2002: *Information Technologies Certificate*, Department of Computer Engineering, Middle East Technical University, Ankara, Turkey.

WORK EXPERIENCE

Year	Place	Enrollment
2009-2010	ARDEN Automotive	Owner
2002-2008	METU Dept. of SSME	Research Assistant

PUBLICATIONS

Berberoğlu, G., Çelebi, Ö., Özdemir, E., Uysal, E., & Yayan, B. (2003). Factors affecting achievement level of Turkish students in the Third International Mathematics and Science Study (TIMSS). *Journal of Educational Sciences and Practice*, 2(3), 1-11.

COMPUTER SKILLS

Programming Languages	**:**	Pascal, C, Java
Web Programming	**:**	HTML, ASP, XML, PHP, JavaScript
RAD Tools	**:**	Borland Delphi, MS Visual Basic
Operating Systems	**:**	Windows 9X – 2000 – XP – Vista, 7; Linux Slackware, Red Hat Linux, Linux Mandrake, Ubuntu (Installation and operating)
Office Tools	**:**	MS Word, MS Excel, MS Access, MS PowerPoint, MS FrontPage, MS Outlook

HOBBIES

Movies, cars, computers