

INVESTIGATION OF THE EFFECT OF STUDENT AND SCHOOL
BACKGROUND VARIABLES, TEACHING AND LEARNING VARIABLES
AND NON-COGNITIVE OUTCOMES ON THE COMPONENTS OF
SCIENTIFIC LITERACY IN PROGRAMME FOR INTERNATIONAL
STUDENT ASSESSMENT (PISA 2015)

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

BY

BURCU ORHAN

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

SEPTEMBER 2020

Approval of the thesis:

**INVESTIGATION OF THE EFFECT OF STUDENT AND SCHOOL
BACKGROUND VARIABLES, TEACHING AND LEARNING VARIABLES
AND NON-COGNITIVE OUTCOMES ON THE COMPONENTS OF
SCIENTIFIC LITERACY IN PROGRAMME FOR INTERNATIONAL
STUDENT ASSESSMENT (PISA 2015)**

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

Prof. Dr. Halil Kalıpçılar
Dean, Graduate School of **Natural and Applied Sciences**

Prof. Dr. Erdinç Çakıroğlu
Head of the Department, **Mathematics and Science Education**

Prof. Dr. Ömer Geban
Supervisor, **Mathematics and Science Education, METU**

Assoc. Prof. Dr. Eren Ceylan
Co-Supervisor, **Elementary Education, Ankara University**

Examining Committee Members:

Prof. Dr. Ayhan Yılmaz
Mathematics and Science Education, Hacettepe University

Prof. Dr. Ömer Geban
Mathematics and Science Education, METU

Prof. Dr. Emine Erdem
Mathematics and Science Education, Hacettepe University

Assoc. Prof. Dr. Ömer Faruk Özdemir
Mathematics and Science Education, METU

Asst. Prof. Dr. Gökhan Öztürk
Mathematics and Science Education, METU

Date: 11.09.2020



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: Burcu Orhan

Signature:

ABSTRACT

INVESTIGATION OF THE EFFECT OF STUDENT AND SCHOOL BACKGROUND VARIABLES, TEACHING AND LEARNING VARIABLES AND NON-COGNITIVE OUTCOMES ON THE COMPONENTS OF SCIENTIFIC LITERACY IN PROGRAMME FOR INTERNATIONAL STUDENT ASSESSMENT (PISA 2015)

Orhan, Burcu

Doctor of Philosophy, Mathematics and Science Education

Supervisor: Prof. Dr. Ömer Geban

Co-Supervisor: Assoc. Prof. Dr. Eren Ceylan

September 2020, 241 pages

This study aims to explore science-related student and school factors and their effects on 15-year-old Turkish students' scientific literacy skills by using the data from Programme for International Student Assessment (PISA 2015). PISA 2015 was implemented with the participation of 187 schools and 5895 students from Turkey.

Student-level variables of the study are out-of-school study time for science, science activities, science self-efficacy, interest in broad science topics, enjoyment of science, instrumental motivation, epistemological beliefs, environmental awareness, environmental optimism, students' expected occupational status, and science learning time. School-level variables include proportion of science teachers by all teachers, proportion of science teachers fully certified, proportion of science teachers with ISCED level 5A and a major in science, total number of science teachers at school, shortage of educational material, shortage of educational staff,

science specific resources, and creative extra-curricular activities. Combined science scores, science knowledge scores and science competencies scores of students were evaluated separately. Data analysis were executed by using hierarchical linear modeling (HLM).

Out-of-school study time for science, science activities and environmental optimism were found to be effecting all scientific literacy subscales negatively. Science learning time, epistemological beliefs, environmental awareness and students' expected occupational status were found to be effecting all scientific literacy subscales positively. Instrumental motivation had no significant effect on scientific literacy of the students. Science self-efficacy had a positive effect on combined science, content knowledge and explaining phenomena scientifically subscales. Interest in broad science topics had a positive effect on interpreting data and evidence scientifically in addition to these subscales. The relationship of enjoyment of science with all subscales except for evaluating and designing scientific inquiry subscale was positive.

Proportion of science teachers by all teachers and science specific resources had positive effect on all scientific literacy, whereas shortage of educational material and total number of science teachers at school effected scientific literacy negatively. Creative extra-curricular activities had positive effect on combined science score, procedural and epistemic knowledge and evaluating and designing scientific inquiry subscales. Proportion of science teachers fully certified, proportion of science teachers with ISCED level 5A and a major in science and shortage of educational staff had no significant effect on scientific literacy.

Keywords: Programme For International Student Assessment (PISA 2015), Scientific Literacy, Hierarchical Linear Modeling (HLM)

ÖZ

ÖĞRENCİ VE OKUL ÖZELLİKLERİ, ÖĞRETME VE ÖĞRENME SÜRECİ İLE İLGİLİ DEĞİŞKENLER VE BİLİŞSEL OLMAYAN ÖĞRENME ÇIKTILARININ PISA 2015 – ULUSLARARASI ÖĞRENCİ DEĞERLENDİRME PROGRAMI FEN OKURYAZARLIĞI ÜZERİNDEKİ ETKİLERİNİN İNCELENMESİ

Orhan, Burcu
Doktora, Matematik ve Fen Bilimleri Eğitimi
Tez Yöneticisi: Prof. Dr. Ömer Geban
Ortak Tez Yöneticisi: Doç. Dr. Eren Ceylan

Eylül 2020, 241 sayfa

Bu çalışmada, Uluslararası Öğrenci Değerlendirme Programı'na (PISA 2015) ait veriler kullanılarak fen ilişkili öğrenci ve okul özelliklerinin Türkiye'deki 15 yaşındaki öğrencilerin fen okuryazarlığına etkisinin araştırılması amaçlanmaktadır. PISA 2015, 187 okuldan 5895 öğrencinin katılımıyla gerçekleştirilmiştir.

Bu çalışmadaki öğrenci düzeyindeki değişkenler; okul dışında fen çalışmaya ayrılan zaman, fen aktiviteleri, fen özyeterliliği, genel fen konularına olan ilgi, fenden keyif alma, araçsak güdülenme, epistemolojik inançlar, çevresel farkındalık, çevresel optimizm, öğrencilerin sahip olmayı umdukları mesleki statü ve fen öğrenme zamanıdır. Okul ilişkili değişkenler ise; fen öğretmenlerinin sayısının tüm öğretmenlerin sayısına oranı, tam sertifikalı fen öğretmenlerinin oranı, ISCED 5A düzeyinde ve fen bilimleri eğitimi almış öğretmenlerin oranı, okuldaki fen öğretmenlerinin toplam sayısı, eğitici materyallerin eksikliği, eğitici personel eksikliği, fene özel kaynaklar ve müfredat dışı yaratıcı aktiviteleri içerir. Öğrencilerin birleşik fen puanları, fen bilgisi puanları ve fen yeterliliği puanları

ayrı ayrı değerlendirilmiştir. Veri analizi hiyerarşik lineer modelleme (HLM) kullanılarak yapılmıştır.

Okul dışında fen çalışmaya ayrılan zaman, fen aktiviteleri ve çevresel optimizmin fen okuryazarlığının tüm boyutlarını negatif yönde etkilediği bulunmuştur. Fen öğrenme zamanı, epistemolojik inançlar, çevresel farkındalık ve öğrencilerin sahip olmayı umdukları mesleki statünün ise fen okuryazarlığının tüm boyutlarını pozitif yönde etkilediği tespit edilmiştir. Araçsal güdülenmenin öğrencilerin fen okuryazarlığı üzerinde önemli bir etkisinin olmadığı gösterilmiştir. Fen özyeterliliğinin birleşik fen puanları, içerik bilgisi ve fenomeni bilimsel olarak açıklama alt boyutlarını pozitif yönde etkilediği bulunmuştur. Genel fen konularına olan ilginin bu alt boyutlara ek olarak veri ve kanıtları bilimsel olarak yorumlama alt boyutunu da pozitif yönde etkilediği tespit edilmiştir. Fenden keyif almanın bilimsel araştırmayı değerlendirme ve tasarlama dışındaki tüm fen okuryazarlığı alt boyutlarını pozitif yönde etkilediği gösterilmiştir.

Fen öğretmenlerinin sayısının tüm öğretmenlerin sayısına oranı ve fene özel kaynaklar fen okuryazarlığının tüm alt boyutlarını pozitif yönde etkilerken eğitici materyallerin eksikliği ve okuldaki fen öğretmenlerinin toplam sayısının fen okuryazarlığının tüm boyutlarını negatif yönde etkilediği tespit edilmiştir. Müfredat dışı yaratıcı aktivitelerin birleşik fen puanı, prosedürel ve epistemik bilgi ve bilimsel araştırmayı değerlendirme ve tasarlama alt boyutlarını pozitif yönde etkilediği bulunmuştur. Tam sertifikalı fen öğretmenlerinin oranı, ISCED 5A düzeyinde ve fen bilimleri eğitimi almış öğretmenlerin oranı ve eğitici personel eksikliğinin fen okuryazarlığı üzerinde önemli bir etkisinin olmadığı saptanmıştır.

Anahtar Kelimeler: Uluslararası Uluslararası Öğrenci Değerlendirme Programı (PISA 2015), Fen Okuryazarlığı, Hiyerarşik Lineer Modelleme (HLM)



To My Parents

ACKNOWLEDGMENTS

I would like to express my deepest gratitude to my supervisor, Prof. Dr. Ömer GEBAN, co-supervisor Assoc. Prof. Dr. Eren CEYLAN and Dr. Savaş PAMUK for their guidance, advice, criticism, encouragements and insight throughout the research.

I would like to thank to TÜBİTAK for financial support throughout the PhD with BİDEP-2211 scholarship provided to PhD students in Turkey.



TABLE OF CONTENTS

ABSTRACT.....	v
ÖZ.....	vii
ACKNOWLEDGMENTS.....	x
TABLE OF CONTENTS.....	xi
LIST OF TABLES.....	xvii
LIST OF ABBREVIATIONS.....	xxii
LIST OF SYMBOLS.....	xxvi
CHAPTERS	
1 INTRODUCTION.....	1
1.1 The Concept of Literacy in the United Nations Educational, Scientific and Cultural Organisation (UNESCO)	2
1.2 Organisation for Economic Co-operation and Development (OECD) and Programme for International Student Assessment (PISA)	4
1.3 Definitions of the Concepts of Literacy in PISA	5
1.3.1 Scientific Literacy Concept in PISA.....	7
1.4 Implementation of the PISA 2015 Study	7
1.5 Purpose of the Study	8
1.6 Significance of the Study	9
1.7 Definition of Important Terms.....	12
2 LITERATURE REVIEW.....	15
2.1 Scientific Literacy and Scientific Literacy Concept in PISA	15
2.1.1 Scientific Literacy Assessment Framework for PISA 2015	16

2.1.1.1	Contexts of Assessment Items	17
2.1.1.2	Scientific Literacy Competencies	19
2.1.1.2.1	The Levels of Science Competencies	20
2.1.1.3	Scientific Knowledge	22
2.1.1.3.1	Content Knowledge	22
2.1.1.3.2	Procedural Knowledge.....	24
2.1.1.3.3	Epistemic Knowledge	24
2.1.1.4	Affective Characteristics of Scientific Literacy Performance ...	25
2.1.2	Assessing Scientific Literacy	28
2.1.2.1	Cognitive Demand	29
2.1.2.2	Test Characteristics	30
2.1.2.3	Response Formats	30
2.1.2.4	Assessment Structure	31
2.2	The Variables in Context Questionnaires in PISA 2015	31
2.2.1	Student Background Variables	38
2.2.2	Non-Cognitive Outcomes	39
2.2.3	Teaching and Learning Variables.....	40
2.2.4	School Policies Variables	42
2.3	Turkish Educational System and Application of PISA 2015 in Turkey .	43
2.4	Studies about Factors Affecting Students' Scientific Performances	45
2.5	Theoretical Background of Hierarchical Linear Modeling and Related Modeling Studies using Hierarchical Linear Modeling	51
3	METHOD.....	57
3.1	Population and Sample	57

3.1.1	Sampling Procedure	58
3.2	Instruments.....	60
3.2.1	Cognitive Level Tests	61
3.2.2	Context Questionnaires.....	62
3.3	Validity and Reliability.....	64
3.3.1	Content-Related Evidence for Validity.....	65
3.3.1.1	Design of Cognitive Level Test and Test Development.....	65
3.3.1.2	Development of Context Questionnaires.....	67
3.3.2	Construct-Related Evidence for Validity.....	68
3.3.3	Reliability.....	69
3.4	Procedures.....	70
3.4.1	Research Questions	71
3.4.2	Threats to Internal Validity	72
3.5	Data Collection	73
3.6	Variables of the Study.....	73
3.6.1	Student-Level Variables.....	75
3.6.2	School-Level Variables.....	75
3.7	Data Analysis	75
3.7.1	Hierarchical Linear Modeling.....	78
3.7.1.1	One-Way ANOVA Model with Random Effects.....	78
3.7.1.2	Means-as-Outcomes Regression Model	80
3.7.1.3	Random-Coefficients Regression Model.....	81
3.7.1.4	Intercepts- and Slopes-as-Outcomes Model	82
3.7.2	Centering the Variables.....	82

3.7.3	Random and Fixed Effects	83
3.7.4	Outcome Variable.....	84
3.7.5	Scaling PISA 2015 Data	85
3.7.6	Hypotheses Testing	85
3.7.7	Weighting of Data	86
3.7.8	Assumptions of Two-Level Hierarchical Linear Model	87
3.8	Missing Data.....	88
3.9	Limitations of the Study	90
4	RESULTS.....	93
4.1	Preliminary Data Analyses	93
4.2	Hierarchical Linear Modeling (HLM) Analyses	99
4.2.1	One-Way Analyses of Variance with Random Intercept Model	99
4.2.1.1	Combined Science.....	100
4.2.1.2	Content Knowledge.....	102
4.2.1.3	Procedural and Epistemic Knowledge	103
4.2.1.4	Explaining Phenomena Scientifically	105
4.2.1.5	Evaluating and Designing Scientific Inquiry	107
4.2.1.6	Interpreting Data and Evidence Scientifically	109
4.2.1.7	Findings from One-Way ANOVA Model.....	111
4.2.2	Means-as-Outcome Regression Model.....	112
4.2.2.1	Combined Science.....	112
4.2.2.2	Content Knowledge.....	115
4.2.2.3	Procedural and Epistemic Knowledge	117
4.2.2.4	Explaining Phenomena Scientifically	120

4.2.2.5	Evaluating and Designing Scientific Inquiry.....	122
4.2.2.6	Interpreting Data and Evidence Scientifically	125
4.2.2.7	Findings from Means-as-Outcome Regression Model	127
4.2.3	Random-Coefficients Regression Model	128
4.2.3.1	Combined Science	129
4.2.3.2	Content Knowledge	132
4.2.3.3	Procedural and Epistemic Knowledge.....	136
4.2.3.4	Explaining Phenomena Scientifically.....	139
4.2.3.5	Evaluating and Designing Scientific Inquiry.....	143
4.2.3.6	Interpreting Data and Evidence Scientifically	146
4.2.3.7	Findings from Random-Coefficients Regression Model	149
4.2.4	Intercepts- and Slopes-as-Outcomes Model	150
4.2.4.1	Combined Science	150
4.2.4.2	Content Knowledge	153
4.2.4.3	Procedural and Epistemic Knowledge.....	156
4.2.4.4	Explaining Phenomena Scientifically.....	159
4.2.4.5	Evaluating and Designing Scientific Inquiry.....	163
4.2.4.6	Interpreting Data and Evidence Scientifically	165
4.2.4.7	Findings from Intercepts- and Slopes-as-Outcomes Model	169
5	DISCUSSION, CONCLUSION, IMPLICATION, AND RECOMMENDATION FOR FURTHER RESEARCH.....	171
5.1	Discussion of the Results	171
5.1.1	School-Level Factors	172
5.1.2	Student-Level Factors	178

5.2	Conclusions	186
5.3	Implications for Policy	191
5.4	Recommendations for Further Research	192
REFERENCES.....		195
APPENDICES		
A.	Questionnaire Items Related to School-Level Variables of the Study	209
B.	Questionnaire Items Related to Student-Level Variables of the Study	215
C.	Technical Characteristic of Student-Level Variables.....	225
D.	Technical Characteristics of School-Level Variables	231
E.	Descriptive Statistics of Outcome Variables	233
CURRICULUM VITAE.....		239

LIST OF TABLES

TABLES

Table 2.1 Contexts of Assessment Items in PISA 2015	18
Table 2.2 PISA 2015 Scientific Competencies	19
Table 2.3 Desired Distribution of Items by Competency	20
Table 2.4 Knowledge of the Content of Science	23
Table 2.5 Distribution of Items by Content of Science.....	24
Table 2.6 Desired Distribution of Items for Knowledge	25
Table 2.7 Measures of Science-Related Non-Cognitive Outcomes Included in the PISA 2015 Main Survey	26
Table 2.8 Motivation in Science and Career Expectations, Science Self-Confidence and Science Learning	28
Table 2.9 Distribution of Items by Depth of Knowledge	30
Table 2.10 Measures Included in the Core Context Assessment for PISA 2015....	33
Table 2.11 Measures of Non-Cognitive Outcomes Included in the PISA 2015 Main Survey	34
Table 2.12 Overview of Nineteen Policy Modules and Their Relevance to Surveys	35
Table 2.13 Modular Structure of the PISA 2015 Context Assessment Design	37
Table 2.14 Scientific Literacy Average Scores by Years	44
Table 3.1 Distribution of the Number of Students in PISA 2015 Turkey Sample .	59
Table 3.2 Item Counts (Field Trial and Main Survey) by Domain and Delivery Mode	67
Table 3.3 Science Item Counts by Framework Category	67
Table 3.4 Reliabilites of the PISA Cognitive Domains and Science Subscales Overall Countries	70
Table 3.5 Reliabilities for Main Cognitive Domains for Turkey.....	70
Table 3.6 Stundet-Level Variables of the Study	76
Table 3.7 School-Level Variables of the Study	77

Table 4.1 Frequency Distribution of Level-1 Variables in the Present Study.....	93
Table 4.2 Expectation Maximization (EM) Estimated Statistics for Student-Level Variables.....	94
Table 4.3 Descriptive Statistics of Level-1 Variabels	95
Table 4.4 Frequency Distribution of Level-2 Variables in the Present Study.....	97
Table 4.5 Expectation Maximization (EM) Estimated Statistics for School-Level Variables.....	97
Table 4.6 Descriptive Statistics of Level-2 Variables	98
Table 4.7 Final Estimation of Fixed Effects of Combined Scientific Literacy Skills of Turkish Students.....	100
Table 4.8 Final Estimation of Fixed Effects of Combined Scientific Literacy Skills of Turkish Students with Robust Errors	101
Table 4.9 Final Estimation of Variance Components of Combined Scientific Literacy Skills of Turkish Students	101
Table 4.10 Final Estimation of Fixed Effects of Content Knowledge Skills of Turkish Students.....	102
Table 4.11 Final Estimation of Fixed Effects of Content Knowledge Skills of Turkish Students with Robust Errors	102
Table 4.12 Final Estimation of Variance Components of Content Knowledge Skills of Turkish Students.....	103
Table 4.13 Final Estimation of Fixed Effects of Procedural and Epistemic Knowledge Skills of Turkish Students	104
Table 4.14 Final Estimation of Fixed Effects of Procedural and Epistemic Knowledge Skills of Turkish Students with Robust Errors.....	104
Table 4.15 Final Estimation of Variance Components of Procedural and Epistemic Knowledge Skills of Turkish.....	105
Table 4.16 Final Estimation of Fixed Effects of Explaining Phenomena Scientifically Skills of Turkish Students	106
Table 4.17 Final Estimation of Fixed Effects of Explaining Phenomena Scientifically Skills of Turkish Students with Robust Errors.....	106

Table 4.18 Final Estimation of Variance Components of Explaining Phenomena Scientifically Skills of Turkish Students.....	107
Table 4.19 Final Estimation of Fixed Effects of Evaluating and Designing Scientific Inquiry Skills of Turkish Students.....	108
Table 4.20 Final Estimation of Fixed Effects of Evaluating and Designing Scientific Inquiry Skills of Turkish Students with Robust Errors	108
Table 4.21 Final Estimation of Variance Components of Evaluating and Designing Scientific Inquiry of Turkish Students	109
Table 4.22 Final Estimation of Fixed Effects of Interpreting Data and Evidence Scientifically Skills of Turkish Students.....	110
Table 4.23 Final Estimation of Fixed Effects of Interpreting Data and Evidence Scientifically Skills of Turkish Students with Robust Errors	110
Table 4.24 Final Estimation of Variance Components of Interpreting Data and Evidence Scientifically Skills of Turkish Students.....	111
Table 4.25 Final Estimation of Fixed Effects for Combined Science	113
Table 4.26 Final Estimation of Variance Components for Combined Science	114
Table 4.27 Final Estimation of Fixed Effects for Content Knowledge	115
Table 4.28 Final Estimation of Variance Components for Content Knowledge ..	117
Table 4.29 Final Estimation of Fixed Effects for Procedural and Epistemic Knowledge	118
Table 4.30 Final Estimation of Variance Components for Procedural and Epistemic Knowledge	119
Table 4.31 Final Estimation of Fixed Effects for Explaining Phenomena Scientifically	120
Table 4.32 Final Estimation of Variance Components for Explaining Phenomena Scientifically	122
Table 4.33 Final Estimation of Fixed Effects for Evaluating and Designing Scientific Inquiry	123
Table 4.34 Final Estimation of Variance Components for Evaluating and Designing Scientific Inquiry	124

Table 4.35 Final Estimation of Fixed Effects for Interpreting Data and Evidence Scientifically	125
Table 4.36 Final Estimation of Variance Components for Interpreting Data and Evidence Scientifically	126
Table 4.37 Final Estimation of Fixed Effects for Combined Science	131
Table 4.38 Final Estimation of Variance Components for Combined Science.....	131
Table 4.39 Final Estimation of Fixed Effects for Content Knowledge	134
Table 4.40 Final Estimation of Variance Components for Content Knowledge...	135
Table 4.41 Final Estimation of Fixed Effects for Procedural and Epistemic Knowledge.....	137
Table 4.42 Final Estimation of Variance Components for Procedural and Epistemic Knowledge.....	138
Table 4.43 Final Estimation of Fixed Effects for Explaining Phenomena Scientifically	141
Table 4.44 Final Estimation of Variance Components for Explaining Phenomena Scientifically	142
Table 4.45 Final Estimation of Fixed Effects for Evaluating and Designing Scientific Inquiry	144
Table 4.46 Final Estimation of Variance Components for Evaluating and Designing Scientific Inquiry	145
Table 4.47 Final Estimation of Fixed Effects for Interpreting Data and Evidence Scientifically	148
Table 4.48 Final Estimation of Variance Components for Interpreting Data and Evidence Scientifically	149
Table 4.49 Estimation of Fixed Effects for Combined Science	152
Table 4.50 Final Estimation of Variance Components for Combined Science.....	153
Table 4.51 Estimation of Fixed Effects for Content Knowledge	155
Table 4.52 Final Estimation of Variance Components for Content Knowledge...	156
Table 4.53 Final Estimation of Fixed Effects for Procedural and Epistemic Knowledge.....	158

Table 4.54 Final Estimation of Variance Components for Procedural and Epistemic Knowledge	159
Table 4.55 Final Estimation of Fixed Effects for Explaining Phenomena Scientifically	161
Table 4.56 Final Estimation of Variance Components for Explaining Phenomena Scientifically	162
Table 4.57 Final Estimation of Fixed Effects for Evaluating and Designing Scientific Inquiry	164
Table 4.58 Final Estimation of Variance Components for Evaluating and Designing Scientific Inquiry	165
Table 4.59 Final Estimation of Fixed Effects for Interpreting Data and Evidence Scientifically	167
Table 4.60 Final Estimation of Variance Components for Interpreting Data and Evidence Scientifically	168
Table 5.1 Fixed Effects of Means-as-Outcomes Model for School-Level Variables	173
Table 5.2 Fixed Effects of Means-as-Outcomes Model for Teaching and Learning Variables	176
Table 5.3 Fixed Effects of Means-as-Outcomes Model for School Policies Variables	177
Table 5.4 Fixed Effects of Random-Coefficients Regression Model for Student-Level Variables	179
Table 5.5 Fixed Effects of Random-Coefficients Regression Model for Student Background Variables.....	181
Table 5.6 Fixed Effects of Random-Coefficients Regression Model for Non-Cognitive Outcomes.....	182
Table 5.7 Fixed Effects of Random-Coefficients Regression Model for Teaching and Learning Variables	185

LIST OF ABBREVIATIONS

ABBREVIATIONS

ANCOVA	Analysis of Covariance
ANOVA	Analysis of Variance
ASST	Attitude Scale for Science and Technology
BSMJ	Students' Expected Occupational Status
CREACTIV	Creative Extra-Curricular Activities
EC	Educational Career Questionnaire
EDUSHORT	Shortage of Educational Material
EGs	Subject Matter Expert Groups
EM	Expectation Maximization
ENVAWARE	Environmental Awareness
ENVOPT	Environmental Optimism
EPIST	Epistemological Beliefs
ESCS	Economic, Social and Cultural Status
HLM	Hierarchical Linear Modeling
ICC	Intraclass Correlation Coefficient
ICT	Information and Communications Technology
INTBRSCI	Interest in Broad Science Topics
INSTSCIE	Instrumental Motivation
IRT	Item Response Theory
ISCED	International Standard Classification of Education

ISEI	International Socio-Economic Index of Occupational Status
JOYSCIE	Enjoyment of Science
LSM	Linear Structural Modeling
MAR	Missing at Random
MCAR	Missing Completely at Random
MNAR	Missing not at Random
MVA	Missing Value Analysis
MDM	Multivariate Data Matrix
MoNE	Ministry of National Education
NPM	National Project Manager
OECD	Organisation for Economic Co-operation and Development
OUTHOURS	Out-of-School Study Time per Week
ÖSYM	Measurement Selection and Placement Center
PAQ	Parent Questionnaire
PGB	PISA Governing Board
PIRLS	Progress in International Reading Literacy Study
PISA	Programme for International Student Assessment
PROSTAT	Proportion of Science Teachers by All Teachers
PROSTCE	Proportion of Science Teachers Fully Certified
PROSTMAS	Proportion of Science Teachers With ISCED Level 5A and A Major in Science
PVSCED	Plausible Value for Evaluating and Designing Scientific Inquiry

PVSCEP	Plausible Value for Explaining Phenomena Scientifically
PVSCID	Plausible Value for Interpreting Data and Evidence Scientifically
PVSCIE	Plausible Value for Combined Science
PVSKCO	Plausible Value for Content Knowledge
PVSKPE	Plausible Value for Procedural and Epistemic Knowledge
QEG	Questionnaire Expert Group
SCIEACT	Science Activities
SCIEEFF	Science Self-Efficacy
SCIERES	Science Specific Resources
SCQ	School Questionnaire
SDG	Sustainable Development Goal
SEI	Students' Expected Occupational Status
SEM	Structural Equation Modeling
SES	Socio-Economic Status
SMINS	Science Learning Time
SMTSL	Students' Motivation toward Science Learning
SPSS	Statistics Package for Social Sciences
STAFFSHORT	Shortage of Educational Staff
STQ	Student Questionnaire
TALIS	Teaching and Learning International Survey
TCQ	Teacher Questionnaire
TIMSS	Trends in International Mathematics and Science Study

TIMSS-R	Third International Mathematics and Science Study-Repeat
TOTST	Total Number of Science Teachers At School
UIS	Institute for Statistics
UNEP	United Nations Environment Programme
UNESCO	Organisation for Economic Co-operation and Development
WLE	Weighted Likelihood Estimate
YKS	Higher Education Institutions Exam



LIST OF SYMBOLS

SYMBOLS

β_0	Expected outcome for observation i , or the intercept
β_1	Expected change in Y_i associated with a unit increase in X_i
u_{0j}	Random effect associated with unit j ; unique increment to the intercept associated with level-2 unit j
u_{1j}	Increment to the slope associated with level-2 unit j
γ_{00}	Average intercept across the level-2 units; grand-mean outcome in the population
γ_{10}	Average regression slope across the level-2 units
Y_i	Outcome variable or dependent variable
X_i	Error term representing the unique effect associated with observation i
σ^2	Within-group variability
τ_{00}	Between-group variability
$Var(u_{0j})$	Unconditional variance in the level-1 intercepts
$Var(u_{1j})$	Unconditional variance in the level-1 slopes
$Cov(u_{0j}, u_{1j})$	Unconditional covariance between the level-1 intercepts and slopes
p	Probability value, level of statistical significance
H_o	Null hypothesis

CHAPTER 1

INTRODUCTION

In an evolving and changing world, education is shaped by the purpose of teaching information to individuals, transferring it to life and adapting to new situations. This can be observed in educational programs, in instructional techniques and methods, and in changes in the assessment tools at the evaluation stage.

Programme for International Student Assessment (PISA), funded by the Organization for Economic Co-operation and Development (OECD), is a research project which measures and evaluates this function of education.

The PISA survey, which has been conducted every three years since 2000, aims to determine the extent to which students in compulsory education in OECD countries and in other participating countries (approximately 90% of the world economy) have the basic knowledge and skills necessary to take their place in the modern society. The target population of the PISA survey is 15-year-old students enrolled in structured education in grades 7 and higher. By PISA, this age is considered to be the smallest age for students to become an adult citizen after schooling, since they are about to complete the compulsory education period in many countries.

Globally, the knowledge and abilities of students in the countries involved in the international studies are compared in order to set standards for improving the level of education by considering the average scores obtained by countries. In this respect, PISA results help to identify the strong and inadequate features of education systems of countries.

About 540.000 students, which represent approximately 29 million students from 72 countries and economies, participated in the PISA 2015 survey. Thirty five of these countries are OECD members.

In Turkey, PISA 2015 was applied as a computer-based assessment with the participation of 187 schools and 5895 students from 61 provinces between 2-22 April

2015 (MoNE, 2016). The results of the PISA 2015 survey, which is the sixth cycle of PISA, were announced by the OECD on December 6, 2016. Science, mathematics, reading, collaborative problem solving and financial literacy are tested in PISA 2015. In PISA, one of the areas of mathematics, science and reading is determined as the main focus every three years. In PISA 2015, scientific literacy was the main focus (OECD, 2017a).

The aim of this study is to explore the student- and school-related factors and their effects on scientific literacy scores of 15-year-old Turkish students by conducting hierarchical linear modeling (HLM) analyses with the data from PISA 2015.

1.1 The Concept of Literacy in the United Nations Educational, Scientific and Cultural Organisation (UNESCO)

Literacy is traditionally known as the ability to read, write and use arithmetic. The contemporary meaning of the term includes the skills about using language, numbers, images, technology and other tools in order to communicate, understand, obtain and acquire useful information (Kirk, 2018).

The United Nations Educational, Scientific and Cultural Organization (UNESCO) plays an important role in building a more peaceful and secure world by supporting co-operation among countries and nations through education, science, culture and communication without considering any race, sex, language or religion differences. For this purpose, UNESCO conducts studies on education, science, culture and communication to transfer and share the knowledge through research, training and teaching activities and to exchange specialized information.

The UNESCO Institute for Statistics (UIS) is the statistical department of UNESCO. The UIS aims to deliver accurate and policy-relevant statistics about education, science, technology and innovation, culture and communication. Every two years, the UIS conducts surveys on literacy, which include questionnaires on educational programmes, educational attainment, students, financial and human resources. The results from different countries can be compared through these questionnaires.

UNESCO has “Sustainable Development Goals (SDGs)”. The main goal of SDG 4 is to provide inclusive and well-qualified education and support opportunities for

lifelong learning for everyone until 2030. The results from the questionnaires conducted by UIS provide information for many SDG 4 indicators. Proficiency in literacy and numeracy are measured as the main subjects of SDG 4 (UNESCO, 2108).

In 1987, UNESCO has defined three levels of literacy. The first level has been defined as basic literacy, the second level as functional literacy, and the third level as multi-functional literacy. The first level, basic literacy is to have the basic literacy skills such as vocalization and understanding sentences. Functional literacy is the ability of the person to use his reading, writing and mathematical knowledge and skills in individual, social and cultural fields. Multi-functional literacy aims at improving the individual's level of understanding and it includes the effort of the individual to contribute to the progress of society by reading and writing. The literate individual has the features such as self-realization, strengthening his creativity, understanding complex issues and having a comprehensive world view (Güneş, 1997).

Initially, the literacy has been simply defined as the skills needed for reading and writing simple statements, and skills needed for some basic arithmetic processes (numeracy). However, UNESCO has improved the definition of literacy as an “ability to identify, understand, interpret, create, communicate and compute, using printed and written materials associated with varying contexts. Literacy involves a continuum of learning in enabling individuals to achieve their goals, to develop their knowledge and potential, and to participate fully in their community and wider society” (UNESCO, 2004, p. 13).

Through PISA and other international studies, UNESCO and Organisation for Economic Co-operation and Development (OECD) aim to contribute to educational inputs and learning outcomes of countries by bringing about improvements in schooling, by preparing young people for adult life and by deepening global interdependence.

1.2 Organisation for Economic Co-operation and Development (OECD) and Programme for International Student Assessment (PISA)

Expenditures and investments made in the field of education are effective on the output of the education and also play an important role in the development of countries. In this context, developed and developing countries are constantly increasing their educational expenditures and investments.

OECD is an international economic organization, which aims to develop and spread policies, that will increase the economic and social welfare of people around the world. OECD follows policies to improve the world economy by supporting sustainable economic growth and employment and by increasing the life standards in the member countries while ensuring financial stability. OECD also aims to contribute to the economic growth of the member countries and the non-member countries in the economic development process and to contribute to the growth of world trade without discrimination and in accordance with international obligations. The OECD makes publications, which include reports, statistics and analyses about the economic growth in the world. The organization analyzes the impact of social policy issues on economic growth and publishes reports and makes policy recommendations to stimulate growth (OECD, 2017a).

Currently, OECD has 36 member countries. The organization has 20 founding members; United States, Austria, Canada, France, the Netherlands, Luxembourg, Germany, Italy, United Kingdom, Belgium, Denmark, Ireland, Greece, Switzerland, Sweden, Spain, Iceland, Norway, Portugal and Turkey. Australia, Chile, Czech Republic, Estonia, Finland, Hungary, Israel, Japan, Korea, Latvia, Lithuania, Mexico, New Zealand, Poland, Slovak Republic and Slovenia have become members of the organization subsequently.

Countries have developed some tools to measure the efficiency of their investments. Each country measures its educational status and educational outcomes according to some evaluation criteria. But these measurements often fail to provide meaningful interpretations. Thus, in particular, the OECD countries are able to achieve more meaningful and realistic implications by making these measures comparatively within themselves. Countries that are members of the OECD have output and

knowledge about very different variables in the field of education. The OECD center obtains this information from all member and partner countries and publishes the outputs comparatively in its reports (About OECD, n.d.).

PISA is an important research project, that provides meaningful data for these reports. PISA research mainly assesses the skills of students in science, mathematics and reading domains. While making this assessment, the main areas are defined through the concept of “literacy”.

The concept of literacy is expressed as the competencies of using, analyzing, making logical conclusions, and communicating effectively while understanding and solving problems that students encounter in different situations in the main subject areas.

Apart from these basic areas, starting from the application of 2012, in each cycle, an innovative field assesses the extent to which students have basic knowledge and abilities. This innovative field has been “creative problem solving” in 2012 and “collaborative problem solving” in 2015.

One of the main areas in each PISA cycle is designated as weighted area. In the PISA 2015 survey, the weighted area was determined as “scientific literacy”. This means that the results of PISA 2015 focus on scientific literacy rather than mathematics literacy and reading literacy skills (OECD, 2017a).

1.3 Definitions of the Concepts of Literacy in PISA

In OECD countries, the literacy concept is extended to include the ability to access technology and information and the ability to assess complex contexts. In the PISA study, literacy is described as the ability of students to use their knowledge and skills in daily life while drawing logical conclusions, to identify and to solve the problems they face in different contexts and to make inferences from what they have learned. Main focus of the PISA 2015 study was science, whereas reading, mathematics and collaborative problem solving were minor assessment fields. The assessment of the financial literacy was dependent on the choices of the participating countries and economies.

Reading literacy is defined as “the capacity of an individual to understand, use, reflect on and engage with written texts in order to achieve one’s goals, to develop one’s knowledge and potential, and to participate in society” (OECD, 2017a). Reading skills focus on students’ skills to apply written knowledge in real life situations. In the PISA 2015 survey, reading skills are described as comprehending, using, thinking about and dealing with written texts to participate in society, to develop their potential and knowledge and to achieve their goals. The concept of the PISA reading skills include understanding a variety of situations that people read, understanding the ways of presentation of written texts (such as printed books, information notes, online forums and news), and understanding how readers use various approaches to comprehend the texts thoroughly. The focus of PISA 2000 and PISA 2009 was reading skills. In the sixth cycle of PISA (PISA 2015), scientific literacy has been identified as the dominant field, so fewer questions (103 questions) were used in the assessment of reading skills and fewer students answered these questions. As a result, findings about general performance levels were included instead of detailed analyzes of the findings of this area (MoNE, 2016).

Mathematical literacy is about measuring the capacity of students to formulate, apply and evaluate mathematics in different situations. Students need to have abilities such as using mathematical concepts, phenomena and procedures while explaining and describing facts. Mathematical literacy helps to understand the role of mathematics while making well-grounded decisions, which are necessary for being a constructive, sensitive and reflective citizen (OECD, 2017a). The PISA study aims to evaluate how effectively students can apply mathematical knowledge in new and unusual contexts. Mathematical literacy has been identified as dominant field in PISA 2003 and PISA 2012. Since scientific literacy was determined to be the dominant field in PISA 2015, fewer questions (83) were used in the assessment of of mathematical literacy, and fewer students answered these questions. As a result, findings about general performance levels were included instead of detailed analyzes of the findings of this area (MoNE, 2016).

1.3.1 Scientific Literacy Concept in PISA

Scientific literacy, one of the fundamental areas of PISA study, is considered as the dominant field in the 2015 implementation. In PISA 2015, scientific literacy is defined as “the ability to engage with science-related ideas and science-related issues as a reflective citizen”. A person who is literate in the field of science is willing to participate in reasoned discourses made in the field of science and technology, which necessitates the abilities such as explaining phenomena scientifically, evaluating and designing scientific enquiry, and interpreting data and evidence scientifically (OECD, 2017a).

The PISA science assessment framework includes following dimensions:

- Competencies (explain phenomena scientifically, evaluate and design scientific enquiry, and interpret data and evidence scientifically);
- Knowledge (content knowledge, procedural knowledge and epistemic knowledge);
- Contexts (personal, local/national and global issues) (OECD, 2017b).

There are also affective factors affecting the scientific literacy skills of the students. Students’ attitudes and tendencies towards the content can influence their level of interest, maintain their participation and motivate them to act.

The scientific literacy term used in PISA assesses what students learn in the field of science as well as what they can do with them and how they can use scientific knowledge in real life situations creatively.

1.4 Implementation of the PISA 2015 Study

PISA 2015, the sixth cycle of PISA, was implemented with approximately 540.000 students, which represent about 29 million students from 72 countries and economies.

PISA 2015 study was the first application of PISA as a computer-based assessment. Countries that do not want to use this method have implemented the paper-pen based assessment. However, for these countries, only the questions used in the previous cycles were used. A pilot study was conducted to investigate whether there was a

significant difference between the two methods used in PISA 2015 application. The studies were conducted with the data obtained from all participating countries and the results did not differ according to the method used.

Fifty seven of the 72 countries, which participated in the PISA 2015 study, performed the application as computer-based assessment and 15 countries applied it as paper-pen based assessment. 66 different booklets for computer-based assessment and 30 different booklets for paper-pen based assessment were generated. The booklets consisted of four chapters and each chapter took 30 minutes to complete. The PISA 2015 survey consisted of two sessions, and 60 minutes was given for each session. There was a 5-10 minute break between sessions. After 15-minute break, which was given after the cognitive test sessions, the students answered the student questionnaire, which took approximately 35 minutes.

In the PISA 2015 survey, the main area of the study was scientific literacy. This means that the PISA 2015 results focus more on science literacy, rather than mathematical literacy and reading literacy skills.

The context questionnaires applied together with the cognitive tests collect data about students' motivations, students' opinions about themselves, psychological characteristics of learning processes, school environments and students' families. These data are used in the interpretation of the data obtained from the cognitive tests. Questionnaires are important components of the PISA study and they provide valuable sources to interpret test results in a more detailed way. In PISA 2015, all countries participated in the student and school questionnaires. Participating in other questionnaires (teacher questionnaire, information and communication technologies questionnaire, educational career questionnaire and parental questionnaire) was dependent on countries' preference (MoNE, 2016).

1.5 Purpose of the Study

The purpose of the present study is to investigate the effect of science-related variables on PISA 2015 combined science, science competencies and science knowledge scores of students by considering some student and school characteristics. The student-level variables are out-of-school study time for science, science

activities, science self-efficacy, interest in broad science topics, enjoyment of science, instrumental motivation, epistemological beliefs, environmental awareness, environmental optimism, students' expected occupational status, and science learning time. The school-level variables are proportion of science teachers by all teachers, proportion of science teachers fully certified, proportion of science teachers with ISCED level 5A and a major in science, total number of science teachers at school, shortage of educational material, shortage of educational staff, science specific resources, and creative extra-curricular activities.

In order to explore the effect of the selected variables on the outcome variable, it was aimed to answer the following research questions by conducting analyses using hierarchical linear modeling. The analyses were repeated for combined science score, and for science competencies and science knowledge subscales.

1. Are there any differences in the scientific literacy skills of Turkish students across schools?
2. What proportions of variances in the scientific literacy scores of Turkish students is explained by school-related variables?
3. Which student characteristics are significantly related to Turkish students' mean combined science and scientific literacy subscales scores?
4. Which school characteristics are significantly related to Turkish students' mean combined science and scientific literacy subscales scores?
5. How do the significant school-level variables effect the relationship between student-level variables and outcome variable?

1.6 Significance of the Study

In the 21st century, people are expected to have the ability to solve complex problems, to communicate effectively in written and oral form, to work in collaboration and to think creatively. Since economic, social, scientific and technological developments have been changing our life styles, countries are aware of the necessity to educate their citizens as technologically and scientifically literate persons. The concept of literacy in PISA is expressed as the scope of interpreting and

solving problems, using knowledge and skills, analyzing, making logical inferences and establishing effective communication (OECD, 2017a).

The extent to which countries have reached their educational goals is highly related to the determination of learning outcomes. In order to improve the achievement of their students, to improve their educational systems, and to provide equal opportunities to students, who differ in their socio-economic and cultural status, countries need to assess their educational outcomes at national or international level.

Large-scale assessments such as PISA, PIRLS and TIMSS provide information on student competencies at the international level. In addition, revealing what successful countries are doing to improve their educational systems, unlike other countries, is another important feature of international studies.

PISA is a very comprehensive study, which provides information about student, school and family characteristics of countries in addition to mathematics, science and reading skills of students. The application area of PISA includes almost the whole world economy, since it is applied in developed and prosperous countries.

Educational policies adopted by the countries have an important place in the education of qualified individuals. Countries need to increase the budget allocated to education, since a country's funding for education is a sign of its level of development. The investment of countries in education affects literacy rate, higher education attendance rate, and thus the level of development of the country.

There are many ways for countries to raise their level of achievement in education to the desired level. The most important of these ways is to increase the budget allocated to education. A country's funding for education means that it invests in the future of the country. Therefore, the budget to be allocated to the education system must be calculated very carefully and systematically.

However, the budget alone is not sufficient to reach the desired level of education systems. Well-developed countries may have problems in their education systems as well. This could be due to student, school and family characteristics of countries. Since PISA provides cross-culturally valid measures, unlike other studies assessing the achievement of particular subjects, the findings of PISA project gives

information about success levels of individual countries and about which factors are related to the success of their students.

Students in the same educational system might have different competency levels. Therefore, it is necessary to explore the relationship between scientific literacy and student and school characteristics in order to compare students having different proficiency levels. PISA 2015 study focuses on the variables that affect the science achievement score. Scientific literacy should be examined to determine the level of learning outcomes in science.

The results from the assessment studies PISA and TIMSS demonstrated that average science and technology scores of Turkish students were lower than that of average countries. This suggested that there is a need to improve students' achievement in science. In order to educate students as scientifically and technologically literate citizens, scientific literacy and the factors affecting it should be examined thoroughly.

This study is important, since it provides information about student- and school-related factors affecting scientific literacy skills. However, there are limited studies, which focus on the subscales of scientific literacy skills. In this study, the subscales of science competencies and science knowledge have been investigated separately in addition to the combined science scores of students.

The results of this study are important for providing opportunities to improve Turkish educational system. Considering the student and school characteristics together in terms of students' scientific literacy performance can provide important information when assessing student achievement.

This study is also important in terms of the analyses carried out. Through hierarchical linear modeling, the effects of significant school-level variables on the relationship between student-level variables and scientific literacy skills can be investigated. In the analyses, both student- and school-related variables were used to assess students' scientific literacy skills. Hence, more precise findings were obtained by hierarchical linear modeling. In addition, the findings obtained from the results of the analyses can be used to contribute to the body of literature and to improve science and technology learning in Turkey.

1.7 Definition of Important Terms

Combined science score: Science achievement score in PISA 2015.

Fixed effects: The only levels of a variable, in which an experimenter is interested in studying, and these variables are constant across individuals.

Grand-mean centered variable: Independent variables, which are centered around a ground mean by subtracting each participant's value on the independent variable from the mean of that variable across the mean of all other participants in the study.

Group-mean centered variable: Independent variables, which are centered around a mean by subtracting the individual's group mean from the individual's score.

Intercepts- and slopes-as-outcomes model: The model of the examination of randomly varying slopes, and modeling the variability in the intercepts and slopes among schools.

Intraclass correlation coefficient: The measure of how strongly units in the same group resemble each other, the measure of the proportion of variance in the outcome that is between groups.

Means-as-outcomes regression model: The model about which level-2 factors effect the differences in the outcome variable.

One-way ANOVA model random-effects: The model of the control of the variation between groups and applied to see how much of the variance is within the groups and how much of it is between the groups.

Plausible value: Intermediate values, which are used for obtaining consistent estimates of population parameters.

Random-coefficients regression model: The model for the effect of variables in the level-1 model on the outcome variables and performed to see the effect of level-1 variables within intra-group variability.

Random effects: The effects that are a subset of the total possible levels of a variable where the experimenter is interested in generalizing the two levels not observed.

Science competencies: The competencies of identifying scientific issue, explaining scientific phenomena and using scientific evidence.

Science knowledge: The knowledge of content, procedural knowledge and epistemic knowledge.





CHAPTER 2

LITERATURE REVIEW

This chapter covers the concept of scientific literacy framework in PISA, the aspects of Turkish educational system and the application of the PISA 2015 study in Turkey. In addition, related literature on factors affecting students' science performance were summarized. Finally, studies related to hierarchical linear modeling were summarized.

2.1 Scientific Literacy and Scientific Literacy Concept in PISA

In our contemporary world, where science and technology is continuously progressing, literacy can be described as understanding and using the complex technology. Being literate in the field of science brings the ability to participate in the discourses made in the field of science and technology within a certain logic, to explain the events scientifically, to establish a scientific inquiry method and to evaluate the data and findings scientifically.

The concept of scientific literacy has gained different dimensions over time. Shen (1975) described three dimensions of science literacy as solving current problems with science (practical), assimilation of science by society (civilian) and science as the main activity of humanity (cultural). According to Durant (1993), science literacy means the combination of cognitive, affective and psychomotor skills related to science. According to Shamos (1995), the meaning of science literacy consists of three components: cultural, functional and true science. According to Köseoğlu et al. (2003), a scientifically literate individual comprehends the nature of science and scientific activities; perceives and uses basic science concepts, principles and theories; seeks solutions to problems within the framework of scientific processes and therefore comprehends the meaning of science-technology and science-

environment relationship, interprets their interactions with the society; and has interests that improve the quality of life.

Humanity encounters challenges in finding adequate food, water and energy sources, fighting with diseases and adapting to changes in the climate (UNEP, 2012). Hence, understanding of science and technology helps a young person to prepare for life (OECD, 2017a).

Science literacy, one of the core areas of PISA, has been considered as the predominant field in the PISA 2015 study. In PISA 2015, scientific literacy is described as the ability to apply the knowledge and scientific processes in daily life rather than having knowledge about science-related subjects.

The science literacy term used in PISA evaluates what students know about science, as well as what they can do with them, and how they can use their knowledge in real life situations. A literate person in the field of science is willing to participate in the discourses made in the field of science and technology within a certain logic. This requires the ability to explain the facts scientifically, to design and evaluate a scientific inquiry method, and to interpret the data and findings scientifically (MoNE, 2016).

The scientific literacy concept in the PISA 2015 framework is explained as the knowledge of science and science-related technology. The scientific literacy includes the knowledge of concepts of science, theories of science and scientific inquiry related procedures and practices. The main objective of the PISA 2015 framework and associated curricula is to educate all young people as citizens with critical thinking skills (OECD, 2017a).

There are also affective factors affecting students' competences in science literacy. Students' attitudes and tendencies towards science can affect their interest levels, continue their participation and motivate them for taking action (MoNE, 2016).

2.1.1 Scientific Literacy Assessment Framework for PISA 2015

Context, knowledge, competency and attitude are the interrelated aspects of the PISA 2015 framework.

Science contexts include personal, local/national and global topics, which could be about current and historical issues.

Scientific knowledge is defined as understanding of the science-related facts, procedures and theories, and it includes content knowledge (knowledge of natural world and technological objects), procedural knowledge (the process of production of ideas) and epistemic knowledge (comprehension of the logic behind the procedures and reasoning for their use).

Scientific competencies are classified as explaining phenomena scientifically, evaluating and designing scientific enquiry, and interpreting data and evidence scientifically.

Scientific attitudes are about students' interest in science and technology, appreciation of the value of scientific methods in the researches, and awareness about environment-related issues.

2.1.1.1 Contexts of Assessment Items

In PISA 2015, scientific knowledge is estimated not only within the framework of the common competencies aspects of the national curricula of participating countries, the assessment also requires successful use of the three scientific literacy competencies in personal, local/national and global settings.

Personal context refers to the situations about self, family and peer groups; local/national context is related to the community; and global context refers to life around the world. Some topics can also be interpreted under historical and technological contexts.

In the PISA 2015 science assessment framework, competencies and knowledge are assessed in specific contexts. These contexts are determined by considering interests and lives of students and they involve the knowledge and understanding that may be acquired by students by the age of 15.

Being scientifically literate plays an important role for the application areas of the assessment items in the science assessment framework (such as health and disease, natural resources, environmental quality, etc.). So, scientific literacy has significant

importance for individuals and countries in sustaining development and quality of life, and in developing public policy. The items in the PISA study are developed according to the cultural and linguistic differences among participating countries (OECD, 2017a). Table 2.1 shows the contexts of assessment items in PISA 2015.

Table 2.1 Contexts of Assessment Items in PISA 2015

	Personal	Local/National	Global
Health and disease	Maintenance of health, accidents, nutrition	Control of disease, social transmission, food choices, community health	Epidemics, spread of infectious diseases
Natural resources	Personal consumption of materials and energy	Maintenance of human populations, quality of life, security, production and distribution of food, energy supply	Renewable and non-renewable natural systems, population growth, sustainable use of species
Environmental quality	Environmentally friendly actions, use and disposal of materials and devices	Population distribution, disposal of waste, environmental impact	Biodiversity, ecological sustainability, control of pollution, production and loss of soil/biomass
Hazards	Risk assessments of lifestyle choices	Rapid changes (e.g. earthquakes, severe weather), slow and progressive changes (e.g. coastal erosion, sedimentation), risk assessment	Climate change, impact of modern communication
Frontiers of science and technology	Scientific aspects of hobbies, personal technology, music and sporting activities	New materials, devices and processes, genetic modifications, health technology, transport	Extinction of species, exploration of space, origin and structure of the universe

Adapted from *PISA 2015: Assessment and Analytical Framework*

2.1.1.2 Scientific Literacy Competencies

In order to understand and engage in science and technology related discussions, three domain-specific competencies are required. Three qualifications for scientific literacy is defined in PISA. Table 2.2 shows the science competencies and the skills expected from the students belonging to these competencies.

Table 2.2 PISA 2015 Scientific Competencies

Scientific Literacy Competencies	Examples for Competencies
Explain phenomena scientifically Recognise, offer and evaluate explanations for a range of natural and technological phenomena demonstrating the ability to:	• Recall and apply appropriate scientific knowledge. • Identify, use and generate explanatory models and representations. • Make appropriate estimates and justify them. • Propose explanatory hypotheses. • Explain the potential implications of scientific knowledge for society.
Evaluate and design scientific enquiry Describe and appraise scientific investigations and propose ways of addressing questions scientifically demonstrating the ability to:	• Identify the question explored in a given scientific study. • Distinguish questions that could be investigated scientifically. • Propose a way of investigating a given question scientifically. • Evaluate ways of exploring a given question scientifically. • Describe and evaluate how scientists ensure the reliability of data, and the objectivity and generalisability of explanations.
Interpret data and evidence scientifically Analyse and evaluate scientific data, claims and arguments in a variety of representations and draw appropriate conclusions, demonstrating the ability to:	• Transform data from one representation to another. • Analyse and interpret data and draw appropriate conclusions. • Identify the assumptions, evidence and reasoning in science-related texts. • Distinguish between arguments that are based on scientific evidence and theory and those based on other considerations. • Evaluate scientific arguments and evidence from different sources (e.g. newspapers, the Internet, journals).

Adapted from *PISA 2015: Assessment and Analytical Framework*

Table 2.3 shows the desired distribution of items by competency in the PISA 2015 science assessment.

Table 2.3 Desired Distribution of Items by Competency

Competency	Percentage of Total Items
Explain phenomena scientifically	40-50
Evaluate and design scientific enquiry	20-30
Interpret data and evidence scientifically	30-40

Adapted from *PISA 2015: Assessment and Analytical Framework*

2.1.1.2.1 The Levels of Science Competencies

In the PISA 2015 study, seven proficiency levels are described for scientific literacy.

Level 6: Students at this level are able to use the content, procedural and epistemic knowledge to present explanatory hypotheses about unusual scientific facts, situations and processes. Students can understand various ideas and concepts in physics, life, space and earth sciences. When interpreting information and findings, the students can distinguish relevant or irrelevant information and obtain extra-curricular information. They can also distinguish scientific and legal information from information based on other opinions. Students can evaluate complex experimental setups, field studies and simulations that can be used in place of each other, and they can justify their selection.

Level 5: Students at this level are able to use abstract scientific opinions to describe uncommon and complicated facts, situations and processes which have versatile causality connections. In order to interpret different experimental designs and to validate their decisions, students may use more complex epistemic knowledge and theoretical knowledge to make predictions or to interpret information. They can evaluate the ways in which a particular question is scientifically investigated. Students can identify the limitations and ambiguity in datasets while evaluating scientific data.

Level 4: Students at this level are able to use or remember more complicated or abstract content knowledge in order to explain not well-known facts, situations and processes. Students can implement experiments with two or more independent

variables in a bounded situation and verify an experimental design by using epistemic and procedural knowledge. They can evaluate the complicated data from a less well-known context. They can make appropriate conclusions and justify their choices.

Level 3: Students at this level are able to use medium-level content knowledge to create descriptions of known facts. They can generate explanations with relevant hints or support in less familiar or more complex situations. They can use the elements of epistemic knowledge or procedural knowledge to apply a basic experiment in a limited context. Students may recognise scientific and non-scientific problems and may identify a finding that supports a scientific expression.

Level 2: Students at this level are able to use basic content and procedural knowledge to make scientific explanations, to evaluate data, and to identify the questions in a basic experimental design. Students can use basic scientific knowledge to define an accurate result in a basic data set. Students can identify basic epistemic knowledge and they can determine questions which can be explored scientifically.

Level 1a: Students at this level are able to use simple content knowledge to distinguish and identify simple scientific phenomena. They can make structured scientific inquiries that do not have more than two variables. The students can determine basic causal and relational connections and they can evaluate graphics and other images which do not require high level cognitive input. They can choose the most appropriate description for data that is known in local and personal contexts.

Level 1b: Students at this level are able to use basic knowledge to distinguish characteristics of a known or basic phenomenon. Students can recognise simple patterns in the data. They can distinguish simple scientific terms, and follow instructions for scientific processes.

In PISA, it is expected that 15-year-old students have basic competencies, which are described at the qualification level 2. The qualification levels below level 2 are defined as sub-qualification levels, while level 5 and level 6 are defined as higher qualification levels (MoNE, 2016).

2.1.1.3 Scientific Knowledge

The knowledge of the facts, concepts, ideas and theories about the natural world that science has established is defined as “content knowledge”. The knowledge of different types of methods and practices that scientists use to identify and recognise the characteristics of scientific knowledge is referred to as “procedural knowledge”. Comprehending the meaning of particular constructs such as hypothesis and theory, describing characteristics which are necessary for building scientific knowledge and understanding the rationale for common activities of scientific enquiry is defined as “epistemic knowledge” (Duschl, 2007).

These three scientific literacy competencies mentioned above require knowledge. Explaining scientific and technological phenomena requires content knowledge. However, evaluating and designing scientific enquiry and interpreting data and evidence scientifically require more than the knowledge of what is known. These competencies also require understanding the way of the establishment of scientific knowledge and how confident it is. The students need both procedural and epistemic knowledge to identify questions that are related to scientific enquiry, to decide whether appropriate procedures are used to ensure decision making, and to distinguish scientific issues from value judgments or economic considerations.

The main focus of PISA 2015 framework is to determine to what extent 15-year-old students are able to display these scientific literacy competencies properly within personal, local/national and global contexts.

2.1.1.3.1 Content Knowledge

The content knowledge questioned in PISA 2015 can be explained as knowledge of the physics, chemistry, biology, earth and space related topics. Items questioning content knowledge are chosen to be related to the real-life conditions, to represent a major science-related concept and to be suitable for the cognitive level of 15-year-old students. Content knowledge is needed to comprehend the natural world and experiences in different contexts. Table 2.4 demonstrates the content knowledge categories with selected examples.

Table 2.4 Knowledge of the Content of Science

Content Knowledge Categories	Examples for Content Knowledge
Physical systems that require knowledge of:	• Structure of matter (e.g. particle model, bonds) • Properties of matter (e.g. changes of state, thermal and electrical conductivity) • Chemical changes of matter (e.g. chemical reactions, energy transfer, acids/bases) • Motion and forces (e.g. velocity, friction) and action at a distance (e.g. magnetic, gravitational and electrostatic forces) • Energy and its transformation (e.g. conservation, dissipation, chemical reactions) • Interactions between energy and matter (e.g. light and radio waves, sound and seismic waves)
Living systems that require knowledge of:	• Cells (e.g. structures and function, DNA, plant and animal) • The concept of an organism (e.g. unicellular and multicellular) • Humans (e.g. health, nutrition, subsystems such as digestion, respiration, circulation, excretion, reproduction and their relationship) • Populations (e.g. species, evolution, biodiversity, genetic variation) • Ecosystems (e.g. food chains, matter and energy flow) • Biosphere (e.g. ecosystem services, sustainability)
Earth and space systems that require knowledge of:	• Structures of the Earth systems (e.g. lithosphere, atmosphere, hydrosphere) • Energy in the Earth systems (e.g. sources, global climate) • Change in Earth systems (e.g. plate tectonics, geochemical cycles, constructive and destructive forces) • Earth's history (e.g. fossils, origin and evolution) • Earth in space (e.g. gravity, solar systems, galaxies) • The history and scale of the universe and its history (e.g. light year, Big Bang theory)

Adapted from *PISA 2015: Assessment and Analytical Framework*

Table 2.5 shows the desired distribution of items by content of science.

Table 2.5 Distribution of Items by Content of Science

System	Percentage of Total Items
Physical	36
Living	36
Earth and space	28
Total	100

Adapted from PISA 2015: Assessment and Analytical Framework

2.1.1.3.2 Procedural Knowledge

Procedural knowledge is about knowing the concepts and processes which are necessary in data collection, analysis and interpretation steps for obtaining reliable and valid data.

The examples listed below are about the general characteristics of procedural knowledge.

- The concept of variables, such as dependent and independent variables.
- The concepts of measurement, such as quantitative and qualitative variables and the use of a scale, such as categorical and continuous variables.
- The ways of assessing and minimising uncertainty, such as repeating and averaging measurements.
- Mechanisms to ensure the replicability and accuracy of data.
- Appropriate representation of data with tables, graphs and charts.
- The use of randomised controlled trials in an experimental design to avoid confounded findings and identify possible causal mechanisms.
- The nature of an appropriate experimental design for a given scientific question.

2.1.1.3.3 Epistemic Knowledge

The students, who have the sense of epistemic knowledge, are able to clearly define the differences between theory, hypothesis, fact and observation concepts. They can suggest a rationale for the features of science-related concepts and processes. The

importance of peer review is appreciated by the students while processing a new knowledge.

Epistemic knowledge can be evaluated when students are asked to analyze and answer a question, which demands more than assessing or concluding directly. They are also asked to justify their conclusions, to decide on the evidence that best supports the hypothesis, and to explain the reason of their choice.

Table 2.6 demonstrates the desired item distribution among content knowledge, procedural knowledge and epistemic knowledge according to the PISA 2015 framework (OECD, 2017a).

Table 2.6 Desired Distribution of Items for Knowledge

Knowledge types	Systems			Total over systems
	Physical	Living	Earth and space	
Content	20-24	20-24	14-18	54-66
Procedural	7-11	7-11	5-9	19- 31
Epistemic	4-8	4-8	2-6	10-22
Total over knowledge types	36	36	28	100

Adapted from *PISA 2015: Assessment and Analytical Framework*

2.1.1.4 Affective Characteristics of Scientific Literacy Performance

In addition to achievement tests, which evaluate the scientific literacy of the students, affective characteristics of the students associated with their academic achievement are assessed in the PISA study as well. In this context, students were asked questions through questionnaires about their science learning motivation, interest to science subjects, instrumental motivation to learn science (for future education or career plans), and science self-efficacy (the level of students being able to effectively achieve science tasks and overcoming difficulties with their own skills) (MoNE, 2016).

Attitudes are the significant part of the construct of scientific literacy, since scientific literacy contains beliefs, motivations, self-efficacy and values of a person. The attitude concept in PISA is based on Klopfer's (1976) affective domain structures in science education. PISA also distinguishes between attitudes towards science and scientific attitudes. Attitudes towards science is about students' interest in scientific issues and activities, while scientific attitudes is about valuing empirical evidence as basic principle of belief.

In the PISA 2015 study, students' attitudes towards science is associated with science and technology interest, environmental awareness, and appreciating the value of scientific methods in researches. These areas are considered as the basis for the measurement of scientific literacy, since a scientifically literate person has positive attitude towards science, values the scientific methods, and has a concern for environment-related issues. Hence, the outcome of compulsory education can be measured by investigating whether students are interested in science-related topics or whether they value science and its applications.

Table 2.7 shows the measures of science-related non-cognitive outcomes in the PISA 2015 study.

Table 2.7 Measures of Science-Related Non-Cognitive Outcomes Included in the PISA 2015 Main Survey

Area	Science-Related
Self	Self-efficacy
Interest, attitudes, and motivation	Interest in broad science topics Enjoyment of science Instrumental motivation
Beliefs and preferences	Epistemological beliefs Environmental awareness Environmental optimism

Adapted from *PISA 2015: Assessment and Analytical Framework*

Self-related beliefs referring to science learning: There is a significant relationship between positive self-efficacy beliefs, learning behaviours and students' science performance (OECD, 2007).

Science learning motivation: Science learning motivation includes interest in broad science topics, enjoyment of science and instrumental motivation (Krapp & Prenzel, 2011). Student engagement and learning activities are affected by the motivation in science learning.

This motivation can be formed by classroom instructions and parental motivation practices (Gottfried et al., 2009; Rakoczy, Klieme, & Pauli, 2008). Instrumental motivation also plays a significant role in the course selection, career choices and performance of the students (Wigfield, Eccles, & Rodriguez, 1998).

Beliefs about science: A strong relationship between epistemological beliefs and students' general values of science and scientific inquiry was shown in the studies (Hofer, 2001; Trautwein & Lüdtke, 2007).

Epistemological beliefs are about considering science as an evolving and changing subject and they are also about how individuals justify knowledge (Conley et al., 2004).

Environmental awareness and environmental optimism are constructs, which are distinct areas of beliefs. In PISA 2006, it was demonstrated that students with socio-economically advantaged background had higher levels of environmental awareness, and environmental awareness was found to be closely related with students' performance in science (OECD, 2007).

Table 2.8 summarizes the measurement and evaluation approaches to science interest, motivation, self-efficacy dimensions and career aspirations.

Table 2.8 Motivation in Science and Career Expectations, Science Self-Confidence and Science Learning

Science Interest	Science Learning Motivation	Science Self-Confidence
Science career aspiration: An indicator based on student responses to the open-ended question “What job do you expect to have at age of 30?”.	Enjoyment of science: An indicator based on student responses to the questions related to science learning interests.	Science self-efficacy: An indicator based on student responses to questions about perceived skills in real-world conditions (eg, understanding and analyzing news reports or participating in debates relevant to science).
Science activities: An indicator based on student responses to the questions about their participation in a range of science-related activities.	Interest in broad science topics: An indicator based on student responses to the questions about the topics such as “biosphere, movement and forces, universe and history, protection from disease”. Instrumental motivation in science learning: An indicator based on student responses to questions about how well science courses in school are beneficial to their education and career plans.	

Adapted from *Ministry of National Education (MoNE), PISA 2015 National Report*

2.1.2 Assessing Scientific Literacy

Cognitive demand, test characteristics, response formats and assessment structures in the PISA 2015 framework are explained as follows.

2.1.2.1 Cognitive Demand

In the PISA 2015 assessment design, cognitive demand is described as the type of mental processes required (Davis & Buckendahl, 2011). While assessing scientific literacy, the levels of cognitive demand are defined for three science competencies. Cognitive demand should not be confused with item difficulty. The difficulty of an item can be high, and it might be related to a not well-known knowledge, however, its cognitive demand may be simply recall. On the other hand, cognitive demand of an item can be high, since it requires relating and evaluating many items of knowledge, and each of them could be easily recalled. Therefore, PISA test instrument provides information both on easier and harder test items and on how students with different ability levels deals with problems having different levels of cognitive demand (Brookhart & Nitko, 2011).

An adapted version of “Webb’s Depth of Knowledge” grid was used in PISA 2015 framework (Webb, 1997). The cognitive framework was designed to assess and report on competencies across the student ability range. In Webb’s Depth of Knowledge Levels, a taxonomy has been formed for cognitive demand according to verbal cues such as analyse, arrange, etc., and depth of knowledge. Depth of knowledge can be categorised as low, medium and high.

Low level is described as carrying out a one-step procedure, such as recalling a fact or concept or recognising a single information from a table or graph.

Medium level is described as using and applying conceptual knowledge for explaining phenomena or selecting correct procedures with two or more steps and organising and interpreting data.

High level is described as analysing complex data, synthesising or evaluating evidence or developing a plan or different approaches for a problem.

Table 2.9 shows the distribution of items according to the depth of knowledge.

Table 2.9 Distribution of Items by Depth of Knowledge

Depth of Knowledge	Percentage of Items
Low	8
Medium	30
High	61
Total	100

Adapted from *PISA 2015: Assessment and Analytical Framework*

Item difficulty is related with the degree of complexity, level of knowledge the item requires, and the cognitive operations required for processing it.

Cognitive demand of an item assessing science achievement can be determined thorough the degree of complexity of knowledge, the level of familiarity of student with the knowledge, and the type of cognitive operations (recall, analyse, evaluate, etc).

2.1.2.2 Test Characteristics

PISA uses unit structure which includes realistic contexts about the complexity of real-life situations. A test unit may be designed with written texts including a table, chart, graph or diagram or with interactive simulations and animations. Items in units are scored independently.

In the PISA 2015 test units, three scientific competencies and three types of scientific knowledge are used. In each test unit multiple competencies and knowledge types are measured. However, individual items in units measure only one competency or one type of knowledge.

2.1.2.3 Response Formats

Three types of items used in the PISA 2015 framework to assess the competencies and scientific knowledge are: simple multiple choice items, which require selection of a single response from other options; complex multiple choice items, which include selection of more than one response from a list, Yes/No questions, and filling blanks; and constructed response items, which require written or drawn responses.

In addition to these response formats, students completed interactive tasks as well. They responded to some items by manipulating the variables in a simulated scientific enquiry.

2.1.2.4 Assessment Structure

The assessment of the scientific literacy was primarily delivered by computers in PISA 2015. New items in PISA 2015 were only available to be conducted through computers. However, paper-based assessment instrument, which was used for the countries that did not prefer computer-based assessment, included only trend items.

Items are organised into sections (clusters), and each of the sections takes 30-minutes. Six clusters have been formed for PISA 2015.

A test form included four clusters taking 30 minutes of testing time. Each student was assigned one two-hour test form. One hour was spent on scientific literacy, the remaining time were used for either one or two of the additional domains of reading, mathematics and collaborative problem solving. For the computer-based assessment of optional financial literacy 60 minutes were given additionally.

Different students took different combinations of test items in the assessment procedure. In the assessment design, personal and global contexts had the same ratio, whereas the number of items related to local/national context was two times higher.

2.2 The Variables in Context Questionnaires in PISA 2015

In PISA, students and their school principals are asked to answer the questionnaire items, and this procedures take about 35 and 45 minutes respectively.

The questionnaires aim to gather information about:

- Economic, social and cultural background of students and their families.
- Students' attitudes towards learning, their habits and life in and outside of school, and their family environment.

- The quality of human and material resources of schools, management, funding and decision-making processes of schools, staffing practices and the schools' curricular emphasis and extra-curricular activities offered.
- Context of instruction, including institutional structures and types, class size, classroom and school climate, and reading activities in class.
- Students' learning aspects such as motivation and interest.

Optional questionnaires in PISA 2015 are parent questionnaire, computer familiarity questionnaire, educational career questionnaire and teacher questionnaire.

The parent questionnaire is mainly about the perceptions and concerns of parents about their child's education, their assistance for at home learning, career expectations of their child and their background.

The computer familiarity questionnaire focuses on students' attitudes towards computer use and how they can accomplish computer tasks, and on the extent to which the information and communications technology (ICT) is available and reachable.

The educational career questionnaire gathers information mainly about how students are prepared for their future career and interruptions in schooling.

The teacher questionnaire asks science teachers about their teaching practices and activities in science classes, the content of a school's science curriculum and its transmission to parents and transformational leadership (OECD, 2017a).

The measures in the core context assessment for PISA and the measures of non-cognitive outcomes in PISA 2015 are represented in Table 2.10 and Table 2.11 respectively.

In PISA, different approaches have been used for survey development. PISA 2009 focuses on the hierarchical structure of educational systems and the efficacy of the education, and PISA 2012 focuses on preschool education. PISA 2015 aims to combine these approaches with new dimensions of policy interest.

PISA 2015 survey framework was constructed with nineteen highly significant policy issues (modules). These modules were categorised under policy areas such as

science education, equality, broader educational outcomes beyond achievement, supportive school context and educational governance (OECD, 2017a).

Table 2.10 Measures Included in the Core Context Assessment for PISA 2015

	Student and School Background	Processes	Non-Cognitive Outcomes
System level		Governance: Decision making, horizontal and vertical differentiation	(aggregated student data)
School level	School location, type and size of school, amount and source of resources (incl. ICT) Social/ethnic/academic composition Class size, teacher qualification	School policies: Programmes offered, admission and grouping policies, allocated learning time, additional learning time and study support, extra-curricular activities, professional development, leadership, parental involvement, assessment/evaluation/accountability policies, school climate (teacher and student behaviour) Teaching and learning: Disciplinary climate, teacher support, cognitive challenge	(aggregated student data)
Student level	Gender, socio-economic status, language and migration background, grade level, pre-primary education, age at school entry	Grade repetition, programme attended, learning time at school (mandatory lessons and additional instruction), out-of-school learning	Domain-general non-cognitive outcomes (e.g. achievement motivation) Domain-specific non-cognitive outcomes (domain-related beliefs and strategies)

Adapted from *PISA 2015: Assessment and Analytical Framework*

Table 2.11 Measures of Non-Cognitive Outcomes Included in the PISA 2015 Main Survey

Area	Science-Related (Module 4)	Domain-General (Modules 6, 10, 11)
Self	Self-efficacy	Test anxiety Well-being in general (life satisfaction) Well-being at school (sense of belonging)
Interest, attitudes, and motivation	Interest in broad science topics Enjoyment of science Instrumental motivation	Achievement motivation
Beliefs and preferences	Epistemological beliefs Environmental awareness Environmental optimism	Collaboration and teamwork dispositions Career aspirations
Technology – ICT		ICT use Interest in ICT Perceived ICT competence Perceived autonomy in using ICT ICT use in social interaction
Behaviour		Health: physical activities Time use: activities before/after school

Adapted from *PISA 2015: Assessment and Analytical Framework*

Table 2.12 demonstrates the relationship of nineteen policy modules with the questionnaires used in PISA 2015 (STQ: Student Questionnaire, SCQ: School Questionnaire, TCQ: Teacher Questionnaire, PAQ: Parent Questionnaire, ICT: Information and Communications Technology Familiarity Questionnaire, EC: Educational Career Questionnaire, ✓: whether this module is applied in the related questionnaire).

Table 2.12 Overview of Nineteen Policy Modules and Their Relevance to Surveys

	STQ	SCQ	TCQ	PAQ	ICT	EC
Policy area: Science education						
1. Teacher qualification and professional knowledge		✓	✓			
2. Science teaching practices	✓	✓	✓			
3. School-level learning environment for science		✓	✓		✓	✓
4. Science-related outcomes: motivation, attitudes, beliefs, strategies	✓				✓	
5. Out-of-school science experience	✓			✓	✓	✓
6. Career aspirations	✓					
Policy area: Equity						
7. Student SES, family and home background	✓			✓	✓	
8. Ethnicity and migration	✓		✓	✓		
9. Educational pathways in early childhood	✓			✓		✓
Policy area: Broader educational outcomes beyond achievement						
10. Domain-general student behaviour and attitudes	✓				✓	
11. Student dispositions related to collaborative problem solving	✓		✓			
Policy area: Supportive school context						
12. Learning time and curriculum	✓	✓	✓			✓
13. School climate: interpersonal relations, trust, expectations	✓	✓				
14. Parental involvement	✓	✓		✓		
15. Leadership and school management		✓	✓			
16. Resources		✓	✓		✓	
Policy area: Educational governance						
17. Locus of control within the school system		✓				
18. Allocation, selection and choice		✓		✓		
19. Assessment evaluation and accountability	✓	✓	✓			

Adapted from *PISA 2015: Technical Report*

PISA 2015 framework assumes that education policy makers in the countries participating the study need to be knowledgeable in four important topics: Non-cognitive outcomes, student background, teaching and learning processes and school policies and educational governance.

Non-cognitive outcomes such as being committed, sharing beliefs and values, being motivated to learn and collaborate and respecting others are significant for individual development, and for success in school and life (Heckman, Stixrud, & Urzua, 2006). In PISA 2015, non-cognitive outcomes are measured within the student questionnaire and ICT familiarity questionnaire.

Socio-economic status, immigrant background and family background related variables need to be analyzed carefully to comprehend how educational opportunities and outcomes are distributed within and across countries. For this purpose, students' economic, social and cultural status (ESCS) need to be assessed and evaluated across the PISA cycles. Since early childhood education has importance for the educational pathway and career, PISA aims to gather information on primary and pre-primary education. The social and academic composition of the school has an effect on learning processes and outcomes. Hence, PISA uses aggregated student data to describe background factors on school-level in addition to structural factors of schools.

Teachers' qualification and teacher support, teaching practices, classroom climate and classroom management, learning time and learning opportunities provide important information about the educational effectiveness (Creemers & Kyriakides, 2008). Data obtained from student and school questionnaires can give information about the learning environment in schools.

The school-level factors, such as professional capacity and professional development, curriculum, school management, school climate and parental involvement foster school effectiveness (Bryk, Allensworth, Easton, Sebring, & Luppescu, 2010). Governance-related issues such as allocation, selection, and assessment and evaluation are used to control school qualities and to observe and support educational development.

In the PISA 2015 study, there are five different policy areas, which include the following 19 policy modules. Table 2.13 shows the modular structure of the context assessment design, which demonstrates these policy areas and modules.

Table 2.13 Modular Structure of the PISA 2015 Context Assessment Design

Student Background			Processes		Non-Cognitive Outcomes
Science-related topics	Family	Education	Actors	Core Processes	Resource Allocation
		5. Out-of-school science experience	1. Teacher qualification and professional knowledge	2. Science teaching practices	12. Learning time and curriculum
			Teaching and learning		
			3. School-level learning environment for science		
General topics	7. Student SES and family	9. Educational pathways in early childhood	14. Parental involvement	13. School climate: interpersonal relations, trust, expectations	16. Resources
	8. Ethnicity and immigration		15. Leadership and school management		
			School policies		
		17. Locus of decision making within the school system	19. Assessment, evaluation and accountability	18. Allocation selection and choice	
			Governance		

Adapted from *PISA 2015: Assessment and Analytical Framework*

The modules related to science are also collected under a separate title as “science-related topics”. According to Table 2.13, modules 1, 2, 3, 4, 5 and 12 are collected under science-related topics. In addition to these modules, career aspirations (module 6) is included in the scope of this thesis, since it is evaluated within the scope of science education policy area (OECD, 2017b).

2.2.1 Student Background Variables

In order to understand children's science learning; in addition to learning that takes place in the school, learning that takes place out-of-school should be considered as well (Eshach, 2007). Students experience an informal culture, which is relevant to the individual outside of school, however the school promotes a public culture, which is relevant to society. Experiences with the physical and natural world, which allow the integration of students' previous experiences into classroom activities, contribute to the organization and teaching of science in the school, as they affect learning outcomes (Schibeci, 1989). In addition to being a source of knowledge, informal experiences provide affective reinforcements, which are necessary for maintaining motivation and interest in the object of the experience (Eagly & Chaiken, 1993).

Homework assignments are a significant part of the out-of-school study time. They offer a chance for teachers to extend the school day and engage students in learning outside the school hours (Marzano & Pickering, 2007). In general, extra study time is used by teachers to assist students to review previous concepts or to prepare for future lessons. Homeworks help students gain a deeper understanding of the content and earn higher grades and achievement scores (Epstein & Van Voorhis, 2001). However, homeworks may be overburdened for students, when the underlying goals of assigning work for students to complete outside the hours of school is not considered. A teacher should assign a homework that bridges the gap between previously covered material and concepts yet to be taught, or helps the student improve study skills (Epstein & Van Voorhis, 2001; Xu, 2005).

Since out-of-school science experience module is determined as a science-related topic in the assessment design, derived variables related to out-of-school science experience module were examined under student background variables. PISA aims to determine in-school and out-of-school learning activities and how much time spent on these activities before and after school in addition to mandatory learning time in schools. This information can be obtained from student and school questionnaires. Information about extra-curricular learning activities and science-related experiences can be gathered from students and school leaders in the school questionnaire. Out-of-school study time for science (OUTHOURS) and science activities (SCIEACT) are

the derived variables examined under this module in the present study (OECD, 2017a).

2.2.2 Non-Cognitive Outcomes

Peoples' attitudes towards science plays an important role in the interest, attention and response of students to science and technology. Science education aims to develop attitudes that lead students to engage with scientific issues. Through the support of such attitudes, scientific and technological knowledge can be applied for personal, local/national and global benefit and students can develop self-efficacy (Bandura, 1997).

Studies have shown that non-cognitive outcomes can predict the success in secondary education and higher education (Singh et al., 2002; Heckman, Stixrud, & Urzua, 2006; Çalışkan, 2008; Chua & Karpudewan, 2017). Hence, PISA aims to investigate relations between non-cognitive outcomes and achievement at the individual, school and country levels.

Epistemological beliefs are viewed as important factors effecting students' achievement with respect to the interpretation of information and knowledge and process of learning. They have been found to be associated with students' learning outcomes and achievement (Hofer, 2001; Trautwein & Lüdtke, 2007; Cano, 2005; Lee, Johanson, & Tsai, 2008).

Citizens should have a certain level of awareness about environmental issues to comprehend the problems and to be aware of their effect on the environment. Studies have shown that scientific knowledge is necessary for appreciating the environment and for comprehending environmental problems (Hadzigeorgiou & Skoumios, 2013; Summers, Kruger, & Childs, 2001). This can be explained with the fact that features of environmental systems and environmental problems related issues can be explained with scientific facts. Scientifically literate people are more likely to have better tendencies about the environment and they tend to be environmentally proactive (Hansla, Gamble, Juliusson, & Gärling, 2008). However, Öztürk (2018) found environmental awareness is positively associated with the scientific literacy; whereas environmental optimism is negatively associated with scientific literacy.

Sulemana et al. (2016) suggested that since students with low socio-economic status do not have the proper knowledge and awareness of the seriousness of environmental problems, that can explain why these students are also more optimistic about the future of the environment.

Since science-related outcomes such as motivation, interest and belief are determined as science-related topics in the assessment design, and since career aspirations is evaluated within the scope of science education policy field, derived variables related to science-related outcomes module were examined under non-cognitive outcomes. Science self-efficacy (SCIEEFF), interest in broad science topics (INTBRSCI), enjoyment of science (JOYSCIE), instrumental motivation (INSTSCIE), epistemological beliefs (EPIST), environmental awareness (ENVAWARE) and environmental optimism (ENVOPT) are the derived variables examined under this module in the present study.

Previous studies reported that students' academic achievement had a minimal affect on the career aspirations of teenage students. They showed that a large number of students held higher level career aspirations regardless of their achievement level, they also found a weak correlation between academic achievement and career aspirations (Rojewski & Yang, 1997; Benbow et al., 1991).

Career aspirations module in the PISA 2015 student questionnaire includes the question about the job students expect to have at the age of 30. Students' expected occupational status (SEI) (BSMJ) variable was examined under this module in the present study (OECD, 2017a).

2.2.3 Teaching and Learning Variables

International comparative researches which point to cross-cultural comparability of the constructs have shown that learning time is closely related to student learning and achievement (OECD, 2011; Martin et al., 2008; Schmidt et al., 2001). However, studies also indicate large differences within and across countries and among different groups of students and schools regarding the learning time (Ghuman & Lloyd, 2010; OECD, 2011). Hence, learning time provided or implemented by the

school, implemented by the teacher, and time used by the students should be distinguished from each other.

PISA 2015 evaluates provided learning time, however there can be loss of time due to teachers' strikes, illness, other teacher or student absenteeism. Time spent on managing classroom, collecting homework or waiting can cause loss of time as well (MacKay, 2009). Time, in which the students actually learn and they are mentally present in the learning environment is defined as engaged learning time.

Learning time includes time spent on additional learning activities within and out-of-school in addition to mandatory schooling time. PISA 2015 aims to evaluate additional learning time by using data from the school, student and educational career questionnaires. Data related to extra-curricular learning activities, time use before and after school every day, and science-related experiences can contribute to comprehend the whole concept (OECD, 2017a).

Since learning time is determined as a science-related topic in the assessment design, derived variables related to this module were examined under teaching and learning variables. Learning time for science (SMINS) is the derived variable examined under this module in the present study.

In addition to learning time, which is explained under student-level variables, the teaching force, which is described in terms of initial education, and beliefs and professional development of teachers, is explained as an important part of teaching and learning variables under school-level variables.

Studies, which investigated the role of the effectiveness of teachers in students' performance and academic achievement have suggested that there is a significant relationship between teacher quality and student achievement. They also suggested that teachers' knowledge and their ability to transfer their knowledge to students affect student achievement (Egalite, Kisida, & Winters, 2015; Marchand & Weber, 2014). Loucks-Horsley and Matsumoto (1999) showed in their study that there is a significant relationship between high-quality professional development and student achievement. Yoon et al. (2007) examined the link between student learning and professional development, and concluded that teachers who receive substantial professional development can improve their students' achievement. However, it was

also stated in the research studies that the level of education of teachers, i.e. teachers' graduate educational attainment have a very limited effect or no significant effect on students' academic achievement (Gage et al., 2017; Badgett et al., 2013; Coenen et al., 2014).

The interaction between teachers and students is affected by the number of students per teacher in a school as well. The number of students in a classroom affects the classroom process, course activities, students' engagement and students' learning. Demir (2009) investigated the effect of student-teacher ratio on academic achievement, and suggested student-teacher ratio is a significant predictor of academic achievement. Small class size gives teachers the opportunity to spend more time on each student and this helps students improve their learning and academic success (Croll & Hastings, 1996). In schools with smaller student teacher ratio, teachers can control the progress of every student and they can provide a more individualized teaching that is more suitable to each student (Johnson, 2011).

Teacher-related factors affecting student learning, such as age and education level of teachers, their initial education and qualifications, their individual beliefs and abilities, and their professional practices on the school-level, such as collaboration and professional development, are evaluated in PISA 2015 framework.

Since teacher qualification and professional knowledge is determined as a science-related topic in the assessment design, derived variables related to this module were examined under teaching and learning variables. Proportion of science teachers by all teachers (PROSTAT), proportion of science teachers fully certified (PROSTCE), proportion of science teachers fully certified (PROSTCE), proportion of science teachers with ISCED level 5A and a major in science (PROSTMAS) and total number of science teachers at school (TOTST) are the derived variables examined under this module in the present study (OECD, 2017a).

2.2.4 School Policies Variables

The results from the studies that investigate the relationship between educational resources of schools and academic achievement suggest that educational resources such as physical and economic resources of schools have significant effect on

academic achievement of students (Card & Krueger, 1996; Greenwald, Hedges, & Laine, 1996). Physical resources can be categorised as people, equipment, physical and economic resources. Adeogun and Osifila (2008) found that there are positive relationships between academic achievement of students and physical, financial and material resources. However, human resources such as teaching and assisting staff are found not to be significantly related to students' academic performance.

Studies have shown that the students' participation in extracurricular science activities organized in the school had a positive effect on their science achievement (Zhang & Tang, 2017; Freeman, 2017). The literature has shown that students who actively participate in extracurricular activities tend to have good grades in their examinations, since they can manage time, strive for excellence, and relieve stress in doing things. In addition, well-grounded extracurricular activities in the school programs have positive influence on students' achievement (Ekeke, 2014).

Since school-level learning environment for science is determined as a science-related topic in the assessment design, derived variables related to this module were examined under school policies variables. Shortage of educational material (STAFFSHORT), shortage of educational staff (EDUSHORT), science specific resources (SCIERES) and creative extra-curricular activities (CREACTIV) are the derived variables examined under this module in the present study (OECD, 2017a).

2.3 Turkish Educational System and Application of PISA 2015 in Turkey

Significant changes have been made in the Turkish educational system since 2012. The 4+4+4 system has been introduced, that is 4 years of primary education, 4 years of secondary school education and 4 years of high school education. Education institutions include public school, private school or open education. Preschool education is a non-compulsory level of education for 36–66-month-old children. Higher education is divided into two-year and four-year institutions and the entrance to higher education is done through exams.

Every student who graduated from public high school, vocational school or open high school can be admitted to higher education through the YKS (Higher Education Institutions Exam) held by ÖSYM (Measurement, Selection and Placement Center).

Higher education is defined as “all education covering at least four semesters at all levels”. Higher education institutions in Turkey are universities, faculties, institutes, colleges, conservatories, vocational schools, and application and research centers.

PISA 2015 study, as the sixth cycle of the survey, was conducted by the Ministry of National Education, General Directorate of Measurement, Evaluation and Examination Service as a computer-based assessment with the participation of 5895 students in Turkey. Preparations made within the scope of the application can be summarized as follows:

- The new test items and questionnaire items developed for PISA 2015 were reviewed and finalized by field experts after translation.
- A number of training meetings (aimed at provincial administrators and school administrators) were organized in order to prepare the selected schools for implementation. Through training materials (brochures, promotional booklets, manuals, etc.) all stakeholders were able to understand PISA and make necessary preparations.
- Testers from the Ministry of National Education have been assigned to help the students to work on the system and to solve problems while students answer the questions on the computer. Thus, the rate of data loss due to different reasons was minimized.
- After the application, the scoring of the open-ended questions, which were answered by the students, was done by the field experts and the data were sent to the international center of PISA.

The changes in PISA average scores in the field of scientific literacy in the last four applications were demonstrated in Table 2.14.

Table 2.14 Scientific Literacy Average Scores by Years

	2006	2009	2012	2015
OECD Average	498	495	501	493
Average of All Countries	478	471	477	465
Average of Turkey	424	454	463	425
Ranking of Turkey	47	42	43	54
Number of Participating Countries	57	65	65	72

Adapted from *Ministry of National Education (MoNE), PISA 2015 National Report*

It can be seen that the average score of science literacy in all countries and in OECD countries increased from PISA 2006 to PISA 2012; however, after PISA 2012, the average scores have decreased in PISA 2015 application. This decrease was statistically significant.

In PISA 2012, the ratio of scientific literacy of Turkish students at level 1 and below (low proficiency) was found as 26.9%. In PISA 2015, this ratio increased to 44.4%. In PISA 2015, the average student rate at level 1 and below in OECD countries is 23.3%. When all countries are included, this rate is 31.4%.

In PISA 2015, the ratio of scientific literacy of Turkish students at level 5 and above (top proficiency) is 0.3%. This ratio is 7.8 % for OECD countries, and 5.3% when all countries are included. In addition, in PISA 2015, the ratio of scientific literacy of Turkish students at level 5 and above is lower than that of in PISA 2012 (MoNE, 2016).

2.4 Studies about Factors Affecting Students' Scientific Performances

Studies related to student and school background variables, teaching and learning variables and non-cognitive outcomes, which effect the scientific literacy skills of students, are represented below.

Alivernini et al. (2010) conducted a multilevel analysis with a four-stage procedure considering the home/student-level nested under the school/teacher-level by using TIMSS 2007 data. The results demonstrated that students performed significantly better in science subjects ($p < 0.01$), if their home educational resources, self-confidence in learning science and the student perception of being safe at school were higher. The findings of the study also showed that the performance of male and native students were higher.

Akpınar et al. (2009) explored the differences by gender and grade level in primary school students' attitudes toward science and examined the relationship between students' attitudes toward science and academic achievement. The participants of the study were 658 primary school students. A "Personal Information Form" and "Attitude Scale for Science and Technology (ASST)", the correlation coefficient of which was 0.89, were administered to the participants. The findings of the study

demonstrated that the level of interest in science of female students were significantly higher than that of male students. However, there were no significant gender differences in terms of other factors such as enjoyment of science, anxiety, and enjoyment of science experiments. There were significant differences among students' grade levels towards science and technology in terms of these factors. In addition, significant positive correlations were found between ASST and academic achievement.

Yetişir (2014) used TIMSS 2011 data in order to investigate the science achievement of eighth graders in Turkey in terms of the relationship of science achievement with the selected student- and classroom-level variables by using a hierarchical linear model to analyze the data. The results of the analysis showed that the variance in science achievement among eighth grade classrooms, which was about 32%, was statistically significant. The results also revealed that, while attitudes towards science and parents' level of education were positively related to science achievement, student engagement had no significant relationship with science achievement. The analysis also demonstrated that, while teacher collaboration and inquiry-related activities did not have a statistically significant effect, class average-engagement and readiness to learn had a significant effect on science achievement.

Papanastasiou and Papanastasiou (2010) aimed to investigate the attitudes of 8th grade students toward science in Australia, Canada, Cyprus, and Korea by using TIMSS data. They suggested a model of family and school influences on students' attitudes. The structural equation model consisted of 3 exogenous constructs – family's educational background, aspiration, school climate and 2 endogenous constructs – teaching and student attitudes towards science. The findings of the study showed that educational background, aspiration, and school climate had a direct effect on attitudes, while teaching was directly affected by aspiration and school climate.

Papanastasiou and Zembylas (2004) investigated the locality of the relationship between attitudes towards science, self-beliefs and science achievement for senior high school students in Australia, Cyprus and the USA by using the Structural Equation Modeling software, AMOS. The data were obtained from the Third International Mathematics and Science Study database. The results of this study

showed the differential effects that science achievement and science attitudes can have on each other, depending on the characteristics of the educational systems within each of country. Investigation of the nature of this relationship revealed that different components of science-related attitudes were differentially related to achievement.

Odom and Bell (2015) explored the association of middle school students' science achievement and attitudes about science with student-reported frequency of teacher lecture demonstrations and student-centered learning. The sample of the study was composed of 602 seventh and eighth grade students enrolled in middle school science. The association of attitudes toward science, student-centered learning, and teacher demonstrations with science achievement were examined through multiple regression. The findings of the study demonstrated that both attitudes toward science and student-centered learning were positively associated with science achievement, and student-centered learning was positively associated with attitude toward science. Teacher demonstrations had a negative association with student achievement, and had no significant association with attitudes toward science. In addition, it was concluded that demonstrations provided insufficient opportunity for students to develop an understanding of the processes of science. It was also found that, observing teacher demonstrations may be valuable, but they are not a substitute for laboratory investigations by students.

In the research of Çalışkan (2008), it was aimed to explore the impact of school- and student-level factors on scientific literacy skills of Turkish students in the PISA 2006 study. The sample of the study was 4942 Turkish students, who were 15-year-old, from 10 primary schools, 88 general high schools and 66 vocational high schools. The results of the study showed that the impact of school- and student-level factors on scientific literacy skills of Turkish students varied from school to school. In addition, PISA index of economic, social and cultural status, general value given to science and science self-efficacy of students effected every aspect of the scientific literacy skills. It was also found that general high school students were more successful than the vocational high school students.

Anıl (2009) aimed to determine the factors affecting the science achievement of 15-year-old Turkish students within the framework of PISA 2006 study. The sample of

the study was 4942 students from Turkey. In data analysis, questions on the student questionnaire were specified through fundamental components paraphrasing factors analysis in order to determine the factors affecting students' science achievement. Procedural multi-regression analysis was used to determine the predicting power of the predictive variables. The results of the study showed that educational level of fathers is the most predictive variable for the 15-year-old students' science achievement in Turkey. In addition, attitude, computer environment, the cultural wealth of the family and educational level of mothers had positive correlation with students' science achievement.

Cavas (2011) aimed to explore sixth to eighth grade Turkish primary students' motivation toward science learning and to determine the factors affecting this. The sample of the study was of 376 students from 5 different primary schools in İzmir. The data were collected through the adapted Students' Motivation toward Science Learning (SMTSL) questionnaire, which was a valid and reliable tool for Turkish students, having 6 subscales which were self-efficacy, science learning value, active learning strategies, performance goal, achievement goal, and learning environment stimulation. In addition, data were collected on students' science attitudes and achievement scores. The results of the study showed that Turkish primary students' science motivation differed significantly in terms of their gender and grade level. Students' motivational level had a considerable impact on their science attitudes and achievement in science.

Chua and Karpudewan (2017) aimed to explore the relationship between lower secondary school students' perceptions about laboratory environment, attitude, and motivation towards learning. A total of 1003 lower secondary school students between 13-15 years old from different schools in the Northern Region of Malaysia constituted the sample of the study. The findings of the study showed that motivation and perceptions of the learning environment predicted attitude towards learning science. In addition, the teaching strategies predicted the students' attitude towards learning science.

Singh et al. (2002) aimed to investigate the effects of motivation, attitude, and academic engagement on 8th grade students' achievement in mathematics and science. As the sample of the study, the nationally representative sample of 8th

graders drawn from the National Education Longitudinal Study 1988, was used. Structural equation models were used to estimate and test the hypothesized relationships of motivation factors, attitude factors, and academic engagement factors on achievement in mathematics and science. The results of the study supported the positive effects of the motivation factors, attitude and academic time on mathematics and science achievement.

Anagün (2011) aimed to determine the impact of some variables about teaching-learning process on students' scientific literacy levels over 15-year-old students in Turkey on the context of PISA 2006. The sample of study was 4942 Turkish students. Structural equation model was used to determine the predicting power of the variables concerning the 15-year-olds' scientific literacy levels. The findings of the study showed that spending time for learning was the most predictive variable for the 15-year-old students' scientific literacy levels in Turkey. In addition, making experiments and inquiry-based learning activities in the teaching-learning process had positive effect on scientific literacy.

Jones et al. (2018) investigated the difference in the science achievement scores of schools, which used different textbooks and labs, when delivering instruction. Analysis of covariance (ANCOVA) and multiple regression were used to analyze the data while controlling the State of Texas Assessment for Academic Readiness (STAAR) eighth grade reading scores to reveal significant differences between classes. The sample and population for this study were predominantly eighth grade Hispanic students in South Texas. The results of ANCOVA demonstrated that classes that used labs with high hand-on experiences with greater direct student participation received higher science scores. In addition, reading scores were significantly related to science scores. Multiple regression findings showed that textbooks and labs predicted student achievement.

Acar and Öğretmen (2012) performed a study in order to investigate the student- and school-level factors affecting the achievement of Turkish students in the 2006 PISA study. School-level variables of the study were the number of computers in the school, the number of computers connected to the Internet and the quality of educational resources of the school; the student-level variables of the study were the gender of the student, the time allocated to learning at school (weekly science

learning time), the time allocated to learning at home (weekly extra-curricular hours) and the index of social-cultural-economic status. Multiple level analysis was conducted by evaluating the results of two different programs. The results of the study showed that as the number of computers with the Internet connection in the school and the quality of educational resources of the school increased, students' science achievement also increased. In addition, a negative relationship was found between the time students devote to learning at home and the science achievement. It was also found that science achievement decreased as students' time spent learning at home increased.

Berberoğlu and Kalender (2005) aimed to investigate the achievement of Turkish students across years, school types and regions by using data from Student Selection Examination and PISA 2003. They have found in their study that when the number of computers with the Internet connection in the school and the quality of educational resources of the school increased, a positive trend was observed in students' science literacy performances. However, if the number of computers in the school increased, it was found that students' science literacy performance decreased. This result could be due the number of computers without the Internet connection.

Atar and Atar (2012) aimed to examine the effects of student centered learning, outfitting classrooms with latest technology and computers by using TIMSS 2007 science achievements data. Two-staged stratified sampling procedure was used in the sampling process, in which schools and classes were sampled in the first and second stage, respectively. TIMSS 2007 study was administered to students, teachers and principles in 145 classes in 145 schools, 28 students from each school participated in the study. Hierarchical linear modeling (HLM) analyses showed that computer access, increase in the socio-economic status (SES) of the families, teacher experience and students' self-confidence in learning science effected students' science achievements positively, whereas inquiry-based science instruction effected students' science achievements negatively. It was also concluded that computer access in classrooms had a positive effect on students' achievement in science.

Spiezia (2010) aimed to investigate the factors affecting students' science achievement score according to PISA 2006 data. The results of the study showed that the frequency of computer use of students had a positive and significant effect on

students' science achievement score. Bayraktar (2010) reported that students' science performance was related to computer and the Internet access according to TIMSS 2007 data. It was also reported in another study that, according to TIMSS 2011 data, students' science achievement was directly proportional to the presence of computers and the Internet at home (Oral & McGivney, 2013).

2.5 Theoretical Background of Hierarchical Linear Modeling and Related Modeling Studies using Hierarchical Linear Modeling

In social research data is structured mostly hierarchical and properties of these groupings in the hierarchical structure may affect the study. Individuals within a group have more common characteristics than those outside the group. Higher-level units can also have their own characteristics. Since multiple measurements are nested within individual, repeated measurement data are hierarchical. Hierarchical linear and non-linear models allow for analyzing the correlation at each level of the hierarchy by considering the variability associated with each level. Unlike traditional methods, model coefficients can be predicted by hierarchical linear models (HLM) at each level. In addition, random effects related to each unit at each level can be estimated.

Multi-level modeling techniques were first introduced by Aitkin and Longford (1986) and a hierarchical linear modeling technique was developed by Bryk and Raudenbush (1992) as one of the multi-level modeling techniques. In hierarchical linear models, each level in the design is represented by its own submodel. This is explained by Raudenbush and Bryk (2002) as "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). HLM is applied for three general research purposes: improved estimation of effects within individual units; modeling cross-level effects; and partitioning of variance and covariance components among levels.

Hierarchical linear modeling, also called "multilevel modeling", "random coefficient modeling", or "mixed effects modeling" has a more sophisticated methodology compared to simple linear regression and multiple linear regression. It is a synthesis of regression and analysis of variance (ANOVA). The intercepts and slopes of the

groups at the lowest level can be used to explain the relationship among variables within each group.

The intercepts and slopes for each group vary in different groups. Through HLM, the variance in the outcome variables can be analyzed at multiple hierarchical levels. However, simple linear and multiple linear regressions do not allow modelling the effects to occur at a multiple level, they occur at a single level. This makes HLM convenient to be used with nested data.

In the hierarchical modeling technique, a statistical model is designed for each stage of the hierarchical structure and then these models are combined to generate the final model. Since, the combined model has a very complex structure and it is hard to interpret in multi-level models, sub-models belonging to stages are used in interpretations. Hierarchical linear models can be considered as two or more level. Although three or four-level hierarchical structures can be used in applications, mostly a two-level hierarchical linear model is used (Raudenbush & Bryk, 2002).

Hierarchical structure can be seen in a large number of educational studies. For example, students are nested within the schools, students are nested within classes, and schools are nested within districts (Hox, 2010). Because of these grouping effects and the interaction between students, students in the same group are more alike than the students in different groups. So, the observation of students within group can no longer be considered statistically independent, which means the traditional statistical approaches, like regression or ANOVA, are not applicable. When the hierarchical nature of data in educational settings cannot be recognized, this can cause unreliable estimation of the effectiveness of schools and their practices, which could lead to misinformed educational policies or practices (Raudenbush & Bryk, 2002).

Hierarchical linear model comes through the conceptual and technical difficulties of multilevel data. Aggregation bias, misestimated standard errors and heterogeneity of regression can be overcome by hierarchical linear models.

The aggregation bias can occur when the variables take on different meanings and therefore may have different effects at different organizational levels. Hierarchical linear models help resolve this confounding situation by facilitating a decomposition

of any observed relationship between variables, such as achievement social class, into separate level-1 and level-2 components.

Misestimated standard errors occur with multilevel data when the dependence among individual responses within the same organization is not taken into account. Hierarchical linear models resolve this problem by incorporating a unique random effect for each organizational unit into the statistical model.

Heterogeneity of regression occur when the relationship between individual characteristics and outcomes vary across organization. Hierarchical linear models enable the investigator to estimate a separate set of regression coefficients for each organizational unit, and then to model variation among the organization in their sets of coefficients as multivariate outcomes to be explained by organizational factor (Raudenbush & Bryk, 2002, p.100).

Another advantage of using hierarchical linear modeling on multilevel data analysis is that HLM is suitable for conducting analyses with plausible values. The software integrates all plausible values and creates the outcome variable (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2019).

Some modeling studies, which were conducted by using hierarchical linear modeling are represented below.

Fonseca (2011) compared Portuguese students' performance in PISA 2006 scientific literacy with that of students from the OECD, Spain, France, the United Kingdom, Turkey, Greece, and the USA. It was aimed to investigate students' science interests, self-efficacy and beliefs about their own competences, and to identify which student variables may help to explain performance in science. The data were obtained from the PISA 2006 students' science performance and from the student questionnaire which measures interests and beliefs. Data analyses were performed by using hierarchical linear modeling. The results of study showed that the ESCS (economic, social and cultural status) effect was strongest in the USA, Finland and UK. The school-wide ESCS effect was superior to that of student ESCS and it was a performance factor for all countries with the exception of Finland. Boys scored higher than girls, however in several countries, gender did not have a significant effect on the score of students. For all countries, the general value of science,

enjoyment of science and science self-efficacy were positively related with performance. In addition, personal value of science was negatively related with performance. The relationship of participation in non-school science activities with performance was either negative or non-significant.

Sun and Bradley (2012) aimed to use hierarchical linear model (HLM) to explore the factors which affect the science achievement of students from Hong Kong. The factors that influence science outcomes were examined from both student and school perspectives by using a two-level hierarchical linear model. Data from PISA 2006 Hong Kong sample were used to explore the factors that impact the science achievement of 15-year-old students. A two-level hierarchical linear model was used to investigate the factors that influence science achievement from both student and school perspectives. At the student-level, the findings of the study showed that male students, students from high SES families, and students whose parents had higher values on the importance of science were more likely to have better achievement in science. At the school-level, the results demonstrated that the variation of science achievement between schools could be caused by differences in school enrollment size and school SES composition.

Lam and Lau (2014) aimed to examine the effect of some factors on the science performances of students from Hong Kong in PISA 2006 study by using hierarchical linear modeling. The analyses showed that student backgrounds as male, at higher grade levels, and born in mainland (when in the same grade) were associated with increased science performance. Among the attitudinal factors, enjoyment of science and self-efficacy in science were associated with increased scientific achievements. However, most of the parental factors did not have significant impacts on achievement after student attitudes were taken into account, only parents' value of science had a small effect. School student intake was found as a strong predictor of school average achievement, as well as a major mediator of the effects of school enrollment size and school socio-economic status.

Ataş and Karadağ (2015) examined the effects of the school-level indicators on overall reading literacy performance of 15-year-old Turkish students within PISA 2015 application by using a two level hierarchical linear model. Gender, socio-economic indicators of family, parents' level of education, student-level indicators

regarding reading skills, and school-level indicators such as school type, number of students and number of teachers at school were integrated to the modeling process. The results of the study showed that, the effect of economic social cultural index (ESCS) and the fathers' educational level were statistically significant, whereas mothers' educational level did not effect reading literacy of students. The effect of school type and the region that the school is located were also statistically significant. In addition, the number of teachers per student had no significant effect on the mean reading literacy scores of the student.

Üstün et al. (2019) aimed to determine the variables affecting the level of scientific literacy of students and schools in Turkey by utilizing the data obtained from PISA 2015. It was aimed to obtain a hierarchical model consisting of variables that predict students' science literacy levels. As a result, a model was constructed with nine student-level variables and four school-level variables, which predicted the science literacy level of the students. The student-level variables were examined in three groups as personal characteristics, variables associated with learning time, and variables associated with teaching-learning process. The school-level variables were grouped as school resources and learning environment in the school. The most prominent variable at the student-level was weekly science learning time, while science specific resources was the most powerful predictor of students' success at the school-level. However, there was a negative correlation between out-of-school-study time and science literacy scores.

Çelebi (2010) aimed to explore the student and school characteristics and their effects on scientific literacy skills of 15-year-old students across Turkey, Canada, and Sweden by using PISA 2006 data. In PISA 2006, Canada was a high performing country, whereas Turkey was far below the average and Sweden had a rank in the average in terms of scientific literacy. In the modeling phase of the study, student-level characteristics were determined by student questionnaire, and school-level characteristics were determined by school questionnaire. The results of the study showed 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 school size were common school factors affecting students' scientific literacy skills in Canada and Sweden. School

characteristics, such as school admittance policies, educational resources, and science promotional activities had impact on scientific literacy in Turkey. 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 three countries.



CHAPTER 3

METHOD

Hierarchical linear models give better results compared to standard models in the presence of multi-level structured data, where the structure consisted of individuals and groups based on the individuals (Raudenbush & Bryk, 2002).

In this study, the effects of student- and school-level variables on the scientific literacy skills of 15-year-old students in Turkey, were analyzed by using a two-level hierarchical linear model.

3.1 Population and Sample

In PISA 2015, the number of 15-year-old Turkish students was determined as 1.324.089, while the eligible population of Turkey was defined as 925.366 students appropriate to participate in the application.

The target population is the abstract population that the researcher wants to reach, but which is difficult to reach and it reflects the ideal choice of the researcher. The accessible population is the concrete population that the researcher can reach, and it has a realistic choice. Stratified random sampling method is preferred in order to describe the characteristics of certain sub-groups and to allow comparisons between them (Karasar, 2007).

The target population of the study was all students, who were aged from 15 years and 3 (completed) months to 16 years and 2 (completed) months when the assessment started.

For PISA 2015 application, in the first stage of the sampling, schools were determined by using stratified random sampling method according to the Classification of Territorial Units for Statistics Level 1, type of education, type of school, location of the school and administrative form of the school. In the second

stage, the students, who would participate in the study from these schools, were selected by random sampling method.

PISA 2015 was implemented with the participation of 187 schools and 5895 students from Turkey. In the PISA 2015 Turkey sample, girls and boys are distributed homogeneously with equal frequencies (2938 girls and 2957 boys). However, the distribution of students depending on the school type is heterogeneous. There are 121 students in basic education (approximately 2% of all students), 3241 students in general secondary education and 2533 students in vocational technical education. Unlike previous cycles of PISA, in PISA 2015 private schools are included in Turkey sample with a proportion of 4%. Out of 187 schools, 7 schools are private and 179 are public (MoNE, 2016). However, there were missing values in the surveys applied. Items not answered by the respondents, multiple or invalid answers, items not valid for the participant and items not reached represent the types of missing values.

The data files for PISA 2015 were downloaded from the website <http://www.oecd.org/PISA/data/2015database/> and the data for Turkey was saved as a separate SPSS data file. Missing values in terms of the variables of the study were treated before running the analyses.

3.1.1 Sampling Procedure

In the PISA 2015 study, the target population is 15-year-old students, who attend classes in grades 7 and higher. The target population includes 15-year-olds, who are enrolled in educational institutions on a full-time or a part-time basis, students in vocational training programmes, students who attend foreign schools within the country and students coming from other countries and attending any of these programmes in the country. However, 15-year-olds who are educated in their homes or at their workplaces and 15-year-olds out of the country were not included in the target population.

The sampling design of PISA assessment framework was a two-stage stratified sample design. In the first-stage, sampling units included individual schools, which have 15-year-old students. The sampling of the schools were done according to PISA

school sampling frame, which has national list of all PISA-eligible schools, by using the function of the estimated number of PISA-eligible 15-year-olds who are enrolled in the school. This is described as systematic probability proportional to size sampling. Before sampling, schools in the sampling frame were assigned to explicit strata, which can be described as mutually exclusive groups based on school characteristics. The second-stage sampling units were the students in the sampled schools. After schools were selected for the sample, a list of 15-year-olds was prepared for each sampled school. School characteristics, such as the region, public or private educational programs or time of instruction were used by the OECD to determine the schools which would participate in the study. Schools may not prefer participating in the study due to cost or scheduling issues. The OECD aimed to represent at least the 95% of the desired target population. After this selection, students were randomly sampled from each school (OECD, 2017b).

In PISA 2015, 187 schools from 61 cities within 12 territories were selected as participants according to the Classification of Territorial Units for Statistics. Table 3.1 shows the distribution of the number of students in PISA 2015 Turkey sample.

Table 3.1 Distribution of the Number of Students in PISA 2015 Turkey Sample

Territory Code	Name of Territory	Number of Students	Percentage of Students
TR1	İstanbul	1070	18.15
TR2	West Marmara	245	4.16
TR3	Aegean	707	11.99
TR4	East Marmara	510	8.65
TR5	West Anatolia	553	9.38
TR6	Mediterranean	817	13.86
TR7	Central Anatolia	334	5.67
TR8	Western Black Sea	303	5.14
TR9	Eastern Black Sea	194	3.29
TRA	Northeastern Anatolia	199	3.38
TRB	Central Eastern Anatolia	276	4.68
TRC	Southeastern Anatolia	687	11.65
	Total	5895	100

Adapted from *Ministry of National Education (MoNE), PISA 2015 National Report*

In the PISA 2015 study, eight different school types were represented with different percentages. These school types are middle school (2%), Anatolian high school (38.1%), science high school (2.1%), social sciences high school (1.4%), fine arts high school (0.7%), vocational and technical high school (36.4%), multiple programmed Anatolian high school (4.1%), and Anatolian imam hatip high school (14.4%). Ninth grade (20.7%) and tenth grade (72.9%) students represented the majority of the PISA 2015 participants, while only 6.4% of the students were in seventh, eighth, eleventh and twelfth grade. Half of the population was boys and the other half was girls (MoNE, 2016).

3.2 Instruments

The instruments used in PISA 2015 include cognitive level tests aiming to measure the academic performance of students with a focus on science literacy and student and school questionnaires evaluating the students as a whole. Cognitive level tests and questionnaires are applied to students in PISA applications. The basis of the application consists of tests to measure cognitive competences in the fields of science literacy, mathematics literacy and reading skills. In PISA 2015 study, all countries participated in student and school questionnaires. Participating in other questionnaires (teacher questionnaire, computer familiarity questionnaire, educational career questionnaire and parent questionnaire) was dependent on countries' preference. Turkey joined only the school questionnaire and the student questionnaire in PISA 2015.

In this study, the data were obtained from the science achievement test and student and school questionnaires applied to 15-year-old Turkish students in PISA 2015. Of the 72 countries that participated in PISA 2015, 57 were assessed on a computer-based basis and 15 were assessed on a paper-based basis. In the PISA 2015 study, 66 different booklets were used for computer-based assessment and 30 different booklets were used for paper-based assessment (MoNE, 2016).

3.2.1 Cognitive Level Tests

The steps in which cognitive tests are generated in the PISA project include; defining the general framework, generation of items, receiving items from countries, evaluation of national items, translation into national languages, implementation of the preliminary application, preparation of the basic application by carrying out the preliminary application and the implementation of the basic application (MoNE, 2016).

Item format is limited with selected-response, open constructed-response and closed constructed-response items, which have different weights in the cognitive field subtests conducted to students within the framework of PISA 2015 application.

Selected-response items ask students to select one or more responses from a list of alternatives. Simple multiple choice questions are structured questions that consist of 4 or 5 answer options, depending on a stem of the question. One of the options is “full right” and the others are incorrect. Complex multiple choice questions are questions that are connected to the stem of the question and are composed of more than one multiple choice questions. The questions come one at a time and are interrelated. In such questions, “partial correct” answers are often not assumed. They call for responses to a series of related “Yes/No” questions that are treated for scoring as a single item, selection of more than one response from a list, completion of a sentence by selecting drop-down choices to fill multiple blanks and “drag-and-drop” responses which allow students to move elements on screen to complete a task of matching, ordering or categorising.

Open constructed-response items have no boundary in answering, depending on the root of the question. They require an extended written response, such as explaining the steps students have taken while finding the answer of a problem.

Closed constructed-response items have boundary in answering. The responses can be single numbers or can include a few words and they can be easily evaluated as correct or incorrect (OECD, 2017a).

In the PISA study, plausible values have been estimated for each field in terms of cognitive field competence at student-level. In the present study, the estimation is done by using ten plausible values for the science performance of each student.

Plausible values are not test scores and should not be treated as such. Plausible values are random numbers that are drawn from the distribution of scores that could be reasonably assigned to each individual. The plausible values contain random error variance components and are not optimal as scores for individuals. The advantage of plausible values is that as a set they are better suited to describing the performance of the population than is a set of scores that are optimal at the individual student-level. The plausible value approach produces consistent estimators of population parameters provided that the imputation model is proportional with the data analytic model. Plausible values are intermediate values that are provided so that consistent estimates of population parameters can be obtained using standard statistical analysis software such as SPSS and SAS (Adams, Wilson, & Wu, 1997). PISA provides ten plausible values per scale or subscale. If an analysis is to be undertaken with one of these ten cognitive scales, then the analysis should be undertaken ten times, once with each of the ten relevant plausible values variables. The results of these ten analyses are averaged and then significance tests, that adjust for variation between the ten sets of results, are computed (Monseur & Adams, 2009).

3.2.2 Context Questionnaires

In addition to information about cognitive skills of students, PISA also needs valid and reliable information about students' non-cognitive outcomes, their socio-economic and cultural backgrounds and institutional characteristics of schools. For this purpose; students, their parents, teachers and school principles are asked to respond to the questionnaire items constructed by PISA. The respondents are expected to be knowledgeable about non-cognitive outcomes, student background, teaching and learning processes, and school policies and educational governance.

Context questionnaires in PISA give information about the individual, school and the context of learning in education at the system, country or economy level. Context questionnaires evaluate non-cognitive outcomes, students' tendencies, and structural processes.

School questionnaires responded by the school principles and student questionnaires responded by all participating students are not optional for countries and economies participating in the study. Educational career questionnaire and information and

communications technology familiarity questionnaire are optional for students. Parent and teacher questionnaires are optional and used for obtaining data from students' parents and teachers.

The questionnaires used in the PISA 2015 study can be explained as follows:

1) Student Questionnaire (Computer-based and paper-based): It was applied to all students participating in the PISA study.

2) School Questionnaire (Computer-based and paper-based): It was applied to all school principals participating in the PISA study.

3) Education Career Questionnaire (Computer based): This optional questionnaire was applied to all students participating in the PISA study in the country or economy that opt to administer it. It contains additional questions about students' educational backgrounds and updates with a focus on additional teaching and learning time.

4) Information and Communication Technology (ICT) Familiarity Questionnaire (Computer-based): This optional questionnaire was applied to all students participating in the PISA study in the country or economy that opt to administer it. It includes questions about the use of electronic devices by students and the availability of ICT at home or school and about the number of such devices. In addition to digital devices, this questionnaire measures students' confidence and attitudes towards ICT.

5) Parent Questionnaire (Paper-based): This optional questionnaire was applied to the parents of all students participating in the application. It contains questions about students' learning contexts, parental support, materials and resources at home, parents' educational status, parents' economical status, interest and attitude towards science.

6) Teacher Questionnaire (Computer based): The main idea behind this optional questionnaire was asking questions directly to the teachers about their competence and teaching and learning strategies. PISA 2015 prepared two different teacher questionnaires, taking into account the main area of science and teacher characteristics and practices (OECD, 2017a).

The development of the questionnaires has been established with the participation of international experts and international co-operation at the OECD headquarters and national centers. The development stages of the questionnaires can be explained as follows:

- i. Translating pre-application materials into the languages of participating countries
- ii. Preliminary application of the questionnaires and cognitive level tests in all participating countries
- iii. Analysis of the preliminary application data
 - a. Detection of the items answered or left blank
 - b. Examining the forms of the questionnaires containing different question formats
 - c. Explanatory and confirmatory factor analyzes
 - d. Validity analysis of items used in different countries according to the structures of items
 - e. Statistical characteristics of items according to classical test theory and item response theory
- iv. Comparison of preliminary application results
- vi. Enabling final forms

The questionnaires included trend indicators or single items. Trend indicators were used to report the trend over time. Most of the questionnaire items measure a latent construct, which cannot be observed directly. They are called “derived variables” such as students’ motivation and economic, social and cultural background. Derived variables are transformed or scaled in order to construct meaningful indices (OECD, 2017b).

3.3 Validity and Reliability

In order to ensure the validity and reliability of the data of the participating countries; the type and the scope of data, data collection tools and data to be measured are formed by experts in the participating countries. In order to reduce cultural and

linguistic differences in all countries, data collection, translation and sampling procedures are carried out carefully. At the same time, longitudinal studies conducted in some countries demonstrates a strong relationship between PISA science performances and career and work life. Through these findings, educational outcomes, which are obtained from the economically developed and developing countries of the world, can be interpreted better. These indicate that the PISA 2015 data have high validity and reliability.

In this section, the development of cognitive assessment instrument, development of questionnaire and quality monitoring were explained.

3.3.1 Content-Related Evidence for Validity

PISA 2015 was applied within the framework determined by the PISA Governing Board (PGB). PGB consists of representatives from all participating countries and economies. The members of PGB are responsible for setting standards and designing assessment instruments and for reporting results. While constructing context questionnaires, PGB experts use both programme policy objectives and international technical expertise. By reviewing the outcomes of the meetings of Subject Matter Expert Groups (EGs) and the Questionnaire Expert Group (QEG), countries and economies can consider about the international validity of the instruments. In order to evaluate the reliability of the assessment materials and educational validity, the cultural and educational contexts of participating countries should be considered.

National Project Managers (NPM) from each participating country implement PISA nationally with internationally agreed common technical and administrative procedures. NPMs are also responsible for developing and validating the international assessment instruments. They are also required to verify and evaluate survey results and analyses (OECD, 2017b).

3.3.1.1 Design of Cognitive Level Test and Test Development

In PISA 2015, optional computer-based modules were used in addition to paper-based modules. Through computer-based modules new aspects of science constructs can be measured with the help of interactive tasks and simulations. In PISA 2015, minor domains were covered and expert groups focused on improved trend

measurements. Since the domain construct coverage is increased through the increase in the number of items for minor domains, systematic measurement error is reduced. This integrated assessment design helped to make an improved comparison across countries by stabilising estimations and the measurement of trends and to make more reliable inferences from the data.

In order to test the invariance of item parameters for the paper-based trend items used in previous cycles, these items were adapted for computer-based assessment. Items having different formats such as multiple-choice or open ended were applied to randomly assigned groups in order to assess the reliability. The scoring was done by computers or expert coders and the PISA Technical Advisory Group evaluated the results.

The newly-developed trend items in minor domains of the PISA 2015 study were proposed and discussed by the Subject Matter Expert Groups (EGs) and adapted for computer-based assessment. The experts also played roles in the design of the instrument development and data analysis processes.

Before applying main survey, a national item review process was conducted. Countries reported the layout and linguistic issues before and after trial data, and errors were corrected centrally by the contractors. After field trial, the test development team recommended the items for the main survey according to country comments, domain coverage of the items, item format and assessment design. Scientific literacy item pool was reviewed by an independent item review and finally approved for the main study item pool. Table 3.2 shows the number of items for each domain in both field trial and main survey. Table 3.3 shows the distribution of 184 scientific literacy items, which represent the distribution recommended in the PISA 2015 framework.

After evaluating technical problems, scoring inconsistencies and application related issues, analyses were conducted. Analyses included item analysis, missing data treatment, coding, scoring and plausible value generation.

Table 3.2 Item Counts (Field Trial and Main Survey) by Domain and Delivery Mode

Domain	Field Trial		Main Survey	
	Paper-based	Computer-based	Paper-based	Computer-based
Science	91	304 (213 new, 91 trend)	85	184 (99 new, 85 trend)
Mathematics	83	82	83	81
Reading	103	103	103	103
Collaborative problem solving	NA	153	NA	117
Financial literacy	NA	43	NA	43

Adapted from *PISA 2015: Technical Report*

Table 3.3 Science Item Counts by Framework Category

	Items	Percent	Framework goal
Competency			
Evaluate and design scientific enquiry	39	21%	20-30%
Explain phenomena scientifically	89	48%	40-50%
Interpret data and evidence scientifically	56	31%	30-40%
Knowledge			
Content	98	53%	54-66%
Epistemic	26	14%	10-22%
Procedural	60	33%	19-31%
System			
Earth and Space	49	27%	28%
Living	74	40%	36%
Physical	61	33%	36%

Adapted from *PISA 2015: Technical Report*

3.3.1.2 Development of Context Questionnaires

PISA 2009 focused on the hierarchical structure of education, while PISA 2012 focused on the educational systems. In the PISA 2015 framework, these approaches are combined with new aspects of policy interest, which is mainly about the educational effectiveness.

PISA 2015 context questionnaire framework consists of 19 policy issues (modules). These modules form the basis of the development of the questionnaire items.

PISA aims to measure trend indicators across cycles, since trend items are the core content of questionnaire material and have higher priority in instrument

development. The PISA context questionnaire framework was guided by the Questionnaire Expert Group (QEG) according to the questionnaire drafts and feedback from countries and economies.

In the development process of questionnaires, all policy modules were represented in several questionnaires.

For quality assurance and comparability of data across countries and economies, newly-developed items were evaluated by the representatives of countries and economies to obtain in-depth feedback. The national review enabled the relevance of specific topics for the educational system. Feedbacks were about whether a topic might be politically acceptable, may cause cultural bias or could have translation and adaptation difficulties. Finally, after national reviews, proposed questions were rephrased or deleted.

Newly-developed questionnaire items were pre-tested with small groups of students and teachers. The time to read and answer the questions was recorded, and respondents were asked whether they understood the question and the answer and whether they have any other comment. And teachers were interviewed about the newly-developed teacher questionnaire items to revise the proposed questionnaire material. Linguistic experts optimized the words in the questionnaire items to be able to translate them into other languages properly by avoiding bias in the meaning of the question.

The questionnaire items were adapted by countries or economies to prevent the problems, which might be related to cultural traditions, national understanding of a question or characteristics of the education system.

3.3.2 Construct-Related Evidence for Validity

PISA aims to develop measurement tools for comparing student backgrounds, attitudes, perceptions and practices across countries. Through the measures derived from these tools (i.e. questionnaire items), differences in student performance can be predicted and information is provided to policy makers for improving educational systems. Hence, the development of methodologies for constructing questionnaires with high validity is an important issue.

PISA intends to measure the same content in different national and cultural contexts. Cultural differences can effect self-reported beliefs and attitudes of students from different countries. So, an appropriate translation procedure and standardised administration is needed for cross-country validity of the constructs.

Psychometric techniques can be used to evaluate the cross-country validity, i.e. the degree of comparability across countries. For this purpose, “internal consistency” of the scales within and between countries can be checked by using Crohnbach’s alpha coefficient. Similar and high values across countries indicate that the measurement across countries has a high reliability. International item parameters, which are estimated by using information from all students in all countries, and language groups in these countries, can be used to ensure the “invariance of item parameters”.

3.3.3 Reliability

In PISA assessment design, marginal reliabilities cannot be calculated for each cognitive domain. Test reliability can be estimated through the explained variance for each cognitive domain by using plausible values.

Table 3.4 demonstrates the median of national reliabilities for the generated scale scores (Cronbach’s alpha) based on all 10 plausible values. It shows the explained variance by the combined IRT and latent regression model (population or conditioning model) across countries.

The reliabilities in the population model are not classical reliability coefficients. It is not possible to compare individual students, since the apparent accuracy of the measures is obtained through statistical adjustments of the estimates based on background data. The plausible value approach is not suitable for comparisons of individuals, however it provides improved behavior of subgroup estimates (Wu, 2005).

Table 3.4 Reliabilites of the PISA Cognitive Domains and Science Subscales Overall Countries

Mode	Domains	Median	S.D.	Max	Min
Computer-Based Assessment	Maths	0.85	0.03	0.90	0.75
	Reading	0.87	0.02	0.90	0.80
	Science	0.91	0.02	0.93	0.82
	CPS	0.78	0.03	0.83	0.70
	Financial literacy	0.83	0.06	0.93	0.72
	Science subscales				
	Explain phenomena scientifically	0.89	0.003	0.91	0.80
	Evaluate and design scientific inquiry	0.87	0.04	0.90	0.71
	Interpret data and evidence scientifically	0.89	0.03	0.92	0.78
	Content	0.89	0.02	0.91	0.81
	Procedural and epistemic	0.90	0.03	0.92	0.78
	Earth and science	0.88	0.03	0.90	0.77
	Living	0.89	0.03	0.91	0.79
	Physical	0.88	0.03	0.91	0.76
Paper-Based Assessment	Maths	0.80	0.05	0.87	0.67
	Reading	0.82	0.04	0.88	0.72
	Science	0.86	0.04	0.92	0.77

Adapted from PISA 2015: Technical Report

Table 3.5 demonstrates the reliabilities for main cognitive domains for Turkey. The reliability coefficient for the plausible value of science knowledge scale is 0.89 for Turkey, indicating a satisfactory reliability of the measure (OECD, 2017b).

Table 3.5 Reliabilities for Main Cognitive Domains for Turkey

Mode	Maths	Reading	Science	CPS	Finacial literacy
Computer-Based Assessment	0.82	0.85	0.89	0.74	-

Adapted from PISA 2015: Technical Report

3.4 Procedures

The research questions of the study and the threats to the internal validity were explained below.

3.4.1 Research Questions

In the present study, it was aimed to investigate the following research questions:

1. Are there any differences in the scientific literacy skills of Turkish students across schools?
2. What proportions of variances in the scientific literacy scores of Turkish students is explained by school-related variables?
3. Which student characteristics are significantly related to Turkish students' mean combined science and scientific literacy subscales scores?
4. Which school characteristics are significantly related to Turkish students' mean combined science and scientific literacy subscales scores?
5. How do the significant school-level variables effect the relationship between student-level variables and outcome variable?

The sub-questions of the present study are:

1. Which school-level variables are associated with differences in combined science scores of Turkish students?
2. Which school-level variables are associated with differences in content knowledge scores of Turkish students?
3. Which school-level variables are associated with differences in procedural and epistemic knowledge scores of Turkish students?
4. Which school-level variables are associated with differences in explaining phenomena scientifically scores of Turkish students?
5. Which school-level variables are associated with differences in evaluating and designing scientific inquiry scores of Turkish students?
6. Which school-level variables are associated with differences in interpreting data and evidence scientifically scores of Turkish students?
7. Which student-level variables are associated with differences in combined science scores of Turkish students?

8. Which student-level variables are associated with differences in content knowledge scores of Turkish students?
9. Which student-level variables are associated with differences in procedural and epistemic knowledge scores of Turkish students?
10. Which student-level variables are associated with differences in explaining phenomena scientifically scores of Turkish students?
11. Which student-level variables are associated with differences in evaluating and designing scientific inquiry scores of Turkish students?
12. Which student-level variables are associated with differences in interpreting data and evidence scientifically scores of Turkish students?

3.4.2 Threats to Internal Validity

The international assessment items in PISA are translated into multiple target languages. The translation of some terms may be harder or some items may function differently in different countries. Same threat can be observed when mode of application is changed from paper to computer-based assessment. Comparison of items across different assessment cycles over years can also be a threat to validity (OECD, 2017b).

In PISA 2015, no treatment occurs. PISA is an assessment study and consists of one-time testing, which does not take more than 3 hours. Therefore history, maturation and subject attrition cannot be considered as threats. The items used in the context questionnaires and in the cognitive tests are not released by PISA and the items are used only one time for each cycle. Since the participants are not familiar with the items, their performance may not differ. Therefore, testing cannot be considered as a threat to this study.

Selection can be a threat, since it is not possible to control all characteristics of students. There could be initial differences among students or oversampling in some countries can be a problem.

Since PISA conducts the application by providing comprehensive instructions and manuals for test administrators and graders, and trained people carry out all

procedures, instrumentation is not a threat in this study. The chance factors of extreme scorers is not an issue for PISA. Since there is no intervention in PISA 2015, regression cannot be considered as a threat to the present study.

3.5 Data Collection

Apart from cognitive tests, in which their academic performance is measured, PISA 2015 study consists of school and student questionnaires, where students are evaluated as a whole.

Cognitive tests aim to measure the students' performance in mathematics literacy, reading skills, and with a particular emphasis on science literacy. 57 of the 72 countries participating in PISA 2015 were subjected to computer-based assessment and 15 of them were paper-based assessment. 66 different booklets were used for computer-based assessment and 30 different booklets were used for paper-pen-based assessment. The booklets were administered in four sets of 30 minutes each. In the PISA 2015 application, which consists of two sessions, 60 minutes are given for each session. There were 5-10 minute breaks between sessions. After the 15-minute break given after the cognitive test sessions, the students answered the student questionnaire, which lasted approximately 35 minutes (MoNE, 2016).

3.6 Variables of the Study

The dependent variables of this study are combined science scores and scores of knowledge (content knowledge and procedural and epistemic knowledge) and competency (explaining phenomena scientifically, evaluating and designing scientific inquiry and interpreting data and evidence scientifically) components of scientific literacy. There were 184 cognitive items in forms of multiple-choice, short closed-constructed response, and open-constructed response formats. For the assessment of explaining phenomena scientifically 89 items, for evaluating and designing scientific inquiry 39 items, and for interpreting data and evidence scientifically 56 items were devoted. For content knowledge 98 items, for epistemic knowledge 26 items and for procedural knowledge 60 items were devoted (OECD, 2017b).

The independent variables of this study are student- and school-level variables, which are mainly about the background and attitudinal factors of students and school characteristics. These variables are obtained from student and school questionnaires.

The questionnaire items in PISA can be indicators, which report changes over time or they can be single items such as gender. There are also many questionnaire items, which measure latent constructs that cannot be observed directly. These items were transformed or scaled in order to construct meaningful indices, and are referred to as “derived variables”. Some of them were trend variables, which report changes over time.

Derived variables can be simple questionnaire indices including one or more items, or they can be based on item response theory scaling, or they can be ESCS (index of economic, social and cultural status) composite scores.

The indices used in PISA summarise responses of students, their parents, teachers or school principals to a series of related questions. In order to validate the comparability of indices across countries and to ensure the theoretically expected behaviour of the indices, structural equation modeling was used (OECD, 2017b).

Simple indices, new scale indices, and trend scale indices were used in PISA 2015. For simple indices, one or more items are recorded or transformed arithmetically in the same way across assessments. For new and trend scale indices, multiple items are scaled by using a two parameter item response model and values of the index correspond to Warm’s weighted likelihood estimates (WLE) (Warm, 1989).

The values of the independent variables were obtained through the answers given to the questions in the school and student questionnaires, and the values of the dependent variables were obtained through the answers given to the questions in the cognitive assessment items. Each module in PISA was represented by the questions in the school and student questionnaires. Since only school and student questionnaires were administered in Turkey, only these surveys were included in the scope of this thesis. The questionnaire items related to school- and student-level variables of the thesis are listed in Appendix A and Appendix B respectively.

3.6.1 Student-Level Variables

Within the scope of the present study, student background variables, non-cognitive outcomes and teaching and learning variables were examined as student-level variables. The independent student-level variables, the questions in the student questionnaire and the related items included in the scope of the thesis are demonstrated in Table 3.6. Technical characteristics of student-level variables are given in Appendix C.

3.6.2 School-Level Variables

Within the scope of the present study, teaching and learning variables and school policies variables were examined as school-level variables. The independent school-level variables (derived variables: simple indices and IRT-based scales), the questions in the school questionnaire and the related items included in the scope of the thesis are demonstrated in Table 3.7. Technical characteristics of school-level variables are given in Appendix D.

3.7 Data Analysis

For this study, a two-level model was constructed with level-1 units (students) and level-2 units (schools). Level-1 units were nested in level-2 units. Submodels of the study 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.

In the present study, hierarchical linear modeling was performed by using HLM 8.00 program for the analysis of the data obtained from the PISA 2015 study. HLM is used instead of standard regression because students are clustered in schools. The violation of the independence of observations assumption is eliminated by HLM analysis by considering that students are clustered in schools. In summary, HLM is a regression analysis that aims to create a model in which the dependent variable is predicted by independent variables at different levels such as student, classroom and school.

Table 3.6 Student-Level Variables of the Study

Derived Variable	Description	Question	Related Item No
A. Student Background Variables			
• Out-of-school science experience (Module 5)			
OUTHOURS (ST071Q01NA)	Out-of-school study time per week	ST071: This school year, approximately how many hours per week do you spend learning in addition to your required school schedule in the following subjects?	ST071Q01NA
SCIEACT	Science activities	ST146: How often do you do these things?	All items in ST146
B. Non-Cognitive Outcomes			
• Science-related outcomes (Motivation, interest, beliefs) (Module 4)			
SCIEEFF	Science self-efficacy	ST129: How easy do you think it would be for you to perform the following tasks on your own?	All items in ST129
INTBRSCI	Interest in broad science topics	ST095: To what extent are you interested in the following <broad science> topics?	All items in ST095
JOYSCIE	Enjoyment of science	ST094: How much do you disagree or agree with the statements about yourself below?	All items in ST094
INSTSCIE	Instrumental motivation	ST113: How much do you agree with the statements below?	All items in ST113
EPIST	Epistemological beliefs	ST131: How much do you disagree or agree with the statements below?	All items in ST131
ENVAWARE	Environmental awareness	ST092: How informed are you about the following environmental issues?	All items in ST092
ENVOPT	Environmental optimism	ST093: Do you think problems associated with the environmental issues below will improve or get worse over the next 20 years?	All items in ST093
• Career Aspirations (Module 6)			
BSMJ	Students' expected occupational status (SEI)	ST114: What kind of job do you expect to have when you are about 30 years old?	All items in ST114
C. Teaching and Learning Variables			
• Learning time and curriculum (Module 12)			
SMINS	Learning time (minutes per week) – <science>	ST059: How many <class periods> per week are you typically required to attend for the following subjects?	ST059Q03TA
		ST061: How many minutes, on average, are there in a <class period>?	ST061Q01NA

Adapted from PISA 2015: Technical Report

Table 3.7 School-Level Variables of the Study

Derived Variable Name	Description	Question	Related Item No
A. Teaching and Learning Variables			
• Teacher qualification and professional knowledge (Module 1)			
PROSTAT	Proportion of science teachers by all teachers	SC018: How many of the following teachers are on the staff of your school? SC019: How many of the following teachers are on the <school science> staff of your school?	SC018Q01TA SC019Q01NA
PROSTCE	Proportion of science teachers fully certified	SC019: How many of the following teachers are on the <school science> staff of your school?	SC019Q02NA
PROSTMAS	Proportion of science teachers with ISCED level 5A and a major in science	SC019: How many of the following teachers are on the <school science> staff of your school?	SC019Q03NA
TOTST	Total number of science teachers at school	SC019: How many of the following teachers are on the <school science> staff of your school?	SC019Q01NA
B. School Policies Variables			
• School-level learning environment for science (Module 3)			
EDUSHORT	Shortage of educational material	SC017: Is your school's capacity to provide instruction hindered by any of the following issues?	SC017Q05NA SC017Q06NA SC017Q07NA SC017Q08NA
STAFFSHORT	Shortage of educational staff	SC017: Is your school's capacity to provide instruction hindered by any of the following issues?	SC017Q01NA SC017Q02NA SC017Q03NA SC017Q04NA
SCIERES	Science specific resources	SC059: Which of the following are true for the science department of your school?	All items in SC059
CREACTIV	Creative extra-curricular activities	SC053: <This academic year>, which of the following activities does your school offer to students in the <national modal grade for 15-year-olds>?	All items in SC053

Adapted from PISA 2015: Technical Report

By using Turkey data, null model (one-way ANOVA with random effects), the level-1 model (random-coefficients regression model) and level-2 model (means-as-outcomes regression model) were created.

In the null model, it is interpreted how meaningful it is for the students to be clustered in schools without including any variable from the first and second levels. In addition, it is aimed to reveal to what extent the student-level variables in level-1 model, and the school-level variables in level-2 model predict the level of scientific literacy of students.

3.7.1 Hierarchical Linear Modeling

Behavioral and social data generally have a hierarchical (nested) structure, that is, the individual subjects of a study may be classified or arranged in groups which themselves have qualities that influence the study. In this case, the individuals can be seen as level-1 units of study, and the groups into which they are arranged are level-2 units. For example, in educational research, students (level-1) are nested in schools (level-2), and the schools are nested in school districts (level-3) and similarly in sociology, individuals (level-1) are nested in neighborhoods (level-2). The analysis of such data requires specialized software. Hierarchical linear models have been developed to allow for the study of relationships at any level in a single analysis, while not ignoring the variability associated with each level of the hierarchy.

For the present study, a two-level model was constructed with its submodels. The submodels, running from the simpler to the more complex, include the one-way ANOVA model with random effects, a regression model with means-as-outcomes, random-coefficients regression model, a model with intercepts- and-slopes-as-outcomes.

3.7.1.1 One-Way ANOVA Model with Random Effects

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

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

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 the error term representing the unique effect associated with observation i .

In 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 effect is:

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

where γ_{00} shows the grand-mean outcome in the population, and u_{0j} represents the random effect associated with unit j , and 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 the one-way ANOVA model with grand mean, γ_{00} ; with a group (level-2) effect, u_{0j} ; and with a person (level-1) effect, r_{ij} . This model is a random-effects model, since the group effects are constructed as randoms. The variance of the outcome is:

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

Estimating the one-way ANOVA model is often useful as a preliminary step in a hierarchical data analysis, since it produces a point estimate and confidence interval for the grand mean, γ_{00} . It also provides information about the outcome variability at each of the two levels. The σ^2 parameter shows the within-group variability, and τ_{00} represents the between-group variability. The hierarchical model is also referred to as

fully unconditional model, since no predictor variables are specified at level-1 and level-2 equations of the model.

The intraclass correlation coefficient is given by formula:

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

and this coefficient measures the proportion of variance in the outcome that is between the level-2 units. It is the average correlation between two randomly chosen units that are in the same group (Raudenbush & Bryk, 2002, p. 23-24).

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 means-as-outcomes regression model 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 W_j is assumed as the only level-2 predictor. The combined model becomes:

$$Y_{ij} = \gamma_{00} + \gamma_{01}W_j + u_{0j} + r_{ij}.$$

Here, the random variable 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. 24-25).

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 (level-1 slope) did not exist in those models.

A major class of applications of hierarchical linear models includes studies in which, level-1 slopes are considered as varying randomly over the population of level-2 units. Random-coefficients regression model is the simplest case of this type. Both the level-1 intercept and one or more level-1 slopes vary randomly in these models, however no attempts is made to predict this variation.

Level-1 equation in a simple model is:

$$Y_{ij} = \beta_{0j} + \beta_{1j}(X_{ij} - \bar{X}_j) + r_{ij}$$

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 ; and

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:

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

where

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

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

$\text{Cov}(u_{0j}, u_{1j}) = \tau_{01}$ = 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 becomes:

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

This combined model implies that the outcome Y_{ij} is a function of the average regression equation, $\gamma_{00} + \gamma_{10}(X_{ij} - \bar{X}_{.j})$ plus a random error having three components: u_{0j} , the random effect of unit j on the mean; $u_{1j}(X_{ij} - \bar{X}_{.j})$, where u_{1j} is the random effect of unit j on the slope β_{1j} ; and the level-1 error, r_{ij} (Raudenbush & Bryk, 2002, p. 26-27).

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 multiple level-2 predictors. Hierarchical linear models may involve multiple level-1 predictors where any combination of random, non-randomly varying, and fixed slopes can be specified (Raudenbush & Bryk, 2002, p. 27-28).

3.7.2 Centering the Variables

In order to evaluate the statistical results by considering the theoretical framework of the research, variables should have precise meaning. In hierarchical linear models, the intercept and slopes in the level-1 model become outcome variables at level-2. To ensure the numerical stability of the value of a variable, a proper location can be chosen in estimating hierarchical linear models.

The meaning of the intercept in the level-1 model depends on the location of the level-1 predictor variables (X s). In order to explain the variation in the intercept β_0 , all level-1 variables were centered around their group means ($X_{ij} - \bar{X}_{.j}$).

Since the numerical instability is less likely, centering of level-2 variables is not as critical as centering of level-1 variables. In general, all of the level-2 variables are centered around their corresponding grand means (Raudenbush & Bryk, 2002).

By mean centering, the results can be easily interpreted, high correlations between level-1 and level-2 variables, cross-level interactions and high correlations between random intercepts and slopes can be removed.

According to Algina and Swaminathan (2011) main effects of the predictors involved in the interaction are affected by centering. Bauer and Curran (2005) recommend group-mean centering of level-1 predictors and grand-mean centering of level-2 to improve computation and interpretation of the main effects when cross-level interactions are tested. Centering level-1 variables around their group means is associated with high reliability.

In this study, level-1 variables were centered around their corresponding group means, while level-2 variables were centered around their grand means.

3.7.3 Random and Fixed Effects

In hierarchical linear models, the decision to take an effect as fixed or random is of great importance, since a wrong decision causes bias in the results. When the slope is non-randomly varying, it varies strictly as a function of the predictor W_j (Raudenbush & Bryk, 2002).

In this study, initially, all effects were accepted as random. As the outputs of HLM program, tables showing fixed effects and variance components are used to decide on the type of the effect. For this purpose, 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, for random effect components having insignificant results, the error terms (u) having insignificant error variances were turned off in order to determine randomly varying slopes. A similar strategy was used to decide on the variables to be added into level-1 and level-2 equations during the model construction.

3.7.4 Outcome Variable

In this study, the outcome variables are combined scientific literacy, science knowledge and science competencies scores of students. In the assessment process, all of the science items are not responded by all students, students responded to different subsets of items. So, scientific literacy scores of the students are not observed scores, they are imputed scores. PISA reports students' performance by using plausible values. Ten plausible values are allocated to each student on each performance scale. Using more than one plausible values enables estimating imputation error. Analyses are conducted for each of these ten plausible values separately. The results of these analyses are aggregated to obtain the final estimates of the statistics and respective standard errors. HLM program helps to make estimations automatically by using plausible values.

Plausible values are multiple imputations of the unobservable latent achievement for each student, so they are not actual test scores. They are random numbers obtained from the marginal posterior distribution. Plausible values include random error variance components and they cannot be used as the scores for individuals. Plausible values are more suitable for describing the performance of the population as a whole rather than estimating individual abilities. (Wu, 2005).

The PISA 2015 student file contains 10 plausible values for each of the scientific literacy subscales. For combined scientific literacy score, the plausible value variables are PV1SCIE to PV10SCIE. For science knowledge subscales; content knowledge and procedural and epistemic knowledge, the plausible values variables are PV1SKCO to PV10SKCO and PV1SKPE to PV10SKPE, respectively. For science competencies subscales; explaining phenomena scientifically, evaluating and designing scientific inquiry, and interpreting data and evidence scientifically, the plausible values variables are PV1SCEP to PV10SCEP and PV1SCED to PV10SCED and PV1SCID to PV10SCID respectively.

Since plausible values cannot be treated as actual scores, ten plausible values for each subscales of scientific literacy are included in the multivariate data matrices (MDM) of the HLM 8.00 computer program. HLM uses the ten plausible values while generating the estimates. The program runs with each of the specified plausible

values internally, and computes their average value and the correct standard errors. Finally, one estimate and a measurement error is generated. This estimate is the average of the specified plausible values. Through this procedure, plausible value data can be treated accurately.

3.7.5 Scaling PISA 2015 Data

In the PISA test design, a subset of items is assigned to each student from a total item pool. This means that different groups of students answer different items but the items in the subsets overlap. Since, students took different sets of items and test forms might have different level of difficulty levels, using the number of correct items while reporting survey results is not appropriate. In addition, item-by-item reporting does not consider the dissimilarities of proficiencies of subgroups who answered these set of items. The mean proficiency of students in a given subpopulation can be estimated with the average percentage of items answered correctly. However this does not give information about variances, i.e. how the skills are distributed within that subpopulation.

Item response theory (IRT) scaling can be used to overcome these limitations. By using the underlying commonalities among the items, regularities of response patterns can be modelled. This regularity can be used to define the item characteristics and the characteristics of students who did not take identical sets of items. With this procedure, the distributions of performance in a population or subpopulation can be described and the relationships between proficiency and student- and school-related factors can be estimated.

In order to obtain more accurate measurement, plausible values are combined with the IRT scaling of the test items with a latent regression model using information from the student context questionnaire in a population model (OECD, 2017b).

3.7.6 Hypotheses Testing

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

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

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

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

One approach to multiparameter test for fixed effects is to test whether a set of level-2 predictors add significant explained variance to an outcome. Multiparameter tests regarding γ are useful for omnibus tests of the relationship between a categorical level-2 predictor and a β_{qj} ; contrasts between categories of a level-2 predictor; examining whether a level-2 characteristic 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 can be formulated as:

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

For multiparameter test of random level-1 coefficients, a comparison is made among two or more components of random-coefficients vector.

In the hierarchical analysis, the researchers need to decide whether level-1 coefficients should be specified as fixed, random or non-randomly varying. In order to determine whether random variation exists, following typical null hypothesis is used for single-parameter test for variance and covariance components:

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

In multiparameter test of variance and covariance components, the most general form of hypothesis testing for variances and covariances is based on a likelihood-ratio test (Raudenbush & Bryk, 2002, p. 57-65).

3.7.7 Weighting of Data

In order to obtain unbiased population parameter estimates, statistical analyses or procedures about the PISA, data should be weighted. Due to stratified sampling, differential participation rates of schools or students required adjustments; and due to

oversampling of some explicit strata for national reporting purposes, students and schools in a particular country did not have the same probability of selection in PISA. Hence, weights are associated to each student and to each school.

PISA database files include weight variables. The weight included in the database is multiplied by the total number of students and divided by the weighted total number of students and the same procedure was conducted for schools. So, the sum of the weights is equal to the number of observations.

3.7.8 Assumptions of Two-Level Hierarchical Linear Model

Comparison of multilevel standard errors with robust standard errors is the main assumption test for hierarchical linear modeling analysis. Considerable differences between these errors reveal violations of an important assumptions of HLM (Hox & Maas, 2004).

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.

In a two-level hierarchical linear model; level-1 error, r_{ij} is assumed to be independent and normally distributed with a mean of zero and a constant level-1 variance, σ^2 . It is also assumed that the level-1 predictors, X_{qij} , are independent of r_{ij} . The vectors of random errors at level-2 are assumed to be multivariate normal, and each has a mean of zero, variance τ_{qq} , and covariance among the random elements. The set of level-2 predictors are assumed to be independent of every u_{qj} and the errors at level-1 and level-2 are assumed to be independent. In addition, it is assumed that the predictors at each level are not correlated with the random effects at the other level (Raudenbush & Bryk, 2002, p. 255).

3.8 Missing Data

Missing data is one of the most commonly seen problems in data analysis step of social sciences research. In multilevel studies, explanatory and outcome variables may be subjected to missingness at any of the levels. It is important to use a multilevel imputation model in multilevel studies, since a single level imputation may cause significant bias.

Missing data could be blank, no response, invalid, not applicable or valid skip. Blank refers to the items, which were not responded, because respondent ended the test early or did not take the questionnaire despite the opportunity. No response or omit refers to items, which were not responded despite the respondent had the opportunity to answer. Invalid refers to the items, in which respondent chose more alternatives than required or the answer was not in an acceptable range or the country requested to suppress the item. Not applicable refers to the responses, which cannot be determined due to misprinting. Valid skip refers to not provided responses, because respondent is told to skip the next item in the previous question (OECD, 2017b).

The amount of missing data, why it is missing and the pattern of missing data are important issues. The pattern of missing data should be examined, since non-randomly missing values may affect the generalizability of results. If only a few data points are missing in a random pattern from a large data set, almost any procedure for handling missing values gives similar results. If only a few cases have missing data with random pattern, deletion can be a good alternative. However, if a lot of data are missing, there can be serious problems. The safest thing to deal with missing data is to test the patterns in it.

Missing data can be MAR (missing at random), MCAR (missing completely at random) or MNAR (missing not at random). IBM SPSS MVA (Missing Value Analysis) is used to determine the patterns of missing values and to replace them in data set. A t-test is examined to see if missingness is related to any of other variables with $\alpha = 0.05$, and it is conducted with variables having at least 5 percent of data missing. The expectation maximization (EM) method is used for determining randomly missing data. EM forms a missing data correlation (or covariance) by assuming the shape of a distribution for the partially missing data and basing

inferences about missing values on the likelihood under that distribution. IBM SPSS MVA performs EM to assess patterns of missing data by providing t-test to predict missingness from other variables in data set and by testing for MCAR. If the result in Little's MCAR test is statistically significant, this means that the pattern of missing values is not random so, MNAR should be considered. In this case, multiple imputation procedure is used. A data set is generated with M subsets (default M=5) with different imputations of missing data in each subset, and the analysis is run (Tabachnick & Fidell, 2013).

There are three options for handling missing data at level 1 in HLM2: listwise deletion of cases when the MDM file is made, listwise deletion of cases when running the analysis, and analysis of multiply-imputed data. At level 2, HLM2 assumes complete data. HLM conducts an automatic listwise deletion of level-2 units with missing data when the MDM file is created.

HLM 8.00 computer program automatically searches the data to find out the variables with missing values; then it estimates a multivariate hierarchical linear model "imputation model" in which all variables having missing values are regressed on all variables having complete data; then it uses the resulting parameter estimates to generate M imputed data sets; then it analyzes each of these in turn and combines the results by using the "Rubin rules". Missing data problem can be solved through multiple, model-based imputation (Rubin, 1987, Little & Rubin, 2002, Schafer, 1997). A multiple imputation procedure produces M "complete" data sets. HLM2 can be applied to these multiply-imputed data to produce appropriate estimates that incorporate the uncertainty resulting from imputation.

HLM conducts a completely automated procedure to handle ignorable missing data in two-level models (Shin & Raudenbush, 2013). Firstly, a two-level HLM model with multiple outcome variables, including a mixture of level-1 and level-2 outcomes is specified. The variables which have missing values are searched, a multivariate imputation model is estimated and multiple imputed data sets are generated by the program. Finally, the analyses are conducted according to the model specified by the user; and the program averages the results by using "Rubin's rules" (Rubin, 1987).

The main process in the multiple imputation is to re-express a desired hierarchical model as the joint distribution of the outcomes and all variables that are subject to

missingness, conditional on all of the covariates that are completely observed, and to estimate the joint model (Raudenbush, Bryk, Cheong, Congdon, & du Toit, 2019).

Bennett (2001) demonstrated that statistical analysis is likely to be biased when more than 10% of data are missing. Hence, the upper limit for deletion of a variable from the data analysis was determined as 10% before the treatment of missing values for the present study. No variables in this thesis were found to have a missing value more than 10% of number of students participated in the study. So, all of them are included in the data analysis.

SPSS data files for student- and school-level factors were extracted from PISA 2015 database. Variables of the study are included in these files. Before building models in HLM program, multivariate data matrices (MDM) file was created by using SPSS files. Student files included student-level variables and plausible values. School files included school-level variables.

Missing values analysis was conducted in order to see the patterns of the missing value. It was found that, the pattern in the missing values was not random (MNAR). So, multiple regression strategy was needed to handle the missing values. For level-1 variables, listwise deletion of cases was conducted in HLM program, which used multiple imputation strategy. No missing values in level-2 files is assumed by HLM. An automatic listwise deletion of level-2 units with missing data was conducted when the MDM file is created.

3.9 Limitations of the Study

The present study is limited to Turkish students in the PISA 2015 study. Hence, the results of this study cannot be generalized to all countries participating in the PISA 2015 study. In addition, generalizing the results to OECD countries or to all students in these countries or to all 15-year-old students is not possible.

Since this study does not explain causalities, the information obtained from this study can explain the effect of student- and school-level factors on the scientific literacy skills, and can describe how significant these effects are. Within the framework of the present study, the variables were limited to the science domain specific variables included in the PISA framework. Hence, the effects of the remaining variables

cannot be explained.

In PISA 2015, a stratified sampling design was used. This approach can cause limited representation of population because of the variance in the identification of stratification variables and because of the variance in the sampling. Due to population coverage limitations, translation, non-response bias and measurement errors, non-sampling errors can occur in the data collection and processing stages. In addition, sampling error can occur due to the variability of the population characteristics and size of the sample.

In the PISA study, the data were obtained from the self-report questionnaires. Hence, the attention of the participants and their honesty have significant importance. It was assumed that the participants' responses reflect their personal opinions and beliefs rather than the opinions, which are believed to be acceptable.

The school questionnaire items are responded by the school principals or administrators. The data obtained from the questionnaires might be unreliable or biased, since the responses might reflect only the perceptions of administrators. In addition, the principals might not have the accurate data or information about the questionnaire items.



CHAPTER 4

RESULTS

In this chapter, the results from the hierarchical linear modeling (HLM) were presented. The results of four HLM models were given for the scientific literacy skills; combined science, scientific knowledge and scientific competencies scales.

4.1 Preliminary Data Analyses

As preliminary data analyses of the study; frequency distributions analysis, descriptive statistics, and missing value analysis were conducted. Analyses were conducted by SPSS (Statistics Package for Social Sciences) computer program version IBM SPSS Statistics 22. Table 4.1 demonstrates the frequency distribution of level-1 variables. HLM does not allow running a parameter estimates analysis with the missing data values. Hence, data cleaning procedure for this data was needed.

Table 4.1 Frequency Distribution of Level-1 Variables in the Present Study

Frequency Distribution of Level-1 Variables					
		This school year, approximately how many hours per week do you spend learning in addition? <School science>	Environmental awareness (WLE)	Index science activities (WLE)	Students' expected occupational status (SEI)
N	Valid	5333	5647	5585	5444
	Missing	562	248	310	451
	Missing (%)	9,5	4,2	5,3	7,7
		Environmental optimism (WLE)	Enjoyment of science (WLE)	Interest in broad science topics (WLE)	Instrumental motivation (WLE)
N	Valid	5661	5634	5393	5595
	Missing	234	261	502	300
	Missing (%)	4,0	4,4	8,5	5,1
		Epistemological beliefs (WLE)	Learning time (minutes per week) - <science>	Science self-efficacy (WLE)	
N	Valid	5610	5619	5596	
	Missing	285	276	299	
	Missing (%)	4,8	4,7	5,1	

The patterns of responses were examined by frequency distributions analysis. The distributions and characteristics of data were explained by descriptive statistics. In the data treatment step, missing data analysis was conducted; this analysis enabled to determine the variables with missing values above the accepted limits.

Missing value analysis was initially done by determining the percentages of missing values from the frequency distribution table. The upper limit for deletion of a variable from the data analysis was determined as 10% before the treatment procedure of missing values, since Bennett (2001) suggested that statistical analysis is likely to be biased when more than 10% of data are missing. The percentage of missing values did not exceed 9,5 for all student-level variables, which is smaller than 10% of total number of students (5985). Hence, 11 student-level variables given in Table 4.1 were included in the final analysis.

Before making the decision about how to handle missing values in these variables, it was necessary to determine the pattern in the missing values. In order to see whether the pattern of missing values is random or not, Little's MCAR test was conducted with IBM SPSS MVA (Missing Value Analysis). Table 4.2 demonstrates expectation maximization (EM) estimated statistics for student-level variables.

Table 4.2 Expectation Maximization (EM) Estimated Statistics for Student-Level Variables

EM Means*					
ENVAWARE	ENVOPT	JOYSCIE	INTBRSCI	INSTSCIE	SCIEEFF
,545843	-,529495	,120156	-,051382	,373218	,333450
EPIST	SCIEACT	BSMJ	SMINS	ST071Q01NA	
-,199295	,687542	63,256904	198,39	5,29	

* = Little's MCAR test: Chi-Square = 2363,188, DF = 1295, Sig. = ,000

According to Table 4.2, Little's MCAR test was found to be significant ($p = 0,000$), meaning that the pattern in missing values was not random. Hence, it was necessary to handle missing values with multiple imputation method. HLM 8.00 computer program estimates a multivariate hierarchical linear imputation model. While

conducting HLM, the missing values in students' data were treated with listwise deletion of cases by conducting multiple imputation procedure as explained in Chapter 3.

Table 4.3 demonstrates the descriptive statistics of level-1 variables.

Table 4.3 Descriptive Statistics of Level-1 Variabels

	N	Minimum	Maximum	Mean	Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic
This school year, approximately how many hours per week do you spend learning in addition? <School science>	5333	0	30	5,34	4,598
Environmental awareness (WLE)	5647	-3,3765	3,2932	,552105	1,4716523
Environmental optimism (WLE)	5661	-1,7932	3,0127	-,534727	1,4463020
Enjoyment of science (WLE)	5634	-2,1154	2,1635	,124894	1,1726052
Interest in broad science topics (WLE)	5393	-2,5813	2,7303	-,065570	1,0191516
Instrumental motivation (WLE)	5595	-1,9301	1,7359	,375385	,9195646
Science self-efficacy (WLE)	5596	-3,7565	3,2775	,336495	1,3225588
Epistemological beliefs (WLE)	5610	-2,7904	2,1552	-,192288	1,1812193
Index science activities (WLE)	5585	-1,7570	3,3617	,687315	1,1566044
Students' expected occupational status (SEI)	5444	10,0000	89,0000	63,358927	16,2464974
Learning time (minutes per week) - <science>	5619	0	800	198,82	108,775

Table 4.3. Descriptive Statistics of Level-1 Variabels (Continued)

	Skewness		Kurtosis	
	Statistic	Std. Error	Statistic	Std. Error
This school year, approximately how many hours per week do you spend learning in addition? <School science>	2,178	,034	7,378	,067
Environmental awareness (WLE)	,128	,033	,144	,065
Environmental optimism (WLE)	,889	,033	-,185	,065
Enjoyment of science (WLE)	-,116	,033	-,395	,065
Interest in broad science topics (WLE)	-,567	,033	1,131	,067
Instrumental motivation (WLE)	-,350	,033	-,078	,065
Science self-efficacy (WLE)	,043	,033	1,531	,065
Epistemological beliefs (WLE)	,153	,033	-,013	,065
Index science activities (WLE)	-,538	,033	,337	,066
Students' expected occupational status (SEI)	-,169	,033	-,632	,066
Learning time (minutes per week) - <science>	,119	,033	,382	,065

Table 4.4 demonstrates the frequency distribution of level-2 variables. Descriptive statistics of outcome variables are given in Appendix C.

Table 4.4 Frequency Distribution of Level-2 Variables in the Present Study

		Shortage of educational material (WLE)	Shortage of educational staff (WLE)	Index proportion of science teachers by all teachers	Index proportion of science teachers fully certified	Index proportion of science teachers with ISCED level 5A and a major in science	Total number of science teachers at school	Creative extra-curricular activities (Sum)	Index science specific resource (Sum)
N	Valid	187	187	186	184	184	186	182	185
	Missing	0	0	1	3	3	1	5	2
	Missing (%)	0	0	0,5	1,6	1,6	0,5	2,7	1,2

According to the Table 4.4, the percentage of missing values did not exceed 2,7 for all school-level variables. Hence, 8 school-level variables given in Table 4.4 were included in the final analysis.

In order to handle missing data in school-level variables, Little's MCAR test was conducted with IBM SPSS MVA (Missing Value Analysis). Table 4.5 demonstrates expectation maximization (EM) estimated statistics for school-level variables.

Table 4.5 Expectation Maximization (EM) Estimated Statistics for School-Level Variables

EM Means*				
EDUSHORT	STAFFSHORT	PROSTAT	PROSTCE	PROSTMAS
0,273453	0,621362	0,130871	0,442460	0,768097
TOTST	CREACTIV	SCIERES		
6,864	1,403793	2,41		

* = Little's MCAR test: Chi-Square = 29,620, DF = 37, Sig. = ,801

According to Table 4.5, Little's MCAR test was found to be non-significant ($p=0,801$), meaning that the pattern in missing values was random. Hence, missing data in school-level variables were treated with listwise deletion procedure in HLM program prior to analysis.

Table 4.6 demonstrates the descriptive statistics of level-2 variables.

Table 4.6 Descriptive Statistics of Level-2 Variables

	N	Minimum	Maximum	Mean	Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic
Shortage of educational material (WLE)	187	-1,2541	3,6096	,273453	1,2861416
Shortage of educational staff (WLE)	187	-1,6823	3,7219	,621362	1,1459639
Index propotion of science teachers by all teachers	186	,0189	,4417	,130902	,0624516
Index proportion of science teachers fully certified	184	0,0000	1,0000	,442373	,4866999
Index proportion of science teachers with ISCED level 5A and a major in science	184	0,0000	1,0000	,768030	,4007453
Total number of science teachers at school	186	1,0	53,0	6,866	5,8949
Creative extra-curricular activities (Sum)	182	0,0000	3,0000	1,412088	1,0085091
Index science specific ressources (Sum)	185	0	8	2,42	2,257

Table 4.6 Descriptive Statistics of Level-2 Variables (Continued)

	Skewness		Kurtosis	
	Statistic	Std. Error	Statistic	Std. Error
Shortage of educational material (WLE)	1,090	,178	,991	,354
Shortage of educational staff (WLE)	-,408	,178	-,150	,354
Index propotion of science teachers by all teachers	1,560	,178	4,297	,355
Index proportion of science teachers fully certified	,230	,179	-1,936	,356
Index proportion of science teachers with ISCED level 5A and a major in science	-1,262	,179	-,305	,356
Total number of science teachers at school	3,628	,178	21,807	,355
Creative extra-curricular activities (Sum)	,113	,180	-1,062	,358
Index science specific ressources (Sum)	,923	,179	-,234	,355

In the present study, two-level multilevel models were analyzed by using HLM computer program version 8.00. For the present study, the level of significance was determined as $p = 0.05$ for hypotheses testing.

4.2 Hierarchical Linear Modeling (HLM) Analyses

The scientific literacy skills of Turkish students were analyzed by using four different models. The results from one-way analyses of variance with random intercept model, means-as-outcome model, random-coefficients regression model and intercepts- and slopes-as-outcomes model were demonstrated as follows.

4.2.1 One-Way Analyses of Variance with Random Intercept Model

In this submodel, also known as the null model, data set is checked, whether there is variation between groups. If there is, it is applied to see how much of the variance is within the groups and how much of it is between the groups. For this purpose, ICC (Intraclass Correlation Coefficient) value is calculated. So, the following research questions can be answered through this analysis: “Are there differences in the students’ scientific literacy skills among schools in Turkey?” and “What proportions of variances in scientific literacy scores in PISA 2015 is explained by school-related factors of Turkish students?”.

The estimated grand means of scientific literacy scores were calculated with 95% confidence interval (C. I.).

In this model, average of 10 plausible values for combined science, for science competencies and for science knowledge was entered as the output variable and level-1 model was tested for each subscales separately. Sigma square and Tau values were used for the analysis. HLM provides an overall estimate of reliability which is the average of school reliabilities.

The student-level equation of this model is:

$$(Y) = \beta_0 + r$$

The school-level equation is:

$$\beta_{0j} = \gamma_{00} + u_0$$

where

Y : the scientific literacy score computed with the imputation of ten plausible values

β_0 : the intercept or the expected literacy score

r : the student-level error

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

u_0 : the random effect associated with the school.

4.2.1.1 Combined Science

The final estimation of fixed effects of combined scientific literacy skills for one-way ANOVA with random effects is represented in Table 4.7.

Table 4.7 Final Estimation of Fixed Effects of Combined Scientific Literacy Skills of Turkish Students

Fixed Effect	Coefficient	Standard error	t-ratio	Approximate d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.516934	4.332891	96.591	176	<0.001

Table 4.7 shows that there are significant differences in students' combined science scores among schools in Turkey. The mean of combined science scores was found as 418.52 (standard error: 4.33), and the confidence interval was calculated as:

$$C. I. = 418.52 \pm 1.96 \times 4.33 = (410.03, 427.01).$$

Table 4.8 shows the final estimation of fixed effects with robust standard errors. The values in this table are compared to final estimation of fixed effects table. Since there is no significant differences between the values in these tables, it appears that the data set complies with the assumptions of HLM.

Table 4.8 Final Estimation of Fixed Effects of Combined Scientific Literacy Skills of Turkish Students with Robust Errors

Fixed Effect	Coefficient	Standard error	t-ratio	Approximate d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.516934	4.320657	96.864	176	<0.001

Table 4.9 demonstrates the final estimation of variance components representing the school means and student-level effects.

Table 4.9 Final Estimation of Variance Components of Combined Scientific Literacy Skills of Turkish Students

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	54.40188	2959.56402	176	5032.35459	<0.001
Student effect, r	53.29332	2840.17824			

According to the results represented in Table 4.9, it can be suggested that students' combined science scores vary among schools significantly ($\chi^2 = 5032.35$, d.f. = 176, $p < 0.050$). The intraclass correlation coefficient is:

$$\rho = \tau_{00} / (\tau_{00} + \sigma^2) = 2959.56 / (2959.56 + 2840.18) = 0.510.$$

This is the average correlation between two randomly chosen students in the same school. It means that 51% of the variance in combined science scores of Turkish students was between-schools and 49% was at the individual level. It also means that 51% of the variance in combined science score is explained by the grouping structure in the population. The results show that school-level factors can explain the differences in combined science scores of Turkish students.

The fact that Chi-square test is significant ($p < 0.050$) shows that level-2 variables need to be added to the model.

The reliability of random level-1 coefficient, β_0 , was found as 0.926, meaning that the mean of the sample is a reliable indicator of the actual school mean.

4.2.1.2 Content Knowledge

The final estimation of fixed effects of content knowledge skills for one-way ANOVA with random effects is represented in Table 4.10.

Table 4.10 Final Estimation of Fixed Effects of Content Knowledge Skills of Turkish Students

Fixed Effect	Coefficient	Standard error	t-ratio	Approximate d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.128741	4.404347	94.935	176	<0.001

Table 4.10 shows that there are significant differences in students' content knowledge scores among schools in Turkey. The mean of content knowledge scores was found as 418.13 (standard error: 4.40), and the confidence interval was calculated as:

$$C. I. = 418.13 \pm 1.96 \times 4.40 = (409.51, 426.75).$$

Table 4.11 shows the final estimation of fixed effects with robust standard errors. The values in this table are compared to final estimation of fixed effects table. Since there is no significant differences between the values in these tables, it appears that the data set complies with the assumptions of HLM.

Table 4.11 Final Estimation of Fixed Effects of Content Knowledge Skills of Turkish Students with Robust Errors

Fixed Effect	Coefficient	Standard error	t-ratio	Approximate d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.128741	4.391760	95.208	176	<0.001

Table 4.12 demonstrates the final estimation of variance components representing the school means and student-level effects.

Table 4.12 Final Estimation of Variance Components of Content Knowledge Skills of Turkish Students

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	55.34402	3062.96080	176	4543.77919	<0.001
Student effect, r	57.28380	3281.43420			

According to the results represented in Table 4.12, it can be suggested that students' content knowledge scores vary among schools significantly ($\chi^2 = 4543.78$, d.f. = 176, $p < 0.050$). The intraclass correlation coefficient is:

$$\rho = \tau_{00} / (\tau_{00} + \sigma^2) = 3062.96 / (3062.96 + 3281.43) = 0.48.$$

This is the average correlation between two randomly chosen students in the same school. It means that 48% of the variance in content knowledge scores of Turkish students was between-schools and 52% was at the individual level. It also means that 48% of the variance in content knowledge score is explained by the grouping structure in the population. The results show that school-level factors can explain the differences in content knowledge scores of Turkish students.

The fact that Chi-square test is significant ($p < 0.050$) shows that level-2 variables need to be added to the model.

The reliability of random level-1 coefficient, β_0 , was found as 0.919, meaning that the mean of the sample is a reliable indicator of the actual school mean.

4.2.1.3 Procedural and Epistemic Knowledge

The final estimation of fixed effects of procedural and epistemic knowledge skills for one-way ANOVA with random effects is represented in Table 4.13.

Table 4.13 Final Estimation of Fixed Effects of Procedural and Epistemic Knowledge Skills of Turkish Students

Fixed Effect	Coefficient	Standard error	t-ratio	Approximate d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	417.690144	4.430767	94.270	176	<0.001

Table 4.13 shows that there are significant differences in students' procedural and epistemic knowledge scores among schools in Turkey. The mean of procedural and epistemic knowledge scores was found as 417.69 (standard error: 4.43), and the confidence interval was calculated as:

$$C. I. = 417.69 \pm 1.96 \times 4.43 = (409.01, 426.37).$$

Table 4.14 shows the final estimation of fixed effects with robust standard errors. The values in this table are compared to final estimation of fixed effects table. Since there is no significant differences between the values in these tables, it appears that the data set complies with the assumptions of HLM.

Table 4.14 Final Estimation of Fixed Effects of Procedural and Epistemic Knowledge Skills of Turkish Students with Robust Errors

Fixed Effect	Coefficient	Standard error	t-ratio	Approximate d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	417.690144	4.418340	94.536	176	<0.001

Table 4.15 demonstrates the final estimation of variance components representing the school means and student-level effects.

Table 4.15 Final Estimation of Variance Components of Procedural and Epistemic Knowledge Skills of Turkish

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	55.24107	3051.57577	176	4938.81741	<0.001
Student effect, r	54.70062	2992.15767			

According to the results represented in Table 4.15, it can be suggested that students' procedural and epistemic knowledge scores vary among schools significantly ($\chi^2=4938.82$, d.f.= 176, $p<0.050$). The intraclass correlation coefficient is:

$$\rho = \tau_{00} / (\tau_{00} + \sigma^2) = 3051.58 / (3051.58 + 2992.16) = 0.50.$$

This is the average correlation between two randomly chosen students in the same school. It means that 50% of the variance in procedural and epistemic knowledge scores of Turkish students was between-schools and 50% was at the individual level. It also means that 50% of the variance in procedural and epistemic knowledge score is explained by the grouping structure in the population. The results show that school-level factors can explain the differences in procedural and epistemic knowledge scores of Turkish students.

The fact that Chi-square test is significant ($p<0.050$) shows that level-2 variables need to be added to the model.

The reliability of random level-1 coefficient, β_0 , was found as 0.925, meaning that the mean of the sample is a reliable indicator of the actual school mean.

4.2.1.4 Explaining Phenomena Scientifically

The final estimation of fixed effects of explaining phenomena scientifically skills for one-way ANOVA with random effects is represented in Table 4.16.

Table 4.16 Final Estimation of Fixed Effects of Explaining Phenomena Scientifically Skills of Turkish Students

Fixed Effect	Coefficient	Standard error	t-ratio	Approximate d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.349814	4.492800	93.116	176	<0.001

Table 4.16 shows that there are significant differences in students' explaining phenomena scientifically scores among schools in Turkey. The mean of explaining phenomena scientifically scores was found as 418.35 (standard error: 4.49), and the confidence interval was calculated as:

$$C. I. = 418.35 \pm 1.96 \times 4.49 = (409.55, 427.15).$$

Table 4.17 shows the final estimation of fixed effects with robust standard errors. The values in this table are compared to final estimation of fixed effects table. Since there is no significant differences between the values in these tables, it appears that the data set complies with the assumptions of HLM.

Table 4.17 Final Estimation of Fixed Effects of Explaining Phenomena Scientifically Skills of Turkish Students with Robust Errors

Fixed Effect	Coefficient	Standard error	t-ratio	Approximate d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.349814	4.480023	93.381	176	<0.001

Table 4.18 demonstrates the final estimation of variance components representing the school means and student-level effects.

Table 4.18 Final Estimation of Variance Components of Explaining Phenomena Scientifically Skills of Turkish Students

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	56.53696	3196.42779	176	4721.23899	<0.001
Student effect, r	57.28771	3281.88161			

According to the results represented in Table 4.18, it can be suggested that students' explaining phenomena scientifically scores vary among schools significantly ($\chi^2=4721.24$, d.f.= 176, $p<0.050$). The intraclass correlation coefficient is:

$$\rho = \tau_{00} / (\tau_{00} + \sigma^2) = 3196.43 / (3196.43 + 3281.88) = 0.49.$$

This is the average correlation between two randomly chosen students in the same school. It means that 49% of the variance in explaining phenomena scientifically scores of Turkish students was between-schools and 51% was at the individual level. It also means that 49% of the variance in explaining phenomena scientifically score is explained by the grouping structure in the population. The results show that school-level factors can explain the differences in explaining phenomena scientifically scores of Turkish students.

The fact that Chi-square test is significant ($p<0.050$) shows that level-2 variables need to be added to the model.

The reliability of random level-1 coefficient, β_0 , was found as 0.922, meaning that the mean of the sample is a reliable indicator of the actual school mean.

4.2.1.5 Evaluating and Designing Scientific Inquiry

The final estimation of fixed effects of evaluating and designing scientific inquiry skills for one-way ANOVA with random effects is represented in Table 4.19.

Table 4.19 Final Estimation of Fixed Effects of Evaluating and Designing Scientific Inquiry Skills of Turkish Students

Fixed Effect	Coefficient	Standard error	t-ratio	Approximate d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	420.576439	4.635297	90.733	176	<0.001

Table 4.19 shows that there are significant differences in students' evaluating and designing scientific inquiry scores among schools in Turkey. The mean of evaluating and designing scientific inquiry scores was found as 420.58 (standard error: 4.64), and the confidence interval was calculated as:

$$C. I. = 420.58 \pm 1.96 \times 4.64 = (411.49, 429.67).$$

Table 4.20 shows the final estimation of fixed effects with robust standard errors. The values in this table are compared to final estimation of fixed effects table. Since there is no significant differences between the values in these tables, it appears that the data set complies with the assumptions of HLM.

Table 4.20 Final Estimation of Fixed Effects of Evaluating and Designing Scientific Inquiry Skills of Turkish Students with Robust Errors

Fixed Effect	Coefficient	Standard error	t-ratio	Approximate d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	420.576439	4.622857	90.978	176	<0.001

Table 4.21 demonstrates the final estimation of variance components representing the school means and student-level effects.

Table 4.21 Final Estimation of Variance Components of Evaluating and Designing Scientific Inquiry of Turkish Students

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	56.45190	3186.81669	176	4842.86165	<0.001
Student effect, r	56.19514	3157.89361			

According to the results represented in Table 4.21, it can be suggested that students' evaluating and designing scientific inquiry scores vary among schools significantly ($\chi^2 = 4842.86$, d.f. = 176, $p < 0.050$). The intraclass correlation coefficient is:

$$\rho = \tau_{00} / (\tau_{00} + \sigma^2) = 3186.82 / (3186.82 + 3157.89) = 0.50.$$

This is the average correlation between two randomly chosen students in the same school. It means that 50% of the variance in evaluating and designing scientific inquiry scores of Turkish students was between-schools and 50% was at the individual level. It also means that 50% of the variance in evaluating and designing scientific inquiry score is explained by the grouping structure in the population. The results show that school-level factors can explain the differences in evaluating and designing scientific inquiry scores of Turkish students.

The fact that Chi-square test is significant ($p < 0.050$) shows that level-2 variables need to be added to the model.

The reliability of random level-1 coefficient, β_0 , was found as 0.924, meaning that the mean of the sample is a reliable indicator of the actual school mean.

4.2.1.6 Interpreting Data and Evidence Scientifically

The final estimation of fixed effects of interpreting data and evidence scientifically skills for one-way ANOVA with random effects is represented in Table 4.22.

Table 4.22 Final Estimation of Fixed Effects of Interpreting Data and Evidence Scientifically Skills of Turkish Students

Fixed Effect	Coefficient	Standard error	t-ratio	Approximate d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	416.068624	4.593107	90.585	176	<0.001

Table 4.22 shows that there are significant differences in students' interpreting data and evidence scientifically scores among schools in Turkey. The mean of interpreting data and evidence scientifically scores was found as 416.07 (standard error: 4.59), and the confidence interval was calculated as:

$$C. I. = 416.07 \pm 1.96 \times 4.59 = (407.07, 425.07).$$

Table 4.23 shows the final estimation of fixed effects with robust standard errors. The values in this table are compared to final estimation of fixed effects table. Since there is no significant differences between the values in these tables, it appears that the data set complies with the assumptions of HLM.

Table 4.23 Final Estimation of Fixed Effects of Interpreting Data and Evidence Scientifically Skills of Turkish Students with Robust Errors

Fixed Effect	Coefficient	Standard error	t-ratio	Approximate d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	416.068624	4.580580	90.833	176	<0.001

Table 4.24 demonstrates the final estimation of variance components representing the school means and student-level effects.

Table 4.24 Final Estimation of Variance Components of Interpreting Data and Evidence Scientifically Skills of Turkish Students

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	56.32113	3172.06976	176	4719.83783	<0.001
Student effect, r	56.60979	3204.66784			

According to the results represented in Table 4.24, it can be suggested that students' interpreting data and evidence scientifically scores vary among schools significantly ($\chi^2= 4719.84$, d.f.= 176, $p<0.050$). The intraclass correlation coefficient is:

$$\rho = \tau_{00} / (\tau_{00} + \sigma^2) = 3172.07 / (3172.07 + 3204.67) = 0.50.$$

This is the average correlation between two randomly chosen students in the same school. It means that 50% of the variance in interpreting data and evidence scientifically scores of Turkish students was between-schools and 50% was at the individual level. It also means that 50% of the variance in interpreting data and evidence scientifically score is explained by the grouping structure in the population. The results show that school-level factors can explain the differences in interpreting data and evidence scientifically scores of Turkish students.

The fact that Chi-square test is significant ($p<0.050$) shows that level-2 variables need to be added to the model.

The reliability of random level-1 coefficient, β_0 , was found as 0.923, meaning that the mean of the sample is a reliable indicator of the actual school mean.

4.2.1.7 Findings from One-Way ANOVA Model

1. Average scientific literacy scores of the schools is statistically significant non-zero (γ_{00}).
2. There is an acceptable level of variance in the scientific literacy score averages of schools. (τ_{00}). The performance of schools were significantly different in their mean combined science score, explaining phenomena scientifically score, evaluating and designing scientific enquiry score, interpreting data and evidence scientifically score,

content knowledge score and procedural and epistemic knowledge score.

3. The intraclass correlation coefficients of combined science, science competencies and science knowledge subscales of schools in Turkey are about 50%. Accordingly, half of the total variability in science literacy can be considered to be caused by the differences between schools and the other half by the differences within the school.

4.2.2 Means-as-Outcome Regression Model

The means-as-outcomes regression submodel gives information about the school-level factors, that effect Turkish students' scientific literacy scores. So, the following research question can be answered through this analysis: "Which school-level variables have significant effect on students' scientific literacy skills in Turkey?".

In this model, students' scientific literacy scores were viewed as varying around their school means. School means were predicted by school-level variables. The school-level factors were added into the level-2 equation. The model was builded by inserting all school-level variable into equation. Then the model was run, the significance of effects of the variables were checked, and the least significant variable was discarded from the model. The model was run again with the remaining variables and this procedure was repeated with the next least significant variable, until all variables in the model become significant. School data file included eight school-level variables. The analysis were repeated for all scientific literacy subscales separately.

4.2.2.1 Combined Science

Level-1 equation of the means-as-outcome regression model is:

$$(Y/PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE, PV6SCIE, PV7SCIE, PV8SCIE, PV9SCIE, PV10SCIE) = \beta_{0j} + r_{ij}$$

Table 4.25 shows the resultant significant factors which are associated with students' combined science scores.

Table 4.25 Final Estimation of Fixed Effects for Combined Science

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.824343	3.724551	112.450	171	<0.001
EDUSHORT, γ_{01}	-9.520449	3.080184	-3.091	171	0.002
PROSTAT, γ_{02}	291.040329	70.809629	4.110	171	<0.001
TOTST, γ_{03}	-1.919039	0.745601	-2.574	171	0.011
CREACTIV, γ_{04}	8.218567	3.921282	2.096	171	0.038
SCIERES, γ_{05}	4.956891	1.857323	2.669	171	0.008

The school-level equation becomes:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(TOTST_j) + \gamma_{04}*(CREACTIV_j) + \gamma_{05}*(SCIERES_j) + u_{0j}$$

where

β_0 : the average of the school means on combined science

γ_{00} : the school mean on combined science

γ_{01} : the differentiating effect of shortage of educational material on school mean of combined science

γ_{02} : the differentiating effect of proportion of science teachers by all teachers on school mean of combined science

γ_{03} : the differentiating effect of total number of science teachers at school on school mean of combined science

γ_{04} : the differentiating effect of creative extra-curricular activities on school mean of combined science

γ_{05} : the differentiating effect of science specific resources on school mean of combined science

u_0 : the residual

The combined model is:

$$(Y/PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE, PV6SCIE, PV7SCIE, PV8SCIE, PV9SCIE, PV10SCIE) = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(TOTST_j) + \gamma_{04}*(CREACTIV_j) + \gamma_{05}*(SCIRES_j) + u_{0j} + r_j$$

The results suggest a significant association between combined science scores of schools (γ_{00} = 418.824, s.e.= 3.725, $p<0.050$) and shortage of educational material (γ_{01} = -9.520, s.e.= 3.080, $p<0.050$); proportion of science teachers by all teachers (γ_{02} = 291.040, s.e.= 70.810, $p<0.050$); total number of science teachers at school (γ_{03} = -1.919, s.e.= 0.746, $p<0.050$); creative extra-curricular activities (γ_{04} = 8.219, s.e.= 3.921, $p<0.050$); science specific resources (γ_{05} = 4.957, s.e.= 1.857, $p<0.050$) for Turkey. These five school-level variables were included in the full model. There was a negative relationship between combined science scores and shortage of educational material and total number of science teachers at school, whereas the relationship was positive for proportion of science teachers by all teachers, creative extra-curricular activities and science specific resources.

Final estimation of variance components for means-as-outcomes model is demonstrated in Table 4.26.

Table 4.26 Final Estimation of Variance Components for Combined Science

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	45.90770	2107. 51726	171	3303.04423	<0.001
Student effect, r	53.30938	2841.89003			

The residual variance between schools in this model (τ_{00} = 2107.52) was found to be smaller than the residual variance in one-way ANOVA model (τ_{00} = 2959.56). The proportion reduction in the variance, which is associated with the level-2 variables; shortage of educational material (EDUSHORT), proportion of science teachers by all teachers (PROSTAT), number of science teachers at school (TOTST), creative extra-curricular activities (CREACTIV) and science specific resources (SCIRES):

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

$$= (2959.56 - 2107.52) / 2959.56$$

$$= 0.288$$

So, EDUSHORT, PROSTAT, TOTST, CREATIV and SCIERES explain 29% of between-school variance in students' combined science scores.

4.2.2.2 Content Knowledge

Level-1 equation of the means-as-outcome regression model is:

$$(Y/PV1SKCO, PV2SKCO, PV3SKCO, PV4SKCO, PV5SKCO, PV6SKCO, PV7SKCO, PV8SKCO, PV9SKCO, PV10SKCO) = \beta_{0j} + r_{ij}$$

Table 4.27 shows the resultant significant factors which are associated with students' content knowledge scores.

Table 4.27 Final Estimation of Fixed Effects for Content Knowledge

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.474181	3.854554	108.566	172	<0.001
EDUSHORT, γ_{01}	-9.806776	3.252106	-3.106	172	0.003
PROSTAT, γ_{02}	289.641197	74.690632	3.878	172	<0.001
TOTST, γ_{03}	-1.704821	0.777709	-2.192	172	0.030
SCIERES, γ_{04}	6.111096	1.896661	3.222	172	0.002

The school-level equation becomes:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(TOTST_j) + \gamma_{04}*(SCIERES_j) + u_{0j}$$

where

β_0 : the average of the school means on content knowledge

γ_{00} : the school mean on content knowledge

γ_{01} : the differentiating effect of shortage of educational material on school mean of content knowledge

γ_{02} : the differentiating effect of proportion of science teachers by all teachers on school mean of content knowledge

γ_{03} : the differentiating effect of total number of science teachers at school on school mean of scientific literacy

γ_{04} : the differentiating effect of science specific resources on school mean of content knowledge

u_0 : the residual

The combined model is:

$$(Y/PV1SKCO, PV2SKCO, PV3SKCO, PV4SKCO, PV5SKCO, PV6SKCO, PV7SKCO, PV8SKCO, PV9SKCO, PV10SKCO) = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(TOTST_j) + \gamma_{04}*(SCIERES_j) + u_{0j} + r_j$$

The results suggest a significant association between content knowledge scores of schools ($\gamma_{00} = 418.474$, s.e. = 3.855, $p < 0.050$) and shortage of educational material ($\gamma_{01} = -9.807$, s.e. = 3.252, $p < 0.050$); proportion of science teachers by all teachers ($\gamma_{02} = 289.641$, s.e. = 74.690, $p < 0.050$); total number of science teachers at school ($\gamma_{03} = -1.705$, s.e. = 0.778, $p < 0.050$) and science specific resources ($\gamma_{04} = 6.111$, s.e. = 1.897, $p < 0.050$) for Turkey. These four school-level variables were included in the full model. There was a negative relationship between content knowledge scores and shortage of educational material and total number of science teachers at school, whereas the relationship was positive for proportion of science teachers by all teachers and science specific resources.

Final estimation of variance components for means-as-outcomes model is demonstrated in Table 4.28.

Table 4.28 Final Estimation of Variance Components for Content Knowledge

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	47.65217	2270.72946	172	3102.08930	<0.001
Student effect, r	57.30351	3283.69258			

The residual variance between schools in this model (τ_{00} = 2270.73) was found to be smaller than the residual variance in one-way ANOVA model (τ_{00} = 3062.96). The proportion reduction in the variance, which is associated the level-2 variables; shortage of educational material (EDUSHORT), proportion of science teachers by all teachers (PROSTAT), number of science teachers at school (TOTST) and science specific resources (SCIERES):

$$\begin{aligned} & \tau_{00} (\text{One-way ANOVA}) - \tau_{00} (\text{Means-as-outcomes}) / \tau_{00} (\text{One-way ANOVA}) \\ &= (3062.96 - 2270.73) / 3062.96 \\ &= 0.259 \end{aligned}$$

So, EDUSHORT, PROSTAT, TOTST and SCIERES explain 26% of between-school variance in students' content knowledge scores.

4.2.2.3 Procedural and Epistemic Knowledge

Level-1 equation of the means-as-outcome regression model is:

$$(Y/PV1SKPE, PV2SKPE, PV3SKPE, PV4SKPE, PV5SKPE, PV6SKPE, PV7SKPE, PV8SKPE, PV9SKPE, PV10SKPE) = \beta_{0j} + r_{ij}$$

Table 4.29 shows the resultant significant factors which are associated with students' procedural and epistemic knowledge scores.

Table 4.29 Final Estimation of Fixed Effects for Procedural and Epistemic Knowledge

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.017576	3.835168	108.996	171	<0.001
EDUSHORT, γ_{01}	-9.920755	3.247061	-3.055	171	0.003
PROSTAT, γ_{02}	285.808795	72.879637	3.922	171	<0.001
TOTST, γ_{03}	-1.846373	0.772840	-2.389	171	0.018
CREACTIV, γ_{04}	8.930716	4.042187	2.209	171	0.028
SCIERES, γ_{05}	4.607940	1.956715	2.355	171	0.020

The school-level equation becomes:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(TOTST_j) + \gamma_{04}*(CREACTIV_j) + \gamma_{05}*(SCIERES_j) + u_{0j}$$

where

β_0 : the average of the school means on procedural and epistemic knowledge

γ_{00} : the school mean on procedural and epistemic knowledge

γ_{01} : the differentiating effect of shortage of educational material on school mean of procedural and epistemic knowledge

γ_{02} : the differentiating effect of proportion of science teachers by all teachers on school mean of procedural and epistemic knowledge

γ_{03} : the differentiating effect of total number of science teachers at school on school mean of procedural and epistemic knowledge

γ_{04} : the differentiating effect of creative extra-curricular activities on school mean of procedural and epistemic knowledge

γ_{05} : the differentiating effect of science specific resources on school mean of procedural and epistemic knowledge

u_0 : the residual

The combined model is:

$$(Y/PV1SKPE, PV2SKPE, PV3SKPE, PV4SKPE, PV5SKPE, PV6SKPE, PV7SKPE, PV8SKPE, PV9SKPE, PV10SKPE) = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(TOTST_j) + \gamma_{04}*(CREACTIV_j) + \gamma_{05}*(SCIRES_j) + u_{0j} + r_j$$

The results suggest a significant association between procedural and epistemic knowledge scores of schools (γ_{00} = 418.018, s.e.= 3.836, $p<0.050$) and shortage of educational material (γ_{01} = -9.921, s.e.= 3.246, $p<0.050$); proportion of science teachers by all teachers (γ_{02} = 285.809, s.e.= 72.880, $p<0.050$); total number of science teachers at school (γ_{03} = -1.846, s.e.= 0.773, $p<0.050$); creative extra-curricular activities (γ_{04} = 8.931, s.e.= 4.042, $p<0.050$) and science specific resources (γ_{05} = 4.608, s.e.= 1.957, $p<0.050$) for Turkey. These five school-level variables were included in the full model. There was a negative relationship between procedural and epistemic knowledge scores and shortage of educational material and total number of science teachers at school, whereas the relationship was positive for proportion of science teachers by all teachers, creative extra-curricular activities and science specific resources.

Final estimation of variance components for means-as-outcomes model is demonstrated in Table 4.30.

Table 4.30 Final Estimation of Variance Components for Procedural and Epistemic Knowledge

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	46.85276	2195.18090	171	3293.09323	<0.001
Student effect, r	54.71836	2994.09888			

The residual variance between schools in this model (τ_{00} = 2195.18) was found to be smaller than the residual variance in one-way ANOVA model (τ_{00} = 3051.58). The proportion reduction in the variance, which is associated with the level-2 variables; shortage of educational material (EDUSHORT), proportion of science teachers by all

teachers (PROSTAT), number of science teachers at school (TOTST), creative extra-curricular activities (CREACTIV) and science specific resources (SCIERES):

$$\begin{aligned} & \tau_{00} \text{ (One-way ANOVA)} - \tau_{00} \text{ (Means-as-outcomes)} / \tau_{00} \text{ (One-way ANOVA)} \\ &= (3051.58 - 2345.08) / 3051.58 \\ &= 0.232 \end{aligned}$$

So, EDUSHORT, PROSTAT, TOTST, CREATIV and SCIERES explain 23% of between-school variance in students' procedural and epistemic knowledge scores.

4.2.2.4 Explaining Phenomena Scientifically

Level-1 equation of the means-as-outcome regression model is:

$$(Y/PV1SCEP, PV2SCEP, PV3SCEP, PV4SCEP, PV5SCEP, PV6SCEP, PV7SCEP, PV8SCEP, PV9SCEP, PV10SCEP) = \beta_{0j} + r_{ij}$$

Table 4.31 shows the resultant significant factors which are associated with students' explaining phenomena scientifically scores.

Table 4.31 Final Estimation of Fixed Effects for Explaining Phenomena Scientifically

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.716105	3.904971	107.226	172	<0.001
EDUSHORT, γ_{01}	-10.272839	3.329064	-3.086	172	0.002
PROSTAT, γ_{02}	312.584807	75.536163	4.138	172	<0.001
TOTST, γ_{03}	-1.862679	0.795902	-2.340	172	0.020
SCIERES, γ_{04}	6.010980	1.895825	3.171	172	0.002

The school-level equation becomes:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(TOTST_j) + \gamma_{04}*(SCIERES_j) + u_{0j}$$

where

β_0 : the average of the school means on explaining phenomena scientifically

γ_{00} : the school mean on explaining phenomena scientifically

γ_{01} : the differentiating effect of shortage of educational material on school mean of explaining phenomena scientifically

γ_{02} : the differentiating effect of proportion of science teachers by all teachers on school mean of explaining phenomena scientifically

γ_{03} : the differentiating effect of total number of science teachers at school on school mean of explaining phenomena scientifically

γ_{04} : the differentiating effect of science specific resources on school mean of explaining phenomena scientifically

u_0 : the residual

The combined model is:

$$(Y/PV1SCEP, PV2SCEP, PV3SCEP, PV4SCEP, PV5SCEP, PV6SCEP, PV7SCEP, PV8SCEP, PV9SCEP, PV10SCEP) = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(TOTST_j) + \gamma_{04}*(SCIRES_j) + u_{0j} + r_j$$

The results suggest a significant association between explaining phenomena scientifically scores of schools (γ_{00} = 418.716, s.e.= 3.905, $p<0.050$) and shortage of educational material (γ_{01} = -10.272, s.e.= 3.329, $p<0.050$); proportion of science teachers by all teachers (γ_{02} = 312.585, s.e.= 75.536, $p<0.050$); total number of science teachers at school (γ_{03} = -1.863, s.e.= 0.796, $p<0.050$); science specific resources (γ_{04} = 6.011, s.e.= 1.896, $p<0.050$) for Turkey. These four school-level variables were included in the full model. There was a negative relationship association between explaining phenomena scientifically scores and shortage of educational material and total number of science teachers at school, whereas the relationship was positive for proportion of science teachers by all teachers and science specific resources.

Final estimation of variance components for means-as-outcomes model is demonstrated in Table 4.32.

Table 4.32 Final Estimation of Variance Components for Explaining Phenomena Scientifically

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	48.42606	2345.08366	172	3172.39454	<0.001
Student effect, r	57.30821	3284.23050			

The residual variance between schools in this model ($\tau_{00}=2,345.08$) was found to be smaller than the residual variance in one-way ANOVA model ($\tau_{00}=3,196.43$). The proportion reduction in the variance, which is associated with the level-2 variables; shortage of educational material (EDUSHORT), proportion of science teachers by all teachers (PROSTAT), number of science teachers at school (TOTST) and science specific resources (SCIRES):

$$\begin{aligned} & \tau_{00} (\text{One-way ANOVA}) - \tau_{00} (\text{Means-as-outcomes}) / \tau_{00} (\text{One-way ANOVA}) \\ &= (3196.43 - 2345.08) / 3196.43 \\ &= 0.266 \end{aligned}$$

So, EDUSHORT, PROSTAT, TOTST and SCIRES explain 27% of between-school variance in students' explaining phenomena scientifically scores.

4.2.2.5 Evaluating and Designing Scientific Inquiry

Level-1 equation of the means-as-outcome regression model is:

$$(Y/PV1SCED, PV2SCED, PV3SCED, PV4SCED, PV5SCED, PV6SCED, PV7SCED, PV8SCED, PV9SCED, PV10SCED) = \beta_{0j} + r_{ij}$$

Table 4.33 shows the resultant significant factors which are associated with students' evaluating and designing scientific inquiry scores.

Table 4.33 Final Estimation of Fixed Effects for Evaluating and Designing Scientific Inquiry

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	420.993357	4.025834	104.573	171	<0.001
EDUSHORT, γ_{01}	-9.777383	3.382135	-2.891	171	0.004
PROSTAT, γ_{02}	302.978863	78.143772	3.877	171	<0.001
TOTST, γ_{03}	-2.307780	0.817783	-2.822	171	0.005
CREACTIV, γ_{04}	8.979185	4.256735	2.109	171	0.036
SCIERES, γ_{05}	5.021903	1.955467	2.568	171	0.011

The school-level equation becomes:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(TOTST_j) + \gamma_{04}*(CREACTIV_j) + \gamma_{05}*(SCIERES_j) + u_{0j}$$

where

β_0 : the average of the school means on evaluating and designing scientific inquiry

γ_{00} : the school mean on evaluating and designing scientific inquiry

γ_{01} : the differentiating effect of shortage of educational material on school mean of evaluating and designing scientific inquiry

γ_{02} : the differentiating effect of proportion of science teachers by all teachers on school mean of evaluating and designing scientific inquiry

γ_{03} : the differentiating effect of total number of science teachers at school on school mean of evaluating and designing scientific inquiry

γ_{04} : the differentiating effect of creative extra-curricular activities on school mean of evaluating and designing scientific inquiry

γ_{05} : the differentiating effect of science specific resources on school mean of evaluating and designing scientific inquiry

u_0 : the residual

The combined model is:

$$(Y/PV1SCED, PV2SCED, PV3SCED, PV4SCED, PV5SCED, PV6SCED, PV7SCED, PV8SCED, PV9SCED, PV10SCED) = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(TOTST_j) + \gamma_{04}*(CREACTIV_j) + \gamma_{05}*(SCIERES_j) + u_{0j} + r_j$$

The results suggest a significant association between evaluating and designing scientific inquiry scores of schools ($\gamma_{00} = 420.993$, s.e. = 4.026, $p < 0.050$) and shortage of educational material ($\gamma_{01} = -9.777$, s.e. = 3.382, $p < 0.050$); proportion of science teachers by all teachers ($\gamma_{02} = 302.979$, s.e. = 78.144, $p < 0.050$); total number of science teachers at school ($\gamma_{03} = -2.308$, s.e. = 0.818, $p < 0.050$); creative extra-curricular activities ($\gamma_{04} = 8.979$, s.e. = 4.257, $p < 0.050$) and science specific resources ($\gamma_{05} = 5.022$, s.e. = 1.956, $p < 0.050$) for Turkey. These five school-level variables were included in the full model. There was a negative relationship between evaluating and designing scientific inquiry scores and shortage of educational material and total number of science teachers at school, whereas the relationship was positive for proportion of science teachers by all teachers, creative extra-curricular activities and science specific resources.

Final estimation of variance components for means-as-outcomes model is demonstrated in Table 4.34.

Table 4.34 Final Estimation of Variance Components for Evaluating and Designing Scientific Inquiry

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	47.62131	2267.78871	171	3200.60450	<0.001
Student effect, r	56.21108	3159.68527			

The residual variance between schools in this model ($\tau_{00} = 2267.79$) was found to be smaller than the residual variance in one-way ANOVA model ($\tau_{00} = 3186.82$). The proportion reduction in the variance, which is associated with the level-2 variables; shortage of educational material (EDUSHORT), proportion of science teachers by all

teachers (PROSTAT), total number of science teachers at school (TOTST), creative extra-curricular activities (CREACTIV) and science specific resources (SCIERES):

$$\begin{aligned} & \tau_{00} \text{ (One-way ANOVA)} - \tau_{00} \text{ (Means-as-outcomes)} / \tau_{00} \text{ (One-way ANOVA)} \\ &= (3186.82 - 2267.79) / 3186.82 \\ &= 0.288 \end{aligned}$$

So, EDUSHORT, PROSTAT, TOTST, CREATIV and SCIERES explain 29% of between-school variance in students' evaluating and designing scientific inquiry scores.

4.2.2.6 Interpreting Data and Evidence Scientifically

Level-1 equation of the means-as-outcome regression model is:

$$(Y/PV1SCID, PV2SCID, PV3SCID, PV4SCID, PV5SCID, PV6SCID, PV7SCID, PV8SCID, PV9SCID, PV10SCID) = \beta_{0j} + r_{ij}$$

Table 4.35 shows the resultant significant factors which are associated with students' interpreting data and evidence scientifically scores.

Table 4.35 Final Estimation of Fixed Effects for Interpreting Data and Evidence Scientifically

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	416.173295	4.104001	101.407	173	<0.001
EDUSHORT, γ_{01}	-9.383878	3.292951	-2.850	173	0.005
PROSTAT, γ_{02}	185.611684	65.202757	2.847	173	0.005
SCIERES, γ_{03}	6.609051	1.912238	3.456	173	<0.001

The school-level equation becomes:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(SCIERES_j) + u_{0j}$$

where

β_0 : the average of the school means on interpreting data and evidence scientifically

γ_{00} : the school mean on interpreting data and evidence scientifically

γ_{01} : the differentiating effect of shortage of educational material on school mean of interpreting data and evidence scientifically

γ_{02} : the differentiating effect of proportion of science teachers by all teachers on school mean of interpreting data and evidence scientifically

γ_{03} : the differentiating effect of science specific resources on school mean of interpreting data and evidence scientifically

u_0 : the residual

The combined model is:

$$(Y/PV1SCID, PV2SCID, PV3SCID, PV4SCID, PV5SCID, PV6SCID, PV7SCID, PV8SCID, PV9SCID, PV10SCID) = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(SCIRES_j) + u_{0j} + r_j$$

The results suggest a significant association between interpreting data and evidence scientifically scores of schools (γ_{00} = 416.173, s.e.= 4.104, $p<0.050$) and shortage of educational material (γ_{01} = -9.384, s.e.= 3.293, $p<0.050$); proportion of science teachers by all teachers (γ_{02} = 185.612, s.e.= 65.203, $p<0.050$) and science specific resources (γ_{03} = 6.609, s.e.= 1.912, $p<0.050$) for Turkey. These three school-level variables were included in the full model. There was a negative relationship between interpreting data and evidence scientifically scores and shortage of educational material, whereas the relationship was positive for proportion of science teachers by all teachers and science specific resources.

Final estimation of variance components for means-as-outcomes model is demonstrated in Table 4.36.

Table 4.36 Final Estimation of Variance Components for Interpreting Data and Evidence Scientifically

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	49.46299	2446.58778	173	3455.23797	<0.001
Student effect, r	56.62615	3206.52034			

The residual variance between schools in this model ($\tau_{00}= 2446.59$) was found to be smaller than the residual variance in one-way ANOVA model ($\tau_{00}=3172.07$). The proportion reduction in the variance, which is associated with the level-2 variables; shortage of educational material (EDUSHORT), proportion of science teachers by all teachers (PROSTAT) and science specific resources (SCIERES):

$$\begin{aligned} & \tau_{00} (\text{One-way ANOVA}) - \tau_{00} (\text{Means-as-outcomes}) / \tau_{00} (\text{One-way ANOVA}) \\ &= (3172.07 - 2446.59) / 3172.07 \\ &= 0.229 \end{aligned}$$

So, EDUSHORT, PROSTAT, SCIERES explain 23% of between-school variance in students' interpreting data and evidence scientifically scores.

4.2.2.7 Findings from Means-as-Outcome Regression Model

1. The combined science score, the science competencies scores and science knowledge scores of Turkish students were predicted by shortage of educational material (EDUSHORT), proportion of science teachers by all teachers (PROSTAT) and science specific resources (SCIERES).
2. Total number of science teachers at school (TOTST) predicted all scientific literacy skills except for interpreting data and evidence scientifically subscale.
3. Combined science score, procedural and epistemic knowledge and evaluating and designing scientific inquiry subscales were predicted by creative extra-curricular activities (CREACTIV) in addition to EDUSHORT, PROSTAT and SCIERES.
4. Shortage of educational material (EDUSHORT) and total number of science teachers at school (TOTST) were negatively related to the combined science score, science competencies scores and science knowledge scores of Turkish students, whereas proportion of science teachers by all teachers (PROSTAT) and science specific resources (SCIERES) were positively related to all scientific literacy skills.
5. Creative extra-curricular activities (CREACTIV) was found to be positively related to combined science score, procedural and epistemic knowledge and evaluating and designing scientific inquiry subscales.

6. The proportions of variances explained by means-as-outcome models of combined science, science competencies and science knowledge change between 23% and 29%.

4.2.3 Random-Coefficients Regression Model

The analysis in the random-coefficients regression submodel is performed to see the effect of variables in the level-1 model on the outcome variables. This model gives information about the student-level factors, that effect Turkish students' scientific literacy skills. So, the following research question can be answered through this analysis: "Which student-level variables have significant effect on students' scientific literacy skills in Turkey?".

Initially, the student-level variables were inserted into equation while building the model, and the model was run. The significance of the effects of the variables were checked, the least significant variable was discarded from the model, and the model was run again with the remaining variables. The purpose of discarding insignificant variables from the analysis is to eliminate the covariance between factors. And this procedure was repeated by discarding the next least significant variable until all variables in the model become significant. At the end of this process, the variables, that were found to be affecting the scientific literacy significantly, were included in the level-1 equation.

Model checking continued with the determination of the random and fixed slopes. The output of variance components tables were used to check the significance of error terms. When all slopes were assumed randomly varying, some insignificant error variances were observed in slopes. While defining random and fixed slopes, random components for level-2 equations were obtained by turning off the insignificant error variances. This was executed by turning off the least significant error variance and continuing the process with the next least significant one until all error variances in the model become significant. Finally, randomly varying slopes were determined. The variables with randomly varying slopes were included in the analysis to determine the variance explained with these factors. For this purpose, the variances of one-way ANOVA model and random-coefficients regression model were compared.

4.2.3.1 Combined Science

Level-1 equation for combined science score becomes:

$$(Y/PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE, PV6SCIE, PV7SCIE, PV8SCIE, PV9SCIE, PV10SCIE) = \beta_{0j} + \beta_{1j}*(ST071Q01_{ij}) + \beta_{2j}*(ENVAWARE_{ij}) + \beta_{3j}*(ENVOPT_{ij}) + \beta_{4j}*(JOYSCIE_{ij}) + \beta_{5j}*(INTBRSCI_{ij}) + \beta_{6j}*(SCIEEFF_{ij}) + \beta_{7j}*(EPIST_{ij}) + \beta_{8j}*(SCIEACT_{ij}) + \beta_{9j}*(BSMJ_{ij}) + \beta_{10j}*(SMINS_{ij}) + r_{ij}.$$

and level-2 equations becomes:

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

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

$$\beta_{2j} = \gamma_{20} + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60j}$$

$$\beta_{7j} = \gamma_{70j}$$

$$\beta_{8j} = \gamma_{80}$$

$$\beta_{9j} = \gamma_{90}$$

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

where

β_{0j} : mean on combined science

β_{1j} : the differentiating effect of out-of-school study time for science

β_{2j} : the differentiating effect of environmental awareness

β_{3j} : the differentiating effect of environmental optimism

β_{4j} : the differentiating effect of enjoyment of science

β_{5j} : the differentiating effect of interest in broad science topics

β_{6j} : the differentiating effect of science self-efficacy

β_{7j} : the differentiating effect of epistemological beliefs

β_{8j} : the differentiating effect of science activities

β_{9j} : the differentiating effect of students' expected occupational status

β_{10j} : the differentiating effect of science learning time (minutes per week)

According to the error variances in the variance components table, slopes of out-of-school study time for science, enjoyment of science, interest in broad science topics, science self-efficacy, epistemological beliefs, science activities and students' expected occupational status were determined as fixed across schools. The slopes of remaining variables were random.

Table 4.37 represents the final estimation of fixed effects in random-coefficients regression model. The student-level variables which effect Turkish students' combined science scores were described according to this table.

Environmental awareness ($\gamma_{20}= 4.597$, $p<0.050$), enjoyment of science ($\gamma_{40}= 3.300$, $p<0.050$), interest in broad science topics ($\gamma_{50}= 2.020$, $p< 0.005$), science self-efficacy ($\gamma_{60}= 1.585$, $p<0.050$), epistemological beliefs ($\gamma_{70}= 4.246$, $p<0.050$), students' expected occupational status ($\gamma_{90}= 0.185$, $p<0.050$) and science learning time (minutes per week) ($\gamma_{100}= 0.075$, $p<0.050$) had positive association with combined science scores; whereas out-of-school study time for science ($\gamma_{10}= -1.173$, $p<0.050$), environmental optimism ($\gamma_{30}= -6.160$, $p<0.050$) and science activities ($\gamma_{80}= -3.431$, $p<0.050$) had negative association with combined science scores of Turkish students.

Table 4.38 represents the estimations of variance components, in which student-level variables with random components are demonstrated.

The relationship between the variables, which have significant error terms in final estimation of variance components table, and outcome variable is effected by at least one of the school-related factors. This is called "mediation effect". In order to examine the effects of significant school-level variables on the relationship between student-level variables and outcome variable, school-level variables are included in the further steps of the analysis.

Table 4.37 Final Estimation of Fixed Effects for Combined Science

Fixed Effect	Coefficient	Standard error	t-ratio	Appr. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.272532	4.332996	96.532	176	<0.001
For ST071Q01slope β_1					
INTRCPT2, γ_{10}	-1.172578	0.195566	-5.996	733	<0.001
For ENVAWARE slope β_2					
INTRCPT2, γ_{20}	4.597292	0.845411	5.438	75	<0.001
For ENVOPT slope β_3					
INTRCPT2, γ_{30}	-6.160010	0.847563	-7.268	70	<0.001
For JOYSCIE slope β_4					
INTRCPT2, γ_{40}	3.295601	0.906688	3.635	149	<0.001
For INTBRSCI slope β_5					
INTRCPT2, γ_{60}	2.020083	0.965768	2.092	323	0.037
For SCIEEFF slope β_6					
INTRCPT2, γ_{60}	1.584517	0.695721	2.278	884	0.023
For EPIST slope β_7					
INTRCPT2, γ_{70}	4.246317	0.855934	4.961	108	<0.001
For SCIEACT slope β_8					
INTRCPT2, γ_{80}	-3.431132	0.951313	-3.607	111	<0.001
For BSMJ slope β_9					
INTRCPT2, γ_{90}	0.184836	0.058980	3.134	241	0.002
For SMINS slope β_{10}					
INTRCPT2, γ_{100}	0.074739	0.011781	6.344	113	<0.001

Table 4.38 Final Estimation of Variance Components for Combined Science

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	54.59592	2980.71445	160	5819.74412	<0.001
ENVAWARE slope, u_2	3.25805	10.61486	160	196.41289	0.026
ENVOPT slope, u_3	3.85493	14.86045	160	211.75346	0.004
SMINS slope, u_{10}	0.06203	0.00385	160	211.72665	0.004
Student effect, r	49.37007	2437.40358			

The variances of one-way ANOVA submodel and random-coefficients regression submodel are compared in order to investigate the variance when these factors are included in the model. So, the proportion index for reduction was calculated as:

$$\begin{aligned} \rho &= \sigma^2(\text{One-Way ANOVA}) - \sigma^2(\text{Random-Coefficients Regression}) / \sigma^2(\text{One-Way ANOVA}) \\ &= 2840.17824 - 2437.40358 / 2840.17824 \\ &= 0.1418 \end{aligned}$$

When the student-level factors; out-of-school study time for science, environmental awareness, environmental optimism, enjoyment of science, interest in broad science topics, science self-efficacy, epistemological beliefs, science activities, students' expected occupational status and science learning time (minutes per week) were included as the predictors of combined science scores, a reduction was observed in within school variance. This finding suggests that, these variables explained 14% of variance in Turkish students' combined science scores.

4.2.3.2 Content Knowledge

Level-1 equation for content knowledge score becomes:

$$\begin{aligned} (Y/PV1SKCO, PV2SKCO, PV3SKCO, PV4SKCO, PV5SKCO, PV6SKCO, \\ PV7SKCO, PV8SKCO, PV9SKCO, PV10SKCO) = & \beta_{0j} + \beta_{1j}*(ST071Q01_{ij}) + \\ & \beta_{2j}*(ENVAWARE_{ij}) + \beta_{3j}*(ENVOPT_{ij}) + \beta_{4j}*(JOYSCIE_{ij}) + \beta_{5j}*(INTBRSCI_{ij}) + \\ & \beta_{6j}*(SCIEEFF_{ij}) + \beta_{7j}*(EPIST_{ij}) + \beta_{8j}*(SCIEACT_{ij}) + \beta_{9j}*(BSMJ_{ij}) + \beta_{10j}*(SMINS_{ij}) + \\ & r_{ij}. \end{aligned}$$

and level-2 equations becomes:

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

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

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60j}$$

$$\beta_{7j} = \gamma_{70j}$$

$$\beta_{8j} = \gamma_{80}$$

$$\beta_{9j} = \gamma_{90}$$

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

where

β_{0j} : mean on content knowledge

β_{1j} : the differentiating effect of out-of-school study time for science

β_{2j} : the differentiating effect of environmental awareness

β_{3j} : the differentiating effect of environmental optimism

β_{4j} : the differentiating effect of enjoyment of science

β_{5j} : the differentiating effect of interest in broad science topics

β_{6j} : the differentiating effect of science self-efficacy

β_{7j} : the differentiating effect of epistemological beliefs

β_{8j} : the differentiating effect of science activities

β_{9j} : the differentiating effect of students' expected occupational status

β_{10j} : the differentiating effect of science learning time (minutes per week)

According to the error variances in the variance components table, slopes of out-of-school study time for science, environmental awareness, enjoyment of science, interest in broad science topics, science self-efficacy, epistemological beliefs, science activities and students' expected occupational status were determined as fixed across schools. The slopes of remaining variables were random.

Table 4.39 represents the final estimation of fixed effects in random-coefficients regression model. The student-level variables which effect Turkish students' content knowledge scores were described according to this table.

Table 4.39 Final Estimation of Fixed Effects for Content Knowledge

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	417.896067	4.397373	95.033	176	<0.001
For <i>ST071Q01</i> slope β_1					
INTRCPT2, γ_{10}	-1.295820	0.221910	-5.839	224	<0.001
For <i>ENVAWARE</i> slope β_2					
INTRCPT2, γ_{20}	5.110416	0.996063	5.131	30	<0.001
For <i>ENVOPT</i> slope β_3					
INTRCPT2, γ_{30}	-7.077865	0.970185	-7.295	43	<0.001
For <i>JOYSCIE</i> slope β_4					
INTRCPT2, γ_{40}	3.976557	1.008286	3.944	103	<0.001
For <i>INTBRSCI</i> slope β_5					
INTRCPT2, γ_{50}	3.057974	1.206526	2.535	61	0.014
For <i>SCIEEFF</i> slope β_6					
INTRCPT2, γ_{60}	2.243468	1.081156	2.075	27	0.048
For <i>EPIST</i> slope β_7					
INTRCPT2, γ_{70}	2.797257	0.918911	3.044	110	0.003
For <i>SCIEACT</i> slope β_8					
INTRCPT2, γ_{80}	-2.819234	1.040070	-2.711	94	0.008
For <i>BSMJ</i> slope β_9					
INTRCPT2, γ_{90}	0.154342	0.070858	2.178	72	0.033
For <i>SMINS</i> slope β_{10}					
INTRCPT2, γ_{100}	0.075423	0.011029	6.839	176	<0.001

Environmental awareness ($\gamma_{20} = 5.110$, $p < 0.050$), enjoyment of science ($\gamma_{40} = 3.977$, $p < 0.050$), interest in broad science topics ($\gamma_{50} = 3.058$, $p < 2.020$), science self-efficacy ($\gamma_{60} = 2.243$, $p < 0.050$), epistemological beliefs ($\gamma_{70} = 2.797$, $p < 0.050$), students' expected occupational status ($\gamma_{90} = 0.154$, $p < 0.050$) and science learning time (minutes per week) ($\gamma_{100} = 0.075$, $p < 0.050$) had positive association with content knowledge scores; whereas out-of-school study time for science ($\gamma_{10} = -1.130$, $p < 0.050$), environmental optimism ($\gamma_{30} = -7.078$, $p < 0.050$) and science activities ($\gamma_{80} = -2.820$, $p < 0.050$) had negative association with content knowledge scores of Turkish students.

Table 4.40 represents the estimations of variance components, in which student-level variables with random components are demonstrated.

Table 4.40 Final Estimation of Variance Components for Content Knowledge

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	55.53445	3084.07493	160	5201.80945	<0.001
<i>ENVOPT</i> slope, u_3	3.75752	14.11894	160	201.06542	0.015
<i>SMINS</i> slope, u_{10}	0.05833	0.00340	160	191.40694	0.045
Student effect, r	53.35115	2846.34535			

The relationship between the variables, which have significant error terms in final estimation of variance components table, and outcome variable is effected by at least one of the school-related factors. This is called “mediation effect”. In order to examine the effects of significant school-level variables on the relationship between student-level variables and outcome variable, school-level variables are included in the further steps of the analysis.

The variances of one-way ANOVA submodel and random-coefficients regression submodel are compared in order to investigate the variance when these factors are included in the model. So, the proportion index for reduction was calculated as:

$$\rho = \frac{\sigma^2(\text{One-Way ANOVA}) - \sigma^2(\text{Random-Coefficients Regression})}{\sigma^2(\text{One-Way ANOVA})}$$

$$= 3281.43420 - 2846.34535 / 3281.43420$$

$$= 0.1326$$

When the student-level factors; out-of-school study time for science, environmental awareness, environmental optimism, enjoyment of science, interest in broad science topics, science self-efficacy, epistemological beliefs, science activities, students’ expected occupational status and science learning time (minutes per week) were included as the predictors of content knowledge scores, a reduction was observed in within school variance. This finding suggests that, these variables explained 13% of variance in Turkish students’ content knowledge scores.

4.2.3.3 Procedural and Epistemic Knowledge

Level-1 equation for procedural and epistemic knowledge score becomes:

$$(Y/PV1SKPE, PV2SKPE, PV3SKPE, PV4SKPE, PV5SKPE, PV6SKPE, PV7SKPE, PV8SKPE, PV9SKPE, PV10SKPE) = \beta_{0j} + \beta_{1j}*(ST071Q01_{ij}) + \beta_{2j}*(ENVAWARE_{ij}) + \beta_{3j}*(ENVOPT_{ij}) + \beta_{4j}*(JOYSCIE_{ij}) + \beta_{5j}*(EPIST_{ij}) + \beta_{6j}*(SCIEACT_{ij}) + \beta_{7j}*(BSMJ_{ij}) + \beta_{8j}*(SMINS_{ij}) + r_{ij}.$$

and level-2 equations becomes:

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

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

$$\beta_{2j} = \gamma_{20} + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60}$$

$$\beta_{7j} = \gamma_{70}$$

$$\beta_{8j} = \gamma_{80} + u_{8j}$$

where

β_{0j} : mean on procedural and epistemic knowledge

β_{1j} : the differentiating effect of out-of-school study time for science

β_{2j} : the differentiating effect of environmental awareness

β_{3j} : the differentiating effect of environmental optimism

β_{4j} : the differentiating effect of enjoyment of science

β_{5j} : the differentiating effect of epistemological beliefs

β_{6j} : the differentiating effect of science activities

β_{7j} : the differentiating effect of students' expected occupational status

β_{8j} : the differentiating effect of science learning time (minutes per week)

According to the error variances in the variance components table, slopes of out-of-school study time for science, enjoyment of science, epistemological beliefs, science activities and students' expected occupational status were determined as fixed across schools. The slopes of remaining variables were random.

Table 4.41 represents the final estimation of fixed effects in random-coefficients regression model. The student-level variables which effect Turkish students' procedural and epistemic knowledge scores were described according to this table.

Table 4.41 Final Estimation of Fixed Effects for Procedural and Epistemic Knowledge

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	417.439648	4.427798	94.277	176	<0.001
For ST071Q01slope β_1					
INTRCPT2, γ_{10}	-1.236604	0.202617	-6.103	692	<0.001
For ENVARE slope β_2					
INTRCPT2, γ_{20}	4.207112	0.978350	4.300	35	<0.001
For ENVOPT slope β_3					
INTRCPT2, γ_{30}	-5.496124	0.907541	-6.056	52	<0.001
For JOYSCIE slope β_4					
INTRCPT2, γ_{40}	3.020286	0.962845	3.137	85	0.002
For EPIST slope β_5					
INTRCPT2, γ_{50}	4.755126	0.890102	5.342	100	<0.001
For SCIEACT slope β_6					
INTRCPT2, γ_{60}	-3.969863	1.001778	-3.963	60	<0.001
For BSMJ slope β_7					
INTRCPT2, γ_{70}	0.189726	0.067720	2.802	75	0.006
For SMINS slope β_8					
INTRCPT2, γ_{80}	0.083007	0.012345	6.724	93	<0.001

Environmental awareness ($\gamma_{20}= 4.207$, $p<0.050$), enjoyment of science ($\gamma_{40}= 3.020$, $p<0.050$), epistemological beliefs ($\gamma_{50}= 4.755$, $p<0.050$), students' expected occupational status ($\gamma_{70}= 0.190$, $p<0.050$) and science learning time (minutes per week) ($\gamma_{80}= 0.083$, $p<0.050$) had positive association with procedural and epistemic knowledge scores; whereas out-of-school study time for science ($\gamma_{10}= -1.237$, $p<0.050$), environmental optimism ($\gamma_{30}= -5.500$, $p<0.050$) and science activities ($\gamma_{60}= -3.970$, $p<0.050$) had negative association with procedural and epistemic knowledge scores of Turkish students.

Table 4.42 represents the estimations of variance components, in which student-level variables with random components are demonstrated.

Table 4.42 Final Estimation of Variance Components for Procedural and Epistemic Knowledge

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	55.44563	3074.21754	160	5606.83351	<0.001
<i>ENVAWARE</i> slope, u_2	3.24237	10.51295	160	191.19292	0.046
<i>ENVOPT</i> slope, u_3	3.79815	14.42595	160	204.43485	0.010
<i>SMINS</i> slope, u_8	0.06260	0.00392	160	210.62062	0.005
Student effect, r	51.11995	2613.24882			

The relationship between the variables, which have significant error terms in final estimation of variance components table, and outcome variable is effected by at least one of the school-related factors. This is called “mediation effect”. In order to examine the effects of significant school-level variables on the relationship between student-level variables and outcome variable, school-level variables are included in the further steps of the analysis.

The variances of one-way ANOVA submodel and random-coefficients regression submodel are compared in order to investigate the variance when these factors are included in the model. So, the proportion index for reduction was calculated as:

$$\rho = \frac{\sigma^2(\text{One-Way ANOVA}) - \sigma^2(\text{Random-Coefficients Regression})}{\sigma^2(\text{One-Way ANOVA})}$$

$$= 2992.15767 - 2613.24882 / 2992.15767$$

$$= 0.12663$$

When the student-level factors; out-of-school study time for science, environmental awareness, environmental optimism, enjoyment of science, epistemological beliefs, science activities, students' expected occupational status and science learning time (minutes per week) were included as the predictors of procedural and epistemic knowledge scores, a reduction was observed in within school variance. This finding suggests that, these variables explained 13% of variance in Turkish students' procedural and epistemic knowledge scores.

4.2.3.4 Explaining Phenomena Scientifically

Level-1 equation for explaining phenomena scientifically score becomes:

$$(Y/PV1SCEP, PV2SCEP, PV3SCEP, PV4SCEP, PV5SCEP, PV6SCEP, PV7SCEP, PV8SCEP, PV9SCEP, PV10SCEP) = \beta_{0j} + \beta_{1j}*(ST071Q01_{ij}) + \beta_{2j}*(ENVAWARE_{ij}) + \beta_{3j}*(ENVOPT_{ij}) + \beta_{4j}*(JOYSCIE_{ij}) + \beta_{5j}*(INTBRSCI_{ij}) + \beta_{6j}*(SCIEEFF_{ij}) + \beta_{7j}*(EPIST_{ij}) + \beta_{8j}*(SCIEACT_{ij}) + \beta_{9j}*(BSMJ_{ij}) + \beta_{10j}*(SMINS_{ij}) + r_{ij}.$$

and level-2 equations becomes:

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

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

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60j}$$

$$\beta_{7j} = \gamma_{70j}$$

$$\beta_{8j} = \gamma_{80}$$

$$\beta_{9j} = \gamma_{90}$$

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

where

β_{0j} : mean on explaining phenomena scientifically

β_{1j} : the differentiating effect of out-of-school study time for science

β_{2j} : the differentiating effect of environmental awareness

β_{3j} : the differentiating effect of environmental optimism

β_{4j} : the differentiating effect of enjoyment of science

β_{5j} : the differentiating effect of interest in broad science topics

β_{6j} : the differentiating effect of science self-efficacy

β_{7j} : the differentiating effect of epistemological beliefs

β_{8j} : the differentiating effect of science activities

β_{9j} : the differentiating effect of students' expected occupational status

β_{10j} : the differentiating effect of science learning time (minutes per week)

According to the error variances in the variance components table, slopes of out-of-school study time for science, environmental awareness, enjoyment of science, interest in broad science topics, science self-efficacy, epistemological beliefs, science activities and students' expected occupational status were determined as fixed across schools. The slopes of remaining variables were random.

Table 4.43 represents the final estimation of fixed effects in random-coefficients regression model. The student-level variables which effect Turkish students' explaining phenomena scientifically scores were described according to this table.

Environmental awareness ($\gamma_{20} = 4.828$, $p < 0.050$), enjoyment of science ($\gamma_{40} = 3.884$, $p < 0.050$), interest in broad science topics ($\gamma_{50} = 3.405$, $p < 2.020$), science self-efficacy ($\gamma_{60} = 2.584$, $p < 0.050$), epistemological beliefs ($\gamma_{70} = 3.105$, $p < 0.050$), students' expected occupational status ($\gamma_{90} = 0.172$, $p < 0.050$) and science learning time (minutes per week) ($\gamma_{100} = 0.071$, $p < 0.050$) had positive association with explaining phenomena scientifically scores; whereas out-of-school study time for science ($\gamma_{10} = -1.217$, $p < 0.050$), environmental optimism ($\gamma_{30} = -7.661$, $p < 0.050$) and

science activities ($\gamma_{80} = -2.787$, $p < 0.050$) had negative association with explaining phenomena scientifically scores of Turkish students.

Table 4.43 Final Estimation of Fixed Effects for Explaining Phenomena Scientifically

Fixed Effect	Coefficient	Standard error	t-ratio	Appr. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.122316	4.488129	93.162	176	<0.001
For <i>ST071Q01</i> slope β_1					
INTRCPT2, γ_{10}	-1.216765	0.244390	-4.979	75	<0.001
For <i>ENVAWARE</i> slope β_2					
INTRCPT2, γ_{20}	4.828280	1.015426	4.755	28	<0.001
For <i>ENVOPT</i> slope β_3					
INTRCPT2, γ_{30}	-7.661022	1.018016	-7.525	34	<0.001
For <i>JOYSCIE</i> slope β_4					
INTRCPT2, γ_{40}	3.884101	1.049986	3.699	72	<0.001
For <i>INTBRSCI</i> slope β_5					
INTRCPT2, γ_{50}	3.405215	1.259805	2.703	47	0.010
For <i>SCIEEFF</i> slope β_6					
INTRCPT2, γ_{60}	2.584243	0.809100	3.194	160	0.002
For <i>EPIST</i> slope β_7					
INTRCPT2, γ_{70}	3.105491	1.018586	3.049	50	0.004
For <i>SCIEACT</i> slope β_8					
INTRCPT2, γ_{80}	-2.787100	1.150521	-2.422	47	0.019
For <i>BSMJ</i> slope β_9					
INTRCPT2, γ_{90}	0.171525	0.063460	2.703	231	0.007
For <i>SMINS</i> slope β_{10}					
INTRCPT2, γ_{100}	0.070511	0.012113	5.821	119	<0.001

Table 4.44 represents the estimations of variance components, in which student-level variables with random components are demonstrated.

Table 4.44 Final Estimation of Variance Components for Explaining Phenomena Scientifically

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	56.73197	3218.51596	160	5442.09268	<0.001
<i>ENVOPT</i> slope, u_3	3.73782	13.97127	160	202.74058	0.012
<i>SMINS</i> slope, u_{10}	0.05749	0.00330	160	195.62452	0.029
Student effect, r	53.20074	2830.31901			

The relationship between the variables, which have significant error terms in final estimation of variance components table, and outcome variable is effected by at least one of the school-related factors. This is called “mediation effect”. In order to examine the effects of significant school-level variables on the relationship between student-level variables and outcome variable, school-level variables are included in the further steps of the analysis.

The variances of one-way ANOVA submodel and random-coefficients regression submodel are compared in order to investigate the variance when these factors are included in the model. So, the proportion index for reduction was calculated as:

$$\begin{aligned} \rho &= \sigma^2(\text{One-Way ANOVA}) - \sigma^2(\text{Random-Coefficients Regression}) / \sigma^2(\text{One-Way ANOVA}) \\ &= 3281.88161 - 2830.31901 / 3281.88161 \\ &= 0.1376 \end{aligned}$$

When the student-level factors; out-of-school study time for science, environmental awareness, environmental optimism, enjoyment of science, interest in broad science topics, science self-efficacy, epistemological beliefs, science activities, students’ expected occupational status and science learning time (minutes per week) were included as the predictors of explaining phenomena scientifically scores, a reduction was observed in within school variance. This finding suggests, these variables explained 14% of variance in Turkish students’ explaining phenomena scientifically scores.

4.2.3.5 Evaluating and Designing Scientific Inquiry

Level-1 equation for evaluating and designing scientific inquiry score becomes:

$$(Y/PV1SCED, PV2SCED, PV3SCED, PV4SCED, PV5SCED, PV6SCED, PV7SCED, PV8SCED, PV9SCED, PV10SCED) = \beta_{0j} + \beta_{1j}*(ST071Q01_{ij}) + \beta_{2j}*(ENVAWARE_{ij}) + \beta_{3j}*(ENVOPT_{ij}) + \beta_{4j}*(EPIST_{ij}) + \beta_{5j}*(SCIEACT_{ij}) + \beta_{6j}*(BSMJ_{ij}) + \beta_{7j}*(SMINS_{ij}) + r_{ij}.$$

and level-2 equations becomes:

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

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

$$\beta_{2j} = \gamma_{200} + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60j}$$

$$\beta_{7j} = \gamma_{70j} + u_{7j}$$

where

β_{0j} : mean on evaluating and designing scientific inquiry

β_{1j} : the differentiating effect of out-of-school study time for science

β_{2j} : the differentiating effect of environmental awareness

β_{3j} : the differentiating effect of environmental optimism

β_{4j} : the differentiating effect of epistemological beliefs

β_{5j} : the differentiating effect of science activities

β_{6j} : the differentiating effect of students' expected occupational status

β_{7j} : the differentiating effect of science learning time (minutes per week)

According to the error variances in the variance components table, slopes of out-of-school study time for science, epistemological beliefs, science activities and

students' expected occupational status were determined as fixed across schools. The slopes of remaining variables were random.

Table 4.45 represents the final estimation of fixed effects in random-coefficients regression model. The student-level variables which effect Turkish students' evaluating and designing scientific inquiry scores were described according to this table.

Environmental awareness ($\gamma_{20}= 5.954$, $p<0.050$), epistemological beliefs ($\gamma_{40}= 4.972$, $p<0.050$), students' expected occupational status ($\gamma_{60}= 0.260$, $p<0.050$) and science learning time (minutes per week) ($\gamma_{70}= 0.077$, $p<0.050$) had positive association with evaluating and designing scientific inquiry scores; whereas out-of-school study time for science ($\gamma_{10}= -1.065$, $p<0.050$), environmental optimism ($\gamma_{30}= -5.608$, $p<0.050$) and science activities ($\gamma_{50}= -3.310$, $p<0.050$) had negative association with evaluating and designing scientific inquiry scores of Turkish students.

Table 4.45 Final Estimation of Fixed Effects for Evaluating and Designing Scientific Inquiry

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	420.353712	4.641181	90.570	176	<0.001
For ST071Q01slope β_1					
INTRCPT2, γ_{10}	-1.064726	0.256621	-4.149	51	<0.001
For ENVAWARE slope β_2					
INTRCPT2, γ_{20}	5.954442	1.240171	4.801	19	<0.001
For ENVOPT slope β_3					
INTRCPT2, γ_{30}	-5.607870	1.006626	-5.571	34	<0.001
For EPIST slope β_4					
INTRCPT2, γ_{40}	4.971984	0.993296	5.006	48	<0.001
For SCIEACT slope β_5					
INTRCPT2, γ_{50}	-3.310218	1.194892	-2.770	26	0.010
For BSMJ slope β_6					
INTRCPT2, γ_{60}	0.259614	0.061642	4.212	318	<0.001
For SMINS slope β_7					
INTRCPT2, γ_{70}	0.077489	0.01156	6.700	176	<0.001

Table 4.46 represents the estimations of variance components, in which student-level variables with random components are demonstrated.

Table 4.46 Final Estimation of Variance Components for Evaluating and Designing Scientific Inquiry

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	56.66329	3210.72787	160	5473.86383	<0.001
<i>ENVAWARE</i> slope, u_2	3.64506	13.28643	160	199.80972	0.018
<i>ENVOPT</i> slope, u_3	3.59506	12.92448	160	199.42377	0.019
<i>SMINS</i> slope, u_7	0.05944	0.00353	160	208.08734	0.006
Student effect, r	56.62582	2769.47642			

The relationship between the variables, which have significant error terms in final estimation of variance components table, and outcome variable is effected by at least one of the school-related factors. This is called “mediation effect”. In order to examine the effects of significant school-level variables on the relationship between student-level variables and outcome variable, school-level variables are included in the further steps of the analysis.

The variances of one-way ANOVA submodel and random-coefficients regression submodel are compared in order to investigate the variance when these factors are included in the model. So, the proportion index for reduction was calculated as:

$$\begin{aligned} \rho &= \sigma^2(\text{One-Way ANOVA}) - \sigma^2(\text{Random-Coefficients Regression}) / \sigma^2(\text{One-Way ANOVA}) \\ &= 3157.89361 - 2769.47642 / 3157.89361 \\ &= 0.1230 \end{aligned}$$

When the student-level factors; out-of-school study time for science, environmental awareness, environmental optimism, epistemological beliefs, science activities, students’ expected occupational status and science learning time (minutes per week) are included as the predictors of evaluating and designing scientific inquiry scores, a reduction was observed in within school variance. This finding suggests that, these

variables explained 12% of variance in Turkish students' evaluating and designing scientific inquiry scores.

4.2.3.6 Interpreting Data and Evidence Scientifically

Level-1 equation for interpreting data and evidence scientifically score becomes:

$$(Y/PV1SCID, PV2SCID, PV3SCID, PV4SCID, PV5SCID, PV6SCID, PV7SCID, PV8SCID, PV9SCID, PV10SCID) = \beta_{0j} + \beta_{1j}*(ST071Q01_{ij}) + \beta_{2j}*(ENVAWARE_{ij}) + \beta_{3j}*(ENVOPT_{ij}) + \beta_{4j}*(JOYSCIE_{ij}) + \beta_{5j}*(INTBRSCI_{ij}) + \beta_{6j}*(EPIST_{ij}) + \beta_{7j}*(SCIEACT_{ij}) + \beta_{8j}*(BSMJ_{ij}) + \beta_{9j}*(SMINS_{ij}) + r_{ij}.$$

and level-2 equations becomes:

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

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

$$\beta_{2j} = \gamma_{20} + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60j}$$

$$\beta_{7j} = \gamma_{70j}$$

$$\beta_{8j} = \gamma_{80}$$

$$\beta_{9j} = \gamma_{90} + u_{9j}$$

where

β_{0j} : mean on interpreting data and evidence scientifically

β_{1j} : the differentiating effect of out-of-school study time for science

β_{2j} : the differentiating effect of environmental awareness

β_{3j} : the differentiating effect of environmental optimism

β_{4j} : the differentiating effect of enjoyment of science

β_{5j} : the differentiating effect of interest in broad science topics

β_{6j} : the differentiating effect of epistemological beliefs

β_{7j} : the differentiating effect of science activities

β_{8j} : the differentiating effect of students' expected occupational status

β_{9j} : the differentiating effect of science learning time (minutes per week)

According to the error variances in the variance components table, slopes of out-of-school study time for science, enjoyment of science, interest in broad science topics, epistemological beliefs, science activities and students' expected occupational status were taken as determined across schools. The slopes of remaining variables were random.

Table 4.47 represents the final estimation of fixed effects in random-coefficients regression model. The student-level variables which effect Turkish students' interpreting data and evidence scientifically scores were described according to this table.

Environmental awareness ($\gamma_{20}= 3.940$, $p<0.050$), enjoyment of science ($\gamma_{40}= 2.869$, $p<0.050$), interest in broad science topics ($\gamma_{50}= 2.466$, $p< 2.020$), epistemological beliefs ($\gamma_{60}= 3.987$, $p<0.050$), students' expected occupational status ($\gamma_{80}= 0.167$, $p<0.050$) and science learning time (minutes per week) ($\gamma_{90}= 0.080$, $p<0.050$) had positive association with interpreting data and evidence scientifically scores; whereas out-of-school study time for science ($\gamma_{10}= -1.293$, $p<0.050$), environmental optimism ($\gamma_{30}= -4.662$, $p<0.050$) and science activities ($\gamma_{70}= -4.651$, $p<0.050$) had negative association with interpreting data and evidence scientifically scores of Turkish students.

Table 4.48 represents the estimations of variance components, in which student-level variables with random components are demonstrated.

The relationship between the variables, which have significant error terms in final estimation of variance components table, and outcome variable is effected by at least one of the school-related factors. This is called "mediation effect". In order to examine the effects of significant school-level variables on the relationship between

student-level variables and outcome variable, school-level variables are included in the further steps of the analysis.

The variances of one-way ANOVA submodel and random-coefficients regression submodel are compared in order to investigate the variance when these factors are included in the model. So, the proportion index for reduction was calculated as:

$$\rho = \sigma^2(\text{One-Way ANOVA}) - \sigma^2(\text{Random-Coefficients Regression}) / \sigma^2(\text{One-Way ANOVA})$$

$$= 3204.66784 - 2839.83227 / 3204.66784$$

$$= 0.11385$$

Table 4.47 Final Estimation of Fixed Effects for Interpreting Data and Evidence Scientifically

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	415.830565	4.598040	90.436	176	<0.001
For ST071Q01slope β_1					
INTRCPT2, γ_{10}	-1.292700	0.243449	-5.310	81	<0.001
For ENVAWARE slope β_2					
INTRCPT2, γ_{20}	3.939889	0.884797	4.453	86	<0.001
For ENVOPT slope β_3					
INTRCPT2, γ_{30}	-4.661551	0.990967	-4.704	47	<0.001
For JOYSCIE slope β_4					
INTRCPT2, γ_{40}	2.869113	1.185341	2.420	38	0.020
For INTBRSCI slope β_5					
INTRCPT2, γ_{60}	2.466051	1.130460	2.181	106	0.031
For EPIST slope β_6					
INTRCPT2, γ_{60}	3.986781	1.411912	2.824	18	0.011
For SCIEACT slope β_7					
INTRCPT2, γ_{70}	-4.651188	0.961939	-4.835	165	<0.001
For BSMJ slope β_8					
INTRCPT2, γ_{80}	0.166602	0.063377	2.629	261	0.009
For SMINS slope β_9					
INTRCPT2, γ_{90}	0.080088	0.012531	6.391	127	<0.001

Table 4.48 Final Estimation of Variance Components for Interpreting Data and Evidence Scientifically

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	56.55325	3198.27024	160	5279.88225	<0.001
<i>EVVAWARE</i> slope, u_2	3.39752	11.54317	160	191.03511	0.047
<i>ENVOPT</i> slope, u_3	4.41826	19.52104	160	211.81726	0.004
<i>SMINS</i> slope, u_9	0.06648	0.00442	160	208.16186	0.006
Student effect, r	53.29008	2839.83227			

When the student-level factors; out-of-school study time for science, environmental awareness, environmental optimism, enjoyment of science, interest in broad science topics, epistemological beliefs, science activities, students' expected occupational status and science learning time (minutes per week) are included as the predictors of interpreting data and evidence scientifically scores, a reduction was observed in within school variance. This finding suggests that, these variables explained 11% of variance in Turkish students' interpreting data and evidence scientifically scores.

4.2.3.7 Findings from Random-Coefficients Regression Model

1. Instrumental motivation (INSTSCIE) had no significant effect on any of the scientific literacy skills of Turkish students.
2. Out-of-school study time for science (ST071Q01), epistemological beliefs (EPIST), environmental awareness (ENVAWARE), environmental optimism (ENVOPT), students' expected occupational status (BSMJ), science learning time (SMINS) and science activities (SCIEACT) effected all scientific literacy skills significantly. Among those variables; out-of-school study time for science, science activities and environmental optimism were negatively associated with all scientific literacy skills of Turkish students, whereas other variables were positively related.
3. Science self-efficacy (SCIEEFF) predicted combined science, content knowledge and explaining phenomena scientifically subscales. Interest in broad science topics (INTBRSCI) predicted interpreting data and evidence scientifically additionally. The relationship between these variables and scores of these subscales was found to be

positive. Enjoyment of science (JOYSCIE) had a positive relationship with all subscales except for evaluating and designing scientific inquiry subscale.

4. The proportion of variances explained by random-coefficients regression models of combined science, science competencies and science knowledge changed between 11% and 14%.

4.2.4 Intercepts- and Slopes-as-Outcomes Model

The intercepts- and slopes-as-outcomes submodel gives information about which school-related variables effect the student-related variables, which have significant effect on scientific literacy skills of Turkish students. In other words, the cross-level interactions of school- and student-level variables are explored. One-way ANOVA with random effects, means-as-outcomes regression, and random-coefficients regression models are incorporated in the full model. In addition, randomly varying school effects (varying slopes) are examined with this submodel. The variability in intercepts and slopes across schools is modeled. So, the following research question can be answered through this analysis: “Which cross-level interactions of school- and student-level variables affect the students’ scientific literacy skills in Turkey?”.

In the present study, the analyses started with the random-coefficients regression, and the school-level variables, which were found to be effecting scientific literacy skills of students in means-as-outcomes model significantly, were included in the analysis. School-level variables were entered for each of the level-2 equations with random effects, and they were checked, whether they have significant effects on the corresponding slopes. The school-level variables having non-significant effect on slopes were discarded from the model starting with the least significant variable. This procedure was repeated with the next least significant variable until all variables in the model become significant.

4.2.4.1 Combined Science

In the intercepts- and slopes-as-outcomes model, level-1 equation for combined science score becomes:

$$(Y/PV1SCIE, PV2SCIE, PV3SCIE, PV4SCIE, PV5SCIE, PV6SCIE, PV7SCIE, PV8SCIE, PV9SCIE, PV10SCIE) = \beta_{0j} + \beta_{1j}*(ST071Q01_{ij}) + \beta_{2j}*(ENVAWARE_{ij}) + \beta_{3j}*(ENVOPT_{ij}) + \beta_{4j}*(JOYSCIE_{ij}) + \beta_{5j}*(INTBRSCI_{ij}) + \beta_{6j}*(SCIEEFF_{ij}) + \beta_{7j}*(EPIST_{ij}) + \beta_{8j}*(SCIEACT_{ij}) + \beta_{9j}*(BSMJ_{ij}) + \beta_{10j}*(SMINS_{ij}) + r_{ij}.$$

and for level-2, the final equations becomes:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(CREACTIV_j) + \gamma_{04}*(SCIRES_j) + u_{0j}$$

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

$$\beta_{2j} = \gamma_{20} + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60j}$$

$$\beta_{7j} = \gamma_{70j}$$

$$\beta_{8j} = \gamma_{80}$$

$$\beta_{9j} = \gamma_{90}$$

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

Table 4.49 represents the final estimations of fixed effects for the full model.

The results showed that mean combined science score across schools was significantly associated with students' combined science scores. Shortage of educational material ($\gamma_{01} = -7.937$, $p < 0.050$) was negatively related to combined science scores of students. On the other hand, proportion of science teachers by all teachers ($\gamma_{02} = 158.574$, $p < 0.050$), creative extra-curricular activities ($\gamma_{03} = 7.963$, $p < 0.050$) and science specific resources ($\gamma_{04} = 4.897$, $p < 0.050$) were positively related to combined science scores of students.

Out-of-school study time for science ($\gamma_{10} = -1.176$, $p < 0.050$), environmental optimism ($\gamma_{30} = -6.170$, $p < 0.050$) and science activities ($\gamma_{80} = -3.538$, $p < 0.050$) had a negative effect on students' combined science scores; however, environmental awareness

($\gamma_{20}= 4.590$, $p<0.050$), enjoyment of science ($\gamma_{40}= 3.276$, $p<0.050$), interest in broad science topics ($\gamma_{50}= 2.191$, $p<0.050$), science self-efficacy ($\gamma_{60}= 1.534$, $p<0.050$), epistemological beliefs ($\gamma_{70}= 4.250$, $p<0.050$), students' expected occupational status ($\gamma_{90}= 0.182$, $p<0.050$) and science learning time (minutes per week) ($\gamma_{100}= 0.076$, $p<0.050$) were positively related to students' combined science scores.

Table 4.49 Estimation of Fixed Effects for Combined Science

Fixed Effect	Coefficient	Standard error	t-ratio	Approx. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.321207	3.805905	109.914	172	<0.001
EDUSHORT, γ_{01}	-7.936848	2.976465	-2.667	172	0.008
PROSTAT, γ_{02}	158.574373	57.479236	2.759	172	0.006
CREACTIV, γ_{03}	7.963057	3.828924	2.080	172	0.039
SCIERES, γ_{04}	4.897224	1.818469	2.693	172	0.008
For <i>ST071Q01</i> slope β_1	-1.176192	0.195306	-6.022	783	<0.001
INTRCPT2, γ_{10}					
For <i>ENVAWARE</i> slope β_2	4.590177	0.847799	5.414	75	<0.001
INTRCPT2, γ_{20}					
For <i>ENVOPT</i> slope β_3	-6.169597	0.849424	-7.263	70	<0.001
INTRCPT2, γ_{30}					
For <i>JOYSCIE</i> slope β_4	3.275891	0.908778	3.605	146	<0.001
INTRCPT2, γ_{40}					
For <i>INTBRSCI</i> slope β_5	2.191372	0.967988	2.264	316	0.024
INTRCPT2, γ_{50}					
For <i>SCIEEFF</i> slope β_6	1.534132	0.697667	2.199	832	0.028
INTRCPT2, γ_{60}					
For <i>EPIST</i> slope β_7	4.250069	0.856514	4.962	107	<0.001
INTRCPT2, γ_{70}					
For <i>SCIEACT</i> slope β_8	-3.538194	0.950475	-3.723	113	<0.001
INTRCPT2, γ_{80}					
For <i>BSMJ</i> slope β_9	0.182090	0.059094	3.081	236	0.002
INTRCPT2, γ_{90}					
For <i>SMINS</i> slope β_{10}	0.076259	0.011751	6.489	109	<0.001
INTRCPT2, γ_{100}					

Table 4.50 represents estimations of variance components for combined science for the full model.

Table 4.50 Final Estimation of Variance Components for Combined Science

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	47.27915	2235.31778	156	4178.68894	<0.001
ENVAWARE slope, u_2	3.30038	10.89254	160	196.02912	0.027
ENVOPT slope, u_3	3.87065	14.98191	160	211.66746	0.004
SMINS slope, u_{10}	0.05967	0.00356	160	211.69450	0.004
Student effect, r	49.38516	2438.89397			

The proportion reduction in variance based on school-related factors variables is:

$$\tau_{qq}(\text{Random-Coefficients Regression}) - \tau_{qq}(\text{Intercepts- and -Slopes}) / \tau_{qq}(\text{Random-Coefficients Regression})$$

$$= 2980.71445 - 2235.31778 / 2980.71445$$

$$= 0.250$$

This finding suggests that, school-level variables in Table 4.49 explained 25% of the variance in combined science scores of Turkish students.

The results indicate that, the effect of environmental awareness, environmental optimism and science learning time (minutes per week) on combined science scores of Turkish students increases, when the proportion of science teachers by all teachers, creative extra-curricular activities and science specific resources increases. On the other hand, the effect of those variables on combined science scores of Turkish students decreases when shortage of educational material increases.

4.2.4.2 Content Knowledge

In the intercepts- and slopes-as-outcomes model, level-1 equation for content knowledge score becomes:

$$(Y/PV1SKCO, PV2SKCO, PV3SKCO, PV4SKCO, PV5SKCO, PV6SKCO, PV7SKCO, PV8SKCO, PV9SKCO, PV10SKCO) = \beta_{0j} + \beta_{1j}*(ST071Q01_{ij}) + \beta_{2j}*(ENVAWARE_{ij}) + \beta_{3j}*(ENVOPT_{ij}) + \beta_{4j}*(JOYSCIE_{ij}) + \beta_{5j}*(INTBRSCI_{ij}) + \beta_{6j}*(EPIST_{ij}) + \beta_{7j}*(SCIEACT_{ij}) + \beta_{8j}*(BSMJ_{ij}) + \beta_{9j}*(SMINS_{ij}) + r_{ij}.$$

and for level-2, the final equations becomes:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(SCIRES_j) + u_{0j}$$

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

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60j}$$

$$\beta_{7j} = \gamma_{70j}$$

$$\beta_{8j} = \gamma_{80}$$

$$\beta_{9j} = \gamma_{90} + u_{9j}$$

Table 4.51 represents the final estimations of fixed effects for the full model.

The results showed that mean content knowledge score across schools was significantly associated with students' content knowledge scores. Shortage of educational material ($\gamma_{01} = -8.091$, $p < 0.050$) was negatively related to content knowledge scores of students. On the other hand, proportion of science teachers by all teachers ($\gamma_{02} = 174.735$, $p < 0.050$) and science specific resources ($\gamma_{03} = 6.067$, $p < 0.050$) were positively related to content knowledge scores of students.

Out-of-school study time for science ($\gamma_{10} = -1.293$, $p < 0.050$), environmental optimism ($\gamma_{30} = -7.072$, $p < 0.050$) and science activities ($\gamma_{70} = -2.213$, $p < 0.050$) had a negative effect on students' content knowledge scores; however, environmental awareness ($\gamma_{20} = 5.466$, $p < 0.050$), enjoyment of science ($\gamma_{40} = 3.969$, $p < 0.050$), interest in broad science topics ($\gamma_{50} = 3.211$, $p < 0.050$), epistemological beliefs ($\gamma_{60} = 2.887$, $p < 0.050$), students' expected occupational status ($\gamma_{80} = 0.152$, $p < 0.050$) and science learning

time (minutes per week) ($\gamma_{90} = 0.077$, $p < 0.050$) were positively related to students' content knowledge scores.

Table 4.51 Estimation of Fixed Effects for Content Knowledge

Fixed Effect	Coefficient	Standard error	t-ratio	Appr. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.011123	3.910261	106.901	173	<0.001
EDUSHORT, γ_{01}	-8.090804	3.196223	-2.531	173	0.012
PROSTAT, γ_{02}	174.734981	60.679419	2.880	173	0.004
SCIERES, γ_{03}	6.067350	1.901104	3.191	173	0.002
For ST071Q01 slope β_1					
INTRCPT2, γ_{10}	-1.292946	0.223154	-5.794	211	<0.001
For ENVAWARE slope β_2					
INTRCPT2, γ_{20}	5.466100	0.965087	5.664	33	<0.001
For ENVOPT slope β_3					
INTRCPT2, γ_{30}	-7.072236	0.974521	-7.257	41	<0.001
For JOYSCIE slope β_4					
INTRCPT2, γ_{40}	3.969388	1.015178	3.910	99	<0.001
For INTBRSCI slope β_5					
INTRCPT2, γ_{50}	3.210962	1.207598	2.659	61	0.010
For EPIST slope β_6					
INTRCPT2, γ_{60}	2.886876	0.919740	3.139	110	0.002
For SCIEACT slope β_7					
INTRCPT2, γ_{70}	-2.212562	1.070901	-2.066	61	0.043
For BSMJ slope β_8					
INTRCPT2, γ_{80}	0.152390	0.071002	2.146	72	0.035
For SMINS slope β_9					
INTRCPT2, γ_{90}	0.076807	0.011159	6.883	176	<0.001

Table 4.52 represents estimations of variance components for content knowledge for the full model.

Table 4.52 Final Estimation of Variance Components for Content Knowledge

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	48.77446	2378.94778	157	3798.07316	<0.001
ENVOPT slope, u_3	3.66096	13.40263	160	202.34477	0.013
SMINS slope, u_9	0.05574	0.00311	160	191.09071	0.047
Student effect, r	53.46893	2858.92603			

The proportion reduction in variance based on school-related factors variables is:

$$\tau_{qq}(\text{Random-Coefficients Regression}) - \tau_{qq}(\text{Intercepts- and -Slopes}) / \tau_{qq}(\text{Random-Coefficients Regression})$$

$$= 3084.07493 - 2378.94788 / 3084.07493$$

$$= 0.2286$$

This finding suggests that, school-level variables in Table 4.51 explained 23% of the variance in content knowledge scores of Turkish students.

The results indicate that, the effect of environmental optimism and science learning time (minutes per week) on content knowledge scores of Turkish students increases, when the proportion of science teachers by all teachers and science specific resources increases. On the other hand, the effect of those variables on content knowledge scores of Turkish students decreases when shortage of educational material increases.

4.2.4.3 Procedural and Epistemic Knowledge

In the intercepts- and slopes-as-outcomes model, level-1 equation for procedural and epistemic knowledge score becomes:

$$(Y/PV1SKPE, PV2SKPE, PV3SKPE, PV4SKPE, PV5SKPE, PV6SKPE, PV7SKPE, PV8SKPE, PV9SKPE, PV10SKPE) = \beta_{0j} + \beta_{1j}*(ST071Q01_{ij}) + \beta_{2j}*(ENVAWARE_{ij}) + \beta_{3j}*(ENVOPT_{ij}) + \beta_{4j}*(JOYSCIE_{ij}) + \beta_{5j}*(EPIST_{ij}) + \beta_{6j}*(SCIEACT_{ij}) + \beta_{7j}*(BSMJ_{ij}) + \beta_{8j}*(SMINS_{ij}) + r_{ij}.$$

and for level-2, the final equations becomes:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(CREACTIV_j) + \gamma_{04}*(SCIERES_j) + u_{0j}$$

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

$$\beta_{2j} = \gamma_{200} + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60j}$$

$$\beta_{7j} = \gamma_{70j}$$

$$\beta_{8j} = \gamma_{800} + u_{8j}$$

Table 4.53 represents the final estimations of fixed effects for the full model.

The results showed that mean procedural and epistemic knowledge score across schools was significantly associated with students' procedural and epistemic knowledge scores. Shortage of educational material ($\gamma_{01} = -8.269$, $p < 0.050$) was negatively related to procedural and epistemic knowledge scores of students. On the other hand, proportion of science teachers by all teachers ($\gamma_{02} = 163.501$, $p < 0.050$), creative extra-curricular activities ($\gamma_{03} = 9.070$, $p < 0.050$) and science specific resources ($\gamma_{04} = 4.550$, $p < 0.050$) were positively related to procedural and epistemic knowledge scores of students.

Out-of-school study time for science ($\gamma_{10} = -1.240$, $p < 0.050$), environmental optimism ($\gamma_{30} = -5.497$, $p < 0.050$) and science activities ($\gamma_{60} = -4.032$, $p < 0.050$) had a negative effect on students' procedural and epistemic knowledge scores; however, environmental awareness ($\gamma_{20} = 4.245$, $p < 0.050$), enjoyment of science ($\gamma_{40} = 3.043$, $p < 0.050$), epistemological beliefs ($\gamma_{50} = 4.766$, $p < 0.050$), students' expected occupational status ($\gamma_{70} = 0.188$, $p < 0.050$) and science learning time (minutes per week) ($\gamma_{80} = 0.085$, $p < 0.050$) were positively related to students' procedural and epistemic knowledge scores.

Table 4.53 Final Estimation of Fixed Effects for Procedural and Epistemic Knowledge

Fixed Effect	Coefficient	Standard error	t-ratio	Approxd.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	417.504648	3.905022	106.915	172	<0.001
EDUSHORT, γ_{01}	-8.268637	3.150541	-2.625	172	0.009
PROSTAT, γ_{02}	163.501441	59.858743	2.731	172	0.007
CREACTIV, γ_{03}	9.070382	3.963287	2.289	172	0.023
SCIERES, γ_{04}	4.549946	1.970161	2.309	172	0.022
For ST071Q01slope β_1					
INTRCPT2, γ_{10}	-1.239474	0.203539	-6.090	615	<0.001
For ENVAWARE slope β_2					
INTRCPT2, γ_{20}	4.249497	0.975948	4.354	35	<0.001
For ENVOPT slope β_3					
INTRCPT2, γ_{30}	-5.496581	0.906414	-6.064	55	<0.001
For JOYSCIE slope β_4					
INTRCPT2, γ_{40}	3.043335	0.962425	3.162	86	0.002
For EPIST slope β_5					
INTRCPT2, γ_{50}	4.766161	0.888022	5.367	102	<0.001
For SCIEACT slope β_6					
INTRCPT2, γ_{60}	-4.032004	0.997578	-4.042	62	<0.001
For BSMJ slope β_7					
INTRCPT2, γ_{70}	0.187858	0.068028	2.761	73	0.007
For SMINS slope β_8					
INTRCPT2, γ_{80}	0.084785	0.012196	6.952	99	<0.001

Table 4.54 represents estimations of variance components for procedural and epistemic knowledge for the full model.

Table 4.54 Final Estimation of Variance Components for Procedural and Epistemic Knowledge

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	48.14963	2318.38665	156	4064.51665	<0.001
<i>ENVAWARE</i> slope, u_2	3.27193	10.70550	160	191.19768	0.046
<i>ENVOPT</i> slope, u_3	3.88557	15.09769	160	204.34569	0.010
<i>SMINS</i> slope, u_8	0.06026	0.00363	160	210.64082	0.005
Student effect, r	51.12829	2614.10215			

The proportion reduction in variance based on school-related factors variables is:

$$\tau_{qq}(\text{Random-Coefficients Regression}) - \tau_{qq}(\text{Intercepts- and -Slopes}) / \tau_{qq}(\text{Random-Coefficients Regression})$$

$$= 3074.21754 - 2318.38665 / 3074.21754$$

$$= 0.246$$

This finding suggests that, school-level variables in Table 4.53 explained 25% of the variance in procedural and epistemic knowledge scores of Turkish students.

The results indicate that, the effect of environmental awareness, environmental optimism and science learning time (minutes per week) on procedural and epistemic knowledge scores of Turkish students increases, when the proportion of science teachers by all teachers, creative extra-curricular activities and science specific resources increases. On the other hand, the effect of those variables on procedural and epistemic knowledge scores of Turkish students decreases when shortage of educational material increases.

4.2.4.4 Explaining Phenomena Scientifically

In the intercepts- and slopes-as-outcomes model, level-1 equation for explaining phenomena scientifically score becomes:

$$(Y/PV1SCEP, PV2SCEP, PV3SCEP, PV4SCEP, PV5SCEP, PV6SCEP, PV7SCEP, PV8SCEP, PV9SCEP, PV10SCEP) = \beta_{0j} + \beta_{1j}*(ST071Q01_{ij}) + \beta_{2j}*(ENVAWARE_{ij}) + \beta_{3j}*(ENVOPT_{ij}) + \beta_{4j}*(JOYSCIE_{ij}) + \beta_{5j}*(INTBRSCI_{ij}) + \beta_{6j}*(SCIEEFF_{ij}) + \beta_{7j}*(EPIST_{ij}) + \beta_{8j}*(SCIEACT_{ij}) + \beta_{9j}*(BSMJ_{ij}) + \beta_{10j}*(SMINS_{ij}) + r_{ij}.$$

and for level-2, the final equations becomes:

$$\beta_{0j} = \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(SCIERES) + u_{0j}$$

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

$$\beta_{2j} = \gamma_{20}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60j}$$

$$\beta_{7j} = \gamma_{70j}$$

$$\beta_{8j} = \gamma_{80}$$

$$\beta_{9j} = \gamma_{90}$$

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

Table 4.55 represents the final estimations of fixed effects for the full model.

The results indicated that mean explaining phenomena scientifically score across schools was significantly associated with students' explaining phenomena scientifically scores. Shortage of educational material ($\gamma_{01} = -8.411$, $p < 0.050$) was negatively related to explaining phenomena scientifically scores of students. On the other hand, proportion of science teachers by all teachers ($\gamma_{02} = 190.830$, $p < 0.050$) and science specific resources ($\gamma_{03} = 6.199$, $p < 0.050$) were positively related to explaining phenomena scientifically scores of students.

Table 4.55 Final Estimation of Fixed Effects for Explaining Phenomena Scientifically

Fixed Effect	Coefficient	Standard error	t-ratio	Appr. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	418.224222	3.979667	105.090	173	<0.001
EDUSHORT, γ_{01}	-8.411177	3.278040	-2.566	173	0.011
PROSTAT, γ_{02}	190.830436	62.539674	3.051	173	0.003
SCIERES, γ_{03}	6.198949	1.881517	3.295	173	0.001
For <i>ST071Q01</i> slope β_1					
INTRCPT2, γ_{10}	-1.222527	0.244457	-5.001	76	<0.001
For <i>ENVAWARE</i> slope β_2					
INTRCPT2, γ_{20}	4.855937	1.015244	4.783	28	<0.001
For <i>ENVOPT</i> slope β_3					
INTRCPT2, γ_{30}	-7.641136	1.020199	-7.490	34	<0.001
For <i>JOYSCIE</i> slope β_4					
INTRCPT2, γ_{40}	3.886367	1.049070	3.705	73	<0.001
For <i>INTBRSCI</i> slope β_5					
INTRCPT2, γ_{50}	3.528221	1.253431	2.815	48	0.007
For <i>SCIEEFF</i> slope β_6					
INTRCPT2, γ_{60}	2.559693	0.811635	3.154	155	0.002
For <i>EPIST</i> slope β_7					
INTRCPT2, γ_{70}	3.135846	1.021286	3.070	50	0.003
For <i>SCIEACT</i> slope β_8					
INTRCPT2, γ_{80}	-2.851544	1.150851	-2.478	47	0.017
For <i>BSMJ</i> slope β_9					
INTRCPT2, γ_{90}	0.168207	0.063383	2.654	240	0.008
For <i>SMINS</i> slope β_{10}					
INTRCPT2, γ_{100}	0.071702	0.012085	5.933	120	<0.001

Out-of-school study time for science (γ_{10} = -1.223, $p<0.050$), environmental optimism (γ_{30} = -7.641, $p<0.050$) and science activities (γ_{80} = -2.852, $p<0.050$) had a negative effect on students' explaining phenomena scientifically scores; however, environmental awareness (γ_{20} = 4.856, $p<0.050$), enjoyment of science (γ_{40} = 3.886, $p<0.050$), interest in broad science topics (γ_{50} = 3.528, $p<0.050$), science self-efficacy

($\gamma_{60} = 2.560$, $p < 0.050$), epistemological beliefs ($\gamma_{70} = 3.136$, $p < 0.050$), students' expected occupational status ($\gamma_{90} = 0.168$, $p < 0.050$) and science learning time (minutes per week) ($\gamma_{100} = 0.072$, $p < 0.050$) were positively related to students' explaining phenomena scientifically scores.

Table 4.56 represents estimations of variance components for explaining phenomena scientifically for the full model.

Table 4.56 Final Estimation of Variance Components for Explaining Phenomena Scientifically

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	49.71265	2471.34728	157	3948.18166	<0.001
<i>ENVOPT</i> slope, u_3	3.66346	13.42091	160	202.54941	0.013
<i>SMINS</i> slope, u_{10}	0.05613	0.00315	160	195.50471	0.029
Student effect, r	53.22645	2833.05472			

The proportion reduction in variance based on school-related factors variables is:

$$\tau_{qq}(\text{Random-Coefficients Regression}) - \tau_{qq}(\text{Intercepts- and -Slopes}) / \tau_{qq}(\text{Random-Coefficients Regression})$$

$$= 3218.51596 - 2471.34728 / 3218.51596$$

$$= 0.23215$$

This finding suggests that, school-level variables in Table 4.55 explained 23% of the variance in explaining phenomena scientifically scores of Turkish students.

The results indicate that, the effect of environmental optimism and science learning time (minutes per week) on explaining phenomena scientifically scores of Turkish students increases, when the proportion of science teachers by all teachers and science specific resources increases. On the other hand, the effect of those variables on explaining phenomena scientifically scores of Turkish students decreases when shortage of educational material increases.

4.2.4.5 Evaluating and Designing Scientific Inquiry

In the intercepts- and slopes-as-outcomes model, level-1 equation for evaluating and designing scientific inquiry score becomes:

$$(Y/PV1SCED, PV2SCED, PV3SCED, PV4SCED, PV5SCED, PV6SCED, PV7SCED, PV8SCED, PV9SCED, PV10SCED) = \beta_{0j} + \beta_{1j}*(ST071Q01_{ij}) + \beta_{2j}*(ENVAWARE_{ij}) + \beta_{3j}*(ENVOPT_{ij}) + \beta_{4j}*(EPIST_{ij}) + \beta_{5j}*(SCIEACT_{ij}) + \beta_{6j}*(BSMJ_{ij}) + \beta_{7j}*(SMINS_{ij}) + r_{ij}.$$

and for level-2, the final equations becomes:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(TOTST_j) + \gamma_{04}*(CREACTIV_j) + \gamma_{05}*(SCIRES_j) + u_{0j}$$

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

$$\beta_{2j} = \gamma_{200} + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60j}$$

$$\beta_{7j} = \gamma_{70j} + u_{7j}$$

Table 4.57 represents the final estimations of fixed effects for the full model.

The results indicated that mean evaluating and designing scientific inquiry score across schools was significantly associated with students' evaluating and designing scientific inquiry scores. Shortage of educational material ($\gamma_{01} = -8.140$, $p < 0.050$) was negatively related to evaluating and designing scientific inquiry scores of students. On the other hand, proportion of science teachers by all teachers ($\gamma_{02} = 242.649$, $p < 0.050$), creative extra-curricular activities ($\gamma_{03} = 9.445$, $p < 0.050$) and science specific resources ($\gamma_{04} = 4.942$, $p < 0.050$) were positively related to evaluating and designing scientific inquiry scores of students.

Table 4.57 Final Estimation of Fixed Effects for Evaluating and Designing Scientific Inquiry

Fixed Effect	Coefficient	Standard error	t-ratio	Appr. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	420.697887	4.049926	103.878	171	<0.001
EDUSHORT, γ_{01}	-8.139517	3.243258	-2.510	171	0.013
PROSTAT, γ_{02}	242.649097	75.297972	3.223	171	0.002
CREACTIV, γ_{03}	9.445178	4.140651	2.281	171	0.024
SCIERES, γ_{04}	4.942116	1.929706	2.561	171	0.011
For ST071Q01slope β_1					
INTRCPT2, γ_{10}	-1.066748	0.256874	-4.153	51	<0.001
For ENVAWARE slope β_2					
INTRCPT2, γ_{20}	5.941688	1.244208	4.775	19	<0.001
For ENVOPT slope β_3					
INTRCPT2, γ_{30}	-5.647552	1.008600	-5.599	34	<0.001
For EPIST slope β_4					
INTRCPT2, γ_{40}	4.978946	0.995128	5.003	48	<0.001
For SCIEACT slope β_5					
INTRCPT2, γ_{50}	-3.412196	1.189078	-2.870	27	0.008
For BSMJ slope β_6					
INTRCPT2, γ_{60}	0.257873	0.061684	4.181	320	<0.001
For SMINS slope β_7					
INTRCPT2, γ_{70}	0.079192	0.011578	6.840	176	<0.001

Out-of-school study time for science ($\gamma_{10} = -1.067$, $p < 0.050$), environmental optimism ($\gamma_{30} = -5.648$, $p < 0.050$) and science activities ($\gamma_{50} = -3.412$, $p < 0.050$) had a negative effect on students' evaluating and designing scientific inquiry scores; however, environmental awareness ($\gamma_{20} = 5.942$, $p < 0.050$), epistemological beliefs ($\gamma_{40} = 4.979$, $p < 0.050$), students' expected occupational status ($\gamma_{60} = 0.258$, $p < 0.050$) and science learning time (minutes per week) ($\gamma_{70} = 0.079$, $p < 0.050$) were positively related to students' evaluating and designing scientific inquiry scores.

Table 4.58 represents estimations of variance components for evaluating and designing scientific inquiry for the full model.

Table 4.58 Final Estimation of Variance Components for Evaluating and Designing Scientific Inquiry

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	48.11639	2315.18685	155	3706.92465	<0.001
<i>ENVAWARE</i> slope, u_2	3.67873	13.53304	160	199.46250	0.018
<i>ENVOPT</i> slope, u_3	3.54102	12.53883	160	199.19313	0.019
<i>SMINS</i> slope, u_7	0.05624	0.00316	160	207.82373	0.007
Student effect, r	52.65831	2772.89712			

The proportion reduction in variance based on school-related factors variables is:

$$\tau_{qq}(\text{Random-Coefficients Regression}) - \tau_{qq}(\text{Intercepts- and -Slopes}) / \tau_{qq}(\text{Random-Coefficients Regression})$$

$$= 3210.72787 - 2315.18685 / 3210.72787$$

$$= 0.2789$$

This finding suggests that, school-level variables in Table 4.57 explained 28% of the variance in evaluating and designing scientific inquiry scores of Turkish students.

The results indicate that, the effect of environmental awareness, environmental optimism and science learning time (minutes per week) on evaluating and designing scientific inquiry scores of Turkish students increases, when the proportion of science teachers by all teachers, creative extra-curricular activities and science specific resources increases. On the other hand, the effect of those variables on evaluating and designing scientific inquiry scores of Turkish students decreases when shortage of educational material increases.

4.2.4.6 Interpreting Data and Evidence Scientifically

In the intercepts- and slopes-as-outcomes model, level-1 equation for interpreting data and evidence scientifically score becomes:

$$(Y/PV1SCID, PV2SCID, PV3SCID, PV4SCID, PV5SCID, PV6SCID, PV7SCID, PV8SCID, PV9SCID, PV10SCID) = \beta_{0j} + \beta_{1j}*(ST071Q01_{ij}) + \beta_{2j}*(ENVAWARE_{ij}) + \beta_{3j}*(ENVOPT_{ij}) + \beta_{4j}*(JOYSCIE_{ij}) + \beta_{5j}*(INTBRSCI_{ij}) + \beta_{6j}*(EPIST_{ij}) + \beta_{7j}*(SCIEACT_{ij}) + \beta_{8j}*(BSMJ_{ij}) + \beta_{9j}*(SMINS_{ij}) + r_{ij}.$$

and for level-2, the final equations becomes:

$$\beta_{0j} = \gamma_{01}*(EDUSHORT_j) + \gamma_{02}*(PROSTAT_j) + \gamma_{03}*(SCIRES_j) + u_{0j}$$

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

$$\beta_{2j} = \gamma_{20} + u_{2j}$$

$$\beta_{3j} = \gamma_{30} + u_{3j}$$

$$\beta_{4j} = \gamma_{40}$$

$$\beta_{5j} = \gamma_{50}$$

$$\beta_{6j} = \gamma_{60j}$$

$$\beta_{7j} = \gamma_{70j}$$

$$\beta_{8j} = \gamma_{80}$$

$$\beta_{9j} = \gamma_{90} + u_{9j}$$

Table 4.59 represents the final estimations of fixed effects for the full model.

The results indicated that mean interpreting data and evidence scientifically score across schools was significantly associated with students' interpreting data and evidence scientifically scores. Shortage of educational material ($\gamma_{01} = -8.592$, $p < 0.050$) was negatively related to interpreting data and evidence scientifically scores of students. On the other hand, proportion of science teachers by all teachers ($\gamma_{02} = 151.080$, $p < 0.050$) and science specific resources ($\gamma_{03} = 6.112$, $p < 0.050$) were positively related to interpreting data and evidence scientifically scores of students.

Table 4.59 Final Estimation of Fixed Effects for Interpreting Data and Evidence Scientifically

Fixed Effect	Coefficient	Standard error	t-ratio	Appr. d.f.	p-value
For INTRCPT1, β_0					
INTRCPT2, γ_{00}	415.966574	4.120353	100.954	173	<0.001
EDUSHORT, γ_{01}	-8.591785	3.172309	-2.708	173	0.007
PROSTAT, γ_{02}	151.080271	62.803228	2.406	173	0.017
SCIERES, γ_{03}	6.111606	1.853189	3.298	173	0.001
For <i>ST071Q01</i> slope β_1					
INTRCPT2, γ_{10}	-1.295167	0.243520	-5.319	82	<0.001
For <i>ENVAWARE</i> slope β_2					
INTRCPT2, γ_{20}	3.934339	0.880595	4.468	90	<0.001
For <i>ENVOPT</i> slope β_3					
INTRCPT2, γ_{30}	-4.658037	0.992433	-4.694	46	<0.001
For <i>JOYSCIE</i> slope β_4					
INTRCPT2, γ_{40}	2.856719	1.188137	2.404	38	0.021
For <i>INTBRSCI</i> slope β_5					
INTRCPT2, γ_{50}	2.611832	1.131259	2.309	107	0.023
For <i>EPIST</i> slope β_6					
INTRCPT2, γ_{60}	3.994998	1.408710	2.836	18	0.011
For <i>SCIEACT</i> slope β_7					
INTRCPT2, γ_{70}	-4.750896	0.966131	-4.917	159	<0.001
For <i>BSMJ</i> slope β_8					
INTRCPT2, γ_{80}	0.163525	0.063590	2.572	251	0.011
For <i>SMINS</i> slope β_9					
INTRCPT2, γ_{90}	0.081538	0.012577	6.483	113	<0.001

Out-of-school study time for science ($\gamma_{10} = -1.295$, $p < 0.050$), environmental optimism ($\gamma_{30} = -4.658$, $p < 0.050$) and science activities ($\gamma_{70} = -4.751$, $p < 0.050$) had a negative effect on students' interpreting data and evidence scientifically scores; however, environmental awareness ($\gamma_{20} = 3.934$, $p < 0.050$), enjoyment of science ($\gamma_{40} = 2.857$, $p < 0.050$), interest in broad science topics ($\gamma_{50} = 2.612$, $p < 0.050$), epistemological beliefs ($\gamma_{60} = 3.995$, $p < 0.050$), students' expected occupational status ($\gamma_{80} = 0.164$,

$p < 0.050$) and science learning time (minutes per week) ($\gamma_{90} = 0.082$, $p < 0.050$) were positively related to students' interpreting data and evidence scientifically scores.

Table 4.60 represents estimations of variance components for interpreting data and evidence scientifically for the full model.

Table 4.60 Final Estimation of Variance Components for Interpreting Data and Evidence Scientifically

Random Effect	Standard Deviation	Variance Component	d.f.	χ^2	p-value
Schools mean, u_0	49.86856	2486.87315	157	3895.87067	<0.001
ENVAWARE slope, u_2	3.39930	11.55526	160	190.65475	0.049
ENVOPT slope, u_3	4.35058	18.92757	160	211.57353	0.004
SMINS slope, u_9	0.06350	0.00403	160	207.95982	0.007
Student effect, r	53.32664	2843.73084			

The proportion reduction in variance based on school-related factors variables is:

$$\tau_{qq}(\text{Random-Coefficients Regression}) - \tau_{qq}(\text{Intercepts- and -Slopes}) / \tau_{qq}(\text{Random-Coefficients Regression})$$

$$= 3198.27024 - 2486.87315 / 3198.27024$$

$$= 0.2224$$

This finding suggests that, school-level variables in Table 4.59 explained 22% of the variance in interpreting data and evidence scientifically scores of Turkish students.

The results indicate that, the effect of environmental awareness, environmental optimism and science learning time (minutes per week) on interpreting data and evidence scientifically scores of Turkish students increases, when the proportion of science teachers by all teachers and science specific resources increases. On the other hand, the effect of those variables on interpreting data and evidence scientifically scores of Turkish students decreases when shortage of educational material increases.

4.2.4.7 Findings from Intercepts- and Slopes-as-Outcomes Model

1. Shortage of educational material (EDUSHORT), proportion of science teachers by all teachers (PROSTAT) and science specific resources (SCIRES) effected the means of schools' all scientific literacy subscale scores.
2. Creative extra-curricular activities (CREACTIV) effected the means of schools' combined science, procedural and epistemic knowledge and evaluating and designing scientific inquiry scores.
3. Shortage of educational material (EDUSHORT) was negatively related to scientific literacy score of students for all scientific literacy subscales.
4. Proportion of science teachers by all teachers (PROSTAT), and science specific resources (SCIRES) were positively related to scientific literacy scores of students for all scientific literacy subscales. The positive effect of proportion of science teachers by all teachers (PROSTAT) on the scientific literacy scores of students was found to be the highest.
5. Creative extra-curricular activities (CREACTIV) was positively related to scientific literacy score of students for combined science, procedural and epistemic knowledge and evaluating and designing scientific inquiry subscales.
6. Out-of-school study time for science (ST071Q01), environmental awareness (ENVAWARE), environmental optimism (ENVOPT), epistemological beliefs (EPIST), science activities (SCIEACT), students' expected occupational status (BSMJ) and science learning time (minutes per week) (SMINS) effected scores of students for all scientific literacy subscales.
7. Enjoyment of science (JOYSCIE) effected scores of students for all scientific literacy subscales except for evaluating and designing scientific inquiry subscale, and the relationship between them was positive.
8. Interest in broad science topics (INTBRSCI) effected scores of students for all scientific literacy subscales except for procedural and epistemic knowledge and evaluating and designing scientific inquiry subscales, and the relationship between them was positive.

9. Science self-efficacy (SCIEEFF) effected scores of students for only combined science and explaining phenomena scientifically subscale, and the relationship between them was positive.

10. Out-of-school study time for science (ST071Q01), environmental optimism (ENVOPT) and science activities (SCIEACT) were negatively related with the scores of students for all scientific literacy subscales.



CHAPTER 5

DISCUSSION, CONCLUSION, IMPLICATION, AND RECOMMENDATION FOR FURTHER RESEARCH

The aim of this study was to explore the effect of student- and school-level factors on 15-year-old students' scientific literacy skills. The student- and school-level factors effecting scientific literacy skills of the students in Turkey were discussed in this chapter. In addition, the consequences of the study were explained and interpreted. Finally the implications of the study and recommendation for further research were explained.

5.1 Discussion of the Results

The present study investigated student and school characteristics that effect students' scientific literacy in Turkey. Hierarchical linear models were used to analyze the Turkish PISA 2015 student and school data. While investigating Turkish students' scientific literacy skills, not only combined science scores were analyzed, but also the science competencies and science knowledge scores and their components were analyzed separately.

The submodels of hierarchical linear modeling were run from the simplest to the more complex in the study. At the beginning of the analyzing of each data, the one-way ANOVA with random effects submodel was used. Afterwards, a regression model with means-as-outcomes model, a random-coefficient regression model and a model with intercepts-and slopes-as-outcomes were executed for combined science score, science competencies scores and science knowledge scores one by one.

The null model or the one-way ANOVA with random effects was used to partition the variances in the outcome variables into within- and between-school components.

The results of null models for all Turkish students and schools were demonstrated in

the tables in the previous chapter. In these tables, the fixed coefficients of combined science, science competencies and science knowledge were found to be significant. The significance of the coefficients were interpreted as the average scores of the students at different schools were different from each other.

The range of average outcomes of the schools, student-level variance, school-level variance and intraclass correlation were also demonstrated in these tables. The intraclass correlation coefficient of combined science was higher than the intraclass correlation coefficients of science competencies and the intraclass correlation coefficients of science knowledge. It can be summarized that the proportion of variances explained by the grouping structure in the combined science was higher than the proportion of variance explained by the grouping structure in the science knowledge and science competencies. In other words, the impact of school characteristics on students' combined science scores was larger than the impact of school characteristics on students' science knowledge and science competencies scores.

5.1.1 School-Level Factors

The means-as-outcomes model was executed to predict variation in the combined science, science competencies and science knowledge by using school-level predictors. The regression model with means-as-outcomes predicts variation in the scientific literacy skills of Turkish high school students by using school-level predictors.

The combined science score, the science competencies scores and the science knowledge scores of Turkish students were predicted by shortage of educational material (EDUSHORT), proportion of science teachers by all teachers (PROSTAT) and science specific resources (SCIRES).

The mean of outcome variables and fixed effects of the school-level predictors in the means-as-outcomes model of combined science, science competencies and science knowledge are displayed in Table 5.1.

Table 5.1 Fixed Effects of Means-as-Outcomes Model for School-Level Variables

		Intercept	EDUSHORT	PROSTAT	TOTST
Combined Science		418.82**	-9.52*	291.04**	-1.92*
Knowledge	Content knowledge	418.47**	-9.81*	289.64**	-1.70*
	Procedural and epistemic knowledge	418.02**	-9.92*	285.81**	-1.85*
	Explaining phenomena scientifically	418.72**	-10.27*	312.58**	-1.86*
Competencies	Evaluating and designing scientific inquiry	420.99**	-9.78*	302.98**	-2.31*
	Interpreting data and evidence scientifically	416.17**	-9.38*	185.61*	-----

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

Table 5.1. Fixed Effects of Means-as-Outcomes Model for School-Level Variables (Continued)

		Intercept	CREACTIV	SCIERES	PROSTCE
Combined science		418.82**	8.22*	4.96*	-----
Knowledge	Content knowledge	418.47**	-----	6.11*	-----
	Procedural and epistemic knowledge	418.02**	8.93*	4.61*	-----
	Explaining phenomena scientifically	418.72**	-----	6.01*	-----
Competencies	Evaluating and designing scientific inquiry	420.99**	8.98*	5.02*	-----
	Interpreting data and evidence scientifically	416.17**	-----	6.61**	-----

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

Table 5.1. Fixed Effects of Means-as-Outcomes Model for School-Level Variables
(Continued)

		Intercept	PROSTMAS	STAFFSHORT
Combined Science		418.82**	-----	-----
Knowledge	Content knowledge	418.47**	-----	-----
	Procedural and epistemic knowledge	418.02**	-----	-----
	Explaining phenomena scientifically	418.72**	-----	-----
Competencies	Evaluating and designing scientific inquiry	420.99**	-----	-----
	Interpreting data and evidence scientifically	416.17**	-----	-----

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

The significant school-level predictors in the means-as-outcomes models of combined science, science competencies and science knowledge are explained below.

1. The school-level variable EDUSHORT (shortage of educational material), (M= 0.27, SD= 1.29), predicts all scientific literacy subscales. Combined science score, science competencies and science knowledge are negatively related to EDUSHORT, meaning that, scientific literacy scores of students decreases with the increase in the shortage of educational materials in schools. Among all school-level variables, EDUSHORT has the highest negative correlation with all science literacy skills.
2. The school-level variable PROSTAT (proportion of science teachers by all teachers), (M= 0.13, SD= 0.06), predicts all scientific literacy subscales. Combined science score, science competencies and science knowledge are positively related to PROSTAT, meaning that, scientific literacy scores of students increases with increase in the proportion of science teachers by all teachers in schools. Among all school-level variables, PROSTAT has the

highest positive correlation with all science literacy skills.

3. The school-level variable TOTST (total number of science teachers at school), ($M= 6.87$, $SD= 5.89$), predicts all scientific literacy subscales except for interpreting data and evidence scientifically. Combined science score, science competencies and science knowledge are negatively related to TOTST, meaning that, scientific literacy scores of students decreases with the increase in the total number of science teachers at school in schools.
4. The school-level variable CREATIV (creative extra-curricular activities), ($M= 1.41$, $SD= 1.01$), predicts combined science, procedural and epistemic knowledge and evaluating and designing scientific inquiry subcales of scientific literacy. These subscales are positively related to CREATIV, meaning that scientific literacy of students within these subscales increases with the creative extra-curricular activities in schools.
5. The school-level variable SCIERES (science specific resources), ($M= 2.42$, $SD= 2.26$), predicts all scientific literacy subscales. Combined science score, science competencies and science knowledge are positively related to SCIERES, meaning that, scientific literacy scores of students increases with increase in the science specific resources in schools.
6. The school-level variables PROSTCE (proportion of science teachers fully certified), PROSTMAS (proportion of science teachers with ISCED level 5A and a major in science) and STAFFSHORT (shortage of educational staff) had no significant effect on Turkish students' scientific literacy performance in the hierarchical linear modeling.

To interpret the models more clearly, all school-level variables used in the models were categorised as:

- i) Teaching and learning variables
- ii) School policies variables

The fixed effects of teaching and learning variables and school policies variables are represented in Table 5.2 and Table 5.3 respectively.

When the school-level variables are adjusted, the average combined science score for some students with average PROSTAT, average TOTST, average EDUSHORT, average SCIERES and CREATIV is 418.82.

The scores of science knowledge and science competencies subscales were found to be close to combined science score. Evaluating and designing scientific inquiry score (420.99) was the highest among all scientific literacy subscales whereas interpreting data and evidence scientifically score (416.17) was the lowest.

In the science knowledge dimension, content knowledge score (418.47) was higher than procedural and epistemic knowledge score (418.02) of Turkish students. In the science competencies dimension, evaluating and designing scientific inquiry score (420.99) was the highest, explaining phenomena scientifically score (418.72) and interpreting data and evidence scientifically score (416.17) followed it respectively.

Table 5.2 Fixed Effects of Means-as-Outcomes Model for Teaching and Learning Variables

		Teaching and Learning Variables		
		Intercept	PROSTAT	TOTST
Combined Science		418.82**	291.04**	-1,92*
Knowledge	Content knowledge	418.47**	289.64**	-1.70*
	Procedural and epistemic knowledge	418.02**	285.81**	-1.85*
	Explaining phenomena scientifically	418.72**	312.58**	-1.86*
Competencies	Evaluating and designing scientific inquiry	420.99**	302.98**	-2.31*
	Interpreting data and evidence scientifically	416.17**	185.61*	-----

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

1. As a teaching and learning variable PROSTAT had the highest correlation

with explaining phenomena scientifically and the lowest correlation with interpreting data and evidence scientifically, and the relationship between them was positive.

2. TOTST had the highest correlation with evaluating and designing scientific inquiry and the lowest correlation with content knowledge, and the relationship between them was negative. TOTST had no significant effect on interpreting data and evidence scientifically subscale.
3. The effect of PROSTAT on the performance of students was higher than the effect of TOTST.
4. Other teaching and learning variables PROSTCE and PROSTMAS had no significant effect on students' performance.

Table 5.3 Fixed Effects of Means-as-Outcomes Model for School Policies Variables

		School Policies Variables			
		Intercept	EDUSHORT	SCIERES	CREACTIV
Combined Science		418.82**	-9.52*	4.96*	8.22*
Knowledge	Content knowledge	418.47**	-9.81*	6.11*	-----
	Procedural and epistemic knowledge	418.02**	-9.92*	4.61*	8.93*
	Explaining phenomena scientifically	418.72**	-10.27*	6.01*	-----
Competencies	Evaluating and designing scientific inquiry	420.99**	-9.78*	5.02*	8.98*
	Interpreting data and evidence scientifically	416.17**	-9.38*	6.61**	-----

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

1. As a school policies variable EDUSHORT had the highest correlation with explaining phenomena scientifically and the lowest correlation with interpreting data and evidence scientifically.

2. SCIERES had the highest correlation with interpreting data and evidence scientifically and the lowest correlation with procedural and epistemic knowledge.
3. CREATIV had the highest correlation with evaluating and designing scientific inquiry and the lowest correlation with combined science. CREATIV had no significant effect on content knowledge, explaining phenomena scientifically and interpreting data and evidence scientifically subscales.
4. The effect of EDUSHORT on students' performance was the highest among all school policies variables. The effect of CREATIV and SCIERES followed EDUSHORT respectively.
5. The school policies variable STAFFSHORT had no significant effect on students' performance.

5.1.2 Student-Level Factors

Table 5.4 shows the student-level variables of the random-coefficients regression models. Each intercept value provides the mean of the outcome variable for someone with average level-1 predictors, meaning the mean of combined science score, the mean of science competencies scores and the mean of science knowledge scores. The rows in the tables include the intercept and the fixed effects of the models.

When the student-level variables are adjusted, the average combined science score for some students with average OUTHOURS, average SCIEACT, average SCIEEFF, average INTBRSCI, average JOYSCIE, average EPIST, average ENVWARE, average ENVOPT, average BSMJ and average SMINS is 418.28.

The scores of science knowledge and science competencies subscales were found to be close to combined science score. Evaluating and designing scientific inquiry score (420.35) was the highest among all scientific literacy subscales, whereas interpreting data and evidence scientifically score (415.83) was the lowest.

Table 5.4 Fixed Effects of Random-Coefficients Regression Model for Student-Level Variables

		Mean of School Means	OUTHOURS (ST071Q01)	ENVAWARE	ENVOPT
Combined Science		418.28**	-1.17**	4.60**	-6.16**
Knowledge	Content knowledge	417.90**	-1.30**	5.11**	-7.08**
	Procedural and epistemic knowledge	417.44**	-1.24**	4.21**	-5.50**
	Explaining phenomena scientifically	418.12**	-1.22**	4.83**	-7.66**
Competencies	Evaluating and designing scientific inquiry	420.35**	-1.06**	5.95**	-5.61**
	Interpreting data and evidence scientifically	415.83**	-1.29**	3.94**	-4.66**

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

Table 5.4 Fixed Effects of Random-Coefficients Regression Model for Student-Level Variables (Continued)

		Mean of School Means	JOYSCIE	INTBRSCI	SCIEEFF	EPIST
Combined Science		418.28**	3.30**	2.02*	1.58*	4.25**
Knowledge	Content knowledge	417.90**	3.98**	3.06*	2.24*	2.80*
	Procedural and epistemic knowledge	417.44**	3.02*	-----	-----	4.76**
	Explaining phenomena scientifically	418.12**	3.88**	3.41*	2.58*	3.11*
Competencies	Evaluating and designing scientific inquiry	420.35**	-----	-----	-----	4.97**
	Interpreting data and evidence scientifically	415.83**	2.87*	2.47*	-----	3.99*

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

Table 5.4 Fixed Effects of Random-Coefficients Regression Model for Student-Level Variables (Continued)

		Mean of School Means	SCIEACT	BSMJ	SMINS	INSTSCIE
Combined Science		418.28**	-3.43**	0.18*	0.07**	-----
Knowledge	Content knowledge	417.90**	-2.82*	0.15*	0.08**	-----
	Procedural and epistemic knowledge	417.44**	-3.97**	0.19*	0.08**	-----
	Explaining phenomena scientifically	418.12**	-2.79*	0.17*	0.07**	-----
Competencies	Evaluating and designing scientific inquiry	420.35**	3.31*	0.26*	0.08**	-----
	Interpreting data and evidence scientifically	415.83**	-4.65**	0.17*	0.08**	-----

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

In the science knowledge dimension, content knowledge score of Turkish students was higher than procedural and epistemic knowledge score. As a part of science competencies dimension, evaluating and designing scientific inquiry score (420.35) of Turkish students was the highest, explaining phenomena scientifically score (418.12) and interpreting data and evidence scientifically score (415.83) followed it respectively.

Among all student-level variables ENVOPT had the highest negative effect on all scientific literacy subscales. OUTHOURS had the least negative effect on all scientific literacy subscales. ENVWARE had the highest positive effect on the combined science, content knowledge, explaining phenomena scientifically and evaluating and designing scientific inquiry subscales. EPIST had the highest positive effect on procedural and epistemic knowledge and interpreting data and evidence scientifically subscales.

To interpret the models more clearly, all student-level variables used in the models were categorised as:

i) Student background variables

ii) Non-cognitive outcomes

iii) Teaching and learning variables

The fixed effects of these student-level variables are represented in Table 5.5, Table 5.6 and Table 5.7 respectively.

Table 5.5 Fixed Effects of Random-Coefficients Regression Model for Student Background Variables

		Student Background Variables		
		Mean of School Means	OUTHOURS (ST071Q01)	SCIEACT
Combined Science		418.28**	-1.17**	-3.43**
Knowledge	Content knowledge	417.90**	-1.30**	-2.82*
	Procedural and epistemic knowledge	417.44**	-1.24**	-3.97**
	Explaining phenomena scientifically	418.12**	-1.22**	-2.79*
	Evaluating and designing scientific inquiry	420.35**	-1.06**	-3.31*
Competencies	Interpreting data and evidence scientifically	415.83**	-1.29**	-4.65**

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

1. As a student background variable, OUTHOURS (out-of-school study time for science), (M= 5.34, SD= 4.60) predicts all scientific literacy subscales. Combined science, scientific knowledge and scientific competencies scores are negatively related to OUTHOURS, meaning that scientific literacy scores of students decrease with the increase in the out-of-school study time for science. OUTHOURS had the highest correlation with content knowledge and the lowest correlation with evaluating and designing scientific inquiry.

2. As a student background variable, SCIEACT (science activities), ($M = 0.69$ $SD = 1.16$) predicts all scientific literacy subscales. Combined science, scientific knowledge and scientific competencies scores are negatively related to SCIEACT, meaning that scientific literacy scores of students decrease with the increase in the science activities. SCIEACT had the highest correlation with interpreting data and evidence scientifically and the lowest correlation with explaining phenomena scientifically.
3. The effect of SCIEACT on the students' performance was higher than the effect of OUTHOURS.

Table 5.6 Fixed Effects of Random-Coefficients Regression Model for Non-Cognitive Outcomes

		Non-Cognitive Outcomes				
		Mean of School Means	SCIEEFF	INTBRSCI	JOYSCIE	INSTSCIE
Combined Science		418.28**	1.58*	2.02*	3.30**	-----
Knowledge	Content knowledge	417.90**	2.24*	3.06*	3.98**	-----
	Procedural and epistemic knowledge	417.44**	-----	-----	3.02*	-----
	Explaining phenomena scientifically	418.12**	2.58*	3.41*	3.88**	-----
Competencies	Evaluating and designing scientific inquiry	420.35**	-----	-----	-----	-----
	Interpreting data and evidence scientifically	415.83**	-----	2.47*	2.87*	-----

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

Table 5.6 Fixed Effects of Random-Coefficients Regression Model for Non-Cognitive Outcomes (Continued)

		Non-Cognitive Outcomes				
		Mean of School Means	EPIST	ENVAWARE	ENVOPT	BSMJ
Combined Science		418.28**	4.25**	4.60**	-6.16**	0.18*
Knowledge	Content knowledge	417.90**	2.80*	5.11**	-7.08**	0.15*
	Procedural and epistemic knowledge	417.44**	4.76**	4.21**	-5.50**	0.19*
	Explaining phenomena scientifically	418.12**	3.11*	4.83**	-7.66**	0.17*
Competencies	Evaluating and designing scientific inquiry	420.35**	4.97**	5.95**	-5.61**	0.26**
	Interpreting data and evidence scientifically	415.83**	3.99*	3.94**	-4.66**	0.17*

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

1. As a student non-cognitive outcome, SCIEEFF (science self-efficacy), (M= 0.34, SD= 1.32) predicts combined science, content knowledge and explaining phenomena scientifically subscales. These subscales are positively related to SCIEEFF, meaning that scientific literacy scores of students increase with the increase in science self-efficacy. SCIEEFF had the highest correlation with explaining phenomena scientifically and the lowest correlation with combined science.
2. As a student non-cognitive outcome, INTBRSCI (interest in broad science topics), (M= -0.66, SD= 1.02) predicts combined science, content knowledge, explaining phenomena scientifically and interpreting data and evidence scientifically subscales. These subscales are positively related to INTBRSCI, meaning that scientific literacy scores of students increase with the increase in interest in broad science topics. INTBRSCI had the highest correlation

with explaining phenomena scientifically and the lowest correlation with combined science.

3. As a student non-cognitive outcome, JOYSCIE (enjoyment of science), ($M = 0.12$, $SD = 1.17$) predicts all scientific literacy subscales except for evaluating and designing scientific inquiry. These subscales are positively related to JOYSCIE, meaning that scientific literacy scores of students increase with the increase in enjoyment of science. JOYSCIE had the highest correlation with content knowledge and the lowest correlation with interpreting data and evidence scientifically.
4. As a student non-cognitive outcome, EPIST (epistemological beliefs), ($M = -0.19$, $SD = 1.18$) predicts all scientific literacy subscales. Combined science, science knowledge and science competencies subscales positively related to EPIST, meaning that scientific literacy scores of students increase with the increase in epistemological beliefs. EPIST had the highest correlation with evaluating and designing scientific inquiry and the lowest correlation with content knowledge.
5. As a student non-cognitive outcome, ENVWARE (environmental awareness), ($M = 0.55$, $SD = 1.47$) predicts all scientific literacy subscales. Combined science, science knowledge and science competencies subscales positively related to ENVWARE, meaning that scientific literacy scores of students increase with the increase in environmental awareness. ENVWARE had the highest correlation with evaluating and designing scientific inquiry and the lowest correlation with interpreting data and evidence scientifically.
6. As a student non-cognitive outcome, ENVOPT (environmental optimism), ($M = -0.53$, $SD = 1.45$) predicts all scientific literacy subscales. Combined science, science knowledge and science competencies subscales negatively related to ENVOPT, meaning that scientific literacy scores of students decrease with the increase in environmental optimism. ENVOPT had the highest correlation with explaining phenomena scientifically and the lowest correlation with interpreting data and evidence scientifically.

7. As a student non-cognitive outcome, BSMJ (students' expected occupational status), ($M= 63.36$, $SD= 16.25$) predicts all scientific literacy subscales. Combined science, science knowledge and science competencies subscales positively related to BSMJ, meaning that scientific literacy scores of students increase with the increase in students' expected occupational status. (BSMJ) had the highest correlation with evaluating and designing scientific inquiry and the lowest correlation with content knowledge.
8. The effect of ENVOPT was the highest among all non-cognitive outcomes, whereas BSMJ had the least effect.
9. INSTSCIE (instrumental motivation), ($M= 0.38$, $SD= 0.92$) had no significant effect on students' performance.

Table 5.7 Fixed Effects of Random-Coefficients Regression Model for Teaching and Learning Variables

Teaching and Learning Variables		
		SMINS
		Mean of School Means
Combined Science		418.28**
Knowledge	Content knowledge	417.90**
	Procedural and epistemic knowledge	417.44**
	Explaining phenomena scientifically	418.12**
Competencies	Evaluating and designing scientific inquiry	420.35**
	Interpreting data and evidence scientifically	415.83**

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

1. As a teaching and learning variable, SMINS (science learning time), ($M=198.82$, $SD=108.78$) predicts all scientific literacy subscales. Combined science, science knowledge and science competencies subscales positively related to SMINS, meaning that scientific literacy scores of students increase with the increase in science learning time. SMINS had the same correlation with content knowledge, procedural and epistemic knowledge, evaluating and designing scientific inquiry and interpreting data evidence scientifically subscales, which is higher than its correlation with combined science and explaining phenomena scientifically subscales.

5.2 Conclusions

Since students are nested within schools, it can be concluded that there is a correlation between school and student characteristics. The student characteristics, which effect learning outcomes, are effected by school membership. The results of the present study showed that some student and school characteristics effected scientific literacy skills of Turkish students significantly. Students' scientific literacy can be predicted by the distribution of the resources among students and schools.

The intraclass correlation coefficient of combined science, science competencies and science knowledge subscales of schools in Turkey is about 50%, meaning that the half of the total variability in science literacy is caused by the differences between schools and the other half by the differences within the school. Hence, it can be concluded that there are significant differences between schools of Turkey with respect to scientific literacy skills of students in all scientific literacy subscales.

The student- and school-level variables of the study were categorised as student background variables, which investigate the distribution of educational opportunities and outcomes within and across countries and the effect of social and academic composition of the school on learning process and outcomes; non-cognitive outcomes, which examine attitudes, beliefs, motivations, aspirations and learning-related behaviours; teaching and learning variables, which investigate teachers' qualification, teacher support, teaching practices, classroom climate, classroom management, learning time and learning opportunities; and school policies variables, which examine professional capacity, professional development and climate of the

school. Following conclusions can be drawn about the effects of these variables on scientific literacy skills of the students.

Student background variables; out-of-school study time for science and science activities were found to be effecting scientific literacy negatively; indicating that students, who spend more time on learning science in addition to required school schedule and perform more often science-related activities such as watching TV programmes, buying books, visiting web sites and reading magazines about science, or attending science clubs, perform lower in scientific literacy scores. The findings about science activities are not consistent with the findings of previous research about out-of-school activities related to science, which state that experiences with the physical and natural world allow the integration of students' previous experiences into classroom activities (Schibeci, 1989; Eshach, 2007; Eagly & Chaiken, 1993). This unexpected result might be due to not identifying and not integrating the critical properties of informal learning to the school-centered science and technology learning properly. School teachers should use the powerful influence of out-of-school experiences on knowledge, motivation, beliefs and attitudes of the students toward science by integrating out-of-school science-related activities to the classroom activities. Out-of-school study time for science includes total hours for homework, additional instruction and private study. Studies have shown that students gain a deeper understanding of the content and, as a result, earn higher grades and achievement scores (Epstein & Van Voorhis, 2001, Sparks, 2011). However, a number of studies demonstrate that science achievement decreases as students' time spent learning at home increases (Acar & Öğretmen, 2012; Üstün et al., 2019). The result that students' science performance decreases when the time devoted to out-of-school study time increases, might be explained with the idea that homeworks may be overburdened for students, when the underlying goals of assigning work for students to complete outside of class time is not considered. Hence, teachers should pay attention to assign homeworks, which bridge the gap between previously covered material and concepts yet to be taught, or help the students develop study skills (Epstein & Van Voorhis, 2001; Xu, 2005).

Non-cognitive outcomes; science self-efficacy, interest in broad science topics, enjoyment of science, epistemological beliefs, environmental awareness and students' expected occupational status were found to be effecting scientific literacy

subscales positively. However, environmental optimism effected scientific literacy of the students negatively and instrumental motivation had no significant effect of scientific literacy of the students. The positive effect of science self-efficacy, interest in broad science topics, enjoyment of science, epistemological beliefs, environmental awareness and students' expected occupational status on scientific literacy is expected, since numerous studies have shown that attitudes towards science effected students' scientific literacy performances significantly (Akpınar et al., 2009; Papanastasiou & Papanastasiou, 2010; Papanastasiou & Zembylas, 2004; Anıl, 2009; Atar & Atar, 2012; Odom & Bell, 2015; Singh et al., 2002; Heckman, Stixrud, & Urzua, 2006; Çalışkan, 2008; Chua & Karpudewan, 2017). The finding that science self-efficacy has a positive effect on scientific literacy is consistent with the studies which demonstrate that students perform significantly better in science subjects if their science self-efficacy is higher (Alivernini et al., 2010; Çalışkan, 2008; Cavas, 2011; Fonseca, 2011; Lam & Lau, 2014). The finding that interest in broad science topics has positive effect on scientific literacy is also consistent with the findings of the previous studies, which suggest that students with high interest in science perform significantly better in science (Akpınar, 2009). The finding that enjoyment of science is associated with increased scientific achievements is consistent with the findings of the previous researches which suggest that enjoyment of science has positive effect on scientific literacy as well (Fonseca, 2011; Lam & Lau, 2006; Çelebi, 2010; Akpınar et al., 2009). This study suggests that epistemological beliefs have positive effect on scientific literacy. This is an expected result, since findings of the previous studies suggest that epistemological beliefs are related to the increasing one's understanding of knowledge and academic achievement (Hofer, 2001; Trautwein & Lüdtke, 2007; Cano, 2005; Lee, Johanson, & Tsai, 2008). The finding that environmental awareness has positive effect on scientific literacy, whereas the effect of environmental optimism is negative, is also expected, since studies have shown that students with higher levels of environmental awareness perform better on science and students, who do not have the proper knowledge and awareness of the seriousness of environmental problems, can be also more optimistic about the future of the environment (Sulemana et al., 2016; Öztürk, 2018; Çelebi, 2010). The results of this study show that instrumental motivation has no significant effect on scientific literacy, and students' expected occupational status has relatively small positive

effect on scientific literacy. Previous studies support these findings, generally a very weak correlation was found between academic achievement and career aspirations (Rojewski & Yang, 1997; Benbow et al., 1991).

Teaching and learning variables; science learning time and proportion of science teachers by all teachers have positive effect on scientific literacy, whereas total number of science teachers at school effects scientific literacy negatively and proportion of science teachers fully certified and proportion of science teachers with ISCED level 5A and a major in science have no significant effect. The finding about science learning time in school is consistent with the findings of previous studies, which have shown that learning time in school is closely related to student learning and achievement (Anagün, 2011; Üstün et al., 2019). The finding that proportion of science teachers by all teachers have positive effect on scientific literacy complies with the findings of previous studies, which suggest that the interaction between teachers and students is affected by the number of students per teacher in a school (Croll & Hastings, 1996; Johnson, 2011; Demir, 2009). However, it was also found that total number of science teachers at school effects scientific literacy negatively. This can be explained with the fact that, as the school enrollment size increases, it is expected that total number of teachers in a school increases as well. Enrollment size is generally associated with class size in a school and it is mainly believed that smaller classes provide a better teaching and learning (Graue, Rauscher, & Sherfinski, 2009). The findings that proportion of science teachers fully certified and proportion of science teachers with ISCED level 5A and a major in science have no significant effect are consistent with the findings of the previous studies, which found that the education level of teachers has no significant effect on the academic achievement of students (Gage et al., 2017; Badgett et al., 2013; Coenen et al., 2014). These findings suggest that, other than having a major in science or being fully certified, teachers' ability to transfer their knowledge to students plays a more important role in the improvement of students' science achievement.

School policies variables; science specific resources and creative extra-curricular activities have positive effect on scientific literacy skills, whereas shortage of educational material has negative effect and shortage of educational staff has no significant effect. The findings that science specific resources have positive effect and shortage of educational material has negative effect on scientific literacy are

consistent with the previous studies which suggest that there are positive relationships between academic achievement of students and physical, financial and material resources (Card & Krueger, 1996; Greenwald, Hedges, & Laine, 1996; Adeogun & Osifila, 2008; Acar & Öğretmen, 2012; Berberoğlu & Kalender, 2005; Atar & Atar, 2012; Spezia, 2010; Jones et al., 2018; Üstün et al., 2019; Çelebi, 2010). The results of this study also demonstrate that shortage of educational staff has no significant effect on scientific literacy. This finding complies with the previous studies, which suggest that the number and quality of teaching and assisting staff are not significantly related to students' academic performance (Adeogun & Osifila, 2008). The finding that creative extra-curricular activities have positive effect on scientific literacy is consistent with the findings of the previous researches, which suggest that the students' participation in extracurricular science activities has positive effect on students' science achievement (Zhang & Tang, 2017; Freeman, 2017; Ekeke, 2014).

In addition, the effects of significant school-level variables on the relationship between student-level variables and scientific literacy skills of students were investigated in this study. The results of the present study indicate that, the effect of environmental optimism and science learning time on all scientific literacy subscale scores of Turkish students increases, when the proportion of science teachers by all teachers and science specific resources increases. On the other hand, the effect of those variables on all scientific literacy subscale scores of Turkish students decreases when shortage of educational material increases. In addition, the effect of environmental awareness on combined science, procedural and epistemic knowledge and evaluating and designing scientific inquiry subscale scores of Turkish students increases, when the proportion of science teachers by all teachers, science specific resources and creative extra-curricular activities increases. The effect of environmental awareness on interpreting data and evidence scientifically subscale scores of Turkish students increases, when the proportion of science teachers by all teachers and science specific resources increases.

Among all scientific literacy subscales, student scores were the highest in the evaluating and designing scientific inquiry subscale, whereas the lowest scores were in the interpreting data and evidence scientifically subscale. These findings suggest that Turkish students are more qualified in describing and appraising scientific

investigations and proposing ways of addressing questions scientifically than in any other scientific literacy subscale. However, they were found to be less qualified in analysing and evaluating data, claims and arguments in a variety of representations and drawing appropriate scientific conclusions compared to other scientific literacy subscales.

Student scores in content knowledge subscale were higher than their scores in procedural and epistemic knowledge subscale. These findings suggest that Turkish students are more qualified in explaining the facts, concepts, ideas and theories about the natural world that science has established. However, they were found to be less qualified in understanding the way of the establishment of scientific knowledge and how confident it is. The students were less qualified in identifying questions that are related to scientific enquiry, deciding whether appropriate procedures are used to ensure decision making, and distinguishing scientific issues from value judgments or economic considerations.

5.3 Implications for Policy

The present study explains the effect of student and school characteristic on educational outcomes of 15-year-old Turkish students and tries to clarify the strengths and weakness of Turkish students' scientific literacy skills. The results of the study suggest the following implications for the policy.

Students with higher scientific literacy skills have access to science specific resources and are provided with more science learning time. In addition, there is a strong positive relationship between the proportion of science teachers by all teachers and scientific literacy skills of Turkish students. However, the shortage of educational material effects the learning outcomes negatively. These findings suggest that to be able to offer a more qualified education for all students, learning opportunities should be provided equally.

Students having more environmental awareness have higher scientific literacy scores, however environmental optimism is negatively related to scientific literacy of the students. According to this result, it can be concluded that a greater awareness of students did not lead to greater optimism. Students, who reported being

knowledgeable about environmental issues, were more likely to consider that these problems would get worse in the future. Schools should provide students with more information, which helps students to develop a deeper understanding of environmental issues, to have the skills to make informed and responsible decisions and to develop a more sophisticated perspective on environmental issues.

Students who have positive feelings about science and appreciate the value of science have higher scientific literacy skills. Non-cognitive outcomes such as science self-efficacy, interest in broad science topics and enjoyment of science are positively related with science performance. Also, epistemological beliefs effect scientific literacy of the students positively. These findings suggest that the learning environments in schools should be designed to encourage the students to actively engage in learning activities in order to promote their general values of science and scientific inquiry. Hence, students should be informed by teachers and schools that scientific literacy is a lifelong process and it requires a deeper understanding of science related contexts.

According to the results of the study, it was found out that out-of-school study time for science and science activities effect scientific literacy of students negatively, however creative extra-curricular activities effect scientific literacy of the students positively. These findings suggest that schools should encourage students to participate in the creative extra-curricular activities such as science clubs and science competitions and to spend time on studying science besides their required school schedule by integrating the critical properties of informal learning to the school-centered science and technology learning properly.

5.4 Recommendations for Further Research

Within the framework of this study, the effects of student- and school-level factors on education outcomes of the Turkish students were explored. However, the results of the present study do not demonstrate the causalities between student- and school-level factors and scientific literacy skills. In order to explain the causalities, experimental studies need to be conducted to examine the contribution of these factors to the scientific literacy. These studies can explain why some factors effect the scientific literacy skills whereas some factors have no effect.

This study is a cross-sectional study. In order to examine the changes in students' scientific literacy skills in relation to student- and school-level variables, longitudinal design is needed. In addition, to assess the effect of cultural differences within the country, more detailed analysis including the regions of the country need to be conducted. Since PISA study is repeated every three years, it can reflect the educational performance of countries through time. Future studies can be conducted by using PISA 2018 study to explore the effect of changing educational policies of Turkey.

PISA allows making country-based comparisons. The educational systems of countries having high performance in PISA rankings can be examined and their practices can be adapted to our educational policies.

Since Turkey opted to answer only student and school questionnaire, data based on the remaining questionnaires cannot be obtained. To examine the effect of other factors and to explore causalities, in-depth interviews with the students, parents or teachers can be used as the supportive data sources in addition to PISA data.

This study was limited to science domain and the effects of student- and school-level variables. Also, sample of the study included 15-year-old students. To explore the effect of these variables in other domains and other grade levels, this study can be replicated in different domains and different grade levels.



REFERENCES

- About OECD*, (n.d.) Retrieved March 4, 2019 from The Organisation for Economic Co-operation and Development Web site: <https://www.oecd.org/about/>.
- Acar, T., & Öğretmen, T. (2012). Analysis of 2006 PISA science performance via multilevel statistical methods. *Education and Science*, 37(163), 178-189.
- Adams, R. J., Wilson, M., & Wu, M. (1997). Multilevel item response models: An approach to errors in variables regression. *Journal of Educational and Behavioral Statistics*, 22(1), 47-76.
- Adeogun, A. A., & Osifila, G. I. (2008). Relationship between educational resources and students' academic performance in Lagos State Nigeria. *International Journal of Educational Management*, 5-6, 144-153.
- Aitkin, M., & Longford, N. (1986). Statistical modelling issues in school effectiveness studies. *Journal of the Royal Statistical Society*, 149(1), 1-43.
- Akpınar, E., Yıldız, E., Tatar, N., & Ergin, Ö. (2009). Students' attitudes toward science and technology: An investigation of gender, grade level, and academic achievement. *Procedia-Social and Behavioral Sciences*, 1(1), 2804-2808.
- Algina, J., & Swaminathan, H. (2011). Centering in two-level nested designs. In J. Hox, & K. Roberts (Eds.), *The handbook of advanced multilevel data analysis* (pp 285-312). New York: Routledge.
- Alivernini, F., Palmerio, L., Vinci, E., & Di Leo, I. (2010). *An analysis of factors affecting pupils' science achievement in Italy*. Paper presented at the fourth IEA (International Association for the Evaluation of Educational Achievement) International Research Conference, Department of Education, University of Gothenburg, July, 2010, Sweden.

- Anagün, Ş. S. (2011). The impact of teaching-learning process variables to the students' scientific literacy levels based on PISA 2006 results. *Education and Science*, 36(162), 84-102.
- Anıl, D. (2009). Factors effecting science achievement of science students in Programme for International Students' Achievement (PISA) in Turkey. *Education and Science*, 34(152), 87-100.
- Atar, H. Y., & Atar, B. (2012). Examining the effects of Turkish education reform on students TIMSS 2007 science achievements. *Educational Sciences: Theory & Practice*, 12(4), 2621-2636.
- Ataş, D., & Karadağ, Ö. (2017). An analysis of Turkey's PISA 2015 results using two-level hierarchical linear modelling. *Journal Of Language and Linguistic Studies*, 13(2), 720-727.
- Badgett, K., Decman, J., & Carman, C. (2013). National implications: The impact of teacher graduate degrees on student math assessments. *National Forum of Teacher Education Journal*, 23(3), 1-18.
- Bandura, A. (1997). *Self-efficacy: The exercise of control*. W.H. Freeman and Company, Macmillan Publishers, New York.
- Bauer, D. J., & Curran, P. J. (2005). Probing interactions in fixed and multilevel regression: Inferential and graphical techniques. *Multivariate Behavioral Research*, 40, 373-400.
- Bayraktar, Ş. (2010). Status of science education in Turkey according to the results of Trends in Mathematics and Science Study (TIMSS 2007): Factors effecting science achievement. *Selçuk University Ahmet Keleşoğlu Journal of Faculty of Education*, 30, 249-270.
- Benbow, C. P., Arjmand, O., & Walberg, H. J. (1991). Educational productivity predictors among mathematically talented students. *Journal of Educational Research*, 84, 215-223.

- Bennett, D. A. (2001). How can I deal with missing data in my study?. *Australian and New Zealand Journal of Public Health*, 25(5), 464-469.
- Berberoğlu, G., & Kalender, İ. (2005). Investigation of student achievement across years, school types and regions: The SSE and PISA analyses. *Educational Sciences and Practice*, 4(7), 21-35.
- Brookhart, S. M., & Nitko, A. J. (2011). Strategies for constructing assessments of higher order thinking skills. In G. Schraw & D.R. Robinson (Eds.), *Assessment of higher order thinking skills* (pp. 327-359). IAP, Charlotte, NC.
- Bryk, A. S., & Raudenbush, S. W. (1992). *Hierarchical linear models: Applications and data analysis methods*. Newbury Park, CA: Sage Publications.
- Bryk, A. S., Allensworth, E., Easton, J. Q., Sebring, P. B., & Luppescu, S. (2010). *Organizing school for improvement: Lessons from Chicago*. The University of Chicago, IL.
- Cano F. (2005). Epistemological beliefs and approaches to learning: Their change through secondary school and their influence on academic performance. *British Journal of Educational Psychology*, 75(2), 203. doi: 10.1348/000709904X22683.
- Card, D., & Krueger, A. (1996). School resources and student outcomes: An overview of the literature and new evidence from North and South Carolina. *Journal of Economic Perspectives*, 10, 31-40.
- Cavas, P. (2011). Factors affecting the motivation of Turkish primary students for science learning. *Science Education International*, 22(1), 31-42.
- Chua, K. H., & Karpudewan, M. (2017). The role of motivation and perceptions about science laboratory environment on lower secondary students' attitude towards science. *Asia-Pacific Forum on Science Learning and Teaching*, 18(2), 1-16.

- Coenen, J., Groot, W., van den Brink, H. M., & van Klaveren, C. (2014). *Teacher characteristics and their effects on student test scores: A best-evidence review*. Maastricht: TIER.
- Conley, A. M., Pintrich, R. P., Vekiri, I., & Harrison, D. (2004). Changes in epistemological beliefs in elementary science students: Epistemological development and its impact on cognition in academic domains. *Contemporary Educational Psychology*, 29(2), 186-204.
- Creemers, B. P. M., & Kyriakides, L. (2008). *The dynamics of educational effectiveness: A contribution to policy, practice and theory in contemporary schools*. Routledge, London.
- Croll, P., & Hastings, N. (Eds.). (1996). *Effective primary teaching –research based classroom strategies*. London: David Fulton.
- 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.
- Çelebi, Ö. (2010). *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)*. Unpublished Doctoral Dissertation, Middle East Technical University, Ankara, Turkey.
- Davis, S. L., & Buckendahl, C.W. (2011). Incorporating cognitive demand in credentialing examinations. In G. Schraw & D. R. Robinson (Eds.), *Assessment of Higher Order Thinking Skills* (pp. 327-359). IAP, Charlotte, NC.
- Demir, C. E. (2009). Factors influencing the academic achievement of the Turkish urban poor. *International Journal of Educational Development*, 29, 17–29.

- Durant, J. R., (1993). What is scientific literacy. In J. R. Durant & J. Gregory (Eds.), *Science and culture in Europe* (pp. 129-137). London: Science Museum.
- Duschl, R. (2007). Science education in three-part harmony: Balancing conceptual, epistemic and social learning goals. *Review of Research in Education*, 32, 268-291. <http://dx.doi.org/10.3102/0091732X07309371>.
- Eagly, A. H. & Chaiken, S. (1993). *The psychology of attitudes*. Forth Worth, TX: Harcourt Brace College Publishers.
- Egalite, A. J., Kisida, B., & Winters, M. A. (2015). Representation in the classroom: The effect of own-race teachers on student achievement. *Economics of Education Review*, 45, 44–52.
- Ekeke, J. T. H. (2014). Students' involvement in extra-curricular activities and their academic performance at the tertiary level. *Atbu Journal of Science, Technology and Education*, 1(2), 73-80.
- Epstein, J. L. & Van Voorhis, F. L. (2001). More than minutes: Teachers' roles in designing homework. *Educational Psychologist*, 36(3), 181–193.
- Eshach, H. (2006). Bridging in-school and out-of-school learning: Formal, non-formal and informal education. *Journal of Science Education and Technology*, 16(2), 171-190.
- Fonseca, J., Valente, M. O., & Conboy, J. (2011). Student characteristics and PISA science performance: Portugal in cross-national comparison. *Procedia-Social and Behavioral Sciences*, 12, 322–329.
- Freeman, R. (2017). *The relationship between extracurricular activities and academic achievement*. Unpublished Doctoral Dissertation, National Louis University, Illinois, Chicago.

- Gage, N. A., Adamson, R., MacSuga-Gage, A. S., & Lewis, T. J. (2017). The relation between the academic achievement of students with emotional and behavioral disorders and teacher characteristics. *Behavioral Disorders*, 43(1), 213-222.
- Ghuman, S., & Lloyd, C. (2010). Teacher absence as a factor in gender inequalities in access to primary schooling in rural Pakistan, *Comparative Education Review*, 54(4), 539-554.
- Gottfried, A. E., Marcoulides, G. A., Gottfried, A. W., & Oliver, P. H. (2009). A latent curve model of parental motivational practices and developmental decline in math and science academic intrinsic motivation. *Journal of Educational Psychology*, 101(3), 729-739.
- Graue, E., Rauscher, E., & Sherfinski, M. (2009). The synergy of class size reduction and classroom quality. *The Elementary School Teacher*, 110(2), 178-201.
- Greenwald, R., Hedges, L. V., & Laine, R. (1996). The effect of school resources on student achievement. *Review of Educational Research*, 66, 361-396.
- Güneş, F. (1997). *Literacy teaching and brain technology*. Ankara: Ocak Publications.
- Hadzigeorgiou, Y., & Skoumios, M. (2013). The development of environmental awareness through school science: Problems and possibilities. *International Journal of Environmental & Science Education*, 8(3), 405-426. doi:10.12973/ijese.2013.212a.
- Hansla, A., Gamble, A., Juliusson, A., & Gärling, T. (2008). The relationships between awareness of consequences, environmental concern, and value orientations. *Journal of Environmental Psychology*, 28(1), 1-9. doi:10.1016/j.jenvp.2007.08.004

Heckman, J. J., Stixrud, J., & Urzua, S. (2006). The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, 24(3), 411-482.

Hofer, B. K. (2001). Personal epistemology research: Implications for learning and teaching. *Educational Psychology Review*, 13, 353-383.

Hox J., Maas C. (2004). Multilevel structural equation models: The limited information approach and the multivariate multilevel approach. In van K. Montfort, J. Oud, & A. Satorra (Eds.), *Recent developments on structural equation models. Mathematical modelling: Theory and applications*, vol 19. Springer, Dordrecht.

Hox, J. (2010). *Multilevel analysis: Techniques and application* (2nd ed.). New York, NY: Routledge.

Johnson, L. (2011). Does class size really matter? *District Administration*, 47(9), 104-105.

Jones, D., Sugalan, A., Mundy, M. A., & Fedynich, L. (2018). Examining the relationship of textbooks and labs on student achievement in eighth grade science. *Administrative Issues Journal Education Practice and Research*, 8(2), 16-30.

Karasar, N. (2007). *Scientific research method*. Ankara: Nobel.

Kirk, R. (2018). *Literacy and Education*. Essex: Ed-Tech Press.

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*. McGraw-Hill Book Company, London.

Köseoğlu, F., Atasoy, B., Kavak, N., Akkuş, H., Budak, E., Tümay, H., Kadayıfçı, H., & Taşdelen, U. (2003). *For constructivist learning environment: How should a science textbook be*. Asil Publishing, Ankara.

- Krapp, A., & Prenzel, M. (2011). Research on interest in science: Theories, methods, and findings. *International Journal of Science Education*, 33(1), 27-50.
- Lam, T. Y. P., & Lau, K. C. (2014). Examining factors affecting science achievement of Hong Kong in PISA 2006 using hierarchical linear modeling. *International Journal of Science Education*, 36(15), 2463-2480.
- Lee, M. H., Johanson, R. E., & Tsai, C. C. (2008). Exploring Taiwanese high school students' conceptions of and approaches to learning science through a structural equation modeling analysis. *Science Education*, 92(2), 191–220.
- Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data* (2nd ed.). Hoboken, NJ: Wiley.
- Loucks-Horsley, S., & Matsumoto, C. (1999). Research on professional development for teachers of mathematics and science: The state of the scene. *School Science and Mathematics*, 99, 258–271.
- MacKay, R. (2009). Remarks on the inefficient use of time in the teaching and learning of mathematics in five secondary schools. In J.H. Meyer & A. van Biljon (Eds.), *Proceedings of the 15th Annual Congress of the Association for Mathematics Education of South Africa (AMESA): Mathematical knowledge for teaching* (pp. 79-85).
- Marchand, J., & Weber, J. (2014). *The labor market and school finance effects of the Texas shale boom on teacher quality and student achievement* (Working Paper No. 2015-15).
- Martin, M. O., Mullis, I. V. S., & Foy, P. (2008). *TIMSS 2007 International science report: Findings from IEA's Trends in International Mathematics and Science Study at the eighth and fourth grades*. Boston College, Chestnut Hill, MA.
- Marzano, R. J. & Pickering, D. J. (2007). The case for and against homework. *Educational Leadership*, 64(6), 74–79.

- Ministry of National Education (MoNE) (2016). *PISA 2015 national report*. Ankara.
- Monseur, C., & Adams, R. J. (2009). Plausible values: How to deal with their limitations. *Journal of applied measurement*, 10(3), 320-34.
- Odom, A. L., & Bell, C. V. (2015). Associations of middle school student science achievement and attitudes about science with student-reported frequency of teacher lecture demonstrations and student-centered learning. *International Journal of Environmental and Science Education*, 10(1), 87–97.
- Oral, I., & McGivney, E. (2013). *The determinants of students' performance and achievement in mathematics and science in Turkey. Analysis of TIMSS 2011*. Education Reform Initiative, Ankara.
- Organisation for Economic Co-operation and Development (2007). *PISA 2006: Science competencies for tomorrow's world: Volume 1: Analysis*. Paris: PISA, OECD Publishing.
- Organisation for Economic Co-operation and Development (2011). *quality time for students: Learning in and out of school*. Paris: PISA, OECD Publishing. <http://dx.doi.org/10.1787/9789264087057-en>.
- Organisation for Economic Co-operation and Development (2017a). *PISA 2015 assessment and analytical framework: Science, reading, mathematics, financial literacy and collaborative problem solving* (Revised Edition). Paris: PISA, OECD Publishing.
- Organisation for Economic Co-operation and Development (2017b). *PISA 2015 technical report*. Paris: PISA, OECD Publishing.
- Organisation for Economic Co-operation and Development (2017c). *PISA 2015 technical background*. Paris: PISA, OECD Publishing.

- Öztürk, Ö. (2018). *Using PISA 2015 data to analyze how the scientific literacy of students from different socioeconomic levels can be predicted by environmental awareness and by environmental optimism*. Unpublished Master's Thesis, Bilkent University, Ankara, Turkey.
- Papanastasiou, E. C., & Zembylas, M. (2004). Differential effects of science attitudes and science achievement in Australia, Cyprus, and the USA. *International Journal of Science Education*, 26(3), 259–280.
- Papanastasiou, C., & Papanastasiou, E. C. (2010). Major influences on attitudes toward science. *International Journal of Phytoremediation*, 21(1), 239–257.
- Rakoczy, K., Klieme, E., & Pauli C. (2008). The impact of the perceived support of basic psychological needs and of the perceived relevance of contents on students' self-determined motivation in mathematics instruction. *Journal of Educational Psychology*, 22(1), 25-35.
- Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (2nd ed.). Thousand Oaks, CA: Sage Publications, Inc.
- Raudenbush, S. W., Bryk, A. S., Cheong, Y. F., Congdon, R., & du Toit M. (2019). *HLM: Hierarchical linear and nonlinear modeling (Version 8) [Computer software]*. Chicago, IL: Scientific Software International.
- Rojewski, J. W., & Yang, B. (1997). Longitudinal analysis of select influences on adolescents occupational aspirations. *Journal of Vocational Behavior*, 51, 375-410.
- Rubin, D. (1987). *Multiple imputation for nonresponse in surveys*. New York: Wiley.
- Schafer, J. (1997). *Analysis of incomplete multivariate data*. London: Chapman & Hall.

- Schibeci, R. (1989). Home, school, and peer group influences on student attitudes and achievement in science. *Science Education*, 73, 13-24.
- Schmidt, W. H., McKnight, C. C., Houang R. T., Wang, H. C., Wiley, D., Cogan, L. S., & Wolfe, R. G. (2001). *Why schools matter: A cross-national comparison of curriculum and learning*. Jossey-Bass, San Francisco, CA.
- Shamos, M. H. (1995). *The myth of scientific literacy*. New Brunswick, NJ: Rutgers University Press.
- Shen, B. S. P. (1975). Scientific literacy and the public understanding of science. In S. B. Day (Eds.), *Communication of Scientific Information* (pp. 44-52). Basel: Karger.
- Shin, Y., & Raudenbush, S. W. (2013). Efficient analysis of Q-level nested hierarchical general linear models given ignorable missing data. *The International Journal of Biostatistics*, 9(1), 109-133.
- 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.
- Sparks, S. (2011). Schools ‘flip’ for lesson model promoted by Khan Academy. *Education Week*, 31(5), 1-14.
- Spiezia, V. (2010). Does computer use increase educational achievements? Student-level evidence from PISA. *OECD Journal: Economic Studies*, 2010(1), 7-7.
- Sulemana, I., James, H. S., & Valdivia, C. B. (2016). Perceived socioeconomic status as a predictor of environmental concern in African and developed countries. *Journal of Environmental Psychology*, 46, 83-95. doi:10.1016/j.jenvp.2016.04.002.

- Summers, M., Kruger, C., & Childs, A. (2001). Understanding the science of environmental issues: Development of a subject knowledge guide for primary teacher education. *International Journal of Science Education*, 23(1), 33-53. doi:10.1080/09500690116990.
- Sun, L., Bradley, K., & Akers, K. (2012). A multi-level model approach to investigating factors impacting science achievement for secondary school students – PISA Hong Kong sample. *International Journal of Science Education*, 34(14), 2107-2125.
- Tabachnick, B. G., & Fidell, L.S. (2013). *Using multivariate statistics* (6th ed.). Pearson, Boston.
- Trautwein, U., & Lüdtke, O. (2007). Epistemological beliefs, school achievement, and college major: A large scale longitudinal study on the impact of certainty beliefs. *Contemporary Educational Psychology*, 32, 348-366.
- United Nations Educational, Scientific and Cultural Organization (UNESCO) Education Sector (2004). *The plurality of literacy and its implications for policies and programmes: Position paper*. Paris: United Nations Educational, Scientific and Cultural Organization.
- United Nations Environment Programme (UNEP) (2012). *21 issues for the 21st century: Result of the UNEP foresight process on emerging environmental issues*. Nairobi, Kenya.
- United Nations Educational Scientific and Cultural Organization (UNESCO) Institute for Statistics (2018). *Quick guide to education indicators for SDG 4*. Montreal, Canada.
- Üstün, U., Özdemir, E., Cansız, M., & Cansız, N. (2019). What are the factors affecting Turkish students' science literacy? A hierarchical linear modelling study using PISA 2015 data. *Hacettepe University Journal of Education*. Advance online publication. doi: 10.16986/HUJE.2019050786.

- Warm, T. A. (1989). Weighted likelihood estimation of ability in item response theory. *Psychometrika*, 54, 427-450.
- Webb, N. L. (1997). *Criteria for alignment of expectations and assessments in mathematics and science education*. Council of Chief State School Officers and National Institute for Science Education Research Monograph, National Institution for Science Education, Washington, DC.
- Wigfield, A., Eccles, J. S., & Rodriguez D. (1998). The development of children's motivation in school contexts. *Review of Research in Education*, 23, 73-118.
- Wu, M. (2005). The role of plausible values in large-scale surveys. *Studies in Educational Evaluation*, 31(2), 114-128.
- Xu, J. (2005). Purposes for doing homework reported by middle and high school students. *Journal of Educational Research*, 99(1), 46-55.
- Yetisir, M. I. (2014). The multilevel effects of student and classroom factors on the science achievement of eighth graders in Turkey. *Education and Science*, 39(172), 108-120.
- Yoon, K. S., Duncan, T., Lee, S. W. Y., Scarloss, B., & Shapley, K. (2007). *Reviewing the evidence on how teacher professional development affects student achievement*. Washington, DC: US Department of Education.
- Zhang, D., & Tang, X. (2017). The influence of extracurricular activities on middle school students' science learning in China. *International Journal of Science Education*, 39(10), 1381-1402.



APPENDICES

A. Questionnaire Items Related to School-Level Variables of the Study

1. Proportion of Science Teachers by All Teachers (PROSTAT)

SC018: How many of the following teachers are on the staff of your school?

Include both full-time and part-time teachers. A full-time teacher is employed at least 90% of the time as a teacher for the full school year. All other teachers should be considered part-time. Regarding the qualification level, please refer only to the teacher's highest qualification level.

(Please enter a number in each space provided. Enter "0" [zero] if there are none.)

		Full-time	Part-time
SC018Q01TA	Teachers in TOTAL		

SC019: How many of the following teachers are on the <school science> staff of your school?

Include both full-time and part-time teachers. A full-time teacher is employed at least 90% of the time as a teacher for the full school year. All other teachers should be considered part-time.

(Please enter a number in each space provided. Enter "0" [zero] if there are none.)

		Full-time	Part-time
SC019Q01NA	<School science> teachers in TOTAL		

2. Proportion of Science Teachers Fully Certified (PROSTCE)

SC019: How many of the following teachers are on the <school science> staff of your school?

Include both full-time and part-time teachers. A full-time teacher is employed at least 90% of the time as a teacher for the full school year. All other teachers should be considered part-time.

(Please enter a number in each space provided. Enter “0” [zero] if there are none.)

	Full-time	Part-time
SC019Q02NA <School science> teachers <fully certified> by <the appropriate authority>		

3. Proportion of Science Teachers with ISCED Level 5A and a Major in Science (PROSTMAS)

SC019: How many of the following teachers are on the <school science> staff of your school?

Include both full-time and part-time teachers. A full-time teacher is employed at least 90% of the time as a teacher for the full school year. All other teachers should be considered part-time.

(Please enter a number in each space provided. Enter “0” [zero] if there are none.)

	Full-time	Part-time
SC019Q03NA <School science> teachers with an <ISCED Level 5A or higher> qualification <with a major> in <school science>		

4. Shortage of Educational Material (EDUSHORT)

SC017: Is your school's capacity to provide instruction hindered by any of the following issues?

(Please select one response in each row.)

		Not at all	Very little	To some extent	A lot
SC017Q05NA	A lack of educational material (e.g. textbooks, IT equipment, library or laboratory material).	☐ ₁	☐ ₂	☐ ₃	☐ ₄
SC017Q06NA	Inadequate or poor quality educational material (e.g. textbooks, IT equipment, library or laboratory material).	☐ ₁	☐ ₂	☐ ₃	☐ ₄
SC017Q07NA	A lack of physical infrastructure (e.g. building, grounds, heating/cooling, lighting and acoustic systems).	☐ ₁	☐ ₂	☐ ₃	☐ ₄
SC017Q08NA	Inadequate or poor quality physical infrastructure (e.g. building, grounds, heating/cooling, lighting and acoustic systems).	☐ ₁	☐ ₂	☐ ₃	☐ ₄

5. Shortage of Educational Staff (STAFFSHORT)

SC017: Is your school's capacity to provide instruction hindered by any of the following issues?

(Please select one response in each row.)

		Not at all	Very little	To some extent	A lot
SC017Q01NA	A lack of teaching staff.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
SC017Q02NA	Inadequate or poorly qualified teaching staff.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
SC017Q03NA	A lack of assisting staff.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
SC017Q04NA	Inadequate or poorly qualified assisting staff.	☐ ₁	☐ ₂	☐ ₃	☐ ₄

6. Total Number of Science Teachers at School (TOTST)

SC019: How many of the following teachers are on the <school science> staff of your school?

Include both full-time and part-time teachers. A full-time teacher is employed at least 90% of the time as a teacher for the full school year. All other teachers should be considered part-time.

(Please enter a number in each space provided. Enter "0" [zero] if there are none.)

	Full-time	Part-time
SC019Q01NA <School science> teachers in		
TOTAL		

7. Science Specific Resources (SCIERES)

SC059: Which of the following are true for the science department of your school?

(Please select one response in each row.)

		Yes	No
SC059Q01NA	Compared to other departments, our school's <school science department> is well equipped.	☐ ₁	☐ ₂
SC059Q02NA	If we ever have some extra funding, a big share goes into improvement of our <school science> teaching.	☐ ₁	☐ ₂
SC059Q03NA	<School science> teachers are among our best educated staff members.	☐ ₁	☐ ₂
SC059Q04NA	Compared to similar schools, we have a well-equipped laboratory.	☐ ₁	☐ ₂
SC059Q05NA	The material for hands-on activities in <school science> is in good shape.	☐ ₁	☐ ₂
SC059Q06NA	We have enough laboratory material that all courses can regularly use it.	☐ ₁	☐ ₂
SC059Q07NA	We have extra laboratory staff that helps support <school science> teaching.	☐ ₁	☐ ₂
SC059Q08NA	Our school spends extra money on up-to-date <school science> equipment.	☐ ₁	☐ ₂

8. Creative Extra-Curricular Activities (CREACTIV)

SC053: <This academic year>, which of the following activities does your school offer to students in the <national modal grade for 15-year-olds>?

(Please select one response in each row.)

		Yes	No
SC053Q01TA	Band, orchestra or choir	☐ ₁	☐ ₂
SC053Q02TA	School play or school musical	☐ ₁	☐ ₂
SC053Q03TA	School yearbook, newspaper or magazine	☐ ₁	☐ ₂
SC053Q04TA	Volunteering or service activities, e.g. <national examples>	☐ ₁	☐ ₂
SC053Q05TA	Science club	☐ ₁	☐ ₂
SC053Q06TA	Science competitions, e.g. <national examples>	☐ ₁	☐ ₂
SC053Q07TA	Chess club	☐ ₁	☐ ₂
SC053Q08TA	Club with a focus on computers/ information and communication technology	☐ ₁	☐ ₂
SC053Q09TA	Art club or art activities	☐ ₁	☐ ₂
SC053Q10TA	Sporting team or sporting activities	☐ ₁	☐ ₂
SC053Q11TA	<country specific item>	☐ ₁	☐ ₂

B. Questionnaire Items Related to Student-Level Variables of the Study

1. Out-of-School Study Time per Week (OUTHOURS)

ST071: This school year, approximately how many hours per week do you spend learning in addition to your required school schedule in the following subjects?

(Please include the total hours for homework, additional instruction and private study.)

(Please move the slider to the number of total hours. Select “0” [zero] if you do not do homework, study or practice for a subject.)

ST071Q01NA <School science>



Slider bar: parking position, range 0–“30 hours per week or more”, step = 1; consistency check/soft reminder for values > 20.

2. Science Learning Time (SMINS)

ST059: How many <class periods> per week are you typically required to attend for the following subjects?

(Please enter a number in each row. Enter “0” [zero] if you have none.)

ST059Q03TA	Number of <class periods> per week in <science>
------------	--

Open text entry full numbers only. Consistency check, if entries are greater than 15.

ST061: How many minutes, on average, are there in a <class period>?

(Please move the slider to the number of minutes per <class period>.)

ST061Q01NA Average minutes in a <class period>



Slider range 0–“120 or more”; consistency check/soft reminder for values smaller than 10 and greater than 80; step 5.

3. Science Activities (SCIEACT)

ST146: How often do you do these things?

(Please select one response in each row.)

		Very often	Regularly	Sometimes	Never or hardly ever
ST146Q01TA	Watch TV programmes about <broad science>	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST146Q02TA	Borrow or buy books on <broad science> topics	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST146Q03TA	Visit web sites about <broad science> topics	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST146Q04TA	Read <broad science> magazines or science articles in newspapers	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST146Q05TA	Attend a <science club>	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST146Q06TA	Simulate natural phenomena in computer programs/ virtual labs	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST146Q07TA	Simulate technical processes in computer programs/ virtual labs	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST146Q08TA	Visit web sites of ecology organisations	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST146Q09TA	Follow news of science, environmental, or ecology organizations via blogs and microblogging	☐ ₁	☐ ₂	☐ ₃	☐ ₄

4. Science Self-Efficacy (SCIEEFF)

ST129: How easy do you think it would be for you to perform the following tasks on your own?

(Please select one response in each row.)

		I could do this easily	I could do this with a bit of effort	I would struggle to do this on my own	I couldn't do this
ST129Q01TA	Recognise the science question that underlies a newspaper report on a health issue.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST129Q02TA	Explain why earthquakes occur more frequently in some areas than in others.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST129Q03TA	Describe the role of antibiotics in the treatment of disease.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST129Q04TA	Identify the science question associated with the disposal of garbage.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST129Q05TA	Predict how changes to an environment will affect the survival of certain species.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST129Q06TA	Interpret the scientific information provided on the labelling of food items.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST129Q07TA	Discuss how new evidence can lead you to change your understanding about the possibility of life on Mars.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST129Q08TA	Identify the better of two explanations for the formation of acid rain.	☐ ₁	☐ ₂	☐ ₃	☐ ₄

5. Interest in Broad Science Topics (INTBRSCI)

ST095: To what extent are you interested in the following <broad science> topics?

(Please select one response in each row.)

		Not interested	Hardly interested	Interested	Highly interested	I don't know what this is
ST095Q04NA	Biosphere (e.g. ecosystem services, sustainability)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
ST095Q07NA	Motion and forces (e.g. velocity, friction, magnetic and gravitational forces)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
ST095Q08NA	Energy and its transformation (e.g. conservation, chemical reactions)	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
ST095Q13NA	The Universe and its history	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5
ST095Q15NA	How science can help us prevent disease	☐ 1	☐ 2	☐ 3	☐ 4	☐ 5

6. Enjoyment of Science (JOYSCIE)

ST094: How much do you disagree or agree with the statements about yourself below?

(Please select one response in each row.)

		Strongly disagree	Disagree	Agree	Strongly agree
ST094Q01NA	I generally have fun when I am learning <broad science> topics.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST094Q02NA	I like reading about <broad science>.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST094Q03NA	I am happy working on <broad science> topics.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST094Q04NA	I enjoy acquiring new knowledge in <broad science>.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST094Q05NA	I am interested in learning about <broad science>.	☐ ₁	☐ ₂	☐ ₃	☐ ₄

7. Instrumental Motivation (INSTSCIE)

ST113: How much do you agree with the statements below?

(Please select one response in each row.)

		Strongly agree	Agree	Disagree	Strongly disagree
ST113Q01TA	Making an effort in my <school science> subject(s) is worth it because this will help me in the work I want to do later on.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST113Q02TA	What I learn in my <school science> subject(s) is important for me because I need this for what I want to do later on.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST113Q03TA	Studying my <school science> subject(s) is worthwhile for me because what I learn will improve my career prospects.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST113Q04TA	Many things I learn in my <school science> subject(s) will help me to get a job.	☐ ₁	☐ ₂	☐ ₃	☐ ₄

8. Epistemological Beliefs (EPIST)

ST131: How much do you disagree or agree with the statements below?

(Please select one response in each row.)

		Strongly disagree	Disagree	Agree	Strongly agree
ST131Q01NA	A good way to know if something is true is to do an experiment.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST131Q03NA	Ideas in <broad science> sometimes change.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST131Q04NA	Good answers are based on evidence from many different experiments.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST131Q06NA	It is good to try experiments more than once to make sure of your findings.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST131Q08NA	Sometimes <broad science> scientists change their minds about what is true in science.	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST131Q11NA	The ideas in <broad science> science books sometimes change.	☐ ₁	☐ ₂	☐ ₃	☐ ₄

9. Environmental Awareness (ENVAWARE)

ST092: How informed are you about the following environmental issues?

(Please select one response in each row.)

		I have never heard of this	I have heard about this but I would not be able to explain what it is really about	I know something about this and could explain the general issue	I am familiar with this and I would be able to explain this well
ST092Q01TA	The increase of greenhouse gases in the atmosphere	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST092Q02TA	The use of genetically modified organisms (<GMO>)	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST092Q04TA	Nuclear waste	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST092Q05TA	The consequences of clearing forests for other land use	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST092Q06TA	Air pollution	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST092Q08TA	Extinction of plants and animals	☐ ₁	☐ ₂	☐ ₃	☐ ₄
ST092Q09TA	Water shortage	☐ ₁	☐ ₂	☐ ₃	☐ ₄

10. Environmental Optimism (ENVOPT)

ST093: Do you think problems associated with the environmental issues below will improve or get worse over the next 20 years?

(Please select one response in each row.)

		Improve	Stay about the same	Get worse
ST093Q01TA	Air pollution	☐ ₁	☐ ₂	☐ ₃
ST093Q03TA	Extinction of plants and animals	☐ ₁	☐ ₂	☐ ₃
ST093Q04TA	Clearing of forests for other land use	☐ ₁	☐ ₂	☐ ₃
ST093Q05TA	Water shortages	☐ ₁	☐ ₂	☐ ₃
ST093Q06TA	Nuclear waste	☐ ₁	☐ ₂	☐ ₃
ST093Q07TA	The increase of greenhouse gases in the atmosphere	☐ ₁	☐ ₂	☐ ₃
ST093Q08TA	The use of genetically modified organisms (<GMO>)	☐ ₁	☐ ₂	☐ ₃

11. Students' Expected Occupational Status (SEI) (BSMJ)

ST114: What kind of job do you expect to have when you are about 30 years old?

ST114Q01TA Please type in the job title. _____



C. Technical Characteristic of Student-Level Variables

1. Student Background Variables

Derived variables related to out-of-school science experience module were explained as student background variables as follows.

1.1. Out-of-School Science Experience (Module 5)

Under this module, out-of-school study time for science (OUTHOURS) and science activities (SCIEACT) were explained as student background variables.

Out-of-school study time for science (OUTHOURS)

In ST071Q01NA, students were asked how much time they spent studying science in addition to their required school schedule.

Science activities (SCIEACT)

Students were asked how often they do science-related activities (ST146) (OECD, 2017c).

2. Non-Cognitive Outcomes

Derived variables related to science-related outcomes and career aspirations modules were explained as non-cognitive outcomes as follows.

2.1. Science-Related Outcomes (Module 4)

Under this module, science self-efficacy (SCIEEFF), interest in broad science topics (INTBRSCI), enjoyment of science (JOYSCIE), instrumental motivation (INSTSCIE), epistemological beliefs (EPIST), environmental awareness (ENVAWARE) and environmental optimism (ENVOPT) were explained as non-cognitive outcomes.

Science self-efficacy (SCIEEFF)

In order to rate the performance of students on some science tasks, students were asked by using a four-point answering scale with categories “I could do this easily”, “I could do this with a bit of effort”, “I would struggle to do this on my own”, and “I couldn’t do this”. The tasks are about recognising newspaper reports on health issues, explaining different frequencies of earthquakes on different areas, explaining how antibiotics treat diseases, describing the effect of environmental changes on the survival of some species, recognising issues about garbage disposal, food items labeling, formation of acid rains and discussing the evidences about the possibility of life on Mars. The responses were reverse-coded, so higher index values and higher difficulty levels indicated higher levels of science self-efficacy. The derived variable SCIEEFF was scaled by using IRT scaling model, which enabled a trend comparison between PISA 2006 and PISA 2015 (OECD, 2017c).

Science self-efficacy (ST129) is a trend question that was taken from PISA 2006 (ID in 2006: ST17). Students were asked to rate how they would perform in different science tasks by using a four-point answering scale with the categories “I could do this easily”, “I could do this with a bit of effort”, “I would struggle to do this on my own”, and “I couldn’t do this”. As a result, the responses had to be reverse-coded so that higher WLEs and higher difficulty correspond to higher levels of science self-efficacy. The derived variable SCIEEFF was scaled by using the IRT scaling model, which allowed for a trend comparison between PISA 2006 and PISA 2015 (OECD, 2017b).

Interest in broad science topics (INTBRSCI) and enjoyment of science (JOYSCIE)

Interest in science was assessed with two scales; students’ enjoyment of science (JOYSCIE) and their interest in broad science topics (INTBRSCI). In order to explore the agreement of students on some science-related statements, students were asked by using a four-point Likert scale with categories “strongly agree”, “agree”, “disagree”, and “strongly disagree”. The derived variable JOYSCIE was

scaled by using the IRT scaling model, which allowed for a trend comparison between PISA 2006 and PISA 2015 at country level. Higher values on the index corresponded to higher agreement levels with these statements. These statements were about whether students have fun when learning, reading or working about science topics and whether they are interested in or they enjoy new science topics (OECD, 2017c).

Instrumental motivation (INSTSCIE)

In order to explore the agreement of students on some science related statements, students were asked by using a four-point Likert scale with categories “strongly agree”, “agree”, “disagree”, and “strongly disagree”. The statements were about the value of learning experience in school for future job and career plans. The derived variable INSTSCIE was scaled by using the IRT scaling model, which enabled a trend comparison between PISA 2006 and PISA 2015 at country level.

Epistemological beliefs (EPIST)

Students were asked by using a four-point Likert scale with the answering categories “strongly agree”, “agree”, “disagree”, and “strongly disagree” about their views on scientific approaches with a question measuring their epistemological beliefs. The derived variable EPIST was scaled by using the IRT scaling model.

Environmental awareness (ENVAWARE) and environmental optimism (ENVOPT)

Students’ awareness of environmental matters (ENVAWARE) and their perception of environmental issues as a concern (ENVOPT) were measured through two trend questions (ST092 and ST093) from PISA 2006. Students were asked about air pollution, extinction of plants and animals and water shortage for ST092 by using a four-point scale with answering categories “I have never heard of this”, “I have heard about this but I would not be able to explain what it is really about”, “I know something about this and could explain the general issue”, and “I am familiar with

this and I would be able to explain this well”. Students were also asked about the increase of greenhouse gases in the atmosphere and the use of genetically modified organisms for ST093 by using a three-point scale with the following categories: “improve”, “stay about the same”, and “get worse”. ST093-items were reverse-coded, so higher index values and higher difficulty levels indicated higher levels of environmental optimism. The derived variables ENVAWARE and ENVOPT were scaled by using the IRT scaling model, which enabled a trend comparison between PISA 2006 and PISA 2015.

2.2. Career Aspirations (Module 6)

Career aspirations module in the PISA 2015 student questionnaire includes the question about the job students expect to have at the age of 30. Students’ expected occupational status (SEI) (BSMJ) variable was examined under this module in the present study (OECD, 2017a).

Students’ expected occupational status (SEI) (BSMJ)

Students were asked to report their expected occupation at age 30 and to describe this job. The responses were coded to four-digit ISCO codes and then recorded into the ISEI (international socio-economic index of occupational status) index, where higher scores of ISEI corresponded to higher levels of expected occupational status (OECD, 2017b).

3. Teaching and Learning Variables

Derived variables related to learning time and curriculum module were explained as teaching and learning variables as follows.

3.1. Learning Time and Curriculum (Module 12)

Under this module, learning time for science (SMINS) was explained as a teaching and learning variable.

Learning time for science (SMINS)

SMINS was computed by using the information about the number of science class periods students attend for science subjects (ST059) and the average minutes in a science class period (ST061).





D. Technical Characteristics of School-Level Variables

1. Teaching and Learning Variables

Derived variables related to teacher qualification and professional knowledge module were explained as teaching and learning variables as follows.

1.1. Teacher Qualification and Professional Knowledge (Module 1)

Under this module, proportion of science teachers by all teachers (PROSTAT), proportion of science teachers fully certified (PROSTCE), proportion of science teachers with ISCED level 5A and a major in science (PROSTMAS) and total number of science teachers at school (TOTST) were explained as teaching and learning variables.

Proportion of science teachers by all teachers (PROSTAT)

PROSTAT was computed by dividing the number of science teachers by the total number of teachers.

Proportion of science teachers fully certified (PROSTCE)

PROSTCE was computed by dividing the number of fully certified science teachers by the total number of teachers.

Proportion of science teachers with ISCED level 5A and a major in science (PROSTMAS)

PROSTMAS was calculated by dividing the number of science teachers with ISCED level 5A and a major in science by the total number of science teachers.

Total number of science teachers at school (TOTST)

TOTST explains the number of all science teachers in a school (OECD, 2017c).

2. School Policies Variables

Derived variables related to school-level learning environment for science module were explained as school policies variables as follows.

2.1. School-Level Learning Environment for Science (Module 3)

Under this module, shortage of educational material (STAFFSHORT), shortage of educational staff (EDUSHORT), science specific resources (SCIERES) and creative extra-curricular activities (CREACTIV) were explained as school policies variables.

Shortage of educational material (STAFFSHORT)

The perceptions of school principals were asked by using a four-point answering scale with categories “not at all”, “very little”, “to some extent”, and “a lot”. STAFFSHORT was derived from items SC017Q01NA, SC017Q02NA, SC017Q03NA, and SC017Q04NA.

Shortage of educational staff (EDUSHORT)

EDUSHORT was scaled by using four items SC017Q05NA, SC017Q06NA, SC017Q07NA, and SC017Q08NA. The items were not reversed for scaling.

Science specific resources (SCIERES)

SCIERES was constructed by summing up the principals' answers to SC059 (yes/no question).

Creative extra-curricular activities (CREACTIV)

This index is about what school principals report about extra-curricular activities schools offered to 15-year old students. CREATIV was computed as the total number of the band, orchestra or choir activities, school play or school musical and art club or art activities (OECD, 2017b).

E. Descriptive Statistics of Outcome Variables

	N	Minimum	Maximum	Mean	Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic
Plausible Value 1 in Science	5895	197,716	707,892	422,45112	77,125510
Plausible Value 2 in Science	5895	186,432	728,629	423,13830	77,518403
Plausible Value 3 in Science	5895	196,499	686,670	421,40192	77,406189
Plausible Value 4 in Science	5895	198,717	709,282	422,28003	77,699549
Plausible Value 5 in Science	5895	171,653	689,673	421,92788	77,358449
Plausible Value 6 in Science	5895	182,170	667,479	421,71557	78,217807
Plausible Value 7 in Science	5895	175,349	681,964	421,31958	76,805191
Plausible Value 8 in Science	5895	183,052	691,170	421,02599	76,959855
Plausible Value 9 in Science	5895	187,266	698,490	422,13211	77,665402
Plausible Value 10 in Science	5895	192,052	686,807	421,68533	78,090112
Plausible Value 1 in Competency Subscale of Science - Explain Phenomena Scientifically	5895	137,141	706,695	420,57179	82,097098
Plausible Value 2 in Competency Subscale of Science - Explain Phenomena Scientifically	5895	160,362	677,208	421,46346	81,658185
Plausible Value 3 in Competency Subscale of Science - Explain Phenomena Scientifically	5895	140,432	708,168	422,38599	82,989239

	N	Minimum	Maximum	Mean	Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic
Plausible Value 4 in Competency Subscale of Science - Explain Phenomena Scientifically	5895	118,476	694,958	422,44291	81,058639
Plausible Value 5 in Competency Subscale of Science - Explain Phenomena Scientifically	5895	126,022	685,401	421,56457	82,563795
Plausible Value 6 in Competency Subscale of Science - Explain Phenomena Scientifically	5895	129,368	685,642	421,17063	82,599055
Plausible Value 7 in Competency Subscale of Science - Explain Phenomena Scientifically	5895	138,052	701,447	421,88843	81,265183
Plausible Value 8 in Competency Subscale of Science - Explain Phenomena Scientifically	5895	162,530	700,074	422,26853	81,754628
Plausible Value 9 in Competency Subscale of Science - Explain Phenomena Scientifically	5895	155,262	703,693	422,14639	82,050816
Plausible Value 10 in Competency Subscale of Science - Explain Phenomena Scientifically	5895	141,750	709,898	421,94518	82,236621
Plausible Value 1 in Competency Subscale of Science - Evaluate and Design Scientific Enquiry	5895	107,305	718,606	423,22500	82,733360
Plausible Value 2 in Competency Subscale of Science - Evaluate and Design Scientific Enquiry	5895	150,238	688,579	422,88834	82,257705
Plausible Value 3 in Competency Subscale of Science - Evaluate and Design Scientific Enquiry	5895	146,651	694,127	424,65850	82,012118

	N	Minimum	Maximum	Mean	Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic
Plausible Value 4 in Competency Subscale of Science - Evaluate and Design Scientific Enquiry	5895	174,080	708,506	424,71377	81,221911
Plausible Value 5 in Competency Subscale of Science - Evaluate and Design Scientific Enquiry	5895	158,802	729,386	422,95098	79,953648
Plausible Value 6 in Competency Subscale of Science - Evaluate and Design Scientific Enquiry	5895	176,396	714,108	424,26602	81,438077
Plausible Value 7 in Competency Subscale of Science - Evaluate and Design Scientific Enquiry	5895	157,689	713,186	421,84340	81,901435
Plausible Value 8 in Competency Subscale of Science - Evaluate and Design Scientific Enquiry	5895	155,648	735,437	422,37306	81,147271
Plausible Value 9 in Competency Subscale of Science - Evaluate and Design Scientific Enquiry	5895	152,456	698,266	425,49246	81,346480
Plausible Value 10 in Competency Subscale of Science - Evaluate and Design Scientific Enquiry	5895	167,952	724,460	425,60131	80,016183
Plausible Value 1 in Competency Subscale of Science - Interpret Data and Evidence Scientifically	5895	151,086	731,226	420,85065	81,966507
Plausible Value 2 in Competency Subscale of Science - Interpret Data and Evidence Scientifically	5895	149,635	702,102	418,43696	81,931014

	N	Minimum	Maximum	Mean	Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic
Plausible Value 3 in Competency Subscale of Science - Interpret Data and Evidence Scientifically	5895	142,699	700,769	418,86227	80,771585
Plausible Value 4 in Competency Subscale of Science - Interpret Data and Evidence Scientifically	5895	147,028	698,626	419,49703	80,319587
Plausible Value 5 in Competency Subscale of Science - Interpret Data and Evidence Scientifically	5895	133,694	686,788	418,22564	81,265679
Plausible Value 6 in Competency Subscale of Science - Interpret Data and Evidence Scientifically	5895	115,783	690,688	420,37812	82,114361
Plausible Value 7 in Competency Subscale of Science - Interpret Data and Evidence Scientifically	5895	168,760	687,238	420,78490	81,155660
Plausible Value 8 in Competency Subscale of Science - Interpret Data and Evidence Scientifically	5895	145,582	685,758	419,33648	80,946330
Plausible Value 9 in Competency Subscale of Science - Interpret Data and Evidence Scientifically	5895	164,826	707,478	420,80967	79,854878

	N	Minimum	Maximum	Mean	Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic
Plausible Value 10 in Competency Subscale of Science - Interpret Data and Evidence Scientifically	5895	145,638	715,133	420,10952	80,841417
Plausible Value 1 in Knowledge Subscale of Science - Content	5895	137,819	750,098	422,83030	82,247731
Plausible Value 2 in Knowledge Subscale of Science - Content	5895	145,219	717,604	421,67771	81,519867
Plausible Value 3 in Knowledge Subscale of Science - Content	5895	152,889	699,557	422,27725	80,906643
Plausible Value 4 in Knowledge Subscale of Science - Content	5895	118,331	729,324	421,48429	81,743613
Plausible Value 5 in Knowledge Subscale of Science - Content	5895	117,801	678,926	421,98597	81,862626
Plausible Value 6 in Knowledge Subscale of Science - Content	5895	143,313	659,416	420,16056	80,527832
Plausible Value 7 in Knowledge Subscale of Science - Content	5895	139,119	701,690	420,46055	81,253686
Plausible Value 8 in Knowledge Subscale of Science - Content	5895	162,954	719,750	420,26939	81,670730
Plausible Value 9 in Knowledge Subscale of Science - Content	5895	176,991	712,627	421,86722	80,482769
Plausible Value 10 in Knowledge Subscale of Science - Content	5895	120,871	742,296	420,87635	80,747319
Plausible Value 1 in Knowledge Subscale of Science - Procedural & Epistemic	5895	163,209	729,430	422,15683	78,847188

	N	Minimum	Maximum	Mean	Std. Deviation
	Statistic	Statistic	Statistic	Statistic	Statistic
Plausible Value 2 in Knowledge Subscale of Science - Procedural & Epistemic	5895	140,095	691,978	421,17990	79,329126
Plausible Value 3 in Knowledge Subscale of Science - Procedural & Epistemic	5895	206,643	672,358	420,74853	77,209452
Plausible Value 4 in Knowledge Subscale of Science - Procedural & Epistemic	5895	196,356	678,091	420,30467	79,023478
Plausible Value 5 in Knowledge Subscale of Science - Procedural & Epistemic	5895	165,648	680,087	420,58086	79,165588
Plausible Value 6 in Knowledge Subscale of Science - Procedural & Epistemic	5895	169,571	667,810	421,52928	79,321250
Plausible Value 7 in Knowledge Subscale of Science - Procedural & Epistemic	5895	182,782	710,151	421,46847	78,757970
Plausible Value 8 in Knowledge Subscale of Science - Procedural & Epistemic	5895	187,909	684,099	422,63162	78,693607
Plausible Value 9 in Knowledge Subscale of Science - Procedural & Epistemic	5895	147,136	718,988	420,64641	80,066684
Plausible Value 10 in Knowledge Subscale of Science - Procedural & Epistemic	5895	162,032	741,374	421,66782	79,478279

CURRICULUM VITAE

PERSONAL INFORMATION

Surname, Name: Orhan, Burcu

Nationality: Turkish (TC)

Date and Place of Birth: 2 July 1987, Ankara

Marital Status: Single

Phone: +90 537 879 75 85

email: burcu.orhan@metu.edu.tr

EDUCATION

Degree	Institution	Year of Graduation
PhD	Middle East Technical University, Mathematics and Science Education (3.9/4.00 GPA)	2020 Spring (Expected)
BS	Anadolu University, Economics (3,18/4,00 GPA)	2013
MS	Hacettepe University, Biology Education (without thesis)	2011
BS	Hacettepe University, Biology Education (3,56/4,00 GPA)	2011
BS	Pädagogische Hochschule Heidelberg (04.2010 – 09.2010)	
High School	Çankaya Atatürk Anadolu High School, Ankara (4,94/5,00 GPA)	2005

VISITING SCHOLAR

04.2010 – 09.2010: As a student of Erasmus Student Exchange Program

WORK EXPERIENCE

Year	Place	Enrollment
2013-Present	Measurement, Selection and Placement and Placement Center (ÖSYM)	Career Specialist

FOREIGN LANGUAGES

Advanced English, Advanced German

COMPUTER SKILLS:

Microsoft Office Tools, SPSS, PubMed/MEDLINE, EndNote

PUBLICATIONS

1. Sözen-Özdoğan, S., Özçakır, B., & Orhan, B. (2019). A case of teacher and student mathematical problem-solving behaviors from the perspective of cognitive-metacognitive framework. *Studia Paedagogica*, 24(4), 221-243.

ADDITIONAL INFORMATION

1. METU Graduate Awards, Course Performance Awards (2016), Middle East Technical University, Secondary Science and Mathematics Education, Highest GPA (3.93/4.00 GPA)
2. Graduation from Hacettepe University Secondary Science and Mathematics Education with first place (2010-2011 İhsan Doğramacı Outstanding Achievement Award)
3. Graduation from Hacettepe University Biology Education with first place

4. Graduation from Mimar Kemal Primary School with first place
5. TÜBİTAK National Scholarship Programme for PhD Students (2011-2016)

HOBBIES

Swimming, playing the piano, travelling.

