TRENDS IN INTERNATIONAL MATHEMATICS AND SCIENCE STUDY

TIMSS

TIMSS 2007 Technical Report

Edited by

John F. Olson

Michael O. Martin

Ina V.S. Mullis

TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

TRENDS IN INTERNATIONAL MATHEMATICS AND SCIENCE STUDY

TIMSS

TIMSS 2007 Technical Report

Edited by

John F. Olson

Michael O. Martin

Ina V.S. Mullis

Contributors

Alka Arora

Juliane Barth

Ebru Erberber

Pierre Foy

Joseph Galia

Ieva Johansone

Marc Joncas

Isaac Li

Barbara Malak

Michael O. Martin

Ina V.S. Mullis

Oliver Neuschmidt

John F. Olson

Christine O'Sullivan

Corinna Preuschoff

Graham Ruddock

TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

Published December 2008, Revised August 2009

TIMSS 2007 Technical Report

Edited by John F. Olson, Michael O. Martin, Ina V.S. Mullis, with contributions from
Alka Arora, Juliane Barth, Ebru Erberber, Pierre Foy, Joseph Galia, Ieva Johansone, Marc Joncas,
Isaac Li, Barbara Malak, Michael O. Martin, Ina V.S. Mullis, Oliver Neuschmidt, John F. Olson,
Christine O'Sullivan, Corinna Preuschoff, Graham Ruddock

Publisher: TIMSS & PIRLS International Study Center,
Lynch School of Education, Boston College

Library of Congress Catalog Card Number: 2008902432

ISBN: 1-889938-50-5

For more information about TIMSS contact:

TIMSS & PIRLS International Study Center

Lynch School of Education

Boston College

Chestnut Hill, MA 02467

United States

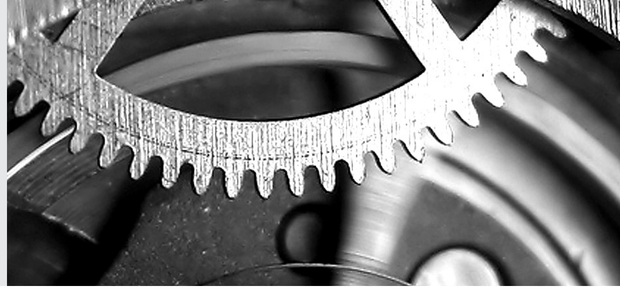
tel: +1-617-552-1600

fax: +1-617-552-1203

e-mail: timss@bc.edu

http://timssandpirls.bc.edu

Contents



Chapter 1 — 1
Overview of TIMSS 2007
Ina V.S. Mullis and Michael O. Martin

Chapter 2 — 13
Developing the TIMSS 2007 Mathematics and Science Assessments and Scoring Guides
Graham J. Ruddock, Christine Y. O'Sullivan, Alka Arora, and Ebru Erberber

Chapter 3 — 45
Developing the TIMSS 2007 Background Questionnaires
Ebru Erberber, Alka Arora, and Corinna Preuschoff

Chapter 4 — 63
Translation and National Adaptations of the TIMSS 2007 Assessment and Questionnaires
Ieva Johansone and Barbara Malak

Chapter 5 — 77
TIMSS 2007 Sample Design
Marc Joncas

Chapter 6 — 93
TIMSS 2007 Survey Operations Procedures
Ieva Johansone and Oliver Neuschmidt

Chapter 7 — 113
Quality Assurance in the TIMSS 2007 Data Collection
Ieva Johansone

Chapter 8 — 141
Creating and Checking the TIMSS 2007 Database
Juliane Barth and Oliver Neuschmidt

Chapter 9 — 153
TIMSS 2007 Sampling Weights and Participation Rates
Marc Joncas

Chapter 10 193

Reviewing the TIMSS 2007 Item Statistics

John F. Olson, Michael O. Martin, Ina V.S. Mullis, Pierre Foy, Ebru Erberber, and Corinna Preuschoff

Chapter 11 225

Scaling the Data from the TIMSS 2007 Mathematics and Science Assessments

Pierre Foy, Joseph Galia, and Isaac Li

Chapter 12 281

Creating the TIMSS 2007 Background Indices

Michael O. Martin and Corinna Preuschoff

Chapter 13 339

The TIMSS 2007 International Benchmarks of Student Achievement in Mathematics and Science

Ina V.S. Mullis, Ebru Erberber, and Corinna Preuschoff

Appendix A 349

Organizations and Individuals Responsible for TIMSS 2007

Appendix B 359

Characteristics of National Samples

Appendix C 439

Country Adaptations to Items and Item Scoring

Appendix D 451

Item Parameters for IRT Analyses of TIMSS 2007 Data

Appendix E 547

Summary Statistics and Standard Errors for Proficiency in Mathematics and Science Content Domains and Cognitive Domains in the Fourth and Eighth Grades

Appendix F 575

Item Descriptions Developed During the TIMSS 2007 Benchmarking

Chapter 1

Overview of TIMSS 2007

Ina V.S. Mullis and Michael O. Martin

1.1 Introduction

IEA is a recognized pioneer of international assessments, having conducted comparative studies of students' academic achievement for 50 years. However, even for IEA and its TIMSS & PIRLS International Study Center at Boston College, TIMSS 2007 was a very complex and ambitious undertaking, involving 67 participants (59 countries and 8 benchmarking entities) in a cooperative, global endeavor to develop and implement a wide-ranging assessment of mathematics and science achievement at fourth and eighth grades, providing a wealth of information about the educational context and current achievement of students in 2007, while measuring trends from earlier cycles of TIMSS in 1995, 1999, and 2003.

The design, development, and implementation of TIMSS 2007 are documented in a series of publications produced at various stages of the project. The *TIMSS 2007 Assessment Frameworks* (Mullis, Martin, Ruddock, O'Sullivan, Arora, & Erberber, 2005) contains the mathematics and science frameworks underlying the assessments at the fourth and eighth grades as well as the contextual framework for the questionnaires, and describes the design of the assessment.

Implementing TIMSS 2007 involved widespread participation from countries around the world, many of whom were collecting the third or fourth cycle of trend data. The *TIMSS 2007 International Mathematics Report* (Mullis, Martin, & Foy, 2008) and the *TIMSS 2007 International Science Report* (Martin, Mullis, & Foy, 2008) summarize fourth-and eighth-grade students' mathematics and science achievement in each of the 59 participating countries and 8 benchmarking participants. The complete TIMSS 2007 database is available on DVD accompanied by the *TIMSS 2007 User Guide for the International Database* (Foy & Olson, 2009).

The *TIMSS 2007 Encyclopedia* (Mullis, Martin, Olson, Berger, Milne, & Stanco, 2008) contains the countries' and benchmarking participants' descriptions of their national contexts for mathematics and science education as well as their mathematics and science curricula. The more qualitative information provided in the *TIMSS 2007 Encyclopedia* is intended to complement both the *TIMSS 2007 International Mathematics Report* and the *TIMSS 2007 International Science Report*.

The purpose of the *TIMSS 2007 Technical Report* is to provide further detail and documentation about the processes underlying the development of the TIMSS 2007 instruments and the methods used in sampling, data collection, scaling, and data analysis. In particular, the *TIMSS 2007 Technical Report* documents the numerous steps and procedures that comprise the rigorous quality assurance program conducted by all those involved, including the TIMSS & PIRLS International Study Center, the IEA Secretariat, the IEA Data Processing and Research Center, Statistics Canada, Educational Testing Service, and the National Research Coordinators and their teams in the participating countries and benchmarking entities.

1.2 Participants in TIMSS 2007

Exhibit 1.1 shows a map of the world identifying the TIMSS 2007 countries and benchmarking participants (regional entities that follow the same assessment procedures as the countries). Exhibit 1.2 lists the TIMSS 2007 participants, and indicates the grade(s) at which they participated and the previous cycles of TIMSS they participated in at that grade. It can be seen that many of the TIMSS 2007 countries and benchmarking participants have data for both the fourth and eighth grades. Exhibit 2 also shows that most TIMSS 2007 participants have trend data and, for each participant, whether it is for two, three, or four points in time—1995, 1999, 2003, and 2007.

TIMSS 2007 was administered near the end of the school year in each country. In countries in the Southern Hemisphere (where the school year typically ends in November or December) the assessment was conducted in October or November 2006. In the Northern Hemisphere, the school year typically ends in June; so in these countries the assessment was conducted in April, May, or June 2007.

1.3 TIMSS 2007 Instruments

The TIMSS 2007 assessment contained 353 items at the fourth grade, including 179 in mathematics and 174 in science. At the eighth grade there were 429 items, 215 in mathematics and 214 items in science. At both grades, the TIMSS 2007 assessment involved assembling the items into 14 blocks of mathematics items and 14 blocks of science items, and then assembling the blocks into 14 booklets, each one including 2 blocks of mathematics items and 2 blocks of science items assembled according to a very careful rotated design. Each student was administered a single booklet. Details about the development process and types of items can be found in Chapter 2.

Chapter 3 contains information about developing the four different types of background questionnaires. In brief, students answered questions pertaining to their home and school environments. The teachers of the sampled students responded to questions about characteristics of the class tested, instructional activities for teaching mathematics or science, the topics covered in students' lessons, and about their education, training, and opportunities for professional development. The principals of schools responded to questions about enrolment and school characteristics, school organization, staffing and resources, and the school environment. The *Curriculum Questionnaire*, a responsibility of the National Research Coordinators, provided data about participants' mathematics and science curricula. As an innovation for TIMSS 2007, the *Curriculum Questionnaire* was administered online.

To increase reliability in reporting background data, the questions in the background questionnaires form a number of scales. These scales and other sets of background questions are used to create background indices for reporting. The methods used to create the TIMSS 2007 background indices are discussed in Chapter 12.

Exhibit 1.1 Countries Participating in TIMSS 2007

Algeria
Armenia
Australia
Austria
Bahrain
Bosnia and Herzegovina
Botswana
Bulgaria
Chinese Taipei
Colombia
Cyprus
Czech Republic
Denmark
Egypt
El Salvador
England
Georgia
Germany
Ghana
Hong Kong SAR
Hungary
Indonesia
Iran, Islamic Rep. of
Israel
Italy
Japan
Jordan
Kazakhstan
Korea, Rep. of
Kuwait
Latvia
Lebanon
Lithuania
Malaysia
Malta
Mongolia
Morocco
Netherlands
New Zealand
Norway
Oman
Palestinian Nat'l Auth.
Qatar
Romania
Russian Federation
Saudi Arabia
Scotland
Serbia
Singapore
Slovak Republic
Slovenia
Sweden
Syrian Arab Republic
Thailand
Tunisia
Turkey
Ukraine
United States
Yemen

Benchmarking Participants

Alberta, Canada
Basque Country, Spain
British Columbia, Canada
Dubai, UAE
Massachusetts, US
Minnesota, US
Ontario, Canada
Quebec, Canada

Exhibit 1.1 Countries Participating in TIMSS 2007 (Continued)

Exhibit 1.2 Countries Participating in TIMSS 1995 Through 2007

Country	Grade 4			Grade 8			
	2007	2003	1995	2007	2003	1999	1995
Algeria	●			●			
Armenia	●	●		●	●		
Australia	●	●	●	●	●	●	●
Austria	●		●				●
Bahrain				●	●		
Bosnia and Herzegovina				●			
Botswana				●	●		
Bulgaria				●	●	●	●
Chinese Taipei	●	●		●	●	●	
Colombia	●			●			●
Cyprus		●	●	●	●	●	●
Czech Republic	●		●	●		●	●
Denmark	●						●
Egypt				●	●		
El Salvador	●			●			
England	●	●	●	●	●	●	●
Georgia	●			●			
Germany	●						●
Ghana				●	●		
Hong Kong SAR	●	●	●	●	●	●	●
Hungary	●	●	●	●	●	●	●
Indonesia				●	●	●	
Iran, Islamic Rep. of	●	●	●	●	●	●	●
Israel			●	●	●	●	●
Italy	●	●	●	●	●	●	●
Japan	●	●	●	●	●	●	●
Jordan				●	●	●	
Kazakhstan	●						
Korea, Rep. of			●	●	●	●	●
Kuwait	●		●	●			●
Latvia	●	●	●		●	●	●
Lebanon				●	●		
Lithuania	●	●		●	●	●	●
Malaysia				●	●	●	
Malta				●			
Mongolia	●			●			
Morocco	●	●		●	●	●	
Netherlands	●	●	●		●	●	●
New Zealand	●	●	●		●	●	●
Norway	●	●	●	●	●		●
Oman				●			
Palestinian Nat'l Auth.				●	●		
Qatar	●			●			
Romania				●	●	●	●
Russian Federation	●	●		●	●	●	●
Saudi Arabia				●	●		

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

Exhibit 1.2 Countries Participating in TIMSS 1995 Through 2007 (Continued)

Country	Grade 4			Grade 8			
	2007	2003	1995	2007	2003	1999	1995
Scotland	●	●	●	●	●		●
Serbia				●	●		
Singapore	●	●	●	●	●	●	●
Slovak Republic	●				●	●	●
Slovenia	●	●	●	●	●	●	●
Sweden	●			●	●		●
Syrian Arab Republic				●	●		
Thailand			●	●		●	●
Tunisia	●	●		●	●	●	
Turkey				●		●	
Ukraine	●			●			
United States	●	●	●	●	●	●	●
Yemen	●	●					
Benchmarking Participants							
Alberta, Canada	●		●			●	●
Basque Country, Spain				●	●		
British Columbia, Canada	●			●		●	
Dubai, UAE	●			●			
Massachusetts, US	●			●		●	
Minnesota, US	●		●	●			●
Ontario, Canada	●	●	●	●	●	●	●
Quebec, Canada	●	●	●	●	●	●	●

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

1.4 Translation Verification

Chapter 4 describes the steps involved in translating the test instruments and background questionnaires from English into numerous different languages. To ensure comparability among translated instruments, participants are given detailed specifications about the process to use in translating the materials, the IEA Secretariat manages a rigorous translation verification process, and the TIMSS & PIRLS International Study Center conducts a verification of final instrument layout before instruments are printed.

1.5 Sample Design, Implementation, and Participation

As explained in Chapter 5, the TIMSS 2007 assessment was administered to carefully drawn probability samples of students from the target populations in each country. The target populations were students enrolled in the fourth grade or eighth grade of formal schooling, counting from the first year of primary school defined by UNESCO's International Standard Classification for Education (UNESCO, 1999). Accordingly, the fourth year or the eighth year of formal schooling should be the fourth grade or eighth grade, respectively, in most countries. However, to avoid testing students who are very young, TIMSS has a policy that the average age of children in the grade tested should not be below 9.5 for the fourth year of schooling or 13.5 for the eighth year of schooling. The basic sampling design was a two-stage stratified cluster design. The first stage consisted of sampling schools, and the second stage consisted of sampling intact classrooms from the target grade in the sampled schools. Typically, at each grade, countries sampled 150 schools and one or two intact classrooms.

Information about the sampling weights and documentation of the participation rates is found in Chapter 9. Most countries achieved the minimum acceptable participation rates—85 percent of both the schools and students, or a combined rate (the product of schools' and students' participation) of 75 percent.

1.6 Survey Operations and Quality Assurance in Data Collection

Each country and benchmarking participant was responsible for carrying out all aspects of data collection and scoring, using standardized procedures explained in a series of survey operations procedure units and various training manuals. The data collection and scoring procedures are described in Chapter 6. In addition, the TIMSS & PIRLS International Study Center together with the IEA Secretariat conducted an independent quality control program. The reports from the Quality Control Monitors provided in Chapter 7 indicated that, in general, the national centers were able to conduct the data collection efficiently, professionally, and in compliance with international procedures.

1.7 The TIMSS 2007 International Database

To ensure comparable, high-quality data for analysis, the IEA Data Processing and Research Center took great care in creating the international database. Once the data were forwarded from the participants, the data underwent an exhaustive cleaning process. As described in Chapter 8, the data were checked and double-checked for consistency within and across countries. The national centers were contacted regularly and given multiple opportunities to review the data for their countries.

Chapter 10 provides details about the process implemented by the TIMSS & PIRLS International Study Center to review item statistics for each achievement item in each country, including scoring reliability data for the constructed-response items—within country, across countries, and for trends. In general, the items exhibited very good psychometric properties in all countries, and the scoring reliability was satisfactory (above 90% agreement in most cases).

1.8 Scaling the Achievement Data and the International Benchmarks

The TIMSS mathematics and science achievement scales were designed to provide reliable measures of student achievement across the trend cycles of the TIMSS assessments, based on the metric established with the 1995 data. As described in Chapter 11, student achievement was summarized using item response theory (IRT) scaling methods. For more accurate estimation of results for subpopulations of students, the TIMSS scaling made use of plausible-value technology. In addition to the overall scales used to estimate student achievement in each assessment including TIMSS 2007 and to measure trends over time, IRT scales were created for each of the content and cognitive domains described in the *TIMSS 2007 Assessment Frameworks*.

Chapter 13 describes the procedures used to report student achievement at the TIMSS International Benchmarks. To describe what performance on the TIMSS achievement scales means in terms of students' mathematics or science proficiency, TIMSS conducted a scale anchoring analysis to describe and interpret student achievement at the Advanced (625), High (550), Intermediate (475), and Low (400) International Benchmarks.

1.9 Ensuring Comparative Validity

In conclusion, a major purpose of the *TIMSS 2007 Technical Report* is to provide detailed documentation about the procedures and methods used by TIMSS to provide internationally comparative data of high quality. This report explains the multi-faceted attention to quality and the many quality assurance steps that were implemented from updating the assessment frameworks for TIMSS 2007 through release of the international database and User's Guide.

TIMSS is dedicated to addressing the classical concerns of high quality measurement—reliability and validity. TIMSS has procedures to ensure that the tests are reliable, that is, that they are constructed with sufficient items to provide reliable measurement, and are accompanied by detailed administration and scoring procedures and supported by extensive training to ensure that the results are not impacted by extraneous factors. Because reliability is not sufficient for good measurement, considerable effort also is dedicated to the validity of the tests, that is, the extent to which inferences drawn from the results can be supported by evidence. For example, does a student with high achievement on the TIMSS mathematics achievement scale actually have a high degree of proficiency in mathematics for an eighth grade student? Validity involves unified agreement in conceptualizing and articulating the constructs of mathematics and science as they apply to fourth- and eighth-grade students, and unified agreement that the items included in the assessments measure those articulations of mathematics and science, respectively.

In addition, as an international study, TIMSS must have comparative validity. For comparative validity, the classical concerns of reliability and validity still apply, but the concepts are extended to encompass the idea that the data should be internationally comparable. That is, that inferences made about achievement differences between countries can be substantiated.

The various chapters of this report describe the TIMSS quality assurance program to ensure comparative validity. Chapter 2 describes how the *TIMSS 2007 Frameworks* were updated through widespread collaboration among the participating countries, and modified to align with current perspectives. The updates were based on surveys of the participating countries and iterative reviews by the National Research Coordinators and experts. Chapter 2 also describes how the items and scoring guides were developed in accordance with the frameworks to assess specified topics, and according to a careful plan for measuring trends. The items were reviewed extensively by experts and the participating countries.

Developing the instruments and operational procedures for TIMSS 2007 involved a full-scale field test that was essential for confirming the appropriateness and comparability of the items. The translation verification process for the TIMSS 2007 assessment is described in Chapter 4. Every effort was made to ensure that the translations were comparable across countries. The data collection and scoring methods are described in Chapter 6, including the complete documentation of the survey operations procedures in manuals and specific training in aspects of data collection. The results of the TIMSS Quality Control Monitoring program conducted as part of test administration are included in Chapter 7.

Chapter 5 describes the efforts taken to ensure sample comparability across countries. Chapter 9 describes the implementation of the sampling procedures. With very few exceptions, countries assessed the correct grade(s), included all of the students in their definition of the target population, kept exclusions to a minimum (lower than 5%), and implemented accurate classroom sampling using the WinW3S software developed by IEA for this purpose. These chapters also describe how each country's sampling procedures must be fully documented, and that the participation rates must be high standards (at least 85% of students and 85% of schools) or be annotated in the international reports.

Chapter 8 addresses the issue of the comparability of the data. It explains how the IEA Data Processing and Research Center (DPC) provides data entry software and variable codebooks to standardize data preparation and conducts extensive training seminars. The IEA DPC checks within each country's data files and across countries for internal consistency and accuracy, and interacts with the country's to resolve data issues.

Subsequent to the field test, and then again, prior to scaling, a thorough review of item statistics was conducted. For every item for every country, the results are scrutinized for discrimination and scoring reliability. Also, the data are reviewed for item-by-country interactions. As described in Chapter 10, for each item TIMSS examined each country's performance on the item in light of its overall performance in 2007, and for trend countries this was done to compare interaction patterns for both assessments.

Chapter 11 describes the scaling methodology for the TIMSS 2007 achievement data, and how the fitted model for each item was checked against the observed data. For trend items, the fit was plotted separately to ensure that the item was a good fit to both sets of assessment data. The scaling was implemented separately for each country and separately for the 30 different scales (overall achievement, content domains, and cognitive domains for the fourth and eighth grades) and all of the results were plotted and checked. Again, with a few rare exceptions for the more difficult domains, the achievement score distributions were very satisfactory and provided an excellent basis for further analysis and reporting the results.

TIMSS devotes considerable effort to ensure that the data can be trusted for important decision making based on comparisons between countries and much of that information is documented in this technical report.

References

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

Martin, M.O., Mullis, I.V.S., & Foy, P. (with Olson, J.F., Erberber, E., Preuschoff, C., & Galia, J.). (2008). *TIMSS 2007 international science report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., & Foy, P. (with Olson, J.F., Preuschoff, C., Erberber, E., Arora, A., & Galia, J.). (2008). *TIMSS 2007 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

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

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

UNESCO Institute for Statistics. (1999). *Operational manual for ISCED-1997 (international standard classification of education)*. Paris: Author.

Chapter 2

Developing the TIMSS 2007 Mathematics and Science Assessments and Scoring Guides

Graham J. Ruddock, Christine Y. O'Sullivan, Alka Arora, and Ebru Erberber

2.1 Overview

The mathematics and science assessments for TIMSS 2007 were developed over a period of 2 years, from January 2005 to November 2006. The process incorporated the expertise of mathematics and science educators and test development specialists from all over the world. In particular, the TIMSS & PIRLS International Study Center worked with the Science and Mathematics Item Review Committee (SMIRC), an international committee of prominent mathematics and science experts who were nominated by participating countries and represented a range of nations and cultures.[1] Their responsibilities were threefold: to review and revise items, check for mathematical and scientific accuracy, and make certain that the items fit the specifications contained within the *TIMSS 2007 Assessment Frameworks* (Mullis, Martin, Ruddock, O'Sullivan, Arora, & Erberber, 2005).

First, the mathematics and science assessment frameworks for TIMSS 2003 were reviewed and updates were made for TIMSS 2007. Because approximately half of the mathematics and science assessment items were released to the public following the publication of the TIMSS 2003 results, a large number of replacement items were newly developed for TIMSS 2007. Item writing was accomplished in large measure by the National Research Coordinators (NRCs), with support and training from the TIMSS & PIRLS

1 See Appendix A for a list of members of the Science and Mathematics Item Review Committee.

International Study Center. Two task forces, one for mathematics and one for science,[2] were convened to assist in managing the resulting pool of items.

This chapter describes the test development process in more detail. An overview is shown in Exhibit 2.1.

Exhibit 2.1 Overview of the TIMSS 2007 Frameworks and Test Development Process

Date(s)		Group and Activity
January	2005	TIMSS & PIRLS International Study Center Review of TIMSS 2003 Assessment Frameworks Propose revisions for TIMSS 2007 Assessment Frameworks to be discussed in the First National Research Coordinators Meeting
February	2005	First National Research Coordinators Meeting (Cairo) Review of proposed changes to TIMSS 2003 Assessment Frameworks
March	2005	TIMSS & PIRLS International Study Center Compile TIMSS 2007 Content Domains and Percentages Survey to be completed by the National Research Coordinators Incorporating results from the survey and the First National Research Coordinators Meeting, prepare draft TIMSS 2007 Assessment Frameworks
April	2005	First Science and Mathematics Item Review Committee Meeting (London) Improve and review draft of TIMSS 2007 Assessment Frameworks
May	2005	Second National Research Coordinators Meeting (Amsterdam) Develop field test item pool using TIMSS Item-writing Guidelines
June-July	2005	TIMSS & PIRLS International Study Center Review, refine, and edit field test items in the pool Develop additional items to cover framework
August	2005	Science and Mathematics Task Forces (Boston) Review and revise field test items Develop additional items to cover framework
August	2005	Second Science and Mathematics Item Review Committee Meeting (Kyoto) Review field test item pool and select preferred and alternate items for field test to be reviewed in the Third National Research Coordinators Meeting
September	2005	TIMSS & PIRLS International Study Center Incorporating improvements and revisions from the Second Science and Mathematics Item Review Committee Meeting, finalize and publish the TIMSS 2007 Assessment Frameworks
November	2005	Third National Research Coordinator Meeting (Ljubljana) Review and approve prefered field test items
November	2005	TIMSS & PIRLS International Study Center Assemble field test item blocks

2 The mathematics task force consisted of Graham Ruddock, TIMSS Mathematics Coordinator; Robert Garden, former TIMSS Mathematics Coordinator; and Mary Lindquist, former president of National Council of Teachers of Mathematics. The science task force consisted of Christine O'Sullivan, TIMSS Science Coordinator; Audrey Champagne, State University of New York at Albany; and Jackie Heaton, University of Aberdeen.

Exhibit 2.1 Overview of the TIMSS 2007 Frameworks and Test Development Process (Continued)

Date(s)		Group and Activity
December	2005	TIMSS & PIRLS International Study Center Finalize field-test instruments and post them on the TIMSS & PIRLS International Study Center website for downloading
January	2006	TIMSS & PIRLS International Study Center Conduct a pilot test of constructed-response items Prepare draft scoring guides for constructed-response items
February	2006	Science and Mathematics Task Forces (Boston) Finalize scoring guides for constructed-response items Develop scoring training materials for the Fourth National Research Coordinators Meeting
March	2006	Fourth National Research Coordinators Meeting (Malta) Field test scoring training
March–April	2006	TIMSS 2007 field test administration
June	2006	TIMSS & PIRLS International Study Center Review field test item statistics Propose items for the main data collection to be discussed in the Third Science and Mathematics Item Review Committee Meeting
July	2006	Third Science and Mathematics Item Review Committee Meeting (St. Petersburg) Review proposed sets of items in the light of field test results
August	2006	Fifth National Research Coordinators Meeting (Riga) Review and approve proposed sets of items for the main data collection
August	2006	TIMSS & PIRLS International Study Center Finalize main survey materials and post them on the TIMSS & PIRLS International Study Center website for downloading
October	2006	TIMSS & PIRLS International Study Center Update scoring guides for constructed-response items Develop scoring training materials for the scoring training for the Southern Hemisphere countries
October	2006	Southern Hemisphere scoring training (Melbourne)
November	2006	TIMSS & PIRLS International Study Center Incorporate revisions from the Southern Hemisphere scoring training, finalize scoring guides for constructed-response items and post them on the TIMSS & PIRLS International Study Center website for downloading

2.2 Updating the TIMSS 2007 Assessment Frameworks

The *TIMSS 2007 Assessment Frameworks* closely resembles its corresponding publication for TIMSS 2003. However, there were three important areas of updates in the assessment for TIMSS 2007.

- In the Mathematics and Science Frameworks, the content domains were presented separately for the fourth and eighth grades, and there was a concerted effort to better reflect fourth-grade curricula. At both grades, there was an effort to consolidate the major content areas and adjust the assessment topic areas and objectives to make them more appropriate and feasible in the context of a large-scale international assessment.
- To increase the potential for analyzing and reporting the mathematics and science results according to cognitive domains, the U.S. National Center for Education Statistics provided support to examine and refine the mathematics cognitive domains used in TIMSS 2003 and report the results.[3] Accordingly, the cognitive domains in the Mathematics and Science Frameworks were updated to enable reporting by cognitive domains in the TIMSS 2007 International Reports.
- The assessment design was updated to increase students' response time in TIMSS 2007 because TIMSS 2003 had been somewhat speeded in some countries. The allocation of blocks to booklets was modified to include fewer blocks in a booklet and to have the design fully balanced. That is, each booklet in TIMSS 2007 included 2 mathematics blocks and 2 science blocks, with half the booklets having the mathematics blocks first and half having the science blocks first. The time provided to students to complete a block was increased—from 12 to 18 minutes at fourth grade and from 15 to 22.5 minutes at eighth grade.

In preparation for the first meeting of the TIMSS 2007 National Research Coordinators, the TIMSS & PIRLS International Study Center staff met with the Mathematics and Science Coordinators to discuss the updates to Mathematics and Science Frameworks and with consultants from Educational Testing Service to discuss the updates to the assessment design.

The proposed updates were discussed during the first NRC meeting in February 2005. The NRCs were pleased with the updated, simplified

3 For details of this cognitive study, see Mullis, I.V.S., Martin, M.O., & Foy, P. (2005). *IEA's TIMSS 2003 international report on achievement in the mathematics cognitive domains*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

design that provided students with increased response time, even though this meant countries measuring trend would be required to participate in a bridge study. In addition to full participation in TIMSS 2007, trend countries would need to administer a subset of TIMSS 2003 booklets in TIMSS 2007 under the TIMSS 2003 conditions. The NRCs also were supportive of the recommendations for updating the content and cognitive domains in the Mathematics and Science Frameworks, and were appreciative of the efforts to report results by the cognitive domains. They recommended conducting a survey to gather further information about how to update the content and cognitive domains.

Following the NRC meeting, TIMSS & PIRLS International Study Center staff distributed the survey questionnaires and asked NRCs to indicate their country's preferences with regard to the content domains (algebra, geometry, measurement, physics, chemistry, etc.) that would characterize mathematics and science in the assessment and the specific topics in each domain that would be addressed by the assessment objectives. The responses to the questionnaires from participating countries showed broad support for the proposed changes to the assessment frameworks and provided valuable guidance in drafting the 2007 frameworks in March 2005. The draft frameworks were reviewed by members of the SMIRC in April 2005, and the final *TIMSS 2007 Assessment Frameworks* were published in September 2005.

For the TIMSS 2007 frameworks, a decision was made to separate the content domains for both mathematics and science by grade because of the increasing complexity of the subject matter and the introduction, at the eighth grade, of topics not covered at the fourth grade. This also allowed the fourth grade content domains to be renamed, where appropriate, to better reflect the content being assessed. The cognitive domains were streamlined based on information stemming from the study of the skills and abilities used to answer mathematics items in the 2003 assessment.

The major revisions in the mathematics content domains were organizational. At fourth grade, the previous number and algebra domains were combined into a revised number domain, while the previous measurement and geometry domains were restructured into a new geometric shapes and measures domain. The 2003 data domain was renamed "data display", better reflecting what was being assessed at this grade. At the topic level, the fourth grade topics were streamlined, and topics where minimal content was assessed were combined with others.

At eighth grade, the 2003 measurement domain was eliminated, and the topics covered were redistributed to geometry (length, area, volume, angle, perimeter, and circumference) or number (time, speed, mass/weight, and temperature).

The cognitive domains for mathematics were reduced from four to three: knowing, applying, and reasoning. The previous two cognitive domains, using concepts and solving routine problems, were split across the three new domains. This change allowed the cognitive structure of the assessment frameworks to be the same for mathematics and science.

The revisions made to the science assessment framework were mostly of an organizational nature. Since the content domains for the fourth and eighth grades were presented separately, the domains of chemistry and physics at the fourth grade were combined into physical science. In addition, at the fourth grade, the topics covered in the environmental science domain were moved to life science and earth science. In TIMSS 2003, environmental science items had been included in these domains, because environmental science was not a reporting strand at the fourth grade. Although the environmental science domain was a reporting strand in 2003 at eighth grade, this strand was eliminated for TIMSS 2007 and its topics moved to biology and earth science, where a new topic area was added—Earth's resources, their use, and conservation. While topic areas at both grades were combined, the content remained essentially the same, with the exception of the topic of sound at the fourth grade. In the 2003 assessment, this topic area was not assessed, however, committee members decided that fourth grade students had a rudimentary knowledge of sound, and this area now should be assessed under physical science.

The topics that made up the science cognitive domains remained essentially the same as those contained within the TIMSS 2003 frameworks. However, the domain names did change so that they would be identical to those of mathematics. Thus, factual knowledge, conceptual understanding, and reasoning and analysis became knowing, applying, and reasoning, respectively. In addition, some of the topics included within the domains were moved from one cognitive domain to another based on current thinking about cognitive processes.

2.3 Mathematics Assessment Framework

The TIMSS 2007 mathematics framework is presented in full in the *TIMSS 2007 Assessment Frameworks*. The basic structure of the mathematics framework is defined by two dimensions, a content and a cognitive dimension, which remains unchanged from TIMSS 2003 and is illustrated in Exhibit 2.2. Exhibit 2.3 shows the topic areas within the dontent domains.

Exhibit 2.2 Target Percentages of the TIMSS 2007 Mathematics Assessment Devoted to Content and Cognitive Domains by Grade Level

Fourth Grade Content Domains	Percentages
Number	50%
Geometric Shapes and Measures	35%
Data Display	15%
Eighth Grade Content Domains	**Percentages**
Number	30%
Algebra	30%
Geometry	20%
Data and Chance	20%

Cognitive Domains	Percentages	
	Fourth Grade	**Eighth Grade**
Knowing	40%	35%
Applying	40%	40%
Reasoning	20%	25%

Exhibit 2.3 Topic Areas Included in the Mathematics Content Domains by Grade Level

Fourth Grade Content Domains	Fourth Grade Topic Areas
Number	Whole numbers
	Fractions and decimals
	Number sentences
	Patterns and relationships
Geometric Shapes and Measures	Lines and angles
	Two- and three-dimensional shapes
	Location and movement
Data Display	Reading and interpreting
	Organizing and representing
Eighth Grade Content Domains	**Eighth Grade Topic Areas**
Number	Whole numbers
	Fractions and decimals
	Integers
	Ratio, proportion, and percent
Algebra	Patterns
	Algebraic Expressions
	Equations/formulas and functions
Geometry	Geometric shapes
	Geometric measurement
	Location and movement
Data and Chance	Data organization and representation
	Data interpretation
	Chance

2.4 Science Assessment Framework

As in mathematics, the science assessment framework is fully detailed in the *TIMSS 2007 Assessment Frameworks*. It is organized along two dimensions: content and cognitive. The content dimension at the fourth grade is made up of three domains: life science, physical science, and earth science. The four content domains at the eighth grade are: biology, chemistry, physics, and earth science. There are three cognitive domains at both fourth and eighth grades: knowing, applying, and reasoning. The target percentages for the content and cognitive domains are shown in Exhibit 2.4 and the topic areas are listed in Exhibit 2.5.

Exhibit 2.4 Target Percentages of the TIMSS 2007 Science Assessment Devoted to Content and Cognitive Domains by Grade Level

Fourth Grade Content Domains	Percentages
Life Science	45%
Physical Science	35%
Earth Science	20%
Eighth Grade Content Domains	**Percentages**
Biology	35%
Chemistry	20%
Physics	25%
Earth Science	20%

Cognitive Domains	Percentages	
	Fourth Grade	**Eighth Grade**
Knowing	40%	30%
Applying	35%	35%
Reasoning	25%	35%

Exhibit 2.5 Topic Areas Included in the Science Content Domains by Grade Level

Fourth Grade Content Domains	Fourth Grade Topic Areas
Life Science	Characteristics and life processes of living things
	Life cycles, reproduction, and heredity
	Interaction with the environment
	Ecosystems
	Human health
Physical Science	Classification and properties of matter
	Physical states and changes in matter
	Energy sources, heat, and temperature
	Light and sound
	Electricity and magnetism
	Forces and motion
Earth Science	Earth's structure, physical characteristics, and resources
	Earth's processes, cycles, and history
	Earth in the solar system
Eighth Grade Content Domains	**Eighth Grade Topic Areas**
Biology	Characteristics, classification, and life processes of organisms
	Cells and their functions
	Life cycles, reproduction, and heredity
	Diversity, adaptation, and natural selection
	Ecosystems
	Human health
Chemistry	Classification and composition of matter
	Properties of matter
	Chemical change
Physics	Physical states and changes in matter
	Energy transformations, heat, and temperature
	Light
	Sound
	Electricity and magnetism
	Forces and motion
Earth Science	Earth's structure and physical features
	Earth's processes, cycles, and history
	Earth's resources, their use, and conservation
	Earth in the solar system and the universe

2.5 Developing Mathematics and Science Items and Scoring Guides

Because approximately half of the fourth- and eighth-grade items from TIMSS 2003 were kept secure to be readministered in the 2007 assessment, blueprints for mathematics and science item development were developed to ensure that the newly developed items met the guidelines laid out in the 2007 assessment framework. The blueprints were created by:

- Estimating the number of items needed in the assessment based on the total score points and percentage of score points in each content domain specified in the framework
- Distributing this number of items across the mathematics and science main topic areas according to their breadth of content
- Accounting for the number of trend items already included in each topic area
- Ensuring coverage of the cognitive domains and the appropriate numbers of multiple-choice and constructed-response items (the frameworks specify approximately 50% each)
- Scaling up the number of items to be developed to allow for attrition during the item selection and field-testing process.

This section describes the test development process and includes a consideration of trend items, development of the international item pool, item review and revision, problem- solving sets, field testing, item selection for the main data collection, development of scoring guides for constructed-response items, and scoring training.

2.5.1 Trend Items

The mathematics and science trend items from 2003 were mapped into the content and cognitive categories described in the TIMSS 2007 frameworks. The results are shown in Exhibits 2.6 and 2.7.

In mathematics at fourth grade , the number of multiple-choice and constructed-response items was about the same in each domain, so the newly developed items needed to maintain this balance. This was not the situation at eighth grade, where multiple-choice items predominated, except in the data and chance domain. For this grade, more new constructed-response items were needed to redress the balance.

Science included both multiple-choice and constructed-response items in each domain. At fourth grade, the number of multiple-choice and constructed-response trend items was about the same. Eighth grade had more multiple-choice items, hence, a larger proportion of constructed-response items needed to be developed for eighth grade.

Exhibit 2.6 Mathematics Trend Items by Grades, Content and Cognitive Domains, and Item Format

Content Domain	Fourth Grade Trend Items			Eighth Grade Trend Items		
	Multiple Choice	Constructed Response	Total	Multiple Choice	Constructed Response	Total
Number	24	27	51	18	10	28
Algebra*	–	–	–	14	8	22
Geometry**	12	9	21	18	9	27
Data and Chance***	5	4	9	6	12	18
Total	**41**	**40**	**81**	**56**	**39**	**95**

Cognitive Domain	Multiple Choice	Constructed Response	Total	Multiple Choice	Constructed Response	Total
Knowing	16	7	23	23	8	31
Applying	18	18	36	26	19	45
Reasoning	7	15	22	7	12	19
Total	**41**	**40**	**81**	**56**	**39**	**95**

* Prealgebraic concepts are included in the Number content domain at the fourth grade.
** Called Geometric Shapes and Measures at the fourth grade
*** Called Data Display at the fourth grade.

Exhibit 2.7 Science Trend Items by Grades, Content and Cognitive Domains, and Item Format

Content Domain	Fourth Grade Trend Items			Eighth Grade Trend Items		
	Multiple Choice	Constructed Response	Total	Multiple Choice	Constructed Response	Total
Biology*	22	10	32	16	18	34
Physical Science	13	16	29	–	–	–
Chemistry	–	–	–	9	5	14
Physics	–	–	–	16	7	23
Earth Science	5	10	15	10	13	23
Total	**40**	**36**	**76**	**51**	**43**	**94**

Cognitive Domain	Multiple Choice	Constructed Response	Total	Multiple Choice	Constructed Response	Total
Knowing	19	13	32	30	9	39
Applying	15	10	25	15	18	33
Reasoning	6	13	19	6	16	22
Total	**40**	**36**	**76**	**51**	**43**	**94**

* Called Life Science at the fourth grade.

2.5.2 Developing the International Item Pool for TIMSS 2007

During the second NRC meeting in May 2005, participants from 60 countries attended the item-writing workshop for TIMSS 2007. The TIMSS & PIRLS International Study Center provided specific instructions on how to write multiple-choice and constructed-response items in accordance with an item-writing manual that had been developed for TIMSS 2007. Participants then were organized into four subgroups—fourth grade mathematics, eighth grade mathematics, grade fourth science, and eighth grade science. These subgroups were further subdivided into content domain groups. The item-writing effort was very successful, yielding a large number of draft items—for mathematics, approximately 210 and 230 items at fourth and eighth grades, respectively, and for science, approximately 200 and 120 items at these respective grades. Also, several countries sent additional items to the TIMSS & PIRLS International Study Center in the weeks that followed.

2.5.3 Item Review and Revision

Following item development, members of the task force reviewed and revised the items and wrote additional items to cover specific areas of the frameworks that had not been addressed. The resulting item pool consisted of 997 items, of which 472 were in mathematics and 525 were in science. The items then were reviewed by the SMIRC. For this review, members of the committee formed two subgroups, a mathematics group and a science group. Members of the subgroups reviewed items for content accuracy, grade appropriateness, and framework fit.

To increase efficiency, the field test blocks were organized to represent the desired assessment as much as possible. So far as the field test was successful, then materials did not have to be reformatted. The replacement items were prepared in parallel blocks, so they could be incorporated into the assessment blocks with minimal disruption. Thus, the field-test item pool was divided into two sets, "preferred" and "alternate", and the items then organized into two sets of item blocks. One set of item blocks contained the "preferred" items and the second set contained the "alternate" items. These blocks of items were reviewed at the third NRC meeting that was held in Slovenia in November 2005. NRCs made suggestions for revising certain items, mostly based on concerns about translation issues. They also raised concerns about the grade appropriateness of a few items, and these were subsequently eliminated.

2.5.4 Item Sets

Several extended reasoning tasks or item sets of related problems were developed by the SMIRC and by members of the task force for inclusion in TIMSS 2007. These were written to specifically measure aspects of the frameworks that were difficult to assess using discrete items and, generally, were classified as reasoning. For these tasks, the number of possible points, typically 3 to 6 points, depended on the requirements for students to successfully complete the task.

In mathematics, the extended tasks involved patterns and their generalization with scenarios based on real-life contexts. Two were included in the final assessment for fourth grade and four for eighth grade.

In science, the extended tasks addressed aspects of science inquiry based on experimental set-ups or student investigations. They were reviewed and revised at each SMIRC and task force meeting. Among the tasks that were field tested, two were included at the fourth grade level and five at the eighth grade level in the final assessment.

2.5.5 Field Test

The newly developed fourth- and eighth-grade items were field tested internationally from March to April 2006. In total, 31 countries participated in the fourth grade field test and 45 countries participated in the eighth grade field test. The field test in each country was administered to a random sample of a minimum of 25 schools. Approximately twice the number of items were field tested than were needed for the TIMSS 2007 assessment. A total of 350 items were included in the fourth grade field test, 192 in mathematics and 158 in science. At the eighth grade, a total of 415 items were included in the field test, 214 in mathematics and 201 in science. Since some constructed-response items contributed 2 score points, this corresponded to a total of 203 score points in mathematics and 180 in science at the fourth grade, and 283 score points in mathematics and 240 in science at the eighth grade.

2.5.6 Item Selection for the TIMSS 2007 Data Collection

The selection of items for the TIMSS 2007 data collection was based on an item analysis of the international results of the field test. Data almanacs containing basic item statistics for each country and internationally were produced, including the following:

- Difficulty levels for each item
- How well items discriminated between high- and low-performing students
- The effectiveness of distractors in multiple-choice items
- The frequency of occurrence of diagnostic codes used in the scoring guides
- Scoring reliability for constructed-response items.

The TIMSS & PIRLS International Study Center and the mathematics and science coordinators reviewed data from the field test in June 2006. Items were selected that discriminated well, had a range of difficulties, and covered the cognitive and content domains, and draft blocks of these items were assembled, including a pool of alternate items.

The SMIRC reviewed these draft blocks in July 2006. Some items were revised slightly and other replaced with items from the alternate pool. The modified draft blocks then were reviewed by the NRCs at the fifth NRC meeting held in August 2006. The NRCs recommended some further changes to items and asked that a few items be replaced. A total of 196 new items at the fourth grade and 240 items at the eighth grade were approved by the NRCs for inclusion in the TIMSS 2007 data collection. The final assessments were made up of 353 items at fourth grade and 429 items at eight grade, including both trend and new items.

Exhibits 2.8 and 2.9 show the distribution of new and trend items in the TIMSS 2007 mathematics and science assessments by item format for fourth and eighth grades, respectively. They reflect the number of individual items and all item subparts included in multipart items.

The percentage of score points contributed to the assessments by constructed-response items for both mathematics and science and at both grade levels ranged from 50 to 55 percent.

Exhibit 2.8 Distribution of New and Trend Items in the TIMSS 2007 by Subject and Item Format – Fourth Grade

Item Format	Number of Items				
	New Items	Trend Items	Total (New+Trend)	Total Score Points	Percentage of Score Points
			Mathematics Items		
Multiple Choice	55	41	96	96	50%
Constructed Response	43	40	83	96	50%
Total Mathematics Items	**98**	**81**	**179**	**192**	
			Science Items		
Multiple Choice	53	40	93	93	48%
Constructed Response	45	36	81	101	52%
Total Science Items	**98**	**76**	**174**	**194**	
			All Items		
Multiple Choice	108	81	189	189	49%
Constructed Response	88	76	164	197	51%
Total Items	**196**	**157**	**353**	**386**	

Exhibit 2.9 Distribution of New and Trend Items in the TIMSS 2007 by Subject and Item Format – Eighth Grade

Item Format	Number of Items				
	New Items	Trend Items	Total (New+Trend)	Total Score Points	Percentage of Score Points
			Mathematics Items		
Multiple Choice	61	56	117	117	49%
Constructed Response	59	39	98	121	51%
Total Mathematics Items	**120**	**95**	**215**	**238**	
			Science Items		
Multiple Choice	56	51	107	107	45%
Constructed Response	64	43	107	133	55%
Total Science Items	**120**	**94**	**214**	**240**	
			All Items		
Multiple Choice	117	107	224	224	47%
Constructed Response	123	82	205	254	53%
Total Items	**240**	**189**	**429**	**478**	

2.5.7 Developing Scoring Guides for Constructed-response Items

Constructed-response items made up approximately half of the total assessment time in the TIMSS 2007 test, with each of these items having its own scoring guide.

2.5.7.1 The TIMSS General Scoring Method

TIMSS 2007, as in previous TIMSS assessments, included constructed-response items and used the same approach to scoring. Constructed-response questions generally are worth 1 or 2 score points, depending on the nature of the task or skills required to complete it. Typically, constructed-response items worth 1 score point require a numerical response in mathematics or a brief descriptive response in science, while those worth 2 score points require students to show their work or provide explanations using words and/or diagrams to demonstrate their conceptual understanding.

The generalized mathematics and science scoring guidelines that were developed and applied for TIMSS 2003 also were used in TIMSS 2007. The scoring guidelines are shown in Exhibit 2.10.

Exhibit 2.10 TIMSS Generalized Scoring Guidelines for Mathematics and Science Constructed-response Items

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

Each constructed-response item has its own scoring guide that utilizes a two-digit scoring scheme to provide diagnostic information. The first digit designates the correctness level of the response: 2 for a 2-point response, 1 for a 1-point response, and 7 for an incorrect response. The second digit, combined with the first, represents a diagnostic code used to identify specific types of approaches, strategies, or common errors and misconceptions. A second digit of 0 to 5 may be used for a predefined international code at each correctness level, while a second digit of 9 corresponds to "other" types of responses that fall within the appropriate correctness level but do not fit any of the predefined international codes. A special code (99) is given for completely blank responses. In general, only a few diagnostic codes are used to track high frequency correct or partial approaches or common misconceptions and errors. In addition to the international codes, second digit codes of 7 and 8 may be used by national centers to monitor specific responses not already captured by the internationally defined codes. The general TIMSS two-digit scoring scheme is summarized in Exhibit 2.11.

Exhibit 2.11 TIMSS Two-digit Scoring Scheme for Constructed-response Items

2-Point Items		1-Point Items	
Correctness Level	**International Code(s)**	**Correctness Level**	**International Code(s)**
Correct Responses	20–25: Category/method #1–#5 29: Other correct method	Correct Responses	10–15: Category/method #1–#5 19: Other correct method
Partial Responses	10–15: Category/method #1–#5 19: Other partial method	Incorrect Responses	70–75: Misconception/error #1–#5 79: Other error
Incorrect Responses	70–75: Misconception/error #1–#5 79: Other error	Blank	99
Blank	99		

2.5.7.2 Developing the TIMSS 2007 Scoring Guides

Scoring guides were written at the same time items were drafted. This helped ensure that the scoring guides captured what the items purported to measure and helped clarify the language of the items. The scoring guides were reviewed with the items by both the task force and members of the SMIRC. After the field test had been assembled, a pilot test of constructed-response items was conducted in English-speaking countries to collect student responses to use in developing scoring training materials for the field test. These responses helped refine the scoring guide, as well as clarify categories. They also helped in item revisions since ambiguities could be recognized by the way in which students responded to items. Selected student responses were included as examples in the scoring guides. Field test scoring training took place in March 2006 at the fourth NRC meeting. Scoring guides were further refined at this meeting.

A review of the field-test item statistics, however, showed that further refinements to the scoring guides could be made. For example, some of the categories were not being used and some guides showed no discrimination between students who received full credit for an item and those who received partial credit. In addition to the data, feedback was received from the scoring trainers in the participating countries about their experiences with the scoring guides. All this information was considered when selecting items for TIMSS 2007.

Where necessary, scoring guides were revised for items chosen for the TIMSS 2007 data collection. They were reviewed again by the SMIRC and NRCs during the final item review. Scoring guides for the trend items remained unchanged from the versions used in 2003.

2.5.7.3 Scoring Training Materials and Procedures

Scoring training for TIMSS 2007 was conducted in October 2006 for countries in the Southern Hemisphere and all of the scoring guides and training materials were finalized at this training. The training was replicated in March 2007 for countries in the Northern Hemisphere. The training materials used and the procedures followed for scoring training were very similar to those for previous TIMSS surveys. Representatives from countries participating in the survey were given information about the TIMSS scoring method and then trained using a selection of items chosen to illustrate the various types of constructed-response items. Materials for the scoring training were posted on the TIMSS & PIRLS International Study Center website prior to the meeting, and participants brought the materials with them. The training materials for each item included the scoring guide, a set of student papers illustrating the different score levels, and a set of practice papers.

The purpose of the scoring training was twofold: to present a model for use within each participating country and to give participants an opportunity to practice and resolve scoring issues with the most difficult items.

The following general procedures were followed while training each item:

- Participants read the item and its scoring guide.
- Trainers discussed the rationale and methodology of the scoring guide.
- Trainers presented and discussed the set of prescored example student responses.
- Participants scored the set of practice student responses.

2.6 Assessment Booklet Design

The TIMSS design for 2007 divided the 353 items at fourth grade and 429 items at eighth grade into 28 item blocks at each grade, 14 mathematics blocks labeled M01 through M14, and 14 science blocks labeled S01 through S14. Each block contained either mathematics items only or science items only. This general block design is shown in Exhibit 2.12 and is the same for each grade level. However, the assessment time was 18 minutes for fourth grade blocks and 22.5 minutes for eighth grade blocks. At fourth and eighth grades, seven blocks (the odd-numbered ones) contained secure items from 2003 that were used to measure trends, and seven blocks (the even-numbered ones) contained items newly developed for TIMSS 2007.

Exhibit 2.12 General Design of the TIMSS Matrix-sampling Blocks

Mathematics Blocks	Source of Items	Science Blocks	Source of Items
M01	Block M05 from TIMSS 2003	S01	Block S14 from TIMSS 2003
M02	New items for TIMSS 2007	S02	New items for TIMSS 2007
M03	Block M06 from TIMSS 2003	S03	Block S05 from TIMSS 2003
M04	New items for TIMSS 2007	S04	New items for TIMSS 2007
M05	Block M07 from TIMSS 2003	S05	Block S06 from TIMSS 2003
M06	New items for TIMSS 2007	S06	New items for TIMSS 2007
M07	Block M08 from TIMSS 2003	S07	Block S07 from TIMSS 2003
M08	New items for TIMSS 2007	S08	New items for TIMSS 2007
M09	Block M11 from TIMSS 2003	S09	Block S08 from TIMSS 2003
M10	New items for TIMSS 2007	S10	New items for TIMSS 2007
M11	Block M12 from TIMSS 2003	S11	Block S11 from TIMSS 2003
M12	New items for TIMSS 2007	S12	New items for TIMSS 2007
M13	Block M14 from TIMSS 2003	S13	Block S12 from TIMSS 2003
M14	New items for TIMSS 2007	S14	New items for TIMSS 2007

In the TIMSS 2007 design, the 28 blocks of items were distributed across 14 student booklets, as shown in Exhibit 2.13. Each booklet consisted of four blocks of items. To enable linking between booklets, each block appears in two booklets. The assessment time for individual students was 72 minutes at fourth grade and 90 minutes at eighth grade, which is comparable to that in the 1995, 1999, and 2003 assessments. The booklets were organized into 2 two-block sessions (Parts I and II), with a break in between each part.

Exhibit 2.13 Booklet Design for TIMSS 2007 – Fourth Grade and Eighth Grade

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

2.6.1 Assembling Item Blocks

The assessment blocks were assembled to create a balance across blocks and booklets with respect to content domain, cognitive domain, and item format. Depending on the exact number of multiple-choice and constructed-response items in each block, the total number of mathematics items in a block ranged from 10–14 at fourth grade and 11–18 at eighth grade. In science, depending on the exact number of multiple-choice and constructed-response items in each block, the total number of science items in a block ranged from 10–14 at fourth grade and 12–18 at eighth grade.

2.6.2 Incorporating Trend Items

At fourth grade, 14 blocks of items from TIMSS 2003 were used in TIMSS 2007—7 from mathematics and 7 from science. These were renumbered as shown in Exhibit 2.14.

At eighth grade in mathematics, one block of items from 1999 and one block containing items from 1999 and 2003 were used in 2007 (see Exhibit 2.15). These are labeled M01 and M03, respectively. The remaining five trend blocks were first administered in 2003. At eighth grade in science, similar to mathematics, one block containing items from 1999 and one block containing items from 1999 and 2003 were used in 2007 and labeled S03 and S05, respectively. The remaining five trend blocks were first administered in 2003.

Exhibit 2.14 TIMSS 2007 Mathematics and Science Blocks – Fourth Grade: Number of Items from Trend Blocks and Score Points by Assessment Year

Block	Number of Items from Trend Blocks*	Score Points by Assessment Year		
Mathematics Blocks		**2003**	**2007**	**Total**
M01	M05(11)	12	0	12
M03	M06(12)	12	0	12
M05	M07(11)	12	0	12
M07	M08(12)	12	0	12
M09	M11(12)	14	0	14
M11	M12(13)	13	0	13
M13	M14(10)	11	0	11
M02, M04, M06, M08, M10, M12, M14	–	0	106	106
Mathematics Total	**81**	**86**	**106**	**192**
Science Blocks				
S01	S14(11)	12	0	12
S03	S05(11)	12	0	12
S05	S06(10)	11	0	11
S07	S07(11)	12	0	12
S09	S08(11)	13	0	13
S11	S11(11)	12	0	12
S13	S12(11)	13	0	13
S02, S04, S06, S08, S10, S12, S14	–	0	109	109
Science Total	**76**	**85**	**109**	**194**
Overall Total	**157**	**171**	**215**	**386**

* The number of items from each trend block is indicated in parentheses.

Exhibit 2.15 TIMSS 2007 Mathematics and Science Blocks – Eighth Grade: Number of Items from Trend Blocks and Score Points by Assessment Year

Block	Number of Items from Trend Blocks*	Score Points by Assessment Year			
Mathematics Blocks		**1999**	**2003**	**2007**	**Total**
M01	M05(13)	16	0	0	16
M03	M06(15)	8	7	0	15
M05	M07(12)	0	17	0	17
M07	M08(15)	0	15	0	15
M09	M11(14)	0	15	0	15
M11	M12(15)	0	16	0	16
M13	M14(11)	0	15	0	15
M02, M04, M06, M08, M10, M12, M14	–	0	0	129	129
Mathematics Total	**95**	**24**	**85**	**129**	**238**
Science Blocks					
S01	S14(12)	0	15	0	15
S03	S05(14)	15	0	0	15
S05	S06(15)	8	7	0	15
S07	S07(12)	0	15	0	14
S09	S08(14)	0	16	0	15
S11	S11(13)	0	15	0	15
S13	S12(14)	0	16	0	16
S02, S04, S06, S08, S10, S12, S14	–	0	0	133	133
Science Total	**94**	**23**	**84**	**133**	**240**
Overall Total	189	47	169	262	478

* The number of items from each trend block is indicated in parentheses.

2.6.3 Alignment with the Mathematics and Science Frameworks

The test development process for TIMSS 2007 resulted in fourth- and eighth-grade assessments that are aligned with the *TIMSS 2007 Assessment Frameworks*. Details of the coverage of the frameworks are presented in the following subsections, for each grade level.

2.6.3.1 Fourth Grade Assessment

Exhibit 2.16 shows the distribution of score points for the fourth grade mathematics assessment by content and cognitive domains. (The mathematics framework target percentages can be seen in Exhibit 2.2.) The content domains were all within 1 percent of the target percentages. The proportion of reasoning items was slightly higher than the target because of the distribution of trend items.

Exhibit 2.16 Distribution of Score Points in the TIMSS 2007 Mathematics Assessment by Content and Cognitive Domains – Fourth Grade

Content Domain	Cognitive Domain			Total Score Points	Percentage of Score Points
	Knowing	Applying	Reasoning		
Number	41	33	24	98	51%
Geometric Shapes and Measures	26	28	11	65	34%
Data Display	6	14	9	29	15%
Total Score Points	**73**	**75**	**44**	**192**	
Percentage of Score Points	**38%**	**39%**	**23%**		

Exhibit 2.17 shows the score point distribution for the fourth grade science assessment by content and cognitive domain. (The science framework target percentages can be seen in Exhibit 2.4.) The percentages for the content domains were met. The percentage of items assessing knowing was higher than the target percentage, and the percentage of items assessing reasoning was lower than the target percentage. This was due in part to some of the items being reclassified based on the movement of several topics from one cognitive domain to another. About 17 percent of the score points in science measured scientific inquiry, and these covered topics from all three content domains.

Exhibit 2.17 Distribution of Score Points in the TIMSS 2007 Science Assessment by Content and Cognitive Domains and Scientific Inquiry Strand – Fourth Grade

Content Domain	Cognitive Domain			Total Score Points	Percentage of Score Points	Scientific Inquiry Score Points
	Knowing	Applying	Reasoning			
Life Science	41	31	13	85	44%	3
Physical Science	24	23	20	67	35%	24
Earth Science	24	14	4	42	21%	6
Total Score Points	**89**	**68**	**37**	**194**		**33**
Percentage of Score Points	**46%**	**35%**	**19%**			**17%**

As specified in the *TIMSS 2007 Assessment Frameworks*, two item types were included in the survey—multiple-choice and constructed-response. Exhibit 2.18 shows the distribution of the fourth grade mathematics and science items by item type and content domain.

Exhibit 2.18 Number of Mathematics and Science Items in TIMSS 2007 by Item Type and Content Domain – Fourth Grade

Content Domain	Multiple Choice	Constructed Response	Total Number of Items
		Mathematics Items	
Number	50	43	93
Geometric Shapes and Measures	32	28	60
Data Display	14	12	26
Total Mathematics Items	**96**	**83**	**179**
		Science Items	
Life Science	42	32	74
Physical Science	35	29	64
Earth Science	16	20	36
Total Science Items	**93**	**81**	**174**

TIMSS reports trends in student achievement in mathematics and science in the content domains of each subject. To facilitate linking to previous assessments, TIMSS 2007 at fourth grade included items from 2003 (TIMSS was not conducted at fourth grade in 1999). The number of score points in mathematics and science contributed by items used in TIMSS 2003 fourth grade can be seen in Exhibit 2.19.

Exhibit 2.19 Number of Score Points in TIMSS 2007 from Each Assessment Year by Mathematics and Science Content Domains – Fourth Grade

Content Domain	From 2003	New in 2007	Total 2007
		Mathematics	
Number	55	43	98
Geometric Shapes and Measures	21	44	65
Data Display	10	19	29
Total in Mathematics	**86**	**106**	**192**
		Science	
Life Science	35	50	85
Physical Science	32	35	67
Earth Science	18	24	42
Total in Science	**85**	**109**	**194**

The block and booklet design for 2007 ensured that the student booklets contained an appropriate balance of mathematics and science content. Exhibit 2.20 shows the number of mathematics and science score points in each fourth grade booklet.

Exhibit 2.20 Number of Score Points in TIMSS 2007 in Each Booklet by Mathematics and Science Content Domains – Fourth Grade

Content Domain	Booklet													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	Mathematics													
Number	15	12	11	12	13	13	14	16	17	16	14	15	14	14
Geometric Shapes and Measures	7	11	11	8	10	12	10	9	9	10	8	6	10	9
Data Display	4	3	5	7	5	3	4	4	4	3	5	4	2	4
Total in Mathematics	**26**	**26**	**27**	**27**	**28**	**28**	**28**	**29**	**30**	**29**	**27**	**25**	**26**	**27**
	Science													
Life Science	10	11	12	13	15	14	14	14	10	9	11	13	13	11
Physical Science	13	10	8	7	7	7	8	10	11	11	10	10	10	12
Earth Science	4	6	8	7	5	7	7	6	6	6	7	6	5	4
Total in Science	**27**	**27**	**28**	**27**	**27**	**28**	**29**	**30**	**27**	**26**	**28**	**29**	**28**	**27**

2.6.3.2 Eighth Grade Assessment

Exhibit 2.21 shows the distribution of score points for the eighth grade mathematics assessment by content and cognitive domain. For both content and cognitive domains, the percentage of score points was within 1 percent of the target percentage (see Exhibit 2.2 for target percentages).

Exhibit 2.21 Distribution of Score Points in the TIMSS 2007 Mathematics Assessment by Content and Cognitive Domains – Eighth Grade

Content Domain	Cognitive Domain			Total Score Points	Percentage of Score Points
	Knowing	Applying	Reasoning		
Number	28	33	11	72	30%
Algebra	33	15	21	69	29%
Geometry	8	29	13	50	21%
Data and Chance	14	21	12	47	20%
Total Score Points	**83**	**98**	**57**	**238**	
Percentage of Score Points	**35%**	**41%**	**24%**		

Exhibit 2.22 shows the distribution of score points across content and cognitive domains in the eighth grade science assessment. The target percentages (see Exhibit 2.4) were met for the content domains, but the percentages for the cognitive domains were higher than the target percent for knowing and applying and lower for reasoning. This was similar to fourth grade and was due in part to reclassification of items based on changes in the cognitive topic areas. Items that measured scientific inquiry accounted for approximately 24 percent of the score points in science. These items covered all three science content domains.

Exhibit 2.22 Distribution of Score Points in the TIMSS 2007 Science Assessment by Content and Cognitive Domains, and Scientific Inquiry Strand – Eighth Grade

Content Domain	Cognitive Domain			Total Score Points	Percentage of Score Points	Scientific Inquiry Score Points
	Knowing	Applying	Reasoning			
Biology	35	31	23	89	37%	20
Chemistry	16	18	12	46	19%	18
Physics	14	32	13	59	25%	14
Earth Science	24	16	4	44	19%	4
Total Score Points	**89**	**97**	**52**	**238**		**56**
Percentage of Score Points	**37%**	**41%**	**22%**			**24%**

Exhibit 2.23 shows the number of multiple-choice and constructed-response science items by their subject area content domains in TIMSS 2007.

Exhibit 2.23 Number of Mathematics and Science Items in TIMSS 2007 by Item Type and Content Domain – Eighth Grade

Content Domain	Multiple Choice	Constructed Response	Total Number of Items
		Mathematics Items	
Number	35	28	63
Algebra	34	30	64
Geometry	31	16	47
Data and Chance	17	24	41
Total Mathematics Items	**117**	**98**	**215**
		Science Items	
Biology	36	40	76
Chemistry	21	21	42
Physics	31	24	55
Earth Science	19	22	41
Total Science Items	**107**	**107**	**214**

To study trends in eighth grade student mathematics and science achievement, TIMSS 2007 included items from the 1999 and 2003 TIMSS assessments. Exhibit 2.24 shows that approximately 10 percent of points for the 2007 assessment in both mathematics and science came from items first administered in 1999, and approximately 30 percent of points came from items first administered in 2003.

Exhibit 2.24 Number of Score Points in TIMSS 2007 from Each Assessment Year by Mathematics and Science Content Domains – Eighth Grade

Content Domain	From 1999	From 2003	New in 2007	Total 2007
		Mathematics		
Number	12	19	41	72
Algebra	1	26	42	69
Geometry	8	22	20	50
Data and Chance	3	18	26	47
Total in Mathematics	**24**	**85**	**129**	**238**
		Science		
Biology	6	36	47	89
Chemistry	4	11	31	46
Physics	9	15	35	59
Earth Science	4	22	20	46
Total in Science	**23**	**84**	**133**	**240**

The number of score points and the distribution of score points across the mathematics and science content domains for each booklet in the eighth-grade assessment is shown in Exhibit 2.25.

Exhibit 2.25 Number of Score Points in TIMSS 2007 in Each Booklet by Mathematics and Science Content Domain – Eighth Grade

Content Domain	Booklet													
	1	2	3	4	5	6	7	8	9	10	11	12	13	14
	Mathematics													
Number	13	14	17	11	8	9	9	8	7	5	6	8	12	17
Algebra	6	5	5	9	10	9	11	11	9	13	13	15	15	7
Geometry	9	7	7	6	7	7	7	9	9	7	7	5	4	8
Data and Chance	6	7	6	10	11	9	6	5	7	8	7	4	4	4
Total in Mathematics	**34**	**33**	**35**	**36**	**36**	**34**	**33**	**33**	**32**	**33**	**33**	**32**	**35**	**36**
	Science													
Biology	14	11	13	12	10	11	11	14	14	12	12	12	15	17
Chemistry	6	7	6	7	7	7	6	5	8	8	7	7	6	5
Physics	6	10	11	8	8	8	9	8	8	9	7	9	10	7
Earth Science	7	5	5	8	8	7	8	8	7	7	7	6	4	5
Total in Science	**33**	**33**	**35**	**35**	**33**	**33**	**34**	**35**	**37**	**36**	**33**	**34**	**35**	**34**
Total Overall	**67**	**66**	**70**	**71**	**69**	**67**	**67**	**68**	**69**	**69**	**66**	**66**	**70**	**70**

2.6.4 Item Release Policy

TIMSS 2007 is the fourth assessment in a series of regular 4-year studies, providing trend data from 1995, 1999, and 2003. As in previous assessments, the design for TIMSS 2007 and beyond (2011, 2015, etc.) provides for retaining some of the items for the measurement of trends and releasing some items into the public domain. In TIMSS 2007, 6 of the 14 assessment blocks in each subject were released after the assessment. The released item blocks include the two blocks containing the items from 1999, two blocks containing trend items from 2003, and two blocks used for the first time in 2007.[4] As item blocks are released, new items will be developed to take their place. Exhibits 2.26 and 2.27 show the number of secure and released items from the TIMSS 2007 assessment for fourth and eighth grades listed by content domain.

4 Because TIMSS did not assess fourth grade students in 1999, the TIMSS 2007 fourth grade released blocks comprise four blocks from 2003 and two from 2007.

Exhibit 2.26 Number of Items in each Mathematics and Science Content Domain by Release Status in TIMSS 2007 – Fourth Grade

Content Domain	Secure	Released	Total
		Mathematics	
Number	55	38	93
Geometric Shapes and Measures	36	24	60
Data Display	14	12	26
Total Mathematics	**105**	**74**	**179**
		Science	
Life Science	45	29	74
Physical Science	38	26	64
Earth Science	20	16	36
Total Science	**103**	**71**	**174**
Total Overall	**208**	**145**	**353**

Exhibit 2.27 Number of Items in each Mathematics and Science Content Domain by Release Status in TIMSS 2007 – Eighth Grade

Content Domain	Secure	Released	Total
		Mathematics	
Number	31	32	63
Algebra	47	17	64
Geometry	25	22	47
Data and Chance	23	18	41
Total Mathematics	**126**	**89**	**215**
		Science	
Biology	47	29	76
Chemistry	26	16	42
Physics	33	22	55
Earth Science	21	20	41
Total Science	**127**	**87**	**214**
Total Overall	**253**	**176**	**429**

References

Mullis, I.V.S., Martin, M.O., & Foy, P. (2005). *IEA's TIMSS 2003 international report on achievement in the mathematics cognitive domains.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

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

Mullis, I.V.S., Martin, M.O., Smith, T.A., Garden, R.A., Gregory, K.D., Gonzalez, E.J., Chrostowki, S.J., & O'Connor, K.M. (2003). *TIMSS assessment frameworks and specifications 2003* (2nd ed.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Chapter 3

Developing the TIMSS 2007 Background Questionnaires

Ebru Erberber, Alka Arora, and Corinna Preuschoff

3.1 Overview

Student learning is influenced by various contextual factors, such as school resources, teacher characteristics, student attitudes, and home environment. To improve student achievement, it is important to understand the educational and social contexts in which students learn and how these relate to their achievement. Therefore, TIMSS 2007 collected a range of contextual information about teaching and learning in mathematics and science by administering background questionnaires at both the fourth and eighth grades. The questionnaires were based on the contextual framework included in the *TIMSS 2007 Assessment Frameworks* (Mullis, Martin, Ruddock, O'Sullivan, Arora, & Erberber, 2005).

This chapter describes the development of the contextual framework and the questionnaires. Four types of background questionnaires—curriculum, school, teacher, and student—organized around the TIMSS curriculum model were used in TIMSS 2007. The curriculum model has three aspects: the intended, implemented, and attained curriculum. These represent, respectively, the countries' mathematics and science curricula students are intended to learn; what is actually is taught in classrooms, including how it is taught and who teaches it; and what students have learned. The curriculum questionnaires asked about the structure and content of the intended curriculum in mathematics and science. The school, teacher, and student questionnaires asked about the mathematics and science content actually taught in classrooms, the instructional approaches used, the organization and resources of schools and classrooms, the preparation of teachers, and experiences and attitudes related to mathematics and science.

3.2 Updating the Contextual Framework for the Background Questionnaires

Just as the mathematics and science frameworks describe the content and cognitive domains to be assessed in those subjects, the contextual framework for TIMSS 2007 identifies the major characteristics of the educational and social contexts to be examined, with a view toward improving student learning in mathematics and science.

In conjunction with updating the TIMSS mathematics and science assessment frameworks for TIMSS 2007, the contextual framework was revised. The process of updating the contextual framework began at the first National Research Coordinators (NRCs) meeting in February 2005. In this meeting, the existing TIMSS 2003 contextual framework (Mullis, Martin, Smith, Garden, Gregory, Gonzales, Chrostowski, & O'Connor, 2003) was reviewed by the NRCs, who offered their suggestions for areas needing strengthening and revision and to identify potential new areas for inclusion in the contextual framework. In general, the NRCs were satisfied with the existing framework and recommended minor modifications.

Based on the suggestions from the NRCs and the TIMSS & PIRLS International Study Center staff, the following revisions were implemented in the TIMSS 2007 contextual framework:

- A section of school demographics–size, location, and characteristics of the student body–was added.
- The section on school environment was broadened to include social climate, such as the values and culture of the students, teachers, and administrators.
- A new section was added on technology, support, and equipment in order to extract information about factors that limit the use of technology in schools.
- A section for teacher evaluation was included.
- In order not to increase the response burden due to new sections being added, sections on classroom climate and students' prior experience were not included.

The revised contextual framework was reviewed one last time by the NRCs at the second NRC meeting in June 2005. NRCs provided additional input on the contextual framework, and, based on their suggestions, minor revisions were made to the framework. In September 2005, the

TIMSS & PIRLS International Study Center published the *TIMSS 2007 Assessment Frameworks* (Mullis, Martin, Ruddock, O'Sullivan, Arora, & Erberber, 2005).

3.3 Updating the TIMSS 2007 Background Questionnaires

TIMSS 2007 included four types of background questionnaires to collect information regarding the contexts in which students learn mathematics and science.

- The *Curriculum Questionnaire* collected information from the participating countries about the organization of the mathematics and science curriculum and the topics intended to be covered up to the fourth and eighth grades. Four versions of this questionnaire were administered: fourth grade mathematics, fourth grade science, eighth grade mathematics, and eighth grade science.
- The *School Questionnaire* asked the students' school principals to provide information about the school contexts and the resources available for mathematics and science instruction. There were separate versions for fourth and eighth grade.
- The *Teacher Questionnaire* collected information from the students' teachers about the teachers' backgrounds, preparation, and professional development. It also asked about instructional activities and collected very detailed information about the subject matter topics taught to students. Because students typically are taught both mathematics and science by the same teacher at the fourth grade, there was a single questionnaire for both subjects. At the eighth grade, there were separate versions for mathematics and science teachers.
- The *Student Questionnaire* addressed students' home and school lives and their experiences learning mathematics and science. There were separate questionnaires for fourth and eighth grade. At the eighth grade, there were different versions for countries where eighth grade science is taught as a single integrated subject and for countries where it is taught as separate subjects (i.e., biology, chemistry, physics, and earth science).

With each assessment cycle, TIMSS has a special committee of experienced NRCs to help guide the process of updating the questionnaires, called the Questionnaire Item Review Committee[1]. Updating the TIMSS 2007 background questionnaires was a collaborative effort among the

1 The members of the TIMSS 2007 Questionnaire Item Review Committee are provided in Appendix A.

TIMSS & PIRLS International Study Center, the NRCs, the Questionnaire Item Review Committee, and the IEA Data Processing and Research Center. The process included a series of reviews of draft questionnaires, a field test of the questionnaires, a review of field test data, and a revision of the field test instruments for use in the main data collection.

The curriculum, school, teacher, and student questionnaires used in TIMSS 2007 were developed based on the TIMSS 2003 questionnaires. While most of the questions were thematically similar in both assessments, some questions from 2003 were eliminated, others were modified, and some new questions were introduced in 2007 to provide additional information in areas deemed important to the study. In general, every effort was made to streamline the questionnaires in order to limit the response burden.

The development work began at the second NRC meeting in June 2005 when NRCs reviewed the TIMSS 2003 questionnaires in conjunction with the draft TIMSS 2007 contextual framework. NRCs thoroughly reviewed the content of the TIMSS 2003 questionnaires and shared comments in light of 1) the reporting of trend results for the 2007 survey and 2) the addition of new questions. As new questions were added, it was important to also eliminate questions to maintain the same level of response burden.

The Questionnaire Item Review Committee first met in September 2005. Committee members reviewed the contents of the questionnaires, in light of the TIMSS 2003 international reports, the TIMSS 2007 contextual framework, and NRC comments. The Questionnaire Item Review Committee members suggested many improvements, as well as ways to reduce response burden by eliminating some questions thought to be less useful for reporting purposes. Where items were used in the TIMSS 2003 reports, these questions were retained, preferably in the same form, in order to measure trends.

One important decision was to begin moving toward online data collection. For the first time, the *TIMSS 2007 Curriculum Questionnaire* was administered in an online format. Since national adaptations were not necessary for the questionnaire and countries completed it in English, the *Curriculum Questionnaire* was best suited for first experiences with the online data-collection process. The online format allowed for gathering more detailed information about educational policies and the implementation of the mathematics and science curriculum. Also, because the approximately 60 NRCs (a relatively small number of respondents) were responsible for the completion of the curriculum questionnaires, it was felt that the various rounds of reviews and the online format would result in good information and that it would not be necessary to include the *Curriculum Questionnaire* as part of the field test.

TIMSS & PIRLS International Study Center staff implemented the revisions suggested by the Questionnaire Item Review Committee and then provided the revised draft questionnaires to NRCs for review at the third TIMSS 2007 NRC meeting in November 2005. NRCs suggested a number of improvements to the questionnaires and these revisions were implemented by TIMSS & PIRLS International Study Center staff during December 2005. In particular, the school, teacher, and student questionnaires were prepared for the field test. The field test questionnaires then were provided to NRCs for translation, production, and administration.[2]

The TIMSS 2007 field test was conducted during March and April 2006. One of the primary purposes of the field test was to check across participating countries whether the questionnaires were appropriate for the measurement purposes for which they were designed. Although the questionnaires were adapted from previous versions, it was necessary to field test them, because there were a number of additions and refinements in the 2007 version. In total, 31 countries participated in the grade 4 field test, and 45 countries participated in the grade 8 field test.

After administering the field test, countries prepared their data files and sent them to the IEA Data Processing and Research Center for checking and cleaning. After the field test data were verified and transformed into the international format, they were sent to the TIMSS & PIRLS International Study Center for analysis and review. To facilitate review of the questionnaire data, the TIMSS & PIRLS International Study Center staff prepared data almanacs for each questionnaire that was field tested. For every country that participated, each almanac displayed student-weighted distributions of responses for each item on the questionnaires. For categorical variables, the weighted percentage of respondents choosing each option was shown together with the corresponding average student achievement in mathematics and science. For questions with numeric responses, the mean, mode, and selected percentiles were given. The almanacs were the basic data summaries that were used by TIMSS & PIRLS International Study Center staff, the Questionnaire Item Review Committee, and NRCs in assessing the quality of the field test instruments and in making suggestions for the instruments to be used in the data collection.

2 See Chapter 4 for more information about the translation and verification processes.

At the second Questionnaire Item Review Committee meeting in July 2006, committee members reviewed the field test results for the school, teacher, and student questionnaires, examining the statistics for each item and determining if there were any anomalies. The committee discussed modifications to some items and potential improvements suggested by the TIMSS & PIRLS International Study Center and finally arrived at a set of recommended changes to be brought before the NRCs at their next meeting. The Questionnaire Item Review Committee also proposed some final refinements to the draft curriculum questionnaires.

At the end of July 2006, TIMSS & PIRLS International Study Center staff prepared draft instruments for the main data collection and documented the recommended changes from the field test version for review by NRCs at the fifth NRC meeting in August 2006. The draft instruments were reviewed by NRCs who recommended several additional improvements. Immediately after the NRC meeting, TIMSS & PIRLS International Study Center staff finalized the instruments, and these were provided to NRCs in August for translation, production, and administration of the TIMSS 2007 data collection. This was held from September through December 2006 in countries participating on the Southern Hemisphere schedule and from March through July 2007 in countries participating on the Northern Hemisphere schedule.

As in TIMSS 2003, school, teacher, and student questionnaires were clearly organized into thematic blocks, each with a heading. The design and layout were updated for the TIMSS 2007 data collection. Parallel questions were used in different questionnaires to measure the same constructs from different sources, and, wherever possible, the wording of these questions was identical.

The content of each TIMSS 2007 background questionnaires follows in the next sections.

3.3.1 Curriculum Questionnaire

The fourth- and eighth-grade curriculum questionnaires for mathematics and science were provided online to NRCs who were asked to supply information about their nation's mathematics and science curricula in the target grades, drawing on the expertise of curriculum specialists and educators in their countries. The curriculum questionnaires were designed to collect basic information about the organization, content, and implementation of the intended mathematics and science curriculum in each country. They also were designed to determine whether the mathematics and science topics included in the TIMSS 2007 assessment were addressed in the country's intended curriculum through the target grade. New emphasis was placed upon policies of assigning homework and parental involvement. The *Curriculum Questionnaire* also asked about country-level policies regarding entry to primary or secondary school, as they related to the students tested in TIMSS 2007.

The four versions of the *Curriculum Questionnaire* were structured the same and were very similar in content, with the mathematics and science versions tailored to the subject matter and grade level, wherever necessary. One notable difference was that the eighth grade science curriculum questionnaire included a question asking whether eighth grade science was taught as a single integrated subject or as separate science subjects. Also, the mathematics versions of the questionnaire collected information about policies on calculator use.

The complete contents of the TIMSS 2007 mathematics and science curriculum questionnaires at fourth and eighth grades are described in Exhibit 3.1.

Exhibit 3.1 Content of the TIMSS 2007 Mathematics and Science Curriculum Questionnaires at the Fourth and Eighth Grades

Item Number				Item Content	Description
Mathematics Grade 4	Mathematics Grade 8	Science Grade 4	Science Grade 8		
1	1	1	1	National curriculum	Whether or not the country has a national mathematics/science currriculum
2	2	2	2	Grade-to-grade structure	Grade-to-grade structure of the primary/lower secondary school mathematics/science curriculum
–	–	–	3	Separate sciences	Whether or not science is taught as separate subjects by eighth grade and the specific subjects and grades taught
3	3	3	4	Year of introduction	Year the current mathematics/science curriculum was introduced
4	4	4	5	Curriculum revision(s)	Whether or not the mathematics/science curriculum is currently under revision
5	5	5	6	Goals, methods, and materials	Goals, objectives, methods, and materials prescribed by mathematics/science curriculum
6	6	–	–	Policy on calculator use	Whether or not the national mathematics curriculum contains statements/policies about the use of calculators
7	7	6	7	Policy on computer use	Whether or not the national mathematics /science curriculum contains statements/policies about the use of computers
8	8	7	8	Emphasis on approaches and processes	How much emphasis the national mathematics/science curriculum places on various instructional approaches and learning processes
9	9	8	9	The teaching of the TIMSS topics	Whether or not the TIMSS mathematics/science topics are included in the national curriculum, the proportion of students intended to be taught the topics, and the grade(s) at which the topics are intended to be taught
10	10	9	10	Differentiation of the curriculum	How the mathematics/science curriculum addresses the issue of students with different levels of ability
11	11	10	11	Form(s) of curriculum	Form(s) the mathematics/science curriculum is made available in
12	12	11	12	Instructional time and homework	Total amount of instructional time, percentage of total instructional time to be devoted to mathematics/science instruction, and whether or not there is a policy to assign mathematics/science homework
13	13	12	13	Remedial instruction	Whether or not there is a policy to provide remedial mathmatics/ science instruction
14	14	13	14	Teaching requirements	Requirements for being a mathematics/science teacher
15	15	14	15	Licensure process	Whether or not there is a process to license or certify mathematics/ science teachers, and what entity licenses the teachers
16	16	15	16	Preservice preparation to teach the curriculum	Whether or not mathematics/science teachers receive specific preparation in how to teach the mathematics/science curriculum as part of preservice education
17	17	16	17	Assistance to implement the curriculum	How do practicing teachers receive assistance to implement the mathematics/science curriculum
18	18	17	18	Communication of curriculum changes to teachers	Methods used to communicate mathematics/science curriculum changes with teachers
19	19	18	19	Communication of curriculum changes to parents	Methods used to communicate mathematics/science curriculum changes with parents
20	20	19	20	Parental involvement	Whether there is a policy to encourage parental involvement
21	21	20	21	Curriculum evaluation	How the implementation of the national curriculum is evaluated
22	22	21	22	Public examinations	Whether or not the country administers examinations in mathematics/science that have consequences for individual students, the authority that administers such examinations, and the grades at which these are given

3.3.2 School Questionnaire

Fourth- and eighth-grade school questionnaires were to be completed by the school principal of each school sampled for the study. They were designed to collect information concerning some of the major factors influencing student achievement in mathematics and science. The fourth- and eighth-grade versions of the questionnaire were nearly identical, with three questions addressing mathematics and science instruction separately at the eighth grade. The *School Questionnaire* was designed to be completed in about 30 minutes.

The complete contents of the TIMSS 2007 school questionnaires at fourth and eighth grades are described in Exhibit 3.2.

Exhibit 3.2 Content of the TIMSS 2007 School Questionnaires at the Fourth and Eighth Grades

Item Number		Item Content	Description
Grade 4	Grade 8		
1	1	Enrollment	Total school enrollment in all grades and in the target grade
2	2	Community size	Size of the community in which the school is located
3	3	Students' background	Percentage of students who come from economically disadvantaged or affluent homes
4	4	Students' native language	Percentage of students whose native language is the language of the test
5	5	Instructional time	Number of days per year and per week the school is open for instruction and number of hours of total instructional time in a typical day
6	6	Principal's time allocation	Percentage of time principal spends on various activities across the school year
7	7	Parental involvement	Whether or not the school asks parents to participate in various activities
8	8	School climate	Principal's perception of teachers' job satisfaction, parental support and involvement, expectations for student achievement, students' desire to do well in school and their regard for school property
9	9	Tracking in mathematics	Whether or not students are grouped by ability in their mathematics classes
10	10	Enrichment/remedial mathematics	Whether or not the school offers enrichment and remedial courses in mathematics
11	11	Tracking in science	Whether or not students are grouped by ability in their science classes
12	12	Enrichment/ remedial science	Whether or not the school offers enrichment and remedial courses in science
13	13	Professional development	Percentage of teachers who participated in various types of professional development activities during the school year
–	14	Teacher evaluation in mathematics	Whether or not the school uses various procedures in evaluating mathematics teachers
–	15	Teacher evaluation in science	Whether or not the school uses various procedures in evaluating science teachers
14	–	Teacher evaluation	Whether or not the school uses various procedures in evaluating teachers
15	16	Teacher vacancies	Difficulty in filling teacher vacancies in mathematics, science, and computer science/ information technology (fourth grade version does not ask about specific subjects)
16	17	Incentives for teachers	Whether or not the school uses incentives to recruit or retain teachers in mathematics, science, and/or other subjects (fourth grade version does not ask about specific subjects)
17	18	Student behavior	Frequency and severity of various problematic student behaviors occurring in the school
18	19	Instructional resources	Degree to which the school's capacity to provide instruction is affected by shortages or inadequacies of various resources
19	20	Science laboratory	Whether or not the school has a science laboratory and assistance for students conducting experiments
20	21	Computers	Number of computers available for educational purposes and proportion of computers with access to the Internet
21	22	Technology support	Whether there is anyone available to help teachers use information and communication technology for teaching and learning

3.3.3 Teacher Questionnaire

The teacher questionnaires were designed to gather information about the classroom contexts for teaching and learning mathematics and science, and about the topics taught in these subjects. For each participating school at the fourth grade, there was one *Teacher Questionnaire* addressed to the classroom teacher of the sampled class. For each sampled school at the eighth grade, a single mathematics class was sampled for the TIMSS 2007 assessment.[3] The mathematics teacher of that class was asked to complete a mathematics teacher questionnaire, and the science teacher(s) of that class was asked to complete a science teacher questionnaire.

Although the general background questions were essentially parallel across versions, questions pertaining to instructional and assessment practices, content coverage, and teachers' views about teaching the subject matter were tailored for mathematics or science. Many questions, such as those related to classroom characteristics and activities and homework and assessment, were specific to the classes sampled for TIMSS.

The TIMSS 2007 teacher questionnaires were designed to take about 45 minutes to complete. Because the fourth grade version includes questions about mathematics and science instruction, other questions that are less relevant at the fourth grade level were eliminated to reduce the response burden. The complete contents of the TIMSS 2007 teacher questionnaires are described in Exhibit 3.3 for the fourth grade and in Exhibit 3.4 for the eighth grade.

3 In some circumstances, it was necessary to sample two classes to yield the desired sample size. See Chapter 5 for more information on sample design.

Exhibit 3.3 Content of the TIMSS 2007 Teacher Questionnaire at the Fourth Grade

Item Number	Item Content	Description
1	Age	Teacher's age
2	Gender	Teacher's gender
3	Teaching experience	Number of years as a teacher
4	Teaching license	Whether or not the teacher has a teaching license or certificate
5	Formal education	Highest level of formal education completed by the teacher
6	Major area of study	Teacher's major area of study during postsecondary education
7	Teacher interactions	Frequency of various types of interactions the teacher has with colleagues
8	School safety	Teacher's perception about school safety
9	School facility	Teacher's perception about the adaquacy of the school facility
10	School climate	Teacher's perception of job satisfaction, parental support and involvement, expectations for student achievement, students' desire to do well in school and their regard for school property
11	Preparation to teach mathematics	How well prepared the teacher feels to teach the topics included in the TIMSS mathematics test
12	Mathematics class size	Number of students in the sampled class for mathematics and number of those in the fourth grade
13	Time spent teaching mathematics	Minutes per week the teacher teaches mathematics to the sampled class
14	Mathematics textbook	Whether or not a textbook(s) is used as a primary or supplementary resource in teaching mathematics
15	Student learning activities in mathematics	Percentage of time students spend doing various learning activities in a typical week of mathematics lessons
16	Calculator use policy	Whether or not the students are permitted to use calculators during mathematics lessons
17	Calculator use	Frequency with which students use calculators for various learning activities in mathematics
18	Computer availability for mathematics	Whether or not the students have access to computers during mathematics lessons and whether or not computers have access to the Internet
19	Computer use in mathematics	Frequency with which students use computers for various learning activities in mathematics
20	Mathematics content-related activities	Frequency with which the teacher asks students to do various content-related activities in mathematics
21	Emphasis on mathematics content areas	Percentage of time spent on mathematics content areas over the course of the year
22	Mathematics topic coverage	When the students were taught the TIMSS mathematics topics, by content area

Exhibit 3.3 Content of the TIMSS 2007 Teacher Questionnaire at the Fourth Grade (Continued)

Item Number	Item Content	Description
23	Mathematics homework	Whether or not the teacher assigns mathematics homework
24	Frequency of mathematics homework	How often the teacher assigns mathematics homework
25	Amount of mathematics homework	Number of minutes it would take an average student to complete a mathematics homework assignment
26	Student factors limiting teaching mathematics	Extent to which the teacher perceives various student factors limit teaching mathematics
27	Professional development in mathematics	Whether the teacher participated in various types of professional development activities for mathematics teaching
28	Preparation to teach science	How well prepared the teacher feels to teach the topics included in the TIMSS science test
29	Science class size	Number of students in the sampled class for science and number of those in the fourth grade
30	Time spent teaching science	Minutes per week the teacher teaches science to the sampled class
31	Computer availability for science	Whether or not the students have access to computers during science lessons and whether or not computers have access to the Internet
32	Computer use in science	Frequency with which students use computers for various learning activities in science
33	Science content-related activities	Frequency with which the teacher asks students to do various content-related activities in science
34	Student learning activities in science	Percentage of time students spend doing various learning activities in a typical week of science lessons
35	Science textbook	Whether or not a textbook(s) is used as a primary or supplementary resource in teaching science
36	Science topic coverage	When students were taught the TIMSS science topics, by content area
37	Science homework	Whether or not the teacher assigns science homework
38	Frequency of science homework	How often the teacher assigns science homework
39	Amount of science homework	Number of minutes it would take an average student to complete a science homework assignment
40	Student factors limiting teaching science	Extent to which the teacher perceives various student factors limit teaching science
41	Professional development in science	Whether or not the teacher participated in various types of professional development activities for science teaching

Exhibit 3.4 Content of the TIMSS 2007 Mathematics and Science Teacher Questionnaires at the Eighth Grade

Item Number		Item Content	Description
Mathematics Teacher Questionnaire	Science Teacher Questionnaire		
1	1	Age	Teacher's age
2	2	Gender	Teacher's gender
3	3	Teaching experience	Number of years as a teacher
4	4	Formal education	Highest level of formal education completed by the teacher
5	5	Major area of study	Teacher's major area of study during postsecondary education
6	6	Teaching license	Whether or not the teacher has a teaching license or certificate
7	7	Preparation to teach	How well prepared the teacher feels to teach the topics included in the TIMSS mathematics/science test
8	8	Teacher interactions	Frequency of various types of interactions the teacher has with colleagues
9	9	Professional development	Whether the teacher participated in various types of professional development activities
10	10	School safety	Teacher's perception about school safety
11	11	School facility	Teacher's perception about the adaquacy of the school facility
12	12	School climate	Teacher's perception of job satisfaction, parental support and involvement, expectations for student achievement, students' desire to do well in school and their regard for school property
13	13	Class size	Number of students in the sampled class
14	14	Time spent teaching subject	Minutes per week the teacher teaches mathematics/science to the sampled class
15	15	Textbook	Whether or not a textbook(s) is used as a primary or supplementary resource
16	16	Student learning activities	Percentage of time students spend doing various learning activities in a typical week
17	17	Content-related activities	Frequency with which the teacher asks students to do various content-related activities in mathematics/science
18	18	Factors limiting teaching	Extent to which the teacher perceives various student and resource factors to limit teaching
19	19	Emphasis on content areas	Percentage of time spent on mathematics/science content areas over the course of the year
20	20	Topic coverage	When students were taught the TIMSS mathematics/science topics, by content area
21	–	Calculator use policy	Whether or not the students are permitted to use calculators during mathematics lessons
22	–	Calculator use	Frequency with which the students use calculators for various learning activities in mathematics
23	21	Computer availability	Whether or not the students have access to computers during mathematics/science lessons and whether or not computers have access to the Internet
24	22	Computer use	Frequency with which the students use computers for various learning activities in mathematics/science
25	23	Homework	Whether or not the teacher assigns mathematics/science homework
26	24	Frequency of homework	How often the teacher assigns mathematics/science homework
27	25	Amount of homework	Number of minutes it would take an average student to complete a mathematics/science homework assignment
28	26	Type of homework	Frequency with which the teacher assigns various types of homework
29	27	Use of homework	How often the teacher uses mathematics/science homework for various purposes
30	28	Sources to monitor progress	Emphasis teacher places on sources to monitor students' progress in mathematics/science
31	29	Assessment	Frequency with which the teacher gives a mathematics/science test or examination
32	30	Question format	Item formats the teacher typically uses in mathematics/science tests or examinations
39	31	Type of questions	Types of questions the teacher uses in mathematics/science tests or examinations

3.3.4 Student Questionnaire

Each student in the sampled fourth- and eighth-grade TIMSS classes completed a *Student Questionnaire*. This questionnaire asked about the student's home background and resources for learning, attitude about mathematics and science, and experiences in learning these subjects. The fourth- and eighth-grade versions of the *Student Questionnaire* were thematically and organizationally similar to each other. While some questions were identical in the two versions, for other questions, the language was simplified in the fourth grade version or the specific content of the question was altered to be appropriate to this grade.

As in TIMSS 1999 and 2003, two versions of the eighth grade questionnaire were used, a *general science* version, intended for countries where eighth grade science is taught as a single integrated subject, and a *separate science subjects* version, intended for countries where eighth grade science is taught as separate subjects (i.e., biology, earth science, chemistry, and physics). Countries administered the version that was consistent with the way in which science instruction was organized at the eighth grade. In the general science version, science-related questions pertaining to students' attitudes and classroom activities were based on single questions asking about "science." Students responded in terms of the "general or integrated science" course they were taking. In the separate science subjects version, the same questions were asked about each science subject area, and students responded with respect to each science course they were taking. This structure accommodated the diverse systems that participated in TIMSS.

The TIMSS 2007 student questionnaires were designed to take about 30 minutes to complete. The complete contents of the TIMSS 2007 student questionnaires are described in Exhibit 3.5 for the fourth grade and in Exhibit 3.6 for the eighth grade.

Exhibit 3.5 Content of the TIMSS 2007 Student Questionnaire at the Fourth Grade

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

Exhibit 3.6 Content of the TIMSS 2007 Student Questionnaire at the Eighth Grade

Item Number		Item Content	Description
General science version	Separate science subjects version		
1	1	Age	Month and year of student's birth
2	2	Gender	Student's gender
3	3	Language	Student's frequency of use of the language of the test at home
4	4	Books in the home	Number of books in the student's home
5	5	Home possessions	Educational resources and general possessions in the student's home
6	6	Parents' education	Highest level of education completed by mother and father
7	7	Educational expectations	Level of education the student expects to complete
8	8	Liking mathematics	How much the student likes and feels competent at mathematics
9	9	Valuing mathematics	Importance and value the student attributes to mathematics
10	10	Learning activities in mathematics	Frequency with which student does various learning activities in mathematics lessons
11	–	Liking science	How much the student likes and feels competent at science
12	–	Valuing science	Importance and value the student attributes to science
13	–	Learning activities in science	Frequency with which student does various learning activities in science lessons
–	11	Study biology	Whether or not the student is studying biology this year
–	12	Liking biology	How much the student likes and feels competent at biology
–	13	Valuing biology	Importance and value the student attributes to biology
–	14	Learning activities in biology	Frequency with which student does various learning activities in biology lessons
–	15	Study earth science	Whether or not the student is studying earth science this year
–	16	Liking earth science	How much the student likes and feels competent at earth science
–	17	Valuing earth science	Importance and value the student attributes to earth science
–	18	Learning activities in earth science	Frequency with which student does various learning activities in earth science lessons
–	19	Study chemistry	Whether or not the student is studying chemistry this year
–	20	Liking chemistry	How much the student likes and feels competent at chemistry
–	21	Valuing chemistry	Importance and value the student attributes to chemistry
–	22	Learning activities in chemistry	Frequency with which student does various learning activities in chemistry lessons
–	23	Study physics	Whether or not the student is studying physics this year
–	24	Liking physics	How much the student likes and feels competent at physics
–	25	Valuing physics	Importance and value the student attributes to physics
–	26	Learning activities in physics	Frequency with which student does various learning activities in physics lessons
14	27	Computers	Whether or not student uses a computer, where student uses it, and frequency with which student uses a computer in mathematics and science
15	28	School climate	Student's affinity for school, perception of other students' motivation in school, and teachers' expectations
16	29	Safety in school	Whether or not the student experienced being the object of problematic behaviors by other students
17	30	Out-of-school activities	Frequency with which student does various nonacademic activities and homework outside of school
18	31	Mathematics homework	Frequency and amount of mathematics homework
19	31	Science homework	Frequency and amount of science homework
20	32	Parents born in country	Whether or not mother and father were born in country
21	33	Student born in country	Whether or not student was born in country and if not, the age at which the student emigrated

References

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

Mullis, I.V.S., Martin, M.O., Smith, T.A., Garden, R.A., Gregory, K.D., Gonzalez, E.J., Chrostowski, S.J., & O'Connor, K.M. (2003). *TIMSS assessment frameworks and specifications 2003* (2nd ed.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Chapter 4

Translation and National Adaptations of the TIMSS 2007 Assessment and Questionnaires

Ieva Johansone and Barbara Malak

4.1 Overview

The international version of the TIMSS 2007 assessment items and background questionnaires was developed and prepared in English by the TIMSS & PIRLS International Study Center, with contributions from the National Research Coordinators (NRCs) of participating countries. The test instruments were subsequently translated by these countries into their languages of instruction, 39 in total. The overarching goal was to create excellent quality translations that were appropriately adapted for the national context and at the same time are internationally comparable.

About one third of the participating countries administered the TIMSS 2007 assessment in more than one language. The most common languages of testing were English (16 countries) and Arabic (14 countries). Because so many countries administered the TIMSS assessment in Arabic, all the test instruments were translated into generic Arabic by CAPSTAN, a linguistic quality control company (Brussels, Belgium).

The TIMSS & PIRLS International Study Center prepared all the test instruments in English and in Arabic using the Adobe InDesign® layout program. The participating countries were provided with detailed instructions on translating and adapting the testing materials. Please refer to Chapter 6 and the TIMSS *2007 Survey Operations Procedures Unit 3, Preparing Materials for the* TIMSS *2007 Data Collection* (TIMSS & PIRLS International Study Center, 2006).

Since high-quality translations were essential to the success of the study, translated assessment items and background questionnaires were subjected to a stringent international translation verification process, organized by the

IEA Secretariat in Amsterdam. This process was intended to make certain that the translated materials were equivalent to the international version through direct comparisons of the two. Each participating country was asked to submit materials for verification prior to both the field test and the main data collection.

Generally, countries complied very well with the requirements for translation verification. Only one country (Saudi Arabia) did not submit their instruments for verification. A number of countries did not submit instruments in languages that were used for a smaller proportion of the sample. For example, Egypt did not submit its English version of the achievement test, and Romania did not submit its Hungarian version of the achievement test and *Student Questionnaire*. Finally, Japan, Mongolia, Morocco, and Latvia didn't submit instruments for translation verification on time and, therefore, verification was not completed until after their instruments were printed.

4.2 TIMSS Instruments to Be Translated

For TIMSS 2007, the following materials required translation for each grade tested:

- 28 blocks of achievement items
- One set of the booklet covers, directions, and instructions
- Background questionnaires for students, teachers, and school principals
- School coordinator and test administrator manuals
- Scoring guides for constructed-response items.

The TIMSS & PIRLS International Study Center provided each country with electronic files containing all of the materials to be translated. Even if each item block appears in two different test booklets, the files were designed in a way that each item block had to be translated only once. The same was true for the cover pages, directions, and instructions that were included in each test booklet. These "component files of the test booklets" were later distributed throughout the achievement booklets.

The item blocks and the background questionnaires underwent the international translation verification process (see section 4.6), whereas the manuals and scoring guides did not.

Countries testing in English and Arabic did not have to translate the instruments but were required to adapt the international version (English) or the generic Arabic version to the vernacular and make adaptations necessary for national reasons. The Arabic-speaking countries had to adapt the generic translation to their national context, comparing introduced adaptations to the international (English) version for international comparability.

4.2.1 Trend Item Blocks

About half of the TIMSS 2007 item blocks were used in TIMSS 2003 and were the basis of the TIMSS trend measurement. Countries that participated in TIMSS 2003 were required to use the same translation for the trend item blocks as was used in 2003.

Some of the trend item blocks also were used in TIMSS 1999. The Czech Republic, Thailand, and Turkey did not participate in TIMSS 2003 but did participate in TIMSS 1999. Thus, these countries used their translations from 1999 for these item blocks.

In some cases, however, countries decided that improvements to some of the translations from 1999 or 2003 were absolutely necessary. These changes were carefully documented and were referenced during data analysis. If the changes seemed to have altered the performance of any item, then this item was not included in trend analyses for this participant.

Additionally, countries that had participated in TIMSS 2003 also administered the four bridging study booklets at each grade. These booklets, labeled B1, B2, B3, and B4, were in fact booklets 5, 6, 11, and 12, respectively, from the TIMSS 2003 assessment. The countries were required to use these booklets as they were administered in TIMSS 2003. Since the translations in these booklets were not to be altered, they were not subject to international translation verification. However, they did undergo layout verification by the TIMSS & PIRLS International Study Center (for more information on the TIMSS & PIRLS International Study Center review of the survey instruments, please refer to Chapter 6).

4.3 Identification of the Target Language

For most participating countries, identifying the target language, the language in which the test instruments would be administered, was straightforward, because they have one dominant language that is used in public and private arenas. Some countries, however, use more than one language of instruction in their education systems. The translation process for these countries was even more challenging, since they had to make sure that the translations were equivalent across languages. In other cases, while there may be one language of instruction, there are other languages that are prominent in other parts of society. For example, Romania administered the achievement test and *Student Questionnaire* in Romanian and Hungarian but the teacher and school questionnaires only in Romanian. Exhibits 4.1 and 4.2 show the languages used by each participant for the various instruments.

Exhibit 4.1 Languages Used for the TIMSS 2007 Grade 4 Test Instruments

Country	Language	Instruments			
		Student Test	Student Questionnaire	Teacher Questionnaire	School Questionnaire
Alberta, Canada	English	●	●	●	●
	French	●	●	●	●
Algeria	Arabic	●	●	●	●
Armenia	Armenian	●	●	●	●
Australia	English	●	●	●	●
Austria	German	●	●	●	●
British Columbia, Canada	English	●	●	●	●
	French	●	●	●	●
Chinese Taipei	Traditional Chinese	●	●	●	●
Colombia	Spanish	●	●	●	●
Czech Republic	Czech	●	●	●	●
Denmark	Danish	●	●	●	●
Dubai, UAE	Arabic	●	●	●	●
	English	●	●	●	●
El Salvador	Spanish	●	●	●	●
England	English	●	●	●	●
Georgia	Georgian	●	●	●	●
Germany	German	●	●	●	●
Hong Kong SAR	Modern Standard Chinese	●	●	●	●
	English	●	●	●	●
Hungary	Hungarian	●	●	●	●
Iran, Islamic Rep. of	Farsi	●	●	●	●
Italy	Italian	●	●	●	●
Japan	Japanese	●	●	●	●
Kazakhstan	Kazakh	●	●	●	●
	Russian	●	●	●	●
Kuwait	Arabic	●	●	●	●
Latvia	Latvian	●	●	●	●
Lithuania	Lithuanian	●	●	●	●
Massachusetts, US	English	●	●	●	●
Minnesota, US	English	●	●	●	●
Mongolia	Mongolian	●	●	●	●
	Kazakh	●	●	●	●
Morocco	Arabic	●	●	●	●
Netherlands	Dutch	●	●	●	●
New Zealand	English	●	●	●	●
Norway	Norwegian	●	●	●	●
Ontario, Canada	English	●	●	●	●
	French	●	●	●	●
Qatar	Arabic	●	●	●	●
	English	●	–	–	–
Quebec, Canada	English	●	●	●	●
	French	●	●	●	●
Russian Federation	Russian	●	●	●	●
Scotland	English	●	●	●	●
Singapore	English	●	●	●	●
Slovak Republic	Slovak	●	●	●	●
	Hungarian	●	●	–	–
Slovenia	Slovene	●	●	●	●
Sweden	Swedish	●	●	●	●
Tunisia	Arabic	●	●	●	●
Ukraine	Ukrainian	●	●	●	●
	Russian	●	●	–	–
United States	English	●	●	●	●
Yemen	Arabic	●	●	●	●

Exhibit 4.2 Languages Used for the TIMSS 2007 Grade 8 Test Instruments

Country	Language	Instruments				
		Student Test	Student Questionnaire	Mathematics Teacher Questionnaire	Science Teacher Questionnaire	School Questionnaire
Algeria	Arabic	●	●	●	●	●
Armenia	Armenian	●	●	●	●	●
Australia	English	●	●	●	●	●
Bahrain	Arabic	●	●	●	●	●
	English	●	●	●	●	●
Basque Country, Spain	Basque	●	●	●	●	●
	Spanish	●	●	●	●	●
Bosnia and Herzegovina	Bosnian	●	●	●	●	●
	Croatian	●	●	●	●	●
	Serbian	●	●	●	●	●
Botswana	English	●	●	●	●	●
British Columbia, Canada	English	●	●	●	●	●
	French	●	●	●	●	●
Bulgaria	Bulgarian	●	●	●	●	●
Chinese Taipei	Traditional Chinese	●	●	●	●	●
Colombia	Spanish	●	●	●	●	●
Cyprus	Greek	●	●	●	●	●
Czech Republic	Czech	●	●	●	●	●
Dubai, UAE	Arabic	●	●	●	●	●
	English	●	●	●	●	●
Egypt	Arabic	●	●	●	●	●
	English	●	–	–	–	–
El Salvador	Spanish	●	●	●	●	●
England	English	●	●	●	●	●
Georgia	Georgian	●	●	●	●	●
Ghana	English	●	●	●	●	●
Hong Kong SAR	Modern Standard Chinese	●	●	●	●	●
	English	●	●	●	●	●
Hungary	Hungarian	●	●	●	●	●
Indonesia	Bahasa Indonesian	●	●	●	●	●
Iran, Islamic Rep. of	Farsi	●	●	●	●	●
Israel	Hebrew	●	●	●	●	●
	Arabic	●	●	●	●	●
Italy	Italian	●	●	●	●	●
Japan	Japanese	●	●	●	●	●
Jordan	Arabic	●	●	●	●	●
Korea, Rep. of	Korean	●	●	●	●	●
Kuwait	Arabic	●	●	●	●	●
Lebanon	English	●	●	●	●	●
	French	●	●	●	●	●
Lithuania	Lithuanian	●	●	●	●	●
Malaysia	Malay	●	●	●	●	●
	English	●	–	●	●	●
Malta	English	●	●	●	●	●
Massachusetts, US	English	●	●	●	●	●
Minnesota, US	English	●	●	●	●	●
Mongolia	Mongolian	●	●	●	●	●
	Kazakh	●	●	●	●	●

Exhibit 4.2 Languages Used for the TIMSS 2007 Grade 8 Test Instruments (Continued)

Country	Language	Instruments				
		Student Test	Student Questionnaire	Mathematics Teacher Questionnaire	Science Teacher Questionnaire	School Questionnaire
Morocco	Arabic	●	●	●	●	●
Norway	Norwegian	●	●	●	●	●
Oman	Arabic	●	●	●	●	●
	English	●	●	●	●	●
Ontario, Canada	English	●	●	●	●	●
	French	●	●	●	●	●
Palestinian Nat'l Auth.	Arabic	●	●	●	●	●
Qatar	Arabic	●	●	●	●	●
	English	●	–	–	–	–
Quebec, Canada	English	●	●	●	●	●
	French	●	●	●	●	●
Romania	Romanian	●	●	●	●	●
	Hungarian	●	●	–	–	–
Russian Federation	Russian	●	●	●	●	●
Saudi Arabia	Arabic	●	●	●	●	●
Scotland	English	●	●	●	●	●
Serbia	Serbian	●	●	●	●	●
Singapore	English	●	●	●	●	●
Slovenia	Slovene	●	●	●	●	●
Sweden	Swedish	●	●	●	●	●
Syrian Arab Republic	Arabic	●	●	●	●	●
Thailand	Thai	●	●	●	●	●
Tunisia	Arabic	●	●	●	●	●
Turkey	Turkish	●	●	●	●	●
Ukraine	Ukrainian	●	●	●	●	●
	Russian	●	●	–	–	–
United States	English	●	●	●	●	●

4.4 Translators and Reviewers

To translate the items and questionnaires, countries were strongly encouraged to hire highly qualified national translators and reviewers. Translators were expected to have an excellent knowledge of both English and the target language, experience in the country's cultural context, and, if possible, experience in the subject matter, preferably at the level of the target grade.

All translations had to be reviewed by a translation reviewer. Reviewers were expected to have experience with students in the target grade (preferably a fourth grade and/or eighth grade mathematics and/or science teacher), experience in a country's cultural context, and an excellent knowledge of both English and the target language.

Countries could employ more than one translator and/or reviewer (per target grade and language) and divide the work, if necessary. However, it was important to ensure the consistency of the translations within and across instruments.

4.5 Translation and Adaptation of the Instruments

Each translator and reviewer was given the international version of the TIMSS 2007 test instruments to be translated. The role of the reviewer was to check that the translation was correct and appropriate for the target population. The reviewer's suggestions then were analyzed by the NRC and incorporated into the translations, if necessary.

The TIMSS & PIRLS International Study Center provided translators and reviewers with directions to follow in translating and adapting the test instruments. The directions were designed to yield translations that were as close as possible to the international (English) version of the survey instruments, while allowing for national adaptations where necessary. In translating and adapting the TIMSS 2007 instruments, translators and reviewers were asked to pay particular attention to the following issues:

- Finding words/terms and phrases in the target language that are equivalent to those in the international version
- Making sure that the essential meaning of the text and reading level do not change
- Making sure that the difficulty level of achievement items does not change
- Ensuring that the translated text has equivalent qualifiers and modifiers appropriate for the target language
- Ensuring that the translated questionnaires ask the same questions as the international version and that national adaptations are made appropriately.

It also was extremely important to keep in mind that these translations were intended for fourth grade and/or eighth grade students and should reflect the language level of this audience. Translators were not permitted to clarify, take out, or add explanations to the source text. At the same time, idiomatic expressions had to be translated appropriately but not necessarily word for word.

Additionally, because unusual results from the TIMSS 2007 field test could have been an indication of errors in translation, each country was asked to check and, if this was the case, to correct the translation for the final TIMSS 2007 test instruments.

4.5.1 Adaptations in Achievement Items

Very few modifications were allowed to the items beyond those necessitated by the translation into the target language. Translators were encouraged to change phrasing and expressions that were not common to the country's national context and not related to the substance of the questions. For example, references to the work week as Monday through Friday could be altered. Also, whenever possible, punctuation or notation (e.g., decimal points), measurement units, expressions of date and time, names of people, places, animals, plants, etc. could be changed to make them equally familiar to all students. However, changes to names had to be similar in length and complexity to the originals.

There were some vocabulary issues that had to be kept in mind when translating the achievement items. Some words were pertinent to the item (e.g., a science item asking to explain "soil erosion"). Some words were not pertinent (e.g., in a mathematics item asking to compare data on "classmates' favorite juice", "mango juice" could be changed to "apple juice"). For multiple-choice items, translators had to pay particular attention to the correspondence between words in the item stems and options. Some items required an exact (verbatim) match between words in the stem and options of an item.

For some items, a nonexisting currency "zed" is used. Countries had to keep this currency or, replace it with another nonexisting currency, in case the word "zed" could not be used in their language.

4.5.2 Adaptations in Background Questionnaires

Unlike the achievement items, there were a few places in the questionnaires where national adaptations were required. Questions or information in carets (< >) had to be replaced with the country-appropriate term. For example, <eighth-grade> in the international version was replaced with "Form III" in the Maltese version. Questions that asked students and teachers about levels of education utilized the ISCED-1997 system. The international versions of the questionnaires provided the generic ISCED levels in carets to be replaced with the educational terms appropriate for each country. For example, <ISCED 3> was replaced with "high school" in the United States version of the questionnaires. NRCs were provided with detailed directions and the Operational Manual for ISCED-1997 (UNESCO, 1999) to assist them in determining the equivalent educational levels in their countries.

In addition to these required adaptations, participants were allowed to add questions of national interest to the questionnaires. Countries were encouraged, however, to add items only at the end of the questionnaires to avoid influencing the responses to the international questions in any way. The country-specific questions were required to appear in the same form as the rest of the questionnaire and to be approved by the TIMSS & PIRLS International Study Center.

4.5.3 Documenting National Adaptations

NRCs were required to document all adaptations made to the international test instruments on the National Adaptations Forms. The forms had to be completed and reviewed at various stages of preparing the national test instruments. NRCs completed Version I of the forms during the internal translation and review process and sent it along with the rest of the materials for international translation verification. After translation verification, NRCs updated the forms (Version II) to reflect any changes resulting from the verification and sent them along with the national instruments for TIMSS & PIRLS International Study Center review. After finalizing the national instruments, NRCs updated the forms again (Version III) for data processing at the IEA Data Processing and Research Center (DPC) and as a final documentation of their national adaptations.

NRCs received detailed instructions on how to complete each version of the National Adaptations Forms. The forms were supplied as an electronic document and treated as a set. It was required that each version be submitted as a single document upon completion.

4.6 International Verification of the Translations

Each translation went through a rigorous verification process that included internal verification of the translations at the national centers, independent verification by an international translation company, and a check by International Quality Control Monitors to determine whether or not the verifier's suggestions had been adapted. As the last step, the TIMSS & PIRLSInternational Study Center reviewed the assembled test instruments from all participating countries. For more information on the checking process used by International Quality Control Monitors and the review process used by TIMSS & PIRLS International Study Center staff, please refer to Chapter 6.

Once the instruments had been translated and internally reviewed, the text of the booklet, cover pages, directions, instructions, item blocks, and background questionnaires were submitted for international translation verification. This process was managed by the IEA Secretariat in Amsterdam, who enlisted the assistance of two independent translation companies to verify translations for each country: Lionbridge (offices in Dublin, Ireland and Brussels, Belgium) and CAPSTAN.

The international translation verifiers for TIMSS were required to have the target language as their first language, formal credentials as translators working in English, be educated at the university level, and live and work in the target country. When the last condition could not be met, verifiers were expected to maintain close contact with the country and its culture.

4.6.1 The Translation Verification Process

The international translation verifiers attended a training seminar where they received general information about the study and the design of the instruments, together with a description of the translation procedures used by the national centers. They also received detailed instructions for reviewing the instruments and registering deviations from their original version.

The primary task of the translation verifiers was to evaluate the accuracy of the translation and adequacy of the national adaptations (reported in the National Adaptation Forms). Their instructions emphasized the importance of maintaining the meaning and the difficulty level of the items, as well as questions included in each of the background questionnaires. Specifically, verifiers had to ensure the following:

- The translation has not affected the meaning or difficulty of the text.
- The items have not been made easier or more difficult when translated or adapted.
- No information has been omitted or added in the translated text.
- The questionnaires contain all the correct questions and answer options.
- The National Adaptations Forms reflect all adaptations planned to be implemented into the national test instruments.

Verifiers also were asked to suggest, if necessary, an alternative that would improve the comparability (i.e., the equivalence between the adapted version and the international source version) and provide the overall evaluation of translation, its accuracy, and cultural relevance.

The verifiers documented any errors or suggested changes either directly in the submitted PDF documents (using the eXPert PDF 4 Professional application) or in a specially created report form (using Microsoft® Word).

To help NRCs understand the comparability of the translated text with the international version, verifiers were asked to assign a "severity code" to any deviations. The following severity codes ranged from 1 (*major change or error*) to 4 (*acceptable change*):

1. *Major change or error*: Examples include the incorrect order of choices in a multiple-choice item, omission of a graphic, omission of an item or question, incorrect translation resulting in the answer being indicated by the item, an incorrect translation that changes the meaning or difficulty of the item or question, and the incorrect order of the questions or items.
2. *Minor change or error*: Examples include spelling errors that do not affect comprehension, misalignment of margins or tabs, inappropriate changes in fonts or font sizes, and discrepancies in the headers and footers of the document.
3. *Suggestion for alternative*: The translation may be adequate, but the verifier suggests a different wording.
4. *Acceptable change*: The change was acceptable and appropriate but was not documented in the National Adaptations Forms.

Additionally, for the countries that also participated in prior cycles of the study, verifiers checked that the translated version of the trend item blocks was identical to the version administered in 1999 or 2003. Any discrepancies were documented in the trend item checklist.

The translation verification feedback (either corrections and comments in the PDF version of the instruments or registered in separate forms in a Word format) was sent to the national centers. NRCs were responsible for reviewing translation verifier's suggestions and revising the test instruments. The NRCs also were asked to complete a Translation Verification Summary Form, providing comments on the verifier's suggestions that they had decided not to implement.

References

TIMSS & PIRLS International Study Center. (2006). *TIMSS 2007 survey operations procedures unit 3: Preparing materials for the TIMSS 2007 data collection*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

UNESCO Institute for Statistics. (1999). *Operational manual for ISCED-1997 (international standard classification of education)*. Paris: Author.

Chapter 5

TIMSS 2007 Sample Design

Marc Joncas

5.1 Overview

This chapter describes the TIMSS 2007 sample design, which consists of a set of specifications for the target and survey populations, sampling frames, survey units, sample selection methods, sampling precision, and sample sizes. The sample design is intended to ensure that the TIMSS 2007 survey data provide accurate and efficient estimates of national student populations. Since measuring trends is a central goal of TIMSS, the sample design also aims to provide accurate measures of changes in student achievement from cycle to cycle. In addition to the sample design, the TIMSS 2007 sampling activities also include estimation procedures for sample statistics and procedures for measuring sampling error. These other components are described in Chapters 9. The basic TIMSS sample design has two stages: schools are sampled with probability proportional to size at the first stage, and one or more intact classes of students from the target grades are sampled at the second stage.

All participants followed the uniform sampling approach specified by the TIMSS 2007 sample design with minimum deviations. This ensured that high quality standards were maintained for all participants, avoiding the possibility that differences between countries in survey results could be attributable to the use of different sampling methodologies. This uniform approach also facilitated an efficient approval process of the national designs by the international project team.

The TIMSS 2007 National Research Coordinator (NRC) of each participating country was responsible for implementing the sample design, including documenting every step of the sampling procedure for approval by the TIMSS & PIRLS International Study Center prior to implementation. To support NRCs in their sampling activities, a series of manuals: the *TIMSS 2007 School Sampling Manual, TIMSS 2007 Survey*

Operations Procedures Unit 2, and *TIMSS 2007 School Coordinator Manual* (TIMSS & PIRLS International Study Center, 2005, 2006a, 2006b) and sampling software (IEA Data Processing and Research Center, 2006) were provided. In addition to these materials, Statistics Canada and the Sampling Unit at the IEA DPC consulted with each country throughout the process.

5.2 TIMSS Target Populations

TIMSS 2007 chose to study achievement in two target populations—the fourth and eighth grade in most countries. Participating countries were free to select either population or both. The target populations can be seen as a collection of units to which the survey results apply. The main groups of interest in TIMSS are student populations (since by-products of the selection methods, schools and classes, also can be considered as populations). The formal definitions of the TIMSS target populations make use of UNESCO's International Standard Classification of Education (ISCED) (UNESCO Institute for Statistics, 1999) in identifying the appropriate target grades:

Fourth grade population. This includes all students enrolled in the grade that represents 4 years of formal schooling, counting from the first year of ISCED Level 1, provided that the mean age at the time of testing is at least 9.5 years. For most countries, the target grade should be the fourth grade or its national equivalent.

Eighth grade population. This includes all students enrolled in the grade that represents 8 years of formal schooling, counting from the first year of ISCED Level 1, provided that the mean age at the time of testing is at least 13.5 years. For most countries, the target grade should be the eighth grade or its national equivalent.

The rationale behind these definitions is as follows:

Since the aim of TIMSS is to improve student learning in mathematics and science, it is crucial to be able to link student achievement to school practices and educational policies, most of which are tied to grade levels. TIMSS grade-level results must be as directly useful as possible for educational purposes.

To be educationally useful, the amount of schooling represented by the grade assessed should be comparable across countries. Therefore, the focus should be on comparing student achievement after the same amount of schooling.

Based on previous cycles of TIMSS and PIRLS, the grades assessed in TIMSS should represent 4 years and 8 years of formal schooling.

The procedure for identifying the first grade from which to begin counting years of schooling should be based on an internationally accepted classification scheme. As mentioned above, such a scheme exists in UNESCO's ISCED.

In IEA studies, the above definitions correspond to what is known as the *international desired target populations*. All students enrolled in the appropriate target grades, regardless of their age, belong to the international desired target populations. All schools of all education subsystems that have students learning full-time in the appropriate target grades are part of the international desired target populations. Schools that do not contain the target grades are automatically excluded from the study. Each participating country was expected to define their *national desired target population*s to correspond as closely as possible to these definitions. In order to measure trends, it was critical that countries that participated in previous TIMSS cycles chose the same target grades for TIMSS 2007 that were used in the previous cycles. Information about the target grades in each country is provided in Chapter 9.

Although countries were expected to include all students in the target grades in their definitions of the population, sometimes it was not possible to include all students who fell under the definition of the international desired target populations. Consequently, based on geographic or linguistic constraints, a country's *national desired target population* excluded some sections of the population occasionally. For example, Lithuania's national desired target populations included only students in Lithuanian-speaking schools, representing respectively, 93 and 92 percent of the fourth and eighth grade international desired populations of students in the country.

Working from the national desired population, each country had to operationalize the definition of its population for sampling purposes and define their *national defined population*. While these national defined target populations ideally should coincide with the national desired target populations, in reality, there may be some regions or school types that cannot be included. All students in the desired populations who are not included in the defined populations are referred to as the excluded populations.

TIMSS participants were expected to ensure that the national defined populations included at least 95 percent of the national desired populations of students. Exclusions (which had to be kept to a minimum) could occur at the school level, within the sampled schools, or both. Although countries were expected to do everything possible to maximize coverage of the national desired populations, *school-level exclusions* sometimes were necessary. Keeping within the 95 percent limit, school-level exclusions from the sampling frame could be for the following reasons:

- Schools were geographically remote.
- They had very few students.
- The curriculum or structure at the school was different from the mainstream education system.
- Schools were specifically for students with special needs.

The difference between these school-level exclusions and those at the previous level is that these schools were included as part of the school sampling frame (i.e., the list of schools to be sampled). They then were eliminated on an individual basis if it was not feasible to include them in the testing.

In many education systems, students with special educational needs are included in ordinary classes or grouped together in special classes within ordinary schools. Due to this fact, another level of exclusions is necessary to reach an effective target population—the population of students who ultimately will be tested. These are called *within-school exclusions* and pertain to students who are unable to be tested for a particular reason but are part of a regular classroom or part of an in-scope school. There are three types of within-school exclusions, which are explained below:

- **Students with intellectual disabilities**. These are students who are considered, in the professional opinion of the school principal or by other qualified staff members, to be intellectually disabled or who have been tested psychologically as such. This includes students who are emotionally or mentally unable to follow even the general instructions of the test. Students should not be excluded solely because of poor academic performance or normal disciplinary problems.

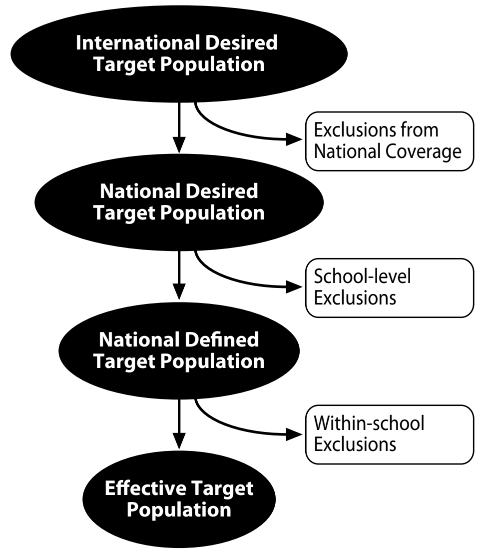
- **Students with functional disabilities.** These are students who are physically disabled in such a way that they cannot perform in the TIMSS testing situation. Functionally disabled students who are able to perform should be included in the testing.
- **Non-native language speakers**. These are students who are unable to read or speak the language(s) of the test and would be unable to overcome the language barrier of the test. Typically, a student who has received less than 1 year of instruction in the language(s) of the test should be excluded, but this definition may need to be adapted in different countries.

Students eligible for within-school exclusion were identified by staff at the schools and still could be administered the test if the school did not want the student to feel out of place during the assessment (though the data from these students were not included in any analyses). Again, it was important to ensure that these populations were as close to the national desired target populations as possible.

If combined school-level and within-school exclusions exceeded 5 percent of the national desired target population, results were annotated in the TIMSS 2007 international reports (Martin, Mullis, & Foy, 2008; Mullis, Martin, & Foy, 2008). Target population coverage and exclusion rates are displayed for each country in Chapter 9. Descriptions of each country's school-level and within-school exclusions can be found in Appendix B.

In any study that utilizes sampling, the population that ultimately participates usually differs slightly from the target population, with some portion of the target population being excluded from the study. A major objective of the TIMSS sampling strategy was to ensure that the effective target population, the population actually sampled by TIMSS, was as close as possible to the international desired population, and to document clearly all excluded populations. Exhibit 5.1 illustrates the relationship between successively more refined definitions of the target population and the excluded populations at each stage.

Exhibit 5.1 Relationship Between the Desired Populations and Exclusions



5.3 Sampling Frames and Survey Units

Once the survey populations were defined, the next step involved building the sampling frames in which all sampling units (grade 4 and/or grade 8 students) within the national defined target populations have a known probability of being sampled. In TIMSS 2007, however, it is important to note that in addition to gathering data on sampled students, a large amount of information also was gathered about their classes and schools, which required other types of sampling units. The intrinsic, hierarchical nature of these nested units necessitated the creation of a sampling frame by stages. Therefore, a two-stage stratified cluster sample design was used, with schools as the first stage and intact classes as the second stage. Because of its large population sizes, it was necessary to include a preliminary sampling stage in the Russian Federation, where regions were sampled first and then schools. Singapore also had a third sampling stage, where students were sampled within classes.

5.3.1 First Stage Sampling Units: Schools

In order to draw school samples that are representative of the student populations, NRCs were asked to provide vital information about all schools (or schools and regions in the Russian Federation) where fourth and eighth grade students could be tested. The following data were required for each school:

- *Measure of size* (MOS): for example, the student enrollment in the target grade, the average student enrollment per grade, the number of classrooms in the target grade, or the total student enrollment in the school.
- *Minimum cluster size* (MCS): the expected number of sampled students per class was required if the number of classrooms in the target grade couldn't be provided. This was calculated as the ratio of the total number of students to the total number of classes for schools having more than one class in the target grade.
- *Variables*: any variables describing school characteristics used for stratification purposes, such as type of school, degree of urbanization codes, or sex of students served by the school.
- *The school sampling probability and status*: information on whether or not that school already was sampled for a study other than TIMSS when overlapping control was required between TIMSS 2007 and other international studies.

In the Russian Federation, a MOS of the regions (preliminary sampling stage) also was required.

5.3.2 Second Stage Sampling Units: Classes

Given the nested nature of the sampling units in TIMSS, listing all classes (along with the class sizes) within sampled schools that agreed to participate in the study was the only requirement for building the class sampling frame. This list included all regular classes, as well as any types of special education classes. Note that within sampled classes, all students were listed. All TIMSS 2007 participating countries had classes as their last stage sampling units except for Singapore, where in addition to classes, students within classes also were sampled through a third sampling stage.

5.4 Sample Selection Method

The student sampling selection method used in TIMSS 2007 is a classic approach that can be found in most sampling textbooks (e.g., Cochran, 1977). The method usually is referred to as a systematic, two-stage *probability proportional-to-size* (PPS) sampling technique, where schools are first sampled and then classes within sampled (and participating) schools. This sampling method is a natural match with the hierarchical nature of the sampling units described above, with classes of students nested within schools. Stratification at the school level was used to complete this technique. Even if a country had a list from which students could be selected directly, this sampling technique, where schools are first sampled and then classes within sampled (and participating) schools, was used for all TIMSS 2007 countries. The only exceptions to this rule were the Russian Federation and Singapore, as mentioned above, which had a three-stage sampling design.

5.4.1 School Stratification

School stratification is the grouping of schools into smaller sampling frames according to information found on the initial sampling frame prior to sampling and may be employed to improve the efficiency of the sample design, to sample sections of the population at different rates, or to ensure adequate representation of specific groups in the sample. School stratification by itself can take two forms: explicit or implicit.

Explicit stratification physically creates smaller sampling frames from which samples of schools and classes ultimately will be drawn. In TIMSS, this type of stratification is used when the usual proportional allocation (i.e., students in certain regions or types of schools are represented in the sample in proportion to their distribution in the population) may not result in adequate representation of some groups of interest in the sample. For example, if a country wanted to make generalizations regarding the science achievement of private sector students, the sampling frame could be split into two strata—public and private sector schools. The sample of schools then could be allocated between the two strata to achieve the desired level of precision in each. In most countries in TIMSS 2007, the school sample allocation among strata was proportional to the number of students found in each stratum. However, it should be noted that even without any stratification, the TIMSS samples represented the different groups found in the population, on average.

Implicit stratification only requires that the school sampling frame be sorted according to some variable(s) prior to sampling and can be nested within explicit stratification. By combining the sorting of the frame with the TIMSS 2007 sampling technique, it is possible to get a sample where students (not schools) are in the same proportions as those found at the population level. When schools from the same implicit stratum tend to have similar behavior, in terms of mathematics and science achievement, implicit stratification also will produce more reliable estimates.

In the basic TIMSS 2007 sample design, all schools in the sampling frame for a country were sorted according to some MOS (see section 5.3.1). If implicit stratification was used, then the sorting by MOS was done within each stratum using a serpentine approach—high to low for the first stratum, followed by low to high for the next, etc. (see the example in Exhibit 5.2).

Exhibit 5.2 MOS Sort Order for Implicit Strata[1]

Implicit Stratum	Sort Order of MOS
1. Rural–Public	High to Low
2. Rural–Private	Low to High
3. Urban–Public	High to Low
4. Urban–Private	Low to High

This way of sorting sampling schools optimizes the chance of choosing a replacement school (see the next section), with a MOS close to that of the originally sampled school it is meant to replace.

5.4.2 Sampling Schools

Schools were sampled using systematic, random sampling with probability proportional to their measures of size. For example, if school A had a MOS value that was twice as large as school B, then School A had twice the chance of being in the sample compared to school B. In the Russian Federation, regions and then schools within sampled regions were sampled following this approach.

To implement the school sampling, schools in each explicit stratum were sorted in order by the implicit stratification variables and within these by the MOS. The measures of size are accumulated from school to school, and a running total, the cumulative measure of size, is recorded next to each school. The cumulative MOS is an indicator of the size of the population of students. Dividing the cumulative MOS by the number of schools to sample gives the sampling interval. In the Russian Federation, the same approach was used to implement the sample of regions. However, no stratification variable was used at the region level.

In order to avoid school sample overlap between TIMSS and another international study (e.g., PISA), where the other study had their sample of schools sampled first, it was necessary to modify the TIMSS school MOS

1 Please refer to the *TIMSS 2007 School Sampling Manual* (TIMSS & PIRLS International Study Center, 2005).

prior to sampling. The technique used for TIMSS is explained below and produced a quasi-PPS school sampling approach. It is a variant of the method originally proposed by Kish and Scott (1971).

Let P_{i1} be the probability of selection of the i^{th} school in sample 1 (already selected prior to TIMSS sampling), and let P_{i2} be its desired PPS probability of selection in TIMSS based on its TIMSS MOS. The i^{th} school in TIMSS with probability P_{i2}' was then selected as follows:

If the i^{th} school was already sampled for the other study,

$$P_{i2}' = Max\left[0, \left(P_{i1} + P_{i2} + 1\right) / P_{i1}\right]$$

If the i^{th} school was not already sampled for the other study,

$$P_{i2}' = Min\left[1, P_{i2} / \left(1 - P_{i1}\right)\right]$$

It is possible to show that over all possible samples, the unconditional probability of selection of the i^{th} school in TIMSS 2007 is P_{i2}. Furthermore, if all of the P_{i1} and P_{i2} are less than 0.5, no school can be sampled twice. However under this approach, the sum over all P_{i2} for a previous given sample (sample 1) is slightly different than n, the desired school sample size for TIMSS. This means that under this approach, there is no control of the sample size even if it is known that it will be n, on average. To get around this problem, an adjustment was done to the P_{i2}' to make them summed to the desired school sample size. This adjustment is given by the following:

$$\tilde{P}_{i2} = \frac{n}{\left(\sum_i P_{i2}'\right)} P_{i2}'$$

With this adjustment, it then was possible to derive a temporary MOS (given by $MOS_i' = \sum MOS_i \bullet \tilde{P}_{i2}' / n$) for each school and use a PPS sampling technique to select the TIMSS sample of schools. Although under this approach, the unconditional probability of the selection of the i^{th} school is not exactly P_{i2}, it is P_{i2} that was used to derive the student weights for TIMSS 2007 (see Chapter 9).

There were three countries and one benchmarking participant that requested control sampling overlap between studies. These are England, the Netherlands, Scotland, and the Canadian province of Alberta.

With systematic PPS sampling, it is possible for a large sampling unit to be selected more than once if its size is greater than the sampling interval. To avoid this situation, all such units were automatically selected by changing each one's MOS to the sampling interval of the associated explicit stratum.

Some schools have so few students that their selection using probability proportional to their size (MOS) becomes problematic. Since the selection of these schools depends on their size, a difference between the number of expected students when drawing the sample and the number of students actually found in the field can contribute substantially to the sampling error. To lessen the impact of this eventuality, any schools with fewer expected students than the average minimum cluster size (MCS) for the explicit stratum were sampled with equal probabilities. For example, if the MCS was 30 students and there were 28 schools with less than 30 students for a total of 476 students, the MOS of these small schools was changed to 476/28 = 17. By doing this, the overall size of the explicit stratum stayed the same, but all small schools had an equal chance of being selected.

The MCS also was used to define very small schools. Whenever a school had an expected number of students less than one quarter of the average MCS, the school was labeled as a very small school. These schools could be excluded, as long as they did not exceed 2 percent of the national desired target population and the overall exclusion rate did not exceed 5 percent.

5.4.3 Replacement Schools

Ideally, response rates always should be 100 percent, and although TIMSS 2007 participants worked hard to achieve this goal, it was anticipated that a 100 percent participation rate would not be possible in all countries. To avoid sample size losses, the TIMSS sampling plan identified, *a priori*, replacement schools for each sampled school. Therefore, if an originally selected school refused to participate in the study, it was possible to replace it with a school that already was identified prior to school sampling. Each originally selected school had up to two pre-assigned replacement schools. In general, the school immediately following the originally selected school on the ordered school sampling frame and the one immediately preceding it were designated as replacement schools. Replacement schools always belong to the same explicit stratum, although they could come from different implicit strata if the originally selected school was either the first or last school of an implicit stratum.

The main objective for having replacement schools in TIMSS 2007 was to ensure adequate sample sizes for analysis of subpopulation differences.

Although the use of replacement schools did not eliminate the risk of bias due to nonresponse, employing implicit stratification and ordering the school sampling frame by size increased the chances that any sampled school's replacements would have similar characteristics. This approach maintains the desired sample size while restricting replacement schools to strata where nonresponse occurred. Since the school frame is ordered by school size, replacement schools also tended to be of the same size as the school they meant to replace. For the field test, replacement schools were used to make sure sample sizes were large enough to validate new items, and no more than one replacement school was assigned per originally selected school.

5.4.4 Sampling Classes

For all participants in TIMSS 2007 except Singapore,[2] intact student classes were the second and final sampling stage, with no student subsampling. This means that all students within sampled classes participated in TIMSS 2007, with the exception of excluded students and students absent the day of the assessment. Classes were selected with equal probability of selection using systematic random sampling. Within each sampled school, all classes of the target grade were listed, and one or more classes were sampled using a random start (different in each sampled school). This method, combined with the PPS sampling method for schools, results in a self-weighting student sample under the following conditions: a) there is a perfect correlation between the school MOS reported in the sampling frame and the actual school size, b) the same number of classes is selected in each school, and c) the MCS is the same for all schools. Given that these conditions were never totally met, student sampling weights varied somewhat from school to school (see Chapter 9 for details about sampling weights).

Within sampled schools, some classes have so few students that it is unreasonable to go through the sampling process and end up with these small classes. Furthermore, small classes tend to increase the risk of unreliable survey estimates. To avoid these problems, a class smaller than half the specified MCS was combined with another class from the same school prior to class sampling.

2 Two classes per school were selected using PPS sampling in Singapore, and 19 students were sampled within each class.

5.5 Sampling Precision and Sample Size

Because TIMSS is fundamentally a study of mathematics and science achievement among fourth and eighth grade students, the precision of survey estimates of student achievement and characteristics was of primary importance. However, TIMSS also reports extensively on school, teacher, and classroom characteristics, so it is necessary to have sufficiently large samples of schools and classes. The TIMSS standards for sampling precision require that all student samples have an effective sample size of at least 400 students for the main criterion variable, which is mathematics and science achievement. In other words, all student samples should yield sampling errors that are no greater than would be obtained from a simple random sample of 400 students.

Given that sampling error, when using simple random sampling, can be expressed as $SE_{SRS} = S / \sqrt{n}$ where S gives the population standard deviation and n the sample size, a simple random sample of 400 students would yield a 95 percent confidence interval for an estimate of a student-level mean of ±10 percent of its standard deviation ($1.96 \bullet S / \sqrt{400}$). Because the TIMSS achievement scale has a standard deviation of 100 points, this translates into a ±10 points confidence limit (or a standard error estimate of approximately 5 points). Similarly, sample estimates of student-level percentages would have a confidence interval of approximately ±5 percentage points.

Notwithstanding these precision requirements, TIMSS required that all student sample sizes should not be less than 4,000 students. This was necessary to ensure adequate sample sizes for analyses where the student population was broken down into many subgroups. For countries involved in the previous TIMSS cycle in 2003, this minimum student sample size was set to 5,150 students in order to compensate for participaton in the TIMSS 2007 Bridging Study. Furthermore, since TIMSS planned to conduct analyses at the school and classroom level in addition to the student level, all school sample sizes were required to be not less than 150 schools, unless a complete census failed to reach this minimum. Under simple random sampling assumptions, a sample of 150 schools yields a 95 percent confidence interval for an estimate of a school-level mean that is ±16 percent of a standard deviation.

Although the TIMSS sampling precision requirements are such that they would be satisfied by a simple random sample of 400 students, sample designs such as the TIMSS 2007 school-and-class design, typically require much larger student samples to achieve the same level of precision. Because students in the same school and even more so in the same class, tend to be more like each other than like other students in the population, sampling a single class of 30 students will yield less information per student than a random sample of students drawn from across all students in the population. TIMSS uses the intraclass correlation, a statistic indicating how much students in a group are similar on an outcome measure, and a related measure known as the design effect to adjust for this "clustering" effect in planning sample sizes.

For countries taking part in TIMSS for the first time in 2007, the following mathematical formulas were used to estimate how many schools should be sampled to achieve an acceptable level of sampling precision:

$$Var_{PPS} = Deff \bullet Var_{SRS} = \frac{Deff \bullet S^2}{n} \cong \frac{\left[1+\rho(mcs-1)\right] \bullet S^2}{n} \cong \frac{\left[1+\rho(mcs-1)\right] \bullet S^2}{a \bullet mcs}$$

where $Deff$ is a compensation factor for using a sample selection method that differs from a simple random sample (also called design effect), S^2 gives the variance of the population, ρ measures the intraclass correlation between clusters, mcs corresponds to the average number of sampled students per class, and a gives the number of schools to sample. Incorporating the precision requirements described earlier into this equation, which translates into $Var_{PPS} = (0.05)^2 \bullet S^2$, gives the number of schools required as:

(1) $$a = 400 \bullet \frac{\left[1+\rho(mcs-1)\right]}{mcs}$$

For planning purposes, the intraclass correlation coefficient usually was set to 0.3 if no other information was available. For example, with a mcs of 20 students and a ρ of 0.3, equation (1) gives 134 schools.

Equation (1) is a model for determining how many schools were required for the TIMSS 2007 sample under the assumption that the standard error of the criterion variable (student mathematics and science achievement) reflects only sampling variance—the usual situation in sample surveys. However, because of its complex matrix-sampling assessment design, standard errors in TIMSS include an imputation error component in addition to the usual sampling error component (see Chapter 11). To keep the standard error within the prescribed precision limits, the number of schools determined by equation (1) has to be increased, as shown in equation (2):

(2) $$a_{irt} = (400 \bullet 0.5) / mcs$$

Continuing the example for a country with a MCS of 20 students, according to this equation (2), 10 schools would have been added to the 134 schools from equation (1), for a total of 144 schools.

For TIMSS 2007 countries that also had participated in TIMSS 2003, the standard errors computed from the 2003 data were reviewed to ensure that the student samples had been large enough to meet the precision requirements in 2003 and would be sufficiently precise to measure trends to 2007. For the several countries falling somewhat short of the sampling requirements not met in 2003, the school sample size for 2007 was increased using the relation that under similar sampling designs, sampling error is inversely proportional to the square root of the sample size. For example, if the sample size in 2003 yielded a standard error of 7 points for an estimate of a mean, the sample size in 2007 was increased by a factor of 2 to provide a standard error of 5 points ($(7/5)^2 \cong 2$). Intraclass correlation coefficients also were calculated for countries that participated in TIMSS 2003. These coefficients were provided in the *TIMSS 2007 School Sampling Manual* (TIMSS & PIRLS International Study Center, 2005).

5.6 Selecting Field-test Samples

Prior to the main data collection, which was conducted from October–November 2006 in Southern Hemisphere countries and from April–May 2007 in Northern Hemisphere countries, TIMSS 2007 conducted a full-scale field test in April 2006 in all participating countries. The field test sample size was approximately 30 schools in each country. Countries were required to draw their field test samples using the same random sampling procedures that they employed for the main samples. This ensured that field test samples closely approximated the main samples. In an attempt to reduce the burden on schools, the field test and main data collection samples of schools were drawn simultaneously, so that a school could be selected for either the field test or the main data collection, but not both. For example, if 150 schools were needed for the main data collection and another 30 schools were needed for the field test, a larger sample of 180 schools was selected using the sampling method described earlier. A systematic subsample of 30 schools then was selected from the 180 schools and assigned to the field test, leaving 150 schools for data collection.[3]

3 In countries where it was necessary to conduct a census of all schools or where the NRC believed that the sampling frame used to draw the combined sample was not appropriate for the data collection, separate sampling frames were provided for the field test and main data collection. In such situations, no attempt was made to minimize the overlap.

References

Cochran, W.G. (1977). Sampling techniques, New York: John Wiley.

IEA. (2006). *Windows Within-school sampling software (WinW3S)* [Computer software and manual]. Hamburg: IEA Data Processing and Research Center.

Kish, L. & Scott, A. (1971). Retaining units after changing strata and probabilities. *Journal of the American Statistical Association,* 66, 461–470.

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

Mullis, I.V.S., Martin, M.O., & Foy, P. (with Olson, J.F., Preuschoff, C., Erberber, E., Arora, A., & Galia, J.). (2008). *TIMSS 2007 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2005). *TIMSS 2007 school sampling manual.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006a). *TIMSS 2007 survey operations procedures unit 2: Contacting schools and sampling classes for TIMSS 2007 assessment.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006b). *TIMSS 2007 school coordinator manual.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

UNESCO Institute for Statistics. (1999). *Operational manual for ISCED-1997 (international standard classification of education)*. Paris: Author.

Chapter 6

TIMSS 2007 Survey Operations Procedures

Ieva Johansone and Oliver Neuschmidt

6.1 Overview

Operationally, TIMSS represents a considerable challenge, and conducting TIMSS 2007 was an ambitious enterprise in each participating country. The contribution that the National Research Coordinators (NRCs) made was crucial to successful assessment administration. In order to assist the NRCs and synchronize activities, internationally standardized survey operations procedures were developed through a collaborative effort between the TIMSS & PIRLS International Study Center, IEA Secretariat, IEA Data Processing and Research Center (DPC), and Statistics Canada. The TIMSS operations were designed to be flexible enough to meet the needs of individual participants, while meeting the high quality standards of IEA. Other IEA studies, the Progress in International Reading Literacy Study (PIRLS) and previous cycles of the TIMSS study, in particular, were used as initial references and the survey operations were refined based on the TIMSS 2007 field test experience.

Guidelines on survey operations procedures for each stage of the assessment, such as contacting schools and sampling classes, preparing materials for data collection, administering the assessment, scoring the assessment, and creating the data files were provided to NRCs. Procedures for quality control and attaining feedback on survey activities also were provided.

6.2 The Role of the National Research Coordinators

In each country and benchmark participant, a research center, under the direction of the NRC, was responsible for the implementation of TIMSS in that country. The NRC was the contact person for all those involved in TIMSS within the country, as well as the representative of the country at the international level. The NRC was responsible for the national decisions regarding TIMSS and, if necessary, implemented and adapted all the internationally agreed-upon procedures for the national context, with guidance from TIMSS and experts from within the country.

6.3 Documentation and Software

The TIMSS 2007 Survey Operations Procedures were disseminated to the NRCs in six units, each accompanied by additional materials, including more specialized manuals and software packages, as necessary. The units and materials were organized and distributed chronologically according to different stages of the study.

The six units and accompanying manuals and software are listed below.

- *Unit 1, Parts 1 and 2: Conducting the TIMSS 2007 Field Test* (TIMSS & PIRLS International Study Center, 2005b)
- *Unit 2: Contacting Schools and Sampling Classes for TIMSS 2007 Assessment* (TIMSS & PIRLS International Study Center, 2006e)
- *Unit 3: Preparing Materials for the TIMSS 2007 Data Collection* (TIMSS & PIRLS International Study Center, 2006f)
- *Unit 4: Administering the TIMSS 2007 Assessment* (TIMSS & PIRLS International Study Center, 2006g)
- *Unit 5: Scoring the TIMSS 2007 Assessment* (TIMSS & PIRLS International Study Center, 2006h)
- *Unit 6: Creating the TIMSS Data Files* (TIMSS & PIRLS International Study Center, 2006i)
- *School Sampling Manual* (TIMSS & PIRLS International Study Center, 2005a) defined the TIMSS 2007 target populations and sampling goals and described the procedures for the sampling of schools.
- *School Coordinator Manual* (TIMSS & PIRLS International Study Center, 2006c) described the role and responsibilities of the School Coordinator as a main contact person within each participating school. The responsibilities included assisting the national center in the identification of classes, teachers, and students; administering

the test and background questionnaires; and keeping test materials secure and confidential at all times while they are in the school.

- *Test Administrator Manual* (TIMSS & PIRLS International Study Center, 2006j) described the role and responsibilities of the Test Administrator, including distribution of the student test instruments according to the student tracking forms, supervising the testing sessions, ensuring the correct timing of the testing sessions, and recording student participation.
- *International and National Quality Control Monitor Manuals* (TIMSS & PIRLS International Center, 2006a, 2006b) provided quality control monitors (QCMs) with information about TIMSS and described their role and responsibilities in the project. The manuals specified the timelines, actions, and procedures that should be followed in order to carry out the international and national quality assurance programs.
- *Scoring Guides for Constructed-response Items* (TIMSS & PIRLS International Center, 2006d) provided detailed and explicit guides used to score each constructed-response item.
- *Windows®* Data Entry Manager Software (*WinDEM) and Manual* (IEA, 2006c) provided for entering, editing, and verifying the TIMSS 2007 data. Along with the software, countries also received codebooks, which described the properties and the layout of the variables to be entered from each TIMSS 2007 assessment instrument.
- *Windows® Within-school Sampling Software (WinW3S) and Manual* (IEA, 2006d) enabled TIMSS 2007 participants to randomly select classes in each sampled school. The software also was used to track school, teacher, student, and student-teacher linkage information; prepare the survey tracking forms; and assign test instruments to students, including printing labels for all the test booklets and questionnaires.
- *Trend Scoring and Reliability Scoring Software and Manual (TSRS)* (IEA, 2006b) provided to document the scoring reliability from one TIMSS cycle to the next. The student responses included in the trend reliability scoring were scanned and provided on individually prepared CDs for each participated country, along with the TSRS software and it’s manual.

- *Cross-country Scoring and Reliability Software and Manual (CCSRS)* (IEA, 2006a) provided to document the reliability of scoring across countries. The responses were scanned and provided on CDs along with the software and it's manual.

In addition to the software manuals, the IEA DPC held two Data Management Seminars to provide training in the use of the WinW3S and WinDEM software.

6.4 Survey Tracking Forms

TIMSS relied on a series of tracking forms to sample classes, assign booklets and questionnaires, and track the participation status of students and their teachers. They facilitated the data collection and data verification process. They also provided information to compute sampling weights and were used to evaluate the quality of the sampling process.

Most of the tracking forms were created automatically by the WinW3S software, then completed by schools and returned to the national centers. There were six different tracking forms in all, which are listed below.

- **School Tracking Form**. This form was sent to national centers by Statistics Canada (Sampling Form 13). It listed the sampled schools and their replacements and included any school information originally provided to Statistics Canada, such as the school ID, school measure of size (MOS), school name, and school contact information.
- **Class Listing Form**. A separate class listing form was created in WinW3S for each sampled school and sent to the School Coordinators for completion. The School Coordinators listed the eligible fourth and/or eighth grade classes in the participating schools and provided details about the classes, such as the class stream, number of students, and names of mathematics teachers.
- **Class Sampling Form**. This form was created in WinW3S as a result of the class sampling procedure. It indicated which classes had been sampled from each school.
- **Student-Teacher Linkage Form**. This form was created in WinW3S for each sampled class and sent to the School Coordinators for completion. The School Coordinators listed the names of the students and their teachers in the sampled classes; students' dates of birth, sex, and exclusion codes; and linked the students to their teachers.

- **Student Tracking Form**. This form was created in WinW3S and sent to the schools with students' test booklets and questionnaires for completion by the Test Administrators during test administration. The Test Administrators used this form to verify the assignment of test instruments to students and indicate student participation.
- **Teacher Tracking Form**. This form was created in WinW3S and sent to the School Coordinators with the *Teacher Questionnaires*. The School Coordinators used this form to indicate the completion of the *Teacher Questionnaires*.

6.5 Contacting Schools and Sampling Classes

One of the essential, first steps in the TIMSS survey activities was to establish good working relationships with the schools that had been sampled to participate in the study (for more information on all sampling procedures, please refer to Chapter 5). NRCs were responsible for contacting these schools and encouraging them to take part in the assessment, which often involved obtaining support from national or regional educational authorities, depending on the national context.

In cooperation with school principals, national centers identified and trained School Coordinators for all participating schools. The School Coordinator could be a teacher or guidance counselor in the school, although the School Coordinator was not allowed to be a teacher of the students who were sampled and who participated in the study. Alternatively, some national centers appointed one of their own members to fill this role. Often this person was responsible for several schools in an area. Each School Coordinator was provided with a *TIMSS 2007 School Coordinator Manual*, which described their responsibilities in detail and encouraged them to contact the NRC if they had any questions.

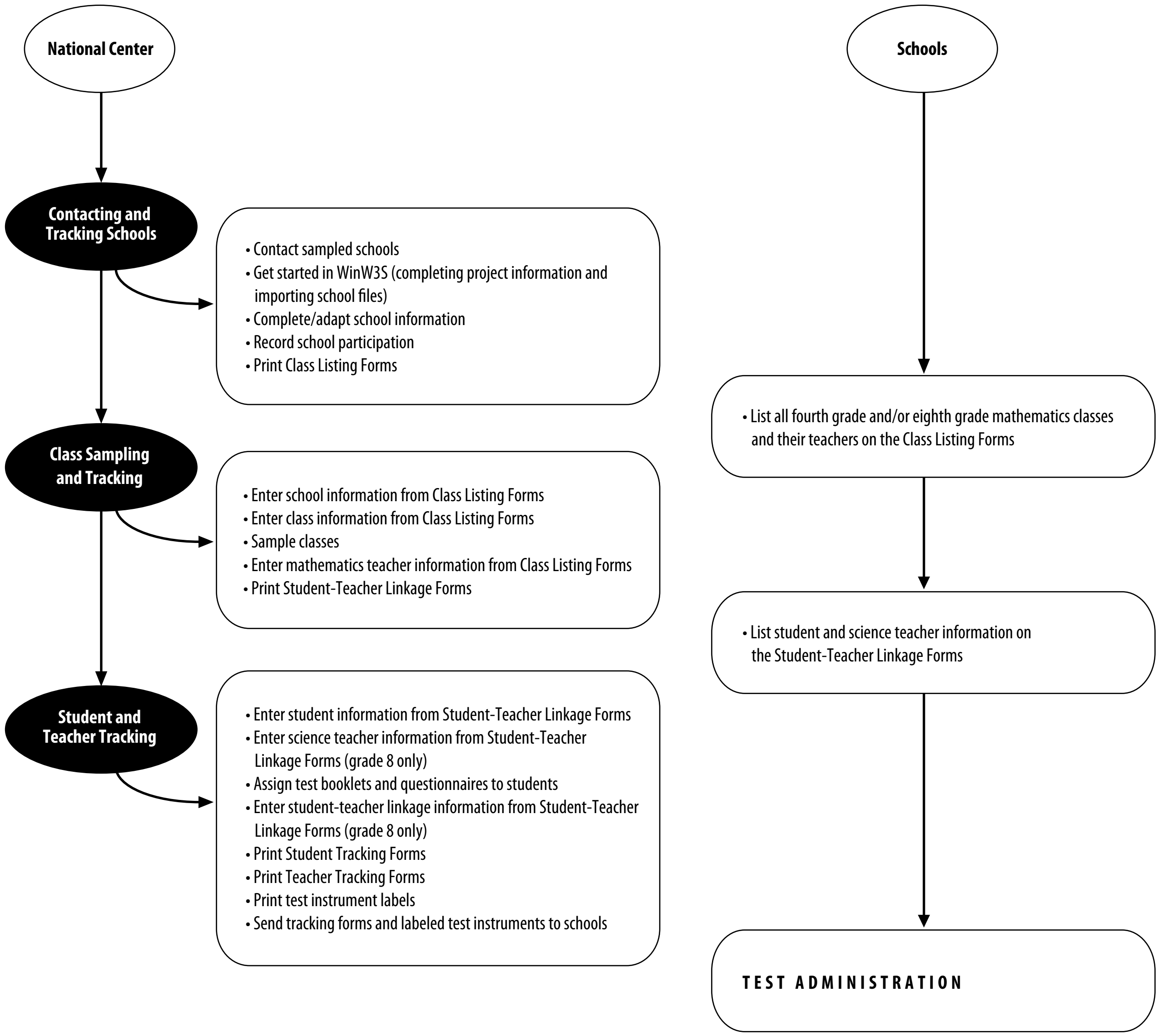
The responsibilities of the School Coordinator included providing the necessary information about their school; coordinating the date, time, and place for testing; identifying and training a Test Administrator; coordinating the completion of the student tracking forms and teacher tracking forms; distributing teacher and school questionnaires; and obtaining parental permission (if necessary). They also ensured that all testing materials were received and kept secure at all times and were returned to the national center after the test administration.

A Class Listing Form was sent to each School Coordinator to provide information on all the eligible fourth grade and/or eighth grade classes in the school. Using this information, the national centers sampled classes within the schools. Intact classes had to be sampled, ensuring that every student in the school was in only one class (course), and no student was in more than one class. Such an organization was necessary for a random sample of classes to result in a representative sample of students. At fourth grade, most countries used the same class for all subjects, including mathematics and science. Accordingly, the fourth grade classroom was the sampling unit. At the eighth grade, students in most countries attended different classes for mathematics and science. For sampling purposes, classrooms usually were defined on the basis of mathematics instruction.

Exhibit 6.1 presents the major activities conducted by the national centers for working with schools to sample classes; track schools, teachers, students, and student-teacher linkage information; and prepare for test administration.

Although all students enrolled in the sampled classes were part of the target population, TIMSS recognized that some student exclusions were necessary because of either some functional or intellectual disability or in cases where there were non-native language speakers. Accordingly, the sampling procedures provided for the exclusion of students with any of several disabilities (for more information on sampling procedures, see Chapter 5). Countries were required to track and account for all excluded students and were cautioned that excluding more than 5 percent of students would lead to their results being annotated in the TIMSS 2007 international reports. It was important that the conditions under which countries excluded students be carefully documented, because the definition of being disabled varied from country to country.

Exhibit 6.1 Procedures for Working with Schools to Prepare for Test Administration



6.5.1 Linking Students to their Teachers and Classes

To track students, teachers, and schools, there was a system in place to assign hierarchical identification codes (IDs). The hierarchical identification numbers that uniquely identified the sampled schools, teachers, and classes were created by the WinW3S software. Exhibit 6.2 represents the hierarchical identification system codes.

Exhibit 6.2 Hierarchical Identification (ID) System

Unit	ID Components	ID Structure	Numeric Example
School	School	CCCC	0001
Class	School + Class within School	CCCCKK	000101
Student	School + Class within School + Student within Class	CCCCKKSS	00010103
Teacher (as an Individual)	School + Teacher within School	CCCCTT	000101
Teacher Link Number	School + Teacher within School + Link Number	CCCCTTLL	00010102

Each mathematics and science teacher of the selected classes (i.e., those listed on the Teacher Tracking Form) was assigned a teacher identification number consisting of the four-digit school number followed by a two-digit teacher number unique within the school. Since a teacher could be teaching both mathematics and science to some or all of the students in a class, it was necessary to have a unique identification number for each teacher linked to a class and a teacher linked to certain students within the class. This was achieved by adding a two-digit link number to the six digits of the teacher identification number and giving a unique eight-digit identification number. This is very important so that during data analyses, each class and student may be linked to a teacher, and student outcomes may be analyzed in relation to teacher-level variables.

6.6 Preparing the Test Instruments for Data Collection

The TIMSS & PIRLS International Study Center provided countries with all the necessary instrument production files, including fonts, style guides, graphic files, and explicit instructions in *TIMSS 2007 Survey Operations Procedures Unit 3* on how to use the materials in order to produce good quality test instruments. The TIMSS achievement booklets and questionnaires were developed using the Adobe®InDesign® layout program.

The overarching goal of the test instrument preparation was to create internationally comparable assessment booklets and background questionnaires that were appropriately adapted for the national context. This began with translating the text of the test instruments from English into the language(s) used in the participating countries.[1] All the translated contents of the test instruments (i.e., item blocks, directions or instructions, cover pages, and questionnaires) were submitted to the IEA for an independent international translation verification, where independent translators provided suggested changes in the texts (for more information on translation and national adaptations of the TIMSS 2007 test instruments, please refer to Chapter 4).

Once the translation verification was completed and any resulting changes implemented into the test instrument production files, the item blocks, cover pages, directions, and instructions had to be distributed throughout the booklets in order to assemble the assessment booklets. Each background questionnaire consisted of a single InDesign file and did not require any assembling.

6.6.1 TIMSS & PIRLS International Study Center Review

Before the test booklets and questionnaires were printed and administered to students, the NRCs were required to submit a print-ready copy of all the test instruments to the TIMSS & PIRLS International Study Center for layout verification and review of national adaptations.

The national test instruments were checked against the international version to identify any deviations. The verification was checked for any discrepancies in pagination, page breaks, item sequence, response options, text formats, graphics, etc.

1 The IEA Secretariat and the TIMSS & PIRLS International Study Center also provided a generic Arabic version of the TIMSS assessment booklets and questionnaires.

The test instruments from the participating countries were not exactly identical due to the changes in text length that often occurred during translation. The international versions, however, were designed with this in mind, and extra space was provided in the margins of the pages to facilitate the use of a longer text and different size paper without extensive changes to the layout of the instruments. All deviations or errors were documented in the layout verification report forms and sent to the NRCs for their consideration. The NRCs were required to comment on whether or not each of the suggested changes was completed, including an explanation if a suggestion was not adapted.

All national adaptations to the international test instruments were documented in the national adaptations forms. During the review, the TIMSS & PIRLS International Study Center checked if the national adaptations in the national background questionnaires influenced the ability to produce internationally comparable data for the affected questions. In some cases, countries had implemented adaptations that were impossible to be recoded later to fit the international data structure. In such cases, suggestions were sent to the NRCs to reconsider their adaptation.

This whole procedure ensured that students experienced the test instruments in the same way, apart from the translation of text.

6.7 Administering the TIMSS 2007 Assessment

Once printed, distributing materials to the schools required careful organization and planning on the part of the NRC. Using labels and the Student Tracking Form produced by WinW3S, each sampled student was assigned one achievement booklet. The test booklets were assigned in a systematic rotation so that each achievement block within the booklets was assigned to an equal number of students. Each student also was assigned a *Student Questionnaire* that was labeled to link it to the achievement booklet. These materials were packaged for each sampled class. In addition, a *Teacher Questionnaire* was assigned and sent for each teacher listed on the Teacher Tracking Form and a *School Questionnaire* for the principal. The packaged materials were sent to the School Coordinator prior to the testing date, who was asked to confirm the receipt of all instruments. The *School Questionnaire* and *Teacher Questionnaire* were then distributed, while the other instruments were kept in a secure room until the testing date.

Each sampled TIMSS class was assigned a Test Administrator whose role was to administer the test along with the *Student Questionnaire*s, according to procedures described in the *Test Administrator Manual*. This person was chosen and trained by the School Coordinator, although, in many cases, the School Coordinator also filled the Test Administrator role. The Test Administrator was responsible for distributing materials to the appropriate students, leading students through the assessment, and timing the sessions accurately. Following the assessment, they administered the *Student Questionnaire*.

The administration of the TIMSS 2007 assessment consisted of two parts. The first part concerned the achievement booklets, which contained two sections. This was followed by the completion of the *Student Questionnaire*. The time allotted for each of these sections was standardized across countries. To complete each part of the achievement test, fourth graders were allowed 36 minutes and eighth graders 45 minutes. There was a required break in between the two parts not exceeding 30 minutes. If a student had completed part 1 or part 2 of the assessment before the allotted time was over, he or she was allowed to review his or her answers or read quietly but was not allowed to leave the testing room. To complete the *Student Questionnaire*, students were given at least 30 minutes and were allowed to continue if extra time was necessary. The Test Administrators were required to document the starting and ending time of each section on the Test Administration Form.

The Test Administrator used the Student Tracking Form to distribute the booklets to the correct students and to document student participation. The School Coordinator used the information on the participation status to calculate the participation rate. If this was below 90 percent in any class, it was the School Coordinator’s responsibility to hold a makeup session for the absent students before returning all of the testing materials to the national center.

The national centers entered the information recorded on the student and teacher tracking forms into WinW3S software.

6.7.1 Quality Control

Considerable effort has been expended in developing standardized materials and procedures so that the data collected in each country for TIMSS will be comparable to the greatest possible extent. In order to further ensure the quality of the TIMSS data, an international quality control program was developed to document data collection activities around the world.

The NRCs were required to nominate an International Quality Control Monitor (QCM) for their country. This person could be a school inspector, a ministry official, or a retired school teacher. He or she had to be fluent in both English and the language(s) spoken in the country. The nomination of a member of the national center, a family member, or a personal friend of the NRC was not allowed.

The QCMs were hired by the IEA and trained by the TIMSS & PIRLS International Study Center. The role and responsibilities of an international QCM also were described in the *International Quality Control Monitor Manual.* The responsibilities included collecting and submitting a number of TIMSS 2007 materials from the national centers to the TIMSS & PIRLS International Study Center. During the test administration, 15 schools per grade tested in each participating country were visited by the QCMs. During their school visits, they noted if any changes were made to the standardized administration script, timing, or procedures. After the TIMSS testing sessions, they interviewed the School Coordinator and/or the Test Administrator about his or her experiences with the TIMSS 2007 assessment. The QCMs also checked whether or not the comments and suggestions made by the international translation verifier had been integrated into the final test instruments.

Additionally, countries were asked to conduct their own quality control procedures in another 10 percent of sampled schools. To assist them, countries were provided with the *National Quality Control Observer Manual*, which could be used to train their observers and modified to suit their national system.

6.8 Scoring the TIMSS 2007 Assessment

The success of assessments containing constructed-response questions depends on the degree to which student responses are scored reliably. Almost half of the TIMSS 2007 assessment items were constructed response, and scoring them in a reliable manner was critical to the quality of the TIMSS 2007 results. This was accomplished through the provision of explicit scoring guides and extensive training in their use, as well as continuous monitoring of the quality of the work.

Two international scoring trainings were held, where the NRCs (or the country representative(s) appointed by the NRC) were trained to score each of the constructed-response items in the TIMSS 2007 assessment. At these trainings, the *TIMSS 2007 Scoring Guides for Constructed-response Items,* which are more thoroughly discussed in Chapter 2, were reviewed and applied to a sample set of student responses that had already been scored. These example papers were actual student answers that came from pilot testing held in several English-speaking countries. They were chosen to represent a range of response types, intended to demonstrate the guides as clearly as possible. Following this, NRCs attempted to apply the scoring guide to a different set of student responses that had not yet been scored. The scores that NRCs gave to these practice papers were shared with the group and any discrepancies discussed. Following the training, NRCs were given a set of the correct scores for these practice papers along with rationales.

NRCs used this information to train their scoring staff on how to apply the scoring guides for constructed-response items. In some cases, NRCs created their own example papers and practice papers from student responses collected in their country.

To prepare for this substantial task, NRCs were provided with suggestions on how to organize, in regards to staff, materials, and procedures, the scoring process. NRCs were encouraged to hire scorers who were attentive to detail and familiar with education, particularly those with a background in mathematics and/or science instruction at the fourth grade and/or eighth grade level. The TIMSS & PIRLS International Study Center also provided guidelines on how to train scorers to accurately and reliably score the constructed-response achievement items.

6.8.1 Documenting Scoring Reliability

In order to demonstrate the quality of the TIMSS data, it was important to document the reliability of the scoring process within countries, across countries, and over time (2003–2007).

To establish the reliability of the scoring within each country, two different scorers independently scored a random sample of 200 responses for each constructed-response item, which corresponded to 100 of each of the 14 test booklets per grade tested. The random sample of test booklets designated to be scored twice was selected by the WinW3S software. The degree of agreement between the scores, assigned by the two scorers, is a measure of the reliability of the scoring process. The scoring procedure recommended by the TIMSS & PIRLS International Study Center blended the scoring of the reliability sample with the normal scoring activity, with both taking place simultaneously in a systematic manner. In collecting the reliability data, Reliability Scoring Sheets were used so that one scorer did not know the scores that the other assigned to the responses.

In order to measure the reliability of the scoring process across countries (cross-country reliability scoring), each country had to have a minimum of two scorers from the TIMSS 2007 scoring team who were able to score student responses written in English. Computing the level of agreement across countries provided information about how consistently the scoring guides were applied from one country to the next. This scoring activity, however, was conducted by participants on the Northern Hemisphere schedule only, since it entailed scoring a set of student responses gathered from the English-speaking countries that participated in TIMSS 2007 on the Southern Hemisphere timeline. The student responses included in the cross-country reliability scoring were scanned by the IEA DPC, stored on CDs, and provided to all countries participating on the Northern Hemisphere timeline, along with the Cross-country Scoring Reliability Software, which was developed by the IEA DPC. The CD also included a manual on how to install and use the software.

The purpose of the trend reliability scoring was to measure how reliable the scoring was from one TIMSS cycle to the next, i.e., from 2003 to 2007. Thus, trend reliability scoring only applied to countries that participated in TIMSS 2003 and submitted their TIMSS 2003 reliability booklets to the IEA DPC to be scanned. Using this approach, scorers for the TIMSS 2007 assessment could score student responses from 2003 and compare their scores to those given in TIMSS 2003. The student responses included in the trend reliability scoring, totaling approximately 10,000 responses per grade tested, were provided on individually prepared CDs for each participating country, along with the software, Trend Scoring and Reliability Software, developed by the IEA DPC. The CD also included a manual on how to install and use the software. At least two different scorers from the TIMSS 2007 scoring team in each country participated in the trend reliability scoring. It was important that the countries that participated in TIMSS 2003 started with the trend reliability scoring prior to all the other TIMSS 2007 scoring activities. The results then were used also as a diagnostic tool to indicate the need for further training. Two scorers independently scored about half of the items provided on the trend reliability scoring CD. Then, NRCs were asked to analyze the results of the agreement between the two scorers, as well as between each of their TIMSS 2007 scorers and the scores that were awarded in 2003. If agreement on any comparison was below 85 percent, retraining of the scorers was required. If agreement was 85 percent or above, countries could continue with the trend reliability scoring and all the other scoring activities.

6.9 Creating the TIMSS 2007 Data Files

As described earlier in this chapter, the IEA DPC provided a Windows-based program called WinDEM to accommodate data entry and data verification. Detailed information on installing and using the program was provided in the *Windows Data Entry Manager Software Manual* accompanying the software. The program worked in conjunction with WinW3S software so that it was not necessary to re-enter tracking information that had been recorded into WinW3S. WinDEM primarily was used for the entry of data from test booklets and questionnaires. The software also offered data and file management capabilities, a convenient checking and editing mechanism, interactive error detection, and reporting and quality-control procedures.

Trainings in using the WinW3S and WinDEM software and in operational procedures of data management were provided to NRCs and/or their data managers by the IEA DPC at various stages of the project, including an extensive 4-day training seminar before the field test and before the TIMSS 2007 data collection.

One of the very important benefits of using WinDEM was that it incorporated the international codebooks describing all variables and their characteristics, thus ensuring that the data files that were produced fulfilled the TIMSS 2007 rules and standards for data entry. There was one codebook for each of the background questionnaires, one for the test booklets, and one for the Reliability Scoring Sheets. Data files for entering the TIMSS data were created based on these codebooks. However, the codebooks had to match the national instruments exactly so that the answers of the respondents could be entered properly. Therefore, any adaptations done to the international instruments also required adaptations of the international codebooks.

The adapted national codebooks then were used for creating the TIMSS 2007 data files within each participating country. Data from the background questionnaires, achievement booklets, and Reliability Scoring Sheets were recorded into WinDEM data files as follows:
For fourth grade:

- *School background file* contained responses from the *School Questionnaire.*
- *Teacher background file* contained responses from the *Teacher Questionnaire.*
- *Student background file* contained responses recorded from the *Student Questionnaire.*
- *Student achievement file* contained responses from the test booklets.
- *Reliability scoring file* contained codes from the constructed-response Reliability Scoring Sheets.

For eighth grade:

- *School background file* contained responses from the *School Questionnaire.*
- *Mathematics teacher background file* contained responses from the *Mathematics Teacher Questionnaire.*
- *Science teacher background file* contained responses from the *Science Teacher Questionnaire.*
- *Student background file* contained responses recorded from the *Student Questionnaire.*
- *Student achievement file* contained responses from the test booklets.
- *Reliability scoring file* contained codes from the constructed-response Reliability Scoring Sheets.

Quality control throughout the data entry process was essential in maintaining accurate data. Therefore, NRCs were responsible for performing periodic reliability checks during the data entry and for applying a series of data verification checks provided by WinDEM software prior to submitting the data files to the IEA DPC. As part of this process, NRCs required their data-entry staff to double enter at least 5 percent of each instrument type to ensure reliability of the data entry process. An error rate of 1 percent or less was acceptable for the background files. An error rate of 0.1 percent or less was required for the student achievement files and the reliability scoring files. If the required agreement was not reached, retraining of the key punchers was required.

Additionally, the data verification module of WinDEM identified a range of problems, such as inconsistencies of identification codes and out-of-range or otherwise invalid codes. WinDEM software also allows for verification of the integrity of the linkage between the students, teachers, and schools entered into the WinDEM data files and tracking of information for those specified in WinW3S.

When all data files had passed the WinDEM quality control checks, they were submitted to the IEA DPC, along with data documentation for further checking and processing. For information on data processing at the IEA DPC, please refer to Chapter 8.

6.9.1 Online Data Collection for Curriculum Questionnaires and Survey Activities Questionnaires

For the first time, in TIMSS 2007, the *Curriculum Questionnaire* and *Survey Activities Questionnaire*[2] were administered online. The online survey system for the questionnaires was developed by the IEA DPC and hosted on its server.

There were many benefits to administering questionnaires via the Internet for a large-scale assessment such as TIMSS. Online data collection saves money and time for printing and distributing the materials. Furthermore, the online administration facilitates data entry, cleaning, and analysis. The responses are directly stored in an MS SQL Server.

Since the *Curriculum Questionnaires* and the *Survey Activities Questionnaires* did not require any national adaptations and were completed in English, unlike the other TIMSS 2007 background questionnaires, they were best suited for the online data collection process.

The purpose of the *Curriculum Questionnaires* was to collect information about the national mathematics and science curriculum at the fourth- and eighth-grade levels. NRCs were asked to complete the questionnaires drawing on the experience of curriculum specialists and educators.

2 The *Survey Activities Questionnaire* replaced the Survey Activities Report, which served the purpose of attaining feedback about survey operations from NRCs for the previous TIMSS cycles.

The purpose of the *Survey Activities Questionnaire* (one per grade tested) was to gather opinions and information about the strength and weaknesses of the TIMSS 2007 assessment materials (e.g., test instruments, manuals, scoring guides, and software) and countries' experiences with the TIMSS 2007 survey operations procedures. NRCs were asked to complete these questionnaires with assistance of their data managers and the rest of the national center staff. The information will be used to improve the quality of the survey activities and materials for future TIMSS cycles.

NRCs were able to familiarize themselves with the content of the online questionnaires prior to completing them online. The TIMSS & PIRLS International Study Center provided countries with PDF versions of the online questionnaires at least 3 months before they were available for completion online.

The individual login information for accessing each questionnaire was sent to the NRCs with Internet links pointing to the location of the online questionnaires. Before submitting the responses to the IEA DPC, NRCs could go back and change their answers if necessary.

6.10 TIMSS 2007 Bridging Study

As a part of the TIMSS 2007 bridging study, countries that had participated in TIMSS 2003 administered four additional booklets per grade. The bridge booklets, labeled B1, B2, B3, and B4, were booklets 5, 6, 11, and 12, respectively, from the TIMSS 2003 assessment.

Operationally, that meant that these countries required additional sample of at least 1,150 students per grade, and the bridge booklets had to be incorporated in all survey operations, including production of the survey instruments, assignment of booklets to students, the scoring of the constructed-response items, and data entry.

The countries were required to use the bridge booklets as they were administered in TIMSS 2003. However, the TIMSS & PIRLS International Study Center provided new covers for the bridge booklets. The procedure for replacing the covers was described in *TIMSS 2007 Survey Operations Procedures Unit 3*. After replacing the covers, countries were required to send these booklets for layout verification, along with their TIMSS 2007 survey instruments. The TIMSS & PIRLS International Study Center reviewed the bridge booklets by comparing them to the booklets administered in 2003.

The assignment of bridge booklets to students was incorporated in the WinW3S software and automated. Specific instructions for students completing any of the bridge booklets were provided in the *Test Administrator Manual*. In order to ease the procedure of scoring the constructed-response items, separate scoring guides for the bridging study were provided by the TIMSS & PIRLS International Study Center. Finally, entering data from the bridge booklets was fully incorporated in the WinDEM software.

6.11 TIMSS 2007 Field Test

The TIMSS 2007 field test was a smaller administration of the TIMSS 2007 assessment, involving approximately 1,400 students per grade tested in each participating country.

The field test was crucial to the development of the instruments for the TIMSS 2007 assessment, particularly the achievement tests. As part of the dissemination of the TIMSS 2003 results, about half of the achievement items were released into the public domain. Items that replaced the released ones were tried out in the field test in order to investigate the psychometric characteristics of the achievement items and make well-informed decisions about the best replacements. The field test involved 14 newly developed item blocks (7 for science and 7 for mathematics), which corresponds to 7 test booklets.

The field test also served the purpose of testing the TIMSS 2007 survey operations procedures in order to avoid any possible problems during the TIMSS 2007 data collection. An essential step towards achieving this goal was to conduct a full-scale field test of all instruments and operational procedures under conditions approximating, as closely as possible, those of the data collection. Additionally, this allowed the NRCs and their staff to become acquainted with the activities and refine their national operations and provide feedback that was used to improve the procedures for the data collection. The field test resulted in some small modifications to survey operations procedures and contributed significantly to the successful execution of TIMSS 2007. The field test was conducted from March–April 2006.

References

IEA. (2006a). *Cross-country scoring reliability software* [Computer software and manual]. Hamburg: IEA Data Processing and Research Center.

IEA. (2006b). *Trend scoring reliability software* [Computer software and manual]. Hamburg: IEA Data Processing and Research Center.

IEA. (2006c). *Windows data entry manager software (WinDEM)* [Computer software and manual]. Hamburg: IEA Data Processing and Research Center,

IEA. (2006d).*Windows within-school sampling software (WinW3S)* [Computer software and manual]. Hamburg: IEA Data Processing and Research Center.

TIMSS & PIRLS International Study Center. (2005a). *TIMSS 2007 school sampling manual*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2005b). *TIMSS 2007 survey operations procedures unit 1, parts 1 and 2: Conducting the* TIMSS *2007 field test*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006a). *TIMSS 2007 international quality control monitor manual*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006b). *TIMSS 2007 national quality control monitor manual*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006c). *TIMSS 2007 school coordinator manual*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006d). *TIMSS 2007 scoring guides for constructed-response items*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006e). *TIMSS 2007 survey operations procedures unit 2: Contacting schools and sampling classes for TIMSS 2007 assessment*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006f). *TIMSS 2007 survey operations procedures unit 3: Preparing materials for the TIMSS 2007 data collection*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006g). *TIMSS 2007 survey operations procedures unit 4: Administering the TIMSS 2007 assessment*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006h). *TIMSS 2007 survey operations procedures unit 5: Scoring the TIMSS 2007 assessment*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006i). *TIMSS 2007 survey operations procedures unit 6: Creating the TIMSS data files*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006j). *TIMSS 2007 test administrator manual.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Chapter 7

Quality Assurance in the TIMSS 2007 Data Collection

Ieva Johansone

7.1 Overview

Considerable effort was made in developing standardized materials and survey operations procedures (see Chapter 6 for more information) in order to ensure the quality of the TIMSS data and make valid comparisons of student achievement across and also within the participating countries. In addition, the TIMSS & PIRLS International Study Center developed an ambitious international quality control program to document data collection activities in the participating countries. To implement this program, the IEA Secretariat, in cooperation with national centers, nominated an international Quality Control Monitor (QCM) in each of the participating countries.

The TIMSS & PIRLS International Study Center conducted an extensive 2-day QCM training on completing the tasks of the TIMSS 2007 international quality control program. The QCMs were introduced to the TIMSS 2007 survey operations procedures, and the design of the test booklets and background questionnaires. During the training, each QCM received the necessary materials for completing their tasks. The materials included a copy of the *TIMSS 2007 International Quality Control Monitor Manual* and *TIMSS 2007 National Quality Control Monitor Manual* (TIMSS & PIRLS International Study Center, 2006a, 2006b), Classroom Observation Record, *TIMSS 2007 Survey Operations Procedures Units 2–4 Manuals (*TIMSS & PIRLS International Study Center, 2006d, 2006e, 2006f), *TIMSS 2007 School Coordinator Manual (*TIMSS & PIRLS International Study Center, 2006c), and *TIMSS 2007 Test Administrator Manual* (TIMSS & PIRLS International Study Center, 2006g).

The major task of the international QCMs was to conduct site visits to a random sample of 15 schools per target grade during test administration in their countries. Where necessary, the QCMs were permitted to recruit one or more assistants in order to efficiently cover the territory and testing timetable. A total of 248 international QCMs and their assistants were trained across the 62 countries (including the four provinces of Canada) that participated in TIMSS 2007. Altogether, these monitors observed 1,371 testing sessions, including 597 for grade 4 and 774 for grade 8. The results of the QCM observations are reported in Section 7.2.

In addition to the international and national quality control programs, the National Research Coordinators (NRCs) were asked to complete the *Survey Activities Questionnaire* (one per grade) about their experiences with the TIMSS 2007 survey operations procedures and the quality of the assessment materials. The main purpose of the questionnaire was to gather opinions and information to be used to further improve the quality of the survey activities and materials for future TIMSS cycles. Section 7.3 summarizes information that reflects the quality of the TIMSS 2007 survey materials and procedures within the participating countries.

7.2 Quality Control Observations of the TIMSS 2007 Test Administration

For each testing session observed, QCMs completed the TIMSS 2007 Classroom Observation Record. The observation record was organized into four sections, listed below, in order to facilitate accurate recording of the test administration's major activities.

Section A: Preliminary Activities of the Test Administrator

Section B: Test Administration Activities

Section C: Summary Observations

Section D: Interview with the School Coordinator and/or Test Administrator

7.2.1 Preliminary Activities of the Test Administrator

Section A of the Classroom Observation Record addressed the extent to which the Test Administrator had prepared for the testing session. QCMs were asked to note the following activities of the Test Administrator: checking the testing materials, reading the administration script, organizing space for the session, and arranging for the necessary equipment.

Exhibits 7.1 and 7.2 summarize the results for Section A. In nearly all testing sessions, Test Administrators observed the proper preparatory procedures. For those few deviations that occurred, QCMs provided reasonable explanations for all the discrepancies. For example, QCMs noted that the main reason some information on student test instruments did not correspond to the Student Tracking Form was that a student had left school and/or a new student had joined the class, which had not been documented on the list. The Test Administrators who did not have a watch with a second hand had a cell phone watch or a classroom clock available to monitor the time remaining in the test sessions. In general, QCMs observed no procedural deviations in test preparations that were severe enough to jeopardize the integrity of the test administration.

Exhibit 7.1 Percentages of QCM Responses for Preliminary Activities of the Test Administrator– Fourth Grade

Question	Yes (%)	No (%)	Not Answered (%)
Had the Test Administrator verified adequate supplies of the test booklets?	97	2	1
Had the Test Administrator familiarized himself or herself with the test administration script prior to the testing?	94	4	2
Did the student identification information on the test booklets and student questionnaires correspond with the Student Tracking Form?	96	3	1
Was there adequate seating space for the students to work without distractions?	94	4	2
Was there adequate room for the Test Administrator to move around during the testing to ensure that student were following directions correctly?	98	2	0
Did the Test Administrator have a watch with a seconds hand (or stopwatch) for accurately timing the testing sessions?	96	3	1

Exhibit 7.2 Percentages of QCM Responses for Preliminary Activities of the Test Administrator– Eighth Grade

Question	Yes (%)	No (%)	Not Answered (%)
Had the Test Administrator verified adequate supplies of the test booklets?	97	2	1
Had the Test Administrator familiarized himself or herself with the test administration script prior to the testing?	96	4	0
Did the student identification information on the test booklets and student questionnaires correspond with the Student Tracking Form?	96	3	1
Was there adequate seating space for the students to work without distractions?	97	3	0
Was there adequate room for the Test Administrator to move around during the testing to ensure that student were following directions correctly?	97	3	0
Did the Test Administrator have a watch with a seconds hand (or stopwatch) for accurately timing the testing sessions?	97	3	0

7.2.2 Assessment Session Activities

Section B of the Classroom Observation Record addressed the activities that took place during the actual assessment session and the administration of the *Student Questionnaire*. The achievement test was administered in two parts with a short break in between. The activities, such as following the Test Administrator script, distributing and collecting test booklets, and making announcements during the testing sessions were reported by the QCMs and are presented in Exhibits 7.3 through 7.8.

Activities during the first part of the testing session are presented in Exhibits 7.3 and 7.4. One of the most important standardizations for the assessment administration was the fact that the test administrator's script was followed in all participating countries. QCMs reported that in almost all of their observations, the Test Administrators followed their script exactly when preparing students, distributing test materials, and reading directions and examples. Of the changes that were made, the majority were considered minor. Changes made to the script were most frequently additions, rather than revisions or deletions. In a very small percentage of all the sessions (5% for grade 4 and 4% for grade 8), the total testing time for Part 1 was not equal to the time allowed. In most sessions, this was because students had completed Part 1 before the allotted time had elapsed. When the allotted time was over, the Test Administrator instructed students to close their test booklets and announced the break to be followed by Part 2 of the test. In 95 percent of the cases for fourth grade and in 96 percent of the cases for eighth grade, the Test Administrator made sure that students stopped working immediately. In most sessions, the room was then either secured or supervised during the break.

Exhibit 7.3 Percentages of QCM Responses for Assessment Session Part 1 – Fourth Grade

Question	Yes (%)	No (%)	Not Answered (%)
Did the test administrator follow the test administrator's script exactly in each of the following tasks?			
Preparing the students	79	18 *(Minor changes)* 2 *(Major changes)*	1
Distributing the materials	90	7 *(Minor changes)* 2 *(Major changes)*	1
Reading the directions	71	26 *(Minor changes)* 2 *(Major changes)*	1
Reading the examples	79	18 *(Minor changes)* 2 *(Major changes)*	1
If the Test Administrator made changes to the script, how would you describe them?			
Additions	26	9	5 *(Not Answered)* 60 *(Not Applicable)*
Revisions	17	15	8 *(Not Answered)* 60 *(Not Applicable)*
Deletions	8	22	10 *(Not Answered)* 60 *(Not Applicable)*
Did the Test Administrator distribute the test booklets according to the booklet assignment on the Student Tracking Form?	97	2	1
Did the Test Administrator record attendance correctly on the Student Tracking Form?	94	1	5
Did the total testing time for Part 1 equal the time allowed?	95	5	0
Did the Test Administrator announce "you have 10 minutes left" prior to the end of Part 1?	93	7	0
Were there any other time remaining announcements made during Part 1?	20	79	1
At the end of Part 1, did the Test Administrator make sure all students had closed their booklets?	95	4	1
Was the total time for the break equal to or less than 30 minutes?	93	3	1
Were the booklets left unattended or unsecured during the break?	7	92	1

Exhibit 7.4 Percentages of QCM Responses for Assessment Session Part 1 – Eighth Grade

Question	Yes (%)	No (%)	Not Answered (%)
Did the test administrator follow the test administrator's script exactly in each of the following tasks?			
Preparing the students	84	13 *(Minor changes)* 2 *(Major changes)*	1
Distributing the materials	91	6 *(Minor changes)* 1 *(Major changes)*	2
Reading the directions	77	18 *(Minor changes)* 3 *(Major changes)*	2
Reading the examples	83	12 *(Minor changes)* 3 *(Major changes)*	2
If the Test Administrator made changes to the script, how would you describe them?			
Additions	14	10	5 *(Not Answered)* 71 *(Not Applicable)*
Revisions	13	11	5 *(Not Answered)* 71 *(Not Applicable)*
Deletions	8	14	7 *(Not Answered)* 71 *(Not Applicable)*
Did the Test Administrator distribute the test booklets according to the booklet assignment on the Student Tracking Form?	98	1	1
Did the Test Administrator record attendance correctly on the Student Tracking Form?	95	2	3
Did the total testing time for Part 1 equal the time allowed?	95	4	1
Did the Test Administrator announce "you have 10 minutes left" prior to the end of Part 1?	94	6	0
Were there any other time remaining announcements made during Part 1?	17	82	1
At the end of Part 1, did the Test Administrator make sure all students had closed their booklets?	96	3	1
Was the total time for the break equal to or less than 30 minutes?	96	4	0
Were the booklets left unattended or unsecured during the break?	4	95	1

Exhibits 7.5 and 7.6 summarize the QCMs' observations during the second part of the testing session. Similar to the timing of Part 1, in a few classrooms, the testing session in Part 2 was shorter than allotted because students had finished the achievement test early. In only two cases, QCMs reported testing sessions of a minute longer.

About 66 percent of the fourth grade Test Administrators and 74 percent of the eighth grade Test Administrators kept to the testing script for signaling a break before administering the student questionnaire. Of those who did make changes, in only 3 percent of the cases, those were reported as major changes. Most had made additions or other minor changes, such as paraphrasing the directions. In 14 percent of the fourth grade QCM observations and 20 percent of the eighth grade QCM observations, students requested additional time to complete the *Student Questionnaire*, which in all cases was granted. Note that the relatively high percentages of QCMs not responding to questions concerning the *Student Questionnaire* administration occurred because some schools chose to administer the questionnaire on a different date. In such cases, QCMs were not required to observe the questionnaire administration.

Exhibit 7.5 Percentages of QCM Responses for Assessment Session Part 2 – Fourth Grade

Question	Yes (%)	No (%)	Not Answered (%)
Was the time spent to restart the testing for Part 2 equal to or less than 5 minutes?	95	4	1
Was the total time for testing in Part 2 correct as indicated in the script?	94	5	1
Did the Test Administrator announce "you have 10 minutes left" prior to the end of Part 2?	87	13	0
Were there any other time remaining announcements made during Part 2?	19	80	1
Were the booklets collected and secured after the assessment session?	93	6	1
When the Test Administrator read the script to end the testing for Part 2, did he/she announce a break to be followed by the Student Questionnaire?	82	9	9
Did the Test Administrator accurately read the script to end the testing and signal a break?	66	21 *(Minor changes)* 3 *(Major changes)*	10
If there were changes, how would you describe them?			
Additions	14	7	3 *(Not Answered)* 76 *(Not Applicable)*
Omissions	8	11	5 *(Not Answered)* 76 *(Not Applicable)*
Did the Test Administrator distribute the Student Questionnaires and give directions as specified in the script?	79	7	14
Did the students ask for additional time to complete the questionnaire?	14	71	15
At the end of the session, prior to dismissing the students, did the Test Administrator thank the students for participating in the study?	80	7	13

Exhibit 7.6 Percentages of QCM Responses for Assessment Session Part 2 – Eighth Grade

Question	Yes (%)	No (%)	Not Answered (%)
Was the time spent to restart the testing for Part 2 equal to or less than 5 minutes?	96	3	1
Was the total time for testing in Part 2 correct as indicated in the script?	94	6	0
Did the Test Administrator announce "you have 10 minutes left" prior to the end of Part 2?	95	5	0
Were there any other time remaining announcements made during Part 2?	13	86	1
Were the booklets collected and secured after the assessment session?	93	6	1
When the Test Administrator read the script to end the testing for Part 2, did he/she announce a break to be followed by the Student Questionnaire?	85	12	3
Did the Test Administrator accurately read the script to end the testing and signal a break?	74	18 *(Minor changes)* 3 *(Major changes)*	5
If there were changes, how would you describe them?			
Additions	10	7	4 *(Not Answered)* 79 *(Not Applicable)*
Omissions	11	6	4 *(Not Answered)* 79 *(Not Applicable)*
Did the Test Administrator distribute the Student Questionnaires and give directions as specified in the script?	85	8	7
Did the students ask for additional time to complete the questionnaire?	20	73	7
At the end of the session, prior to dismissing the students, did the Test Administrator thank the students for participating in the study?	85	9	6

Exhibits 7.7 and 7.8 provide observations on student compliance with instructions and the alignment of the scripted instructions with their implementation. The results show that in almost all of the sessions, students complied well or very well with the instructions to stop working. In most cases, the dismissal of students had been very orderly or somewhat orderly.

Exhibit 7.7 Percentages of QCM Responses for Student Cooperation at the End of the Assessment Sessions – Fourth Grade

Question	Very Well (%)	Well (%)	Fairly Well (%)	Not well at all (%)	Not Answered (%)
When the Test Administrator ended Part 1, how well did the student comply with the instruction to stop work?	84	13	2	0	1
When the Test Administrator ended Part 2, how well did the student comply with the instruction to stop work?	86	12	2	0	0

Question	Very orderly (%)	Somewhat orderly (%)	Not orderly at all (%)	Not Answered (%)
How orderly was the dismissal of the students?	69	17	1	13

Exhibit 7.8 Percentages of QCM Responses for Student Cooperation at the End of the Assessment Sessions – Eighth Grade

Question	Very Well (%)	Well (%)	Fairly Well (%)	Not well at all (%)	Not Answered (%)
When the Test Administrator ended Part 1, how well did the student comply with the instruction to stop work?	78	18	2	1	1
When the Test Administrator ended Part 2, how well did the student comply with the instruction to stop work?	76	19	3	1	1

Question	Very orderly (%)	Somewhat orderly (%)	Not orderly at all (%)	Not Answered (%)
How orderly was the dismissal of the students?	68	26	2	4

7.2.3 General Observations

Section C of the Classroom Observation Record referred to the general observations by QCMs during the testing sessions, including their overall impressions of the test administration, how well the Test Administrator monitored students, and any unusual circumstances that arose during the testing session (e.g., student refusal to participate, defective instrumentation, emergency situations, and cheating).

The results presented in Exhibits 7.9 through 7.12 show that, for most testing sessions, no problems were observed. In almost all cases, Test Administrators addressed students' questions adequately and as instructed in the *Test Administrator Manual*. In 10 percent of the cases, QCMs reported evidence of students attempting to cheat on the test. However, when asked to explain the situation, QCMs generally indicated that students were merely looking around at their neighbors to see whether their test booklets were indeed different. Because the TIMSS 2007 test design involves 14 different booklets for each of the two target grades, students were unlikely to have the same booklet as their neighbors.

In the few sessions where a defective test instrument was detected, the Test Administrator almost always replaced the instrument appropriately. In the very few cases where a student refused to take the test, it was because parental permission for participation was denied. In one case, a student refused to complete the second part of the test. In 15 percent of the observed fourth grade testing sessions and in 10 percent of the observed eighth grade testing sessions, a student left the room for an “emergency” during the testing session. In such cases, Test Administrators were instructed that they should collect the student’s test booklet, and give it back after he or she returned. However, in two cases, students did not return to the class at all, and in almost all the other cases, the student had already completed the test and, thus, it was not necessary to receive the test booklet back after returning to the classroom. In five cases, students became ill and did not return to the testing at all, and, in all the remaining cases, students were instructed to close their booklets and leave them on their tables while being out of the classroom.

QCMs reported no cases where students were not orderly and cooperative during the testing sessions for the fourth grade and only 1 percent for the eighth grade. There were very few cases where students’ orderliness or cooperation was less than perfect or very good. In all such cases, Test Administrators managed to control the situation. QCMs reported that the overall quality of all testing sessions was good, very good, or, in 58 percent of the cases for the fourth grade and 49 percent of the cases for the eighth grade, testing sessions were excellent.

Exhibit 7.9 Percentages of QCM Responses for General Observations – Fourth Grade

Question	Yes (%)	No (%)	Not Answered (%)
During the testing sessions did the Test Administrator walk around the room to be sure students were working on the correct section of the test and/or behaving properly?	97	3	0
Did the Test Administrator address students' questions appropriately?	96	2	2
Did you see any evidence of students attempting to cheat on the tests (e.g., by copying from a neighbor)?	10	90	0
Were any defective test books detected and replaced before the testing began?	2	97	1
Were any defective test books detected and replaced after the testing began?	2	94	4
If any defective test books were replaced, did the Test Administrator replace them appropriately?	3	0	1 *(Not Answered)* 96 *(Not Applicable)*
Did any students refuse to take the test either prior to the testing or during the testing?	2	97	1
If a student refused, did the Test Administrator accurately follow the instructions for excusing the student (collect the test book and record the incident on the Student Tracking Form)?	1	1	1 *(Not Answered)* 97 *(Not Applicable)*
Did any students leave the room for an "emergency" during the testing?	15	83	2
If a student left the room for an emergency during the testing, did the Test Administrator address the situation appropriately (collect the test booklet, and if re-admitted, return the test booklet)?	11	3	3 *(Not Answered)* 83 *(Not Applicable)*

Exhibit 7.10 Percentages of QCM Responses for General Observations – Eighth Grade

Question	Yes (%)	No (%)	Not Answered (%)
During the testing sessions did the Test Administrator walk around the room to be sure students were working on the correct section of the test and/or behaving properly?	96	3	1
Did the Test Administrator address students' questions appropriately?	97	2	1
Did you see any evidence of students attempting to cheat on the tests (e.g., by copying from a neighbor)?	10	89	1
Were any defective test books detected and replaced before the testing began?	4	95	1
Were any defective test books detected and replaced after the testing began?	5	92	3
If any defective test books were replaced, did the Test Administrator replace them appropriately?	6	1	1 *(Not Answered)* 92 *(Not Applicable)*
Did any students refuse to take the test either prior to the testing or during the testing?	3	96	1
If a student refused, did the Test Administrator accurately follow the instructions for excusing the student (collect the test book and record the incident on the Student Tracking Form)?	2	0	2 *(Not Answered)* 96 *(Not Applicable)*
Did any students leave the room for an "emergency" during the testing?	10	87	2
If a student left the room for an emergency during the testing, did the Test Administrator address the situation appropriately (collect the test booklet, and if re-admitted, return the test booklet)?	8	2	3 *(Not Answered)* 87 *(Not Applicable)*

Exhibit 7.11 Percentages of QCM Responses for Observations of Student Behavior – Fourth Grade

Question	Extremely (%)	Moderately (%)	Somewhat (%)	Hardly (%)	Not answered (%)
To what extent would you describe the students as orderly and cooperative?	76	21	2	0	1

	No, there were no late students (%)	No, they were not admitted (%)	Yes, but before testing began (%)	Yes, after testing began (%)	Not answered (%)
Were any late students admitted to the testing room?	91	2	3	3	1

	Excellent (%)	Very good (%)	Good (%)	Fair (%)	Poor (%)	Not answered (%)
In general, how would you describe the overall quality of the testing session?	58	29	9	2	0	2

Exhibit 7.12 Percentages of QCM Responses for Observations of Student Behavior – Eighth Grade

Question	Extremely (%)	Moderately (%)	Somewhat (%)	Hardly (%)	Not Answered (%)
To what extent would you describe the students as orderly and cooperative?	65	29	4	1	1

	No, There Were No Late Students (%)	No, They Were Not Admitted (%)	Yes, but Before Testing Began (%)	Yes, After Testing Began (%)	Not Answered (%)
Were any late students admitted to the testing room?	87	2	4	5	2

	Excellent (%)	Very Good (%)	Good (%)	Fair (%)	Poor (%)	Not Answered (%)
In general, how would you describe the overall quality of the testing session?	49	33	12	4	0	2

7.2.4 Interview with the Test Administrator and/or School Coordinator

As the final step of each observation, the QCMs conducted an interview with the Test Administrator and/or School Coordinator. Details of the interview were recorded in Section D of the Classroom Observation Record. The interview addressed activities, such as shipment of assessment materials, arrangements for test administration, responsiveness of the NRC to queries, necessity for make-up sessions, and, as a validation of within-school sampling procedures, organization of classes in the school.

The results presented in Exhibits 7.13 and 7.14 show that overall, School Coordinators considered the TIMSS 2007 administration in their schools a success. Mistakes that did occur tended to be minor and could easily be remedied. There were only a few cases where shipments of test materials had something missing, and, in all such cases, they were resolved before the testing date.

In order to better estimate the time needed to complete the *Teacher Questionnaires*, QCMs asked if the current estimate of 45 minutes was appropriate. From all cases where *Teacher Questionnaires* already were completed, 55 percent of the fourth grade School Coordinators and 65 percent of the eighth grade School Coordinators reported that the estimate of 45 minutes was about right. Twenty-one percent of the fourth grade School Coordinators and 15 percent of the eighth grade School Coordinators reported that the questionnaires took longer, and about 10 percent (per each grade) said that they took less time to complete.

In more than half of the cases, School Coordinators indicated that students were given special instructions, motivational talks, or incentives by a school official or the classroom teacher prior to testing.

In 11 percent of the observed fourth grade classes and 15 percent of the observed eighth grade classes, the School Coordinator anticipated that a make-up session was needed, and most of them were sure that they would be conducting one.

Because the sampling of classes requires a complete list of all classes in the school at the target grade, QCMs were asked to verify that the class list did indeed include all classes. This was more confusing for the eighth grade due to some very complicated ways of organizing courses in some of the countries. In spite of complicated course structures, almost all School Coordinators reported that the complete list of classes had been documented and all students appeared in one and only one of these classes. Additional comments from School Coordinators showed that some were very confused by the question itself, commenting that they sent a list of all classes to the national center, but only one or two classes were selected to participate. Therefore, a small percentage of them answered that there were students at the grade level who did not have a chance to participate.

A tribute to the planning and implementation of TIMSS 2007 was the fact that 90 percent of respondents said they would be willing to serve as a School Coordinator in future international assessments. Furthermore, the results in Exhibits 7.15 and 7.16 suggest that the majority of School Coordinators believed the testing sessions went very well and that school staff members had mostly positive attitudes towards the TIMSS testing.

Exhibit 7.13 Receipt of Materials and Test Administration, Percentages of Responses from QCM Interviews with the Test Administrator and/or School Coordinator – Fourth Grade

Question	Yes (%)	No (%)	Not Answered (%)
Prior to the assessment day did you have time to check your shipment of materials from your TIMSS National Coordinator?	86	12	2
Did you receive the correct shipment of the following items?			
School Coordinator Manual	89	4	7
Test Administrator Manual	93	5	2
Student Tracking Forms	99	1	0
Test booklets	95	4	1
Student Questionnaires	94	5	1
Teacher Questionnaires	98	1	1
School Questionnaire	98	2	0
Test Administration Form	97	1	2
Teacher Tracking Form	91	7	2
Envelopes or boxes addressed to the National Center for the purpose of returning the materials after the assessment	82	16	2
Was the National Coordinator responsive to your questions or concerns?	76	5	19
Was the estimated time of 45 minutes to complete the Teacher Questionnaires a correct estimate?	55	21 *(Took longer)* 10 *(Took less time)*	14
Were you satisfied with the accommodations (testing room) you were able to arrange for the testing?	98	1	1
Do you anticipate that a makeup session will be required at your school?	11	85	4
If you anticipate a makeup session, do you intend to conduct one?	10	0	5 *(Not Answered)* 85 *(Not Applicable)*
Did the students receive any special instructions, a motivational talk, or incentives to prepare them for the assessment?	53	46	1
Is this a complete list of the classes in this grade in this school?	93	4	3
To the best of your knowledge, are there any students in this grade level who are not in any of these classes?	4	93	3
To the best of your knowledge, are there any students in this grade level in more than one of these classes?	1	96	3
If there was another international assessment, would you be willing to serve as a School Coordinator?	90	7	3

Exhibit 7.14 Receipt of Materials and Test Administration, Percentages of Responses from QCM Interviews with the Test Administrator and/or School Coordinator – Eighth Grade

Question	Yes (%)	No (%)	Not Answered (%)
Prior to the assessment day did you have time to check your shipment of materials from your TIMSS National Coordinator?	82	12	6
Did you receive the correct shipment of the following items?			
School Coordinator Manual	85	5	10
Test Administrator Manual	87	7	6
Student Tracking Forms	94	2	4
Test booklets	90	5	5
Student Questionnaires	90	5	5
Teacher Questionnaires	94	2	4
School Questionnaire	94	1	5
Test Administration Form	90	5	5
Teacher Tracking Form	90	5	5
Envelopes or boxes addressed to the National Center for the purpose of returning the materials after the assessment	78	17	5
Was the National Coordinator responsive to your questions or concerns?	84	4	12
Was the estimated time of 45 minutes to complete the Teacher Questionnaires a correct estimate?	65	15 *(Took longer)* 9 *(Took less time)*	11
Were you satisfied with the accommodations (testing room) you were able to arrange for the testing?	94	3	3
Do you anticipate that a makeup session will be required at your school?	15	80	5
If you anticipate a makeup session, do you intend to conduct one?	12	1	7 *(Not Answered)* 80 *(Not Applicable)*
Did the students receive any special instructions, a motivational talk, or incentives to prepare them for the assessment?	61	37	2
Is this a complete list of the classes in this grade in this school?	89	6	5
To the best of your knowledge, are there any students in this grade level who are not in any of these classes?	5	92	3
To the best of your knowledge, are there any students in this grade level in more than one of these classes?	3	93	4
If there was another international assessment, would you be willing to serve as a School Coordinator?	90	6	4

Exhibit 7.15 Overall Impressions, Percentages of Responses from QCM Interviews with the Test Administrator and/or School Coordinator – Fourth Grade

Question	Very Well, No Problems (%)	Satisfactorily, Few Problems (%)	Unsatisfactorily, Many Problems (%)	Not Answered (%)
Overall, how would you say the session went?	85	14	0	1

	Positive (%)	Neutral (%)	Negative (%)	Not Answered (%)
Overall, how would you rate the attitude of the other school staff members towards the survey?	73	22	3	2

	Worked Well (%)	Needs Improvement (%)	Not Answered (%)
Overall, do you feel the School Coordinator Manual worked well or does it need improvement?	83	8	9

Exhibit 7.16 Overall Impressions, Percentages of Responses from QCM Interviews with the Test Administrator and/or School Coordinator – Eighth Grade

Question	Very Well, No Problems (%)	Satisfactorily, Few Problems (%)	Unsatisfactorily, Many Problems (%)	Not Answered (%)
Overall, how would you say the session went?	82	16	1	1

	Positive (%)	Neutral (%)	Negative (%)	Not Answered (%)
Overall, how would you rate the attitude of the other school staff members towards the survey?	75	21	2	2

	Worked Well (%)	Needs Improvement (%)	Not Answered (%)
Overall, do you feel the School Coordinator Manual worked well or does it need improvement?	82	8	10

7.3 Survey Activities Questionnaire

The *Survey Activities Questionnaire* was designed to elicit information about NRCs experiences in preparing for and conducting the TIMSS 2007 data collection, with a focus on identifying and selecting samples, translating the test instruments, assembling and printing the test materials, packing and shipping the test materials, scoring constructed-response items, entering and verifying data, implementing the national quality assurance program, and suggesting improvements in the process. To make this data collection more efficient, the questionnaire was administered to the NRCs online.

This section reports information gathered from the *Survey Activities Questionnaire*, reflecting the quality of the TIMSS 2007 survey materials and procedures in the participating countries.

7.3.1 Sampling

The first part of the *Survey Activities Questionnaire* asked questions about sampling schools and classes within the sampled schools. Exhibits 7.17 and 7.18 show that nearly all countries did not have problems selecting their samples using the manuals provided by the TIMSS & PIRLS International Study Center. Only two countries did not use the Windows® Within-school Sampling Software (WinW3S) provided by the IEA Data Processing and Research Center (DPC) to select classes. In these cases, countries chose to use their own software, because they felt their experience using this software would make the process more efficient.

A small number of NRCs encountered organizational constraints in their systems that necessitated a deviation from the sample design. In each case, the Statistics Canada sampling expert was consulted to ensure that the altered design remained compatible with TIMSS standards. In one of the cases that requested a deviation in their sampling design (Qatar), no school sampling was necessary because the TIMSS sample included the entire target population.

Exhibit 7.17 Numbers of NRC Responses to the Survey Activities Questionnaire–Sampling – Fourth Grade

Question	Yes	No	Not Answered
Were you able to select your sample of schools and classes within schools using the manuals provided by the TIMSS & PIRLS International Study Center?	36	0	3
Did you use the Windows Within-School Sampling Software provided by the IEA Data Processing and Research Center to sample classes within schools?	34	2	3
Were there any conditions or organizational constraints that necessitated deviations from the basic TIMSS sampling design?	9	27	3

Exhibit 7.18 Numbers of NRC Responses to the Survey Activities Questionnaire–Sampling – Eighth Grade

Question	Yes	No	Not Answered
Were you able to select your sample of schools and classes within schools using the manuals provided by the TIMSS & PIRLS International Study Center?	44	0	4
Did you use the Windows Within-School Sampling Software provided by the IEA Data Processing and Research Center to sample classes within schools?	44	0	4
Were there any conditions or organizational constraints that necessitated deviations from the basic TIMSS sampling design?	10	34	4

7.3.2 Translating the Test Instruments

Exhibits 7.19 and 7.20 provide NRCs answers to questions about translating the test instruments. In translating the survey instruments, NRCs generally reported using their own staff or a combination of their staff and outside experts. Almost all NRCs reported that they had gone through the process of external translation verification of the assessment items and background questionnaires organized by the IEA Secretariat. Dubai, United Arab Emirates, reported that they used the survey instruments from Qatar that had already gone through the process of verification.

Exhibit 7.19 Numbers of NRC Responses to the Survey Activities Questionnaire – Translating the Test Instruments – Fourth Grade

Question	Own Staff	Outside Translator(s)	Outside Reviewer(s)	Combination	Not Answered
Did you use your own staff or outside experts to translate the mathematics assessment items?	12	5	1	15	6
Did you use your own staff or outside experts to translate the science assessment items?	12	5	1	15	6
Did you use your own staff or outside experts to translate the background questionnaires?	16	6	1	10	6

	Yes	No	Not Answered
Did you go through the process of external translation verification of the assessment items by the IEA?	32	1	6
Did you go through the process of external translation verification of the background questionnaires by the IEA?	33	1	5

Exhibit 7.20 Numbers of NRC Responses to the Survey Activities Questionnaire – Translating the Test Instruments – Eighth Grade

Question	Own Staff	Outside Translator(s)	Outside Reviewer(s)	Combination	Not Answered
Did you use your own staff or outside experts to translate the mathematics assessment items?	15	4	0	24	5
Did you use your own staff or outside experts to translate the science assessment items?	15	4	0	24	5
Did you use your own staff or outside experts to translate the background questionnaires?	19	6	1	17	5

	Yes	No	Not Answered
Did you go through the process of external translation verification of the assessment items by the IEA?	41	2	5
Did you go through the process of external translation verification of the background questionnaires by the IEA?	41	2	5

7.3.3 Assembling and Printing the Test Instruments

The NRCs were asked to answer some questions about assembling and printing the test materials, as well as issues related to checking the materials and securely storing them. The results in Exhibits 7.21 and 7.22 show that all NRCs answered that they were able to assemble the test booklets according to the instructions provided, and only one country did not go through the process of external layout verification of the test booklets by the TIMSS & PIRLS International Study Center. Nearly all countries conducted the recommended quality control checks during the printing process. The most common errors that countries detected during the printing process were missing pages and wrong page order. The NRCs were able to fix all of the systematic errors before sending the tests for administration.

All countries reported that they followed procedures to protect the security of the tests during assembly and printing. One country was concerned that there could be a breach of security because so many different people were involved in the study, even though they all were asked to sign a nondisclosure agreement.

Exhibit 7.21 Numbers of NRC Responses to the Survey Activities Questionnaire – Assembling and Printing the Test Instruments – Fourth Grade

Question	Yes	No	Not Answered
Were you able to assemble the test booklets according to the instructions provided by the TIMSS & PIRLS International Study Center?	34	0	5
Were you able to assemble the background questionnaires according to the instructions provided by the TIMSS & PIRLS International Study Center?	34	0	5
Did you go through the process of external layout verification of the survey instruments by the TIMSS & PIRLS International Study Center?	32	2	5
Did you conduct the quality assurance procedures for checking the survey instruments during the printing process?	32	1	6
If errors were detected, what was the nature of the errors?			
Poor print quality	7	26	6
Pages missing	7	26	6
Page order	9	24	6
Upside down pages	2	30	7
Did you follow procedures to protect the security of the survey materials during the assembly and printing process?	33	0	6
Did you discover any potential breaches of security?	0	33	6

Exhibit 7.22 Numbers of NRC Responses to the Survey Activities Questionnaire – Assembling and Printing the Test Instruments – Eighth Grade

Question	Yes	No	Not Answered
Were you able to assemble the test booklets according to the instructions provided by the TIMSS & PIRLS International Study Center?	43	0	5
Were you able to assemble the background questionnaires according to the instructions provided by the TIMSS & PIRLS International Study Center?	43	0	5
Did you go through the process of external layout verification of the survey instruments by the TIMSS & PIRLS International Study Center?	42	1	5
Did you conduct the quality assurance procedures for checking the survey instruments during the printing process?	40	3	5
If errors were detected, what was the nature of the errors?			
Poor print quality	7	35	6
Pages missing	18	24	6
Page order	17	25	6
Upside down pages	9	33	6
Did you follow procedures to protect the security of the survey materials during the assembly and printing process?	43	0	5
Did you discover any potential breaches of security?	1	42	5

7.3.4 Packing and Shipping the Testing Materials

Some questions in the questionnaire addressed the extent to which NRCs detected errors in the testing materials as they were packed for shipping to School Coordinators. As shown in Exhibits 7.23 and 7.24, a few errors were found in the materials. All errors that were discovered before distribution were remedied. In cases where errors were found after distribution, they usually were very minor and could be remedied by school coordinators. In more severe cases, the provided replacement materials were used.

Exhibit 7.23 Numbers of NRC Responses to the Survey Activities Questionnaire – Packing and Shipping the Testing Materials – Fourth Grade

Question	No Errors, or Not Used	Errors Found Before Distribution	Errors Found After Distribution	Errors Found Before And After Distribution	Not Answered
In packing the assessment materials for shipment to schools, did you detect any errors in any of the following items?					
Test booklets	15	5	11	3	5
Student Questionnaires	26	4	3	1	5
Student Tracking Forms	30	0	3	1	5
Test Booklet ID labels	29	2	2	1	6
Student Questionnaire ID labels	29	2	2	0	6
Sequencing of Booklets or Student Questionnaires	24	3	6	1	5
Teacher Questionnaires	32	1	1	0	5
Teacher Tracking Forms	29	1	3	0	6
School Questionnaires	32	1	1	0	5
Test Administrator Manual	32	0	2	0	5
School Coordinator Manual	32	0	2	0	5
Return Labels	32	0	1	0	6
Self-addressed postcards for test dates	31	0	0	0	8

Exhibit 7.24 Numbers of NRC Responses to the Survey Activities Questionnaire – Packing and Shipping the Testing Materials – Eighth Grade

Question	No Errors, or Not Used	Errors Found Before Distribution	Errors Found After Distribution	Errors Found Before And After Distribution	Not Answered
In packing the assessment materials for shipment to schools, did you detect any errors in any of the following items?					
Test booklets	20	7	13	4	4
Student Questionnaires	34	7	3	0	4
Student Tracking Forms	40	1	2	1	4
Test Booklet ID labels	38	3	2	1	4
Student Questionnaire ID labels	39	3	2	0	4
Sequencing of Booklets or Student Questionnaires	32	4	6	2	4
Teacher Questionnaires	41	3	0	0	4
Teacher Tracking Forms	41	2	0	1	4
School Questionnaires	41	3	0	0	4
Test Administrator Manual	41	1	1	1	4
School Coordinator Manual	42	1	0	1	4
Return Labels	41	0	2	0	5
Self-addressed postcards for test dates	40	0	0	0	8

7.3.5 Scoring Constructed-response Items

The *Survey Activities Questionnaire* collected information from NRCs about preparation for scoring and scoring the constructed-response items. The scoring process was an ambitious effort, requiring recruiting and training scoring staff to score student responses including independent double scoring to verify scoring reliability. Exhibits 7.25 and 7.26 indicate that almost all NRCs understood the procedures of within-country reliability scoring, trend-reliability scoring, and cross-country reliability scoring, as explained in the manuals provided by the TIMSS & PIRLS International Study Center. Three countries had time-, money-, and language-related (English was used for this activity) problems completing the cross-country reliability scoring task. Countries on the Northern Hemisphere timeline did not participate in the cross-country reliability scoring activity, because most of them (the ones testing in English) supplied student responses used by all the other countries.

Note that the question on trend-reliability scoring procedures did not apply to countries that did not participate in TIMSS 2003.

Exhibit 7.25 Numbers of NRC Responses to the Survey Activities Questionnaire – Scoring Constructed-response Items – Fourth Grade

Question	Own Staff	Teachers	University Students	Combination of the Above	Other	Not Answered
Who primarily scored your constructed-response mathematics assessment items?	2	11	4	13	5	4
Who primarily scored your constructed-response science assessment items?	2	11	5	12	5	4

Question	Yes	No	Not Answered
Did you understand the procedure of reliability scoring, as explained in the manuals provided by the TIMSS & PIRLS International Study Center?	34	1	4
Did you understand the trend-reliability scoring procedure, as explained in the manuals provided by the TIMSS & PIRLS International Study Center?	19	4	*Not Applicable*
Did you understand thecross-country reliability scoring procedure, as explained in the manuals provided by the TIMSS & PIRLS International Study Center?	25	6	8

Exhibit 7.26 Numbers of NRC Responses to the Survey Activities Questionnaire – Scoring Constructed-response Items – Eighth Grade

Question	Own Staff	Teachers	University Students	Combination of the Above	Other	Not Answered
Who primarily scored your constructed-response mathematics assessment items?	2	15	2	14	11	4
Who primarily scored your constructed-response science assessment items?	2	15	2	14	11	4

Question	Yes	No	Not Answered
Did you understand the procedure of reliability scoring, as explained in the manuals provided by the TIMSS & PIRLS International Study Center?	42	1	5
Did you understand the trend-reliability scoring procedure, as explained in the manuals provided by the TIMSS & PIRLS International Study Center?	28	4	*Not Applicable*
Did you understand thecross-country reliability scoring procedure, as explained in the manuals provided by the TIMSS & PIRLS International Study Center?	31	7	10

7.3.6 Data Entry and Verification

Exhibits 7.27 and 7.28 report that most countries entered the data from a percentage of test booklets twice as a verification procedure. The estimated proportion of booklets to be entered twice ranged from 5 to 30 percent.

Only one NRC reported having concerns about establishing a secure storage area for the returned tests after data entry.

Exhibit 7.27 Results of the Survey Activities Questionnaire – Data Entry and Verification – Fourth Grade

Question	Yes	No	Not Answered
Did you enter a percentage of test booklets twice as a verification procedure?	20	14	5
Did you use the Windows Data Entry Manager software provided by the IEA Data Procesing Center and Research to enter your test instrument data?	30	5	4
Were the returned tests stored in a secure area after scoring and data entry until the original documents could be discarded?	35	0	4

Question	Own Staff	External Data Entry Firm	Combination of the Above	Other	Not Answered
Who primarily entered the TIMSS data for your country?	11	6	8	9	5

Exhibit 7.28 Results of the Survey Activities Questionnaire – Data Entry and Verification – Eighth Grade

Question	Yes	No	Not Answered
Did you enter a percentage of test booklets twice as a verification procedure?	27	16	5
Did you use the Windows Data Entry Manager software provided by the IEA Data Procesing Center and Research to enter your test instrument data?	40	3	5
Were the returned tests stored in a secure area after scoring and data entry until the original documents could be discarded?	41	1	6

Question	Own Staff	External Data Entry Firm	Combination of the Above	Other	Not Answered
Who primarily entered the TIMSS data for your country?	17	9	10	7	5

7.3.7 National Quality Assurance Program

As part of the national quality assurance activities, NRCs were required to send National Quality Control Observers to 10 percent of the participating schools in order to observe the test administration and document compliance with prescribed procedures. The last section of the *Survey Activities Questionnaire* addressed preparation for and implementation of the national quality assurance program.

As shown in Exhibits 7.29 and 7.30, almost all the national centers used the *National Quality Control Monitor Manual* provided by the TIMSS & PIRLS International Study Center in order to conduct their quality assurance program. The on-site quality control observations were conducted either by an external agency, members of the national center, or in some cases, other professionals, such as inspectors, retired teachers, mathematics and science supervisors, or ministry representatives.

Exhibit 7.29 Numbers of NRC Responses to the Survey Activities Questionnaire – National Quality Assurance Program – Fourth Grade

Question	An External Agency	Members of the National Center	A Combination of the Above	Other	Not Answered
Who did the classroom observations?	2	9	9	15	4

Question	Yes	No	Not Answered
When conducting your own quality assurance program, did you use the National Quality Control Monitor Manual provided by the TIMSS & PIRLS International Study Center?	34	1	4

Exhibit 7.30 Numbers of NRC Responses to the Survey Activities Questionnaire –National Quality Assurance Program – Eighth Grade

Question	An External Agency	Members of the National Center	A Combination of the Above	Other	Not Answered
Who did the classroom observations?	4	10	12	16	6

Question	Yes	No	Not Answered
When conducting your own quality assurance program, did you use the National Quality Control Monitor Manual provided by the TIMSS & PIRLS International Study Center?	39	3	6

References

TIMSS & PIRLS International Study Center. (2006a). TIMSS *2007 international quality control monitor manual*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006b). TIMSS *2007 national quality control monitor manual*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006c). TIMSS *2007 school coordinator manual*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006d). TIMSS *2007 survey operations procedures unit 2: Contacting schools and sampling classes*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006e). TIMSS *2007 survey operations procedures unit 3: Preparing materials for data collection*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006f). TIMSS *2007 survey operations procedures unit 4: Administering the* TIMSS *2007 assessment*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

TIMSS & PIRLS International Study Center. (2006g). TIMSS *2007 test administrator manual*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Chapter 8

Creating and Checking the TIMSS 2007 Database

Juliane Barth and Oliver Neuschmidt

8.1 Overview

This chapter describes the TIMSS 2007 data checking and database creation procedures implemented by the IEA Data Processing and Research Center (DPC), the TIMSS & PIRLS International Study Center, Statistics Canada, and the national centers of participating countries. The overriding concerns were to ensure that all information in the database conformed to the internationally defined data structure, national adaptations to questionnaires were reflected appropriately in the codebooks and documentation, and all variables used for international comparisons were comparable across countries. Quality control measures were applied throughout the process to assure the quality and accuracy of the TIMSS data.

8.2 Steps Taken to Confirm the Integrity of the TIMSS 2007 International Database

The following summarizes the steps taken at all institutions to confirm the integrity of the international database. First, the IEA DPC was responsible for checking the data files from each country, applying standard cleaning rules to verify the accuracy and consistency of the data, and documenting electronically any deviations from the international file structure. Any queries were addressed to the national centers, and modifications were made to the data files as necessary. After all modifications had been applied, all data were processed and checked again. This process of editing the data, checking the reports, and implementing corrections was repeated as many times as necessary until all data were consistent and comparable within and between countries.

When the national files had been checked, the IEA DPC provided national univariate and reliability statistics to the national centers, while the TIMSS & PIRLS International Study Center provided them with data almanacs containing international univariate statistics and national item statistics so that National Research Coordinators (NRCs) could examine their data from an international perspective. This was one of the most important checks in terms of ensuring the international comparability of the data. A particular statistic may seem plausible in a national context, but it may be an outlier when comparing data across countries in an international context. Any such instances were investigated and, if necessary, addressed by either recoding the affected variables or removing them from the international database.

Once verified and in the international file format, the achievement data were sent to the TIMSS & PIRLS International Study Center where basic item statistics were produced and reviewed. At the same time, the IEA DPC sent data files containing information on the participation of schools and students in each country's sample to Statistics Canada. This information, together with data provided by the NRC tracking forms and the software designed to standardize operations and tasks, was used by Statistics Canada to calculate sampling weights, population coverage, and school and student participation rates.[1]

When the review of the item statistics was completed and Statistics Canada finalized the computation of sampling weights, the TIMSS & PIRLS International Study Center conducted the IRT scaling and generated proficiency scores in mathematics and science for each participating student. The scaling methods and procedures are described in Chapter 11. Once the sampling weights and the proficiency scores had been verified at the TIMSS & PIRLS International Study Center, they were sent to the IEA DPC for inclusion in the international database and for distribution to the national centers.

8.3 Data Checking at the IEA Data Processing and Research Center

As described in Chapter 6, each participating country was responsible for entering their TIMSS 2007 data into the appropriate data files and submitting these files to the IEA DPC, where they underwent an exhaustive process of checking and editing—a process known as data cleaning. To facilitate the data cleaning process, countries were requested to provide the IEA DPC with

1 See Chapter 5 for details about the TIMSS 2007 sampling design.

detailed documentation of their data, in addition to the data files themselves. This data documentation included copies of all original survey tracking forms, copies of the national versions of test booklets and questionnaires, and completing the *Survey Activities Questionnaire,* an Internet-based questionnaire about countries' data collection activities (TIMSS, 2005-2006). To ensure that all national adaptations to the survey instruments were fully documented, countries also were required to submit National Adaptation Forms (NAFs).

Countries also were asked to send the IEA DPC the sample of test booklets selected for double-scoring the constructed-response items (approximately 1,400 booklets per population). The student responses to constructed-response items in these booklets are digitally scanned and preserved for use in the next cycle of TIMSS in 2011, when they will be rescored by TIMSS 2011 scoring staff to monitor consistency in scoring practices between TIMSS 2007 and 2011.

8.3.1 Quality Control in Data Cleaning

TIMSS is a very large and complex study with very demanding standards for data quality. This requires an extensive set of interrelated data checking and cleaning procedures. To ensure that all procedures were conducted in the correct sequence, that no special requirements were overlooked, and that the cleaning process was implemented independently of the persons in charge, the following steps were undertaken:

- Before their use with real data, all data-cleaning programs were thoroughly tested using simulated data sets containing all possible problems and inconsistencies.
- All incoming data and documents were registered in a specific database. The date of arrival was recorded, along with any specific issues meriting attention.
- The cleaning was organized following strict rules. Deviations from the cleaning sequence were not possible, and the scope for involuntary changes to the cleaning procedures was minimal.
- All corrections to a country's data files were listed in a country-specific cleaning report.
- Occasionally, it was necessary to make changes to a country's data files. Every such "manual" correction was logged using a specially developed editing program (SAS-ManCorr), which recorded all

changes and allowed IEA DPC staff to undo changes or to redo the whole manual cleaning process automatically at a later stage of the cleaning.

- Once the data cleaning was completed for a country, all cleaning steps were repeated from the beginning to detect any problems that might have been inadvertently introduced during the cleaning process.
- IEA DPC staff worked closely with the national centers, and at different steps of the cleaning process, countries were provided with the processed data files and accompanying documentation and statistics, allowing them to thoroughly review and correct any inconsistencies detected (see section 8.4).
- All national adaptations that countries recorded in their documentation were verified against the structure of the national data files. All deviations from the international data structure that were detected were recorded in a National Adaptation Database in the *TIMSS 2007 User Guide* (Foy & Olson, 2009). Whenever possible, national deviations were recoded to follow the international data structure. However, if international comparability could not be assured, the corresponding data was removed from the international database.

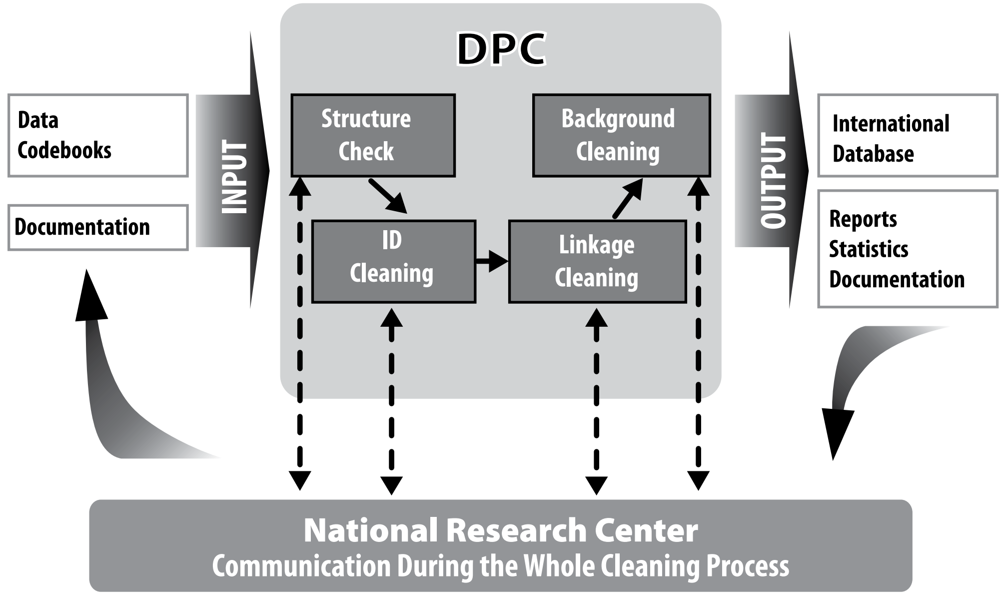
8.3.2 Preparing National Data Files

The main objective of the data cleaning process was to ensure that the data adhered to international formats; school, teacher, and student information could be linked between different survey files; and the data accurately and consistently reflected the information collected within each country.

The program-based data cleaning consisted of the following steps, which are shown in Exhibit 8.1 and explained in the following sections:

- Documentation and structure check
- Identification variable (ID) cleaning
- Linkage check
- Resolving inconsistencies in background questionnaire data.

Exhibit 8.1 Overview of Data Processing at the DPC



8.3.3 Documentation and Structure Check

For each country, data cleaning began with an exploration of its data file structures and a review of its data documentation: National Adaptation Forms, Student Tracking Forms, Student-Teacher Linkage Forms, Teacher Tracking Forms, and Test Administration Forms. Most countries sent all required documentation along with their data, which greatly facilitated the data checking.

At the beginning of the cleaning process, the tracking information and sampling information captured in the WinW3S database was combined with the WinDEM data files containing the corresponding survey instrument data (see Chapter 6 for more information).

The first checks implemented at the IEA DPC looked for differences between the international file structure and the national file structures. Some countries made adaptations (such as adding national variables or omitting or modifying international variables) to their background questionnaires. The extent and nature of such changes differed across the countries: some countries administered the questionnaires without any changes (apart from the translations), whereas other countries inserted items or options within existing international variables or added entirely new national variables. To keep track of any adaptations, NRCs were asked to complete National Adaptation Forms as they adapted the international codebooks. Where necessary, the IEA DPC modified the structure of the country's data to ensure that the resulting data remained comparable between countries.

As part of this standardization process, since direct correspondence between the data collection instruments and the data files was no longer necessary, the file structure was rearranged from a booklet-oriented model designed to facilitate data entry to an item-oriented layout more suited to data analysis. Variables created purely for verification purposes during data entry were dropped at this time, and a provision was added for new variables necessary for analysis and reporting (i.e., reporting variables, derived variables, sampling weights, and achievement scores).

After each data file matched the international standard, as specified in the international codebooks, a series of standard cleaning rules were applied to the files. This was conducted using software developed at the IEA DPC that could identify and, in many cases, correct inconsistencies in the data. Each problem was recorded in a database, identified by a unique problem number, and included a description of the problem and the action taken by the program or by the IEA DPC staff.

Where problems could not be rectified automatically, they were reported to the responsible NRC so that the original data collection instruments and tracking forms could be checked to trace the source of the errors. Wherever possible, staff at the IEA DPC suggested a remedy and asked the NRCs to either accept it or propose an alternative. Data files then were updated to reflect the solutions agreed upon. Where the NRC could not solve problems by inspecting the instruments or forms, a general cleaning rule was applied to the files to rectify this. After all automatic updates had been applied, remaining corrections to the data files were applied directly by keyboard, using a specially developed editing program (SAS-ManCorr).

8.3.4 Identification Variable (ID) Cleaning

Each record in a data file should have a unique identification number. The existence of records with duplicate ID numbers in a file implies an error of some kind. If two records share the same ID number, and contained exactly the same data, one of the records was deleted and the other remained in the database. If the records contained different data apart from the ID numbers and it was impossible to identify which record contained the "true data," both records were removed from the database. The IEA DPC tried to keep such losses at a minimum, and in only a few cases were data actually deleted.

The ID cleaning focused on the student background questionnaire file, which contained most of the critical ID variables. Apart from the unique student ID number, variables pertaining to the student participation and exclusion status, as well as the dates of birth and dates of testing used to calculate age at the time of testing were important to check. The Student Tracking Forms[2] were essential in resolving any anomalies, as was close cooperation with NRCs (since, in most cases, the Student Tracking Forms were completed in the country's official language). After cleaning, databases created from the WinW3S program containing information about student participation and exclusion were sent to Statistics Canada, where they were used to calculate students' participation rates, exclusion rates, and student sampling weights.

8.3.5 Linkage Check

In TIMSS, data about students and their schools and teachers appeared in several different files, so that it was crucial that the records from these files link together correctly to provide meaningful data for analysis and reporting. The linkage was implemented through a hierarchical ID numbering system incorporating a school, class, and student component[3] and was cross-checked against the tracking forms. It was necessary that students' entries in the achievement file and student background file were matched correctly; that the student entries in the reliability scoring file matched of the student entries in the achievement file; that the teachers were linked to the correct students; and that the schools were linked to the correct teachers and students.

8.3.6 Resolving Inconsistencies in Background Questionnaire Data

The number of inconsistent and implausible responses in background files varied from country-to-country, but no country's data were completely free of inconsistent responses. Treatment of these responses was determined on a question-by-question basis, using available documentation to make an informed decision. All background questionnaire data were checked for consistency among the responses given. For example, question number 1(a) in the *School Questionnaire* asked for the total school enrollment (number of students) in all grades, while 1(b) asked for the enrollment in the target grade only. Clearly, the number given for 1(b) should not exceed the number given for 1(a). All such inconsistencies that were detected were flagged and

2 Tracking forms were used to record the sampling of schools, classes, teachers, and students (also see Chapter 6).

3 The ID number of a higher level is included in the ID number of a lower sampling level. The class ID includes the school ID, and the student ID includes the class ID (e.g., student 1220523 may be described as student 23 of class 05 in school 122).

the NRCs asked to investigate. Those cases that could not be corrected or where the data made no sense were recoded to "Omitted".

Filter questions, which appear in some questionnaires, were used to direct the respondent to a particular section of the questionnaire. Filter questions and the dependent questions that follow were subject to the following cleaning rules: If the answer to the filter question was "No" or "Not applicable" and the dependent questions were answered, then the filter question was recoded to "Yes".

Split variable checks were applied to questions where the answer was coded into several variables. For example, question 5 in the *Student Questionnaire* listed a number of home possessions and asked the student to check all that applied. Student responses were captured in a series of nine variables, each one coded as "Yes" if the corresponding possession was checked and "No" if left unchecked. Occasionally, students checked the "Yes" boxes but left the "No" boxes unchecked or missing. Since in these cases, it was clear that the unchecked boxes actually meant "No," these were recoded accordingly.

8.3.7 National Cleaning Documentation

NRCs received a detailed report (IEA, 2007) of all problems identified in their data and the steps applied to correct them. These included the following:

- Documentation of any data problems detected by the cleaning program and the steps applied to resolve them
- A record of all deviations from the international data collection instruments and the international file structure.

Additionally, the IEA DPC provided each NRC with revised data files incorporating all agreed-upon edits, updates, and structural modifications. The revised files included a range of new variables that could be used for analytic purposes. For example, the student files included nationally standardized scores in mathematics and science that could be used in national analyses to be conducted before the international database became available.

8.3.8 Handling of Missing Data

When the TIMSS data were entered using WinDEM, two types of entries were possible: valid data values and missing data values. Missing data can be assigned a value of omitted or not administered during data entry. At the IEA

DPC, additional missing codes were applied to the data to be used for further analyses. In the international database, four missing codes are used:

- **Not administered**. The respondent was not administered the actual item. He or she had no chance to read and answer the question (assigned both during data entry and data processing).
- **Omitted**. The respondent had a chance to answer the question but did not do so. This code also was used for responses that were not interpretable in both the background and the achievement files (assigned both during data entry and data processing).
- **Logically not applicable**. The respondent answered a preceding filter question in a way that made the following dependent questions not applicable to him or her (assigned during data processing only).
- **Not reached** (only used in the achievement files). This code indicates those items not reached by the students due to a lack of time (assigned during data processing only).

8.4 Data Products

Data products sent to NRCs by the IEA DPC and the TIMSS & PIRLS International Study Center included both data almanacs and data files.

8.4.1 Data Almanacs and Item Statistics

Each country received a set of data almanacs or summaries, produced by the TIMSS & PIRLS International Study Center. These contained weighted summary statistics for each participating country on each variable included in the survey instruments. The data almanacs were sent to participating countries for review. When necessary, they were accompanied by specific questions about the data presented in them. They also were used by the TIMSS & PIRLS International Study Center during the data review and in the production of the reporting exhibits. Also, the IEA DPC produced a set of preliminary scoring reliability statistics for each country containing summary statistics at the item level on the percent of agreement between scorers.

8.4.2 Versions of the National Data Files

Building the international database was an iterative process. The IEA DPC provided each NRC with a new version of their country's data files whenever a major step in data processing was completed. This also guaranteed that NRCs had a chance to review their data and run their own checks to validate the data files. Before the TIMSS international database was published, several versions of the data files were sent to each country. Each country received its own data only. The first version was sent as soon as the data could be regarded as "clean" concerning identification codes and linkage issues. These first files contained nationally standardized achievement scores calculated by the IEA DPC using a Rasch-based scaling method. Documentation, with a list of the cleaning checks and corrections made in the data, was included to enable the NRC to review the cleaning process. Another version of the data files was sent to countries when the weights and international achievement scores were available and had been merged in the files, together with the data almanacs. This was done after all exhibits of the TIMSS international reports had been verified and final updates to the data files implemented, and enabled the NRCs to replicate the results presented in the international reports.

8.4.3 The TIMSS 2007 International Database

The international database incorporated all national data files. Data processing at the IEA DPC ensured that:

- Information coded in each variable was internationally comparable.
- National adaptations were reflected appropriately in all variables.
- Questions that were not internationally comparable were removed from the database.
- All entries in the database could be linked to the appropriate respondent—student, teacher, or principal.
- Sampling weights and student achievement scores were available for international comparisons.

In a joint effort of the IEA DPC and the TIMSS & PIRLS International Study Center, a National Adaptations Database containing all adaptations to questionnaires made by individual countries and documenting how they were handled was constructed. The meaning of country-specific items also can be found in this database, as well as recoding requirements by the TIMSS & PIRLS International Study Center. Information contained in this database is provided in the *TIMSS 2007 User Guide for the International Database* (Foy & Olson, 2009) upon release of the TIMSS 2007 data.

References

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

IEA. (2007). *General cleaning documentation V11*. Hamburg: IEA Data Processing and Research Center.

TIMSS (2005-2006). *TIMSS 2007 survey operations procedures, units 1-6.* Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Chapter 9

TIMSS 2007 Sampling Weights and Participation Rates

Marc Joncas

9.1 Overview

Rigorous sampling of schools and students was a key component of the TIMSS 2007 project. Implementing the sampling plan was the responsibility of the National Research Coordinator (NRC) in each participating country. NRCs were supported in this endeavor by TIMSS 2007 sampling consultants, Statistics Canada, and the Samplin g Unit of the IEA Data Processing and Research Center (DPC). Sampling consultants conducted the school sampling for most countries and trained NRCs in selecting probability samples of students and using the *Windows® Within-school Sampling Software (WinW3S)* (2006) provided by the IEA DPC. As an essential part of their sampling activities, NRCs were responsible for providing detailed documentation describing their national sampling plans (sampling data, school sampling frames, and school sample selections). The documentation for each TIMSS participant was reviewed and completed by the sampling consultants, including details on coverage and exclusion levels, stratification variables, sampling, participation rates, and variance estimates. The TIMSS & PIRLS International Study Center and the TIMSS 2007 Sampling Referee, Dr. Keith Rust of Westat, Inc., used this information to evaluate the quality of the samples.

This chapter gives a summary of the major characteristics of the national samples, along with a description of how sampling weights and participation rates were calculated for TIMSS 2007. School, classroom, and student participation rates for each country also are presented. More detailed summaries of the sample design for each country, including details of population coverage and exclusions, stratification variables, and schools' sampling allocations, are provided in Appendix B.

9.2 Sampling Implementation

9.2.1 Target Populations

As described in Chapter 5, TIMSS 2007 chose to study achievement in two target populations, and participating countries were free to select either population or both. The *international target populations* for TIMSS were defined as the grade that represented 4 or 8 years of schooling, counting from the first year of primary or elementary schooling, unless this would result in an average student age of less than 9.5 years for the lower grade or 13.5 for the higher grade.

Exhibits 9.1 and 9.2 present the grades identified as the target grades for sampling by each country, together with the number of years of formal schooling the grades represent and the average age of students in the target grade that were sampled for TIMSS at the time of testing for fourth and eighth grades, respectively. For most countries, the target grades did indeed turn out to be the grades with 4 and 8 years of schooling. In England, Malta, New Zealand, and Scotland, children begin primary school at age 5, and therefore, these countries assessed students in the fifth or ninth year of schooling. Their students were still among the youngest in TIMSS 2007. In Bosnia and Herzegovina, students from the five regions of the Republika Srpska had 9 years of schooling, compared to 8 years for the rest of the country, due to the early school-entry age (at age 6, compared to age 7 for the other regions). Finally, Kuwait and the non-Indian schools of Dubai, UAE[1] also tested in the fifth and ninth grade in October 2007 due to late data collection.

1 The school year for the Indian schools starts in April, and students under that schedule were tested at the end of their school year (grade 4 or grade 8). All other students start their school year in September and were tested at the beginning of their school year (grade 5 or grade 9).

Exhibit 9.1 National Grade Definitions – Fourth Grade

Country	Country's Name for Grade Tested	Years of Formal Schooling*	Average Age at Time of Testing
Algeria	Four year primary	4	10.2
Armenia	Grade 4	4	10.6
Australia	Year 4	4	9.9
Austria	Fourth grade / Last grade of primary education	4	10.3
Chinese Taipei	Elementary school, grade 4	4	10.2
Colombia	Fourth grade	4	10.4
Czech Republic	Grade 4	4	10.3
Denmark	Grade 4	4	11.0
El Salvador	Fourth grade of basic education	4	11.0
England	Year 5	5	10.2
Georgia	Grade 4	4	10.1
Germany	Grade 4	4	10.4
Hong Kong SAR	Primary 4	4	10.2
Hungary	Fourth grade	4	10.7
Iran, Islamic Rep. of	Fourth grade of primary school	4	10.2
Italy	Grade 4 (IV class of primary school)	4	9.8
Japan	Fourth grade at the elementary school	4	10.5
Kazakhstan	Fourth grade (1st stage of basic education)	4	10.6
Kuwait	Grade 5 (Primary)	4	10.2
Latvia	Grade 4	4	11.0
Lithuania	Grade 4	4	10.8
Morocco	Grade 4 primary school	4	10.6
Netherlands	Grade 6 (the first year of kindergarten is grade 1)	4	10.2
New Zealand	Year 5 (year 1 is equivalent to Kindergarten)	4.5–5.5	10.0
Norway	Grade 4	4	9.8
Qatar	Fourth grade	4	9.7
Russian Federation	Fourth grade	4	10.8
Scotland	Primary 5 (P5)	5	9.8
Singapore	Primary 4	4	10.4
Slovak Republic	Fourth grade	4	10.4
Slovenia	Grade 4	4	9.8
Sweden	Grade 4	4	10.8
Tunisia	Fourth grade of basic school	4	10.2
Ukraine	Grade 4	4	10.3
United States	Grade 4 of elementary school	4	10.3
Yemen	Grade 4	4	11.2
Benchmarking Participants			
Alberta, Canada	Grade 4	4	9.8
British Columbia, Canada	Grade 4	4	9.8
Dubai, UAE	Grade 4 or Grade 5	4	10.0
Massachusetts, US	Fourth grade	4	10.3
Minnesota, US	Fourth grade	4	10.3
Ontario, Canada	Grade 4	4	9.8
Quebec, Canada	Second year of second cycle	4	10.1

* Represents years of schooling counting from the first year of ISCED Level 1.

Exhibit 9.2 National Grade Definitions – Eighth Grade

Country	Country's Name for Grade Tested	Years of Formal Schooling*	Average Age at Time of Testing
Algeria	Second year of middle school	8	14.5
Armenia	Grade 8	8	14.9
Australia	Year 8	8	13.9
Bahrain	Second Intermediate	8	14.1
Bosnia and Herzegovina	Final grade (grade 8 and grade 9)	8 or 9	14.7
Botswana	Form One	8	14.9
Bulgaria	Grade 8	8	14.9
Chinese Taipei	Junior high school, grade 8	8	14.2
Colombia	Eighth grade	8	14.5
Cyprus	B Gymnasium	8	13.8
Czech Republic	Grade 8	8	14.4
Egypt	Preparatory 2	8	14.1
El Salvador	Eighth grade of basic education	8	15.0
England	Year 9	9	14.2
Georgia	Grade 8	8	14.2
Ghana	Junior secondary school II (JSS II)	8	15.8
Hong Kong SAR	Secondary 2	8	14.4
Hungary	Eighth grade	8	14.6
Indonesia	Grade 8	8	14.3
Iran, Islamic Rep. of	Third year in guidance school	8	14.2
Israel	Eighth grade	8	14.0
Italy	Grade 8 (III Media)	8	13.9
Japan	Second grade at the lower secondary school	8	14.5
Jordan	Grade 8	8	14.0
Korea, Rep. of	Grade 2 of middle school	8	14.3
Kuwait	Ninth grade (Intermediate)	8	14.4
Lebanon	Grade 8 of the basic educational level	8	14.4
Lithuania	Grade 8	8	14.9
Malaysia	Form 2 (Grade 8)	8	14.3
Malta	Form 3 (Grade 9)	9	14.0
Morocco	Second year collegial	8	14.8
Norway	Grade 8	8	13.8
Oman	Grade 8	8	14.3
Palestinian Nat'l Auth.	Eighth grade	8	14.0
Qatar	Grade 8	8	13.9
Romania	Grade 8	8	15.0
Russian Federation	Eighth grade	7 or 8	14.6
Saudi Arabia	Second year of middle school	8	14.4
Scotland	Secondary 2 (S2)	9	13.7
Serbia	Eighth grade	8	14.9
Singapore	Secondary 2	8	14.4
Slovenia	Grade 8	7 or 8	13.8
Sweden	Grade 8	8	14.8
Syrian Arab Republic	Grade 8	8	13.9
Thailand	Middle school grade 2	8	14.3
Tunisia	Eighth year of basic school	8	14.5
Turkey	Eighth grade	8	14.0
Ukraine	Grade 8	8	14.2
United States	Grade 8	8	14.3
Benchmarking Participants			
Basque Country, Spain	Second course of secondary compulsory education	8	14.1
British Columbia, Canada	Grade 8	8	13.9
Dubai, UAE	Grade 8 or Grade 9	8	14.2
Massachusetts, US	Eighth grade	8	14.2
Minnesota, US	Eighth grade	8	14.3
Ontario, Canada	Grade 8	8	13.8
Quebec, Canada	Secondary II (cycle one)	8	14.2

* Represents years of schooling counting from the first year of ISCED Level 1.

9.2.2 Population Coverage and Exclusions

Exhibits 9.3 and 9.4 summarize population coverage and exclusions for the TIMSS 2007 target populations. National coverage of the international target population was generally comprehensive, with some exceptions. For example, at the fourth grade (Exhibit 9.3), Georgia (tested only students taught in Georgian), Kazakhstan (students taught in Kazakh or Russian), Latvia (students taught in Latvian), and Lithuania (students taught in Lithuanian) chose a national target population that was less than the international target population. Since coverage was below 100 percent, the results for these countries were footnoted in the TIMSS 2007 international reports. At eighth grade, as shown in Exhibit 9.4, all countries except Georgia (tested only students taught in Georgian), Lithuania (students taught in Lithuanian), and Serbia (did not include Kosovo) sampled from 100 percent of the international target population. Since coverage was below 100 percent for these countries, the results were footnoted in the TIMSS 2007 international reports.

Bulgaria presents an unusual case since its eighth grade exclusion statistics differ between mathematics and science. Because a number of schools in Bulgaria do not teach science at the eighth grade, students sampled in those schools were not administered the science part of the assessment and consequently became part of the excluded population for science. The entries for Bulgaria in eighth grade exhibits in this chapter represent the population of students assessed in mathematics. The figures for science are presented in a footnote.

Exhibit 9.3 Coverage of TIMSS 2007 Target Population – Fourth Grade

Country	International Target Population		Exclusions from National Target Population		
	Coverage	Notes on Coverage	School-level Exclusions	Within-sample Exclusions	Overall Exclusions
Algeria	100%		2.1%	0.0%	2.1%
Armenia	100%		2.7%	0.7%	3.4%
Australia	100%		1.3%	2.7%	4.0%
Austria	100%		1.3%	3.7%	5.0%
Chinese Taipei	100%		0.2%	2.5%	2.8%
Colombia	100%		1.3%	0.8%	2.1%
Czech Republic	100%		4.4%	0.5%	4.9%
Denmark	100%		2.0%	2.1%	4.1%
El Salvador	100%		1.4%	0.9%	2.3%
England	100%		1.6%	0.5%	2.1%
Georgia	85%	Students taught in Georgian	2.3%	2.5%	4.8%
Germany	100%		1.2%	0.2%	1.3%
Hong Kong SAR	100%		4.9%	0.5%	5.4%
Hungary	100%		2.6%	1.7%	4.4%
Iran, Islamic Rep. of	100%		2.9%	0.0%	3.0%
Italy	100%		0.1%	5.3%	5.3%
Japan	100%		0.4%	0.6%	1.1%
Kazakhstan	94%	Students taught in Kazakh or Russian	2.2%	3.1%	5.3%
Kuwait	100%		0.0%	0.0%	0.0%
Latvia	72%	Students taught in Latvian	4.2%	0.4%	4.6%
Lithuania	93%	Students taught in Lithuanian	2.2%	3.1%	5.4%
Morocco	100%		1.4%	0.0%	1.4%
Netherlands	100%		3.7%	1.0%	4.8%
New Zealand	100%		2.8%	2.6%	5.4%
Norway	100%		1.9%	3.3%	5.1%
Qatar	100%		1.5%	0.2%	1.8%
Russian Federation	100%		2.2%	1.5%	3.6%
Scotland	100%		2.6%	1.9%	4.5%
Singapore	100%		1.5%	0.0%	1.5%
Slovak Republic	100%		1.4%	1.9%	3.3%
Slovenia	100%		0.8%	1.3%	2.1%
Sweden	100%		2.0%	1.1%	3.1%
Tunisia	100%		2.7%	0.2%	2.9%
Ukraine	100%		0.6%	0.0%	0.6%
United States	100%		0.0%	9.2%	9.2%
Yemen	100%		1.9%	0.1%	2.0%
Benchmarking Participants					
Alberta, Canada	100%		2.0%	5.7%	7.6%
British Columbia, Canada	100%		2.2%	6.9%	9.2%
Dubai, UAE	100%		4.2%	1.2%	5.4%
Massachusetts, US	100%		0.0%	10.4%	10.4%
Minnesota, US	100%		0.0%	8.3%	8.3%
Ontario, Canada	100%		0.6%	5.7%	6.3%
Quebec, Canada	100%		2.1%	4.3%	6.4%

Exhibit 9.4 Coverage of TIMSS 2007 Target Population – Eighth Grade

Country	International Target Population		Exclusions from National Target Population		
	Coverage	Notes on Coverage	School-level Exclusions	Within-sample Exclusions	Overall Exclusions
Algeria	100%		0.1%	0.0%	0.1%
Armenia	100%		2.7%	0.5%	3.3%
Australia	100%		0.6%	1.2%	1.9%
Bahrain	100%		1.4%	0.1%	1.5%
Bosnia and Herzegovina	100%		0.4%	1.1%	1.5%
Botswana	100%		0.0%	0.1%	0.1%
Bulgaria	100%		2.2%	1.3%	3.4%
Chinese Taipei	100%		0.1%	3.3%	3.3%
Colombia	100%		1.5%	0.1%	1.6%
Cyprus	100%		0.0%	2.5%	2.5%
Czech Republic	100%		4.3%	0.3%	4.6%
Egypt	100%		0.1%	0.4%	0.5%
El Salvador	100%		1.2%	1.6%	2.8%
England	100%		2.0%	0.3%	2.3%
Georgia	85%	Students taught in Georgian	2.3%	1.6%	3.9%
Ghana	100%		0.9%	0.0%	0.9%
Hong Kong SAR	100%		3.7%	0.1%	3.8%
Hungary	100%		2.6%	1.4%	3.9%
Indonesia	100%		3.4%	0.0%	3.4%
Iran, Islamic Rep. of	100%		0.5%	0.0%	0.5%
Israel	100%		14.5%	8.3%	22.8%
Italy	100%		0.0%	4.9%	5.0%
Japan	100%		0.6%	2.9%	3.5%
Jordan	100%		0.2%	1.8%	2.0%
Korea, Rep. of	100%		1.2%	0.5%	1.6%
Kuwait	100%		0.0%	0.3%	0.3%
Lebanon	100%		1.4%	0.0%	1.4%
Lithuania	92%	Students taught in Lithuanian	1.4%	2.7%	4.2%
Malaysia	100%		3.3%	0.0%	3.3%
Malta	100%		0.8%	2.1%	2.9%
Morocco	100%		0.1%	0.0%	0.1%
Norway	100%		0.9%	1.7%	2.6%
Oman	100%		0.3%	0.9%	1.2%
Palestinian Nat'l Auth.	100%		0.1%	0.9%	1.0%
Qatar	100%		0.6%	0.2%	0.8%
Romania	100%		1.5%	0.3%	1.8%
Russian Federation	100%		1.1%	1.2%	2.3%
Saudi Arabia	100%		0.4%	0.1%	0.5%
Scotland	100%		1.3%	0.4%	1.7%
Serbia	80%	Serbia without Kosovo	2.9%	3.9%	6.8%
Singapore	100%		1.8%	0.0%	1.8%
Slovenia	100%		0.9%	1.0%	1.9%
Sweden	100%		2.1%	1.6%	3.6%
Syrian Arab Republic	100%		0.6%	0.0%	0.6%
Thailand	100%		3.4%	0.0%	3.4%
Tunisia	100%		0.0%	0.0%	0.0%
Turkey	100%		2.1%	0.5%	2.6%
Ukraine	100%		0.2%	0.0%	0.2%
United States	100%		0.0%	7.9%	7.9%
Benchmarking Participants					
Basque Country, Spain	100%		1.2%	3.0%	4.2%
British Columbia, Canada	100%		2.8%	15.0%	17.7%
Dubai, UAE	100%		4.2%	0.8%	5.0%
Massachusetts, US	100%		0.0%	8.4%	8.4%
Minnesota, US	100%		0.0%	7.5%	7.5%
Ontario, Canada	100%		0.4%	5.8%	6.2%
Quebec, Canada	100%		1.5%	12.1%	13.6%

Note: In Bulgaria, the figures shown above are for eighth grade mathematics. The figures for the eighth grade science population are as follows: 100%, 2.2%, 18.2%, and 20.3%, respectively.

TIMSS & PIRLS International Study Center
Lynch School of Education, Boston College

Within the national target population, it was possible to exclude certain types of schools, such as very small or very remote schools and certain students, such as those who had a disability that prevented them from participating in the assessment. For the most part, school-level exclusions consisted of schools for students with disabilities and very small or remote schools. Occasionally, schools were excluded for other reasons, as documented in Appendix B. Within-school exclusions generally consisted of students with disabilities or students who could not be assessed in the language of the test (Appendix B gives more details about the exclusions for each participant in TIMSS 2007). For most participants, the overall percentage of excluded students (combining school and within-school levels) was less than 5 percent. However, at fourth grade, the United States along with almost all of the benchmarking participants (the U.S. states of Massachusetts and Minnesota and the Canadian provinces of Québec, Ontario, Alberta, and British Columbia) have exclusions accounting for between 5 and 10 percent of the national target population. At eighth grade, Serbia, the United States, and the U.S. states of Massachusetts and Minnesota, along with the Canadian province of Ontario, have exclusions accounting for between 5 and 10 percent of the national target population. Only Israel and the Canadian provinces of Québec and British Columbia had exclusions exceeding 10 percent. Results for participants with more than 5 percent exclusions were annotated in the international reports. Note that some TIMSS participants had no within-school exclusions.

9.2.3 General Sampling Approach

The basic sample design used in TIMSS 2007 is known as a two-stage stratified cluster design, with the first stage consisting of a sample of schools, and the second stage having a sample of intact classrooms (usually mathematics classes) from the target grades in the sampled schools. While all participants adopted this basic two-stage design, there were some acceptable variations, as follows. The Russian Federation introduced a preliminary stage (first sampling regions). Singapore also added a third sampling stage—subsampling students within classrooms rather than selecting intact classes. Finally, the Basque Country, Spain had a frame of split schools by type (Castilian, Basque, or mixed) and the first stage consisted of a sample of school/type entities rather than schools. As a result, some schools appeared in the sample up to three times (see Section 9.3.1). The reason for this deviation from the general sampling design was to optimize the sampling results by school type.

For countries participating in TIMSS 2007, school stratification was used to enhance the precision of the survey results. Many participants employed explicit stratification, where the complete school sampling frame was divided into smaller sampling frames according to some criterion, such as region, to ensure a predetermined number of schools sampled in each stratum. For example, Australia divided its sampling frame into eight states and territories to ensure equal precision in the survey results between states and between the two territories (see Appendix B for stratification information for each country). Stratification also could be done implicitly, a procedure by which schools in a sampling frame were sorted according to a set of stratification variables prior to sampling. For example, Australia employed implicit stratification by school type (Government, Catholic, Independent) and school location (metropolitan area or elsewhere) within each explicit stratum. Regardless of the other stratification variables used, all countries used implicit stratification by a measure of size of the school.

All countries used a systematic (random start, fixed interval) probability-proportional-to-size (PPS) sampling approach to sample schools. Note that when this method is combined with an implicit stratification procedure, the allocation of schools in the sample is approximately proportional to the size of the implicit strata. Within sampled schools, classes were sampled using a systematic random start method in all countries except Singapore, where classes were sampled with a systematic PPS approach and students within classes were sampled with equal probability using a systematic random start method.

About half of the countries sampled 150 schools, which was the minimum required to meet the TIMSS sampling standards. Most countries sampled one or two classrooms per sampled school. Details on the sampling of schools and classrooms for each country are provided in Appendix B.

The TIMSS 2007 sample designs were implemented in an acceptable manner by all participating countries except Morocco (eighth grade) and Mongolia. Both adopted classroom sampling procedures that did not meet the TIMSS sampling standards and, therefore, could not be approved by the TIMSS & PIRLS International Study Center. For Morocco, schools where the classroom sampling was not implemented correctly were eliminated from the sample, reducing the participation rate. As a result, data for Morocco for eighth grade appear at the bottom of all tables in the international reports. In addition to sampling irregularities, Mongolia had problems implementing and documenting sampling operations in the field. As a result, data for this country were summarized in an appendix to the international reports.

9.2.4 Target Population Sizes

Exhibits 9.5 and 9.6 show the number of schools and students in each participant's target population,[2] based on the sampling frame used to select the TIMSS 2007 sample, as well as the number of sampled schools and students that participated in the study and an estimate of the student population size based on the student sample. The sample figures were derived using sampling weights (see Section 9.3). The population size estimate based on the sampling frame did not take into account the portion of the population excluded within schools and made no adjustment for changes in the population between the date when the information in the sampling frame was collected and the date of the TIMSS 2007 data collection—usually a 2-year interval. Nevertheless, a comparison of the two estimates of the population size can be seen as a check on the sampling procedure. In most cases, the estimated population size closely matched the population size from the sampling frame.

2 After school level exclusions.

Exhibit 9.5 Population and Sample Sizes – Fourth Grade

Country	Population		Sample			Average Age at Time of Testing
	Schools	Students	Schools	Students	Est. Pop.	
Algeria	13,767	719,784	149	4,223	609,356	10.2
Armenia	1,332	55,289	148	4,079	38,614	10.6
Australia	6,755	266,540	229	4,108	233,914	9.9
Austria	3,236	90,422	196	4,859	85,156	10.3
Chinese Taipei	2,512	318,160	150	4,131	308,536	10.2
Colombia	38,591	926,735	142	4,801	946,135	10.4
Czech Republic	3,391	96,768	144	4,235	90,676	10.3
Denmark	1,789	67,179	137	3,519	59,331	11.0
El Salvador	4,558	161,459	148	4,166	146,032	11.0
England	15,304	608,118	143	4,316	578,564	10.2
Georgia	2,059	46,061	144	4,108	46,056	10.1
Germany	18,364	801,257	246	5,200	805,112	10.4
Hong Kong SAR	599	68,244	126	3,791	69,095	10.2
Hungary	2,897	107,693	144	4,048	96,917	10.7
Iran, Islamic Rep. of	47,562	1,248,474	224	3,833	1,081,972	10.2
Italy	7,651	555,976	170	4,470	535,617	9.8
Japan	19,645	1,188,308	148	4,487	1,149,805	10.5
Kazakhstan	6,475	240,140	141	3,990	222,389	10.6
Kuwait	210	27,529	150	3,803	25,721	10.2
Latvia	647	13,448	146	3,908	13,323	11.0
Lithuania	1,135	37,900	156	3,980	33,213	10.8
Morocco	18,526	657,196	184	3,894	600,010	10.6
Netherlands	6,599	186,869	141	3,349	168,143	10.2
New Zealand	1,778	56,372	220	4,940	55,115	10.0
Norway	2,236	60,750	145	4,108	58,011	9.8
Qatar	114	7,190	114	7,019	7,240	9.7
Russian Federation	47,611	1,331,118	206	4,464	1,211,412	10.8
Scotland	1,896	58,071	139	3,929	54,981	9.8
Singapore	177	49,363	177	5,041	49,376	10.4
Slovak Republic	1,998	56,648	184	4,963	53,646	10.4
Slovenia	428	17,576	148	4,351	17,025	9.8
Sweden	3,636	112,057	155	4,676	93,999	10.8
Tunisia	3,939	185,746	150	4,134	175,182	10.2
Ukraine	11,600	368,230	144	4,292	355,822	10.3
United States	72,670	4,049,655	257	7,896	3,367,262	10.3
Yemen	10,835	417,535	144	5,811	414,308	11.2
Benchmarking Participants						
Alberta, Canada	1,060	40,148	146	4,037	35,741	9.8
British Columbia, Canada	1,236	45,723	150	4,153	40,742	9.8
Dubai, UAE	136	13,234	97	3,064	13,597	10.0
Massachusetts, US	1,020	72,459	47	1,747	61,595	10.3
Minnesota, US	949	59,789	50	1,846	51,652	10.3
Ontario, Canada	3,646	152,833	188	3,496	127,754	9.8
Quebec, Canada	1,810	88,710	186	3,885	76,767	10.1

Exhibit 9.6 Population and Sample Sizes – Eighth Grade

Country	Population		Sample			Average Age at Time of Testing
	Schools	Students	Schools	Students	Est. Pop.	
Algeria	3,891	624,353	149	5,447	656,405	14.5
Armenia	1,332	55,289	148	4,689	50,218	14.9
Australia	2,417	270,116	228	4,069	255,699	13.9
Bahrain	74	11,667	74	4,230	11,370	14.1
Bosnia and Herzegovina	569	45,579	150	4,220	37,754	14.7
Botswana	214	40,115	150	4,208	38,859	14.9
Bulgaria	2,309	78,729	163	4,019	74,387	14.9
Chinese Taipei	888	316,997	150	4,046	307,288	14.2
Colombia	10,034	648,634	148	4,873	641,920	14.5
Cyprus	67	9,500	67	4,399	9,237	13.8
Czech Republic	2,669	124,325	147	4,845	115,466	14.4
Egypt	8,179	1,342,127	233	6,582	1,059,228	14.1
El Salvador	2,626	109,671	145	4,063	90,302	15.0
England	3,886	636,732	137	4,025	583,214	14.2
Georgia	2,059	46,061	135	4,178	52,447	14.2
Ghana	7,589	346,289	163	5,294	338,472	15.8
Hong Kong SAR	455	83,267	120	3,470	82,514	14.4
Hungary	2,968	118,049	144	4,111	107,073	14.6
Indonesia	29,701	2,799,024	149	4,203	3,026,953	14.3
Iran, Islamic Rep. of	29,956	1,475,368	208	3,981	1,262,265	14.2
Israel	805	97,132	146	3,294	83,931	14.0
Italy	5,824	602,185	170	4,408	551,089	13.9
Japan	10,708	1,201,082	146	4,312	1,153,745	14.5
Jordan	1,691	108,856	200	5,251	110,338	14.0
Korea, Rep. of	2,727	696,156	150	4,240	683,289	14.3
Kuwait	163	23,827	158	4,091	23,926	14.4
Lebanon	1,574	63,755	136	3,786	59,668	14.4
Lithuania	1,021	49,887	142	3,991	45,023	14.9
Malaysia	1,930	429,048	150	4,466	443,398	14.3
Malta	60	5,260	59	4,670	4,943	14.0
Morocco	1,636	368,656	131	3,060	359,911	14.8
Norway	1,070	62,348	139	4,627	58,806	13.8
Oman	722	56,569	146	4,752	50,834	14.3
Palestinian Nat'l Auth.	1,130	94,376	148	4,378	92,608	14.0
Qatar	67	7,332	66	7,184	7,429	13.9
Romania	6,099	251,054	149	4,198	203,652	15.0
Russian Federation	42,188	2,140,032	210	4,472	1,298,236	14.6
Saudi Arabia	6,271	332,479	165	4,243	370,822	14.4
Scotland	418	64,812	129	4,070	59,252	13.7
Serbia	1,310	81,275	147	4,045	77,540	14.9
Singapore	164	50,904	164	4,599	50,872	14.4
Slovenia	428	19,138	148	4,043	19,066	13.8
Sweden	1,531	125,478	159	5,215	117,344	14.8
Syrian Arab Republic	3,756	270,389	150	4,650	260,481	13.9
Thailand	9,481	844,336	150	5,412	802,663	14.3
Tunisia	804	176,555	150	4,080	169,108	14.5
Turkey	16,112	1,163,830	146	4,498	1,091,653	14.0
Ukraine	12,184	479,467	146	4,424	482,176	14.2
United States	46,112	4,219,262	239	7,377	3,445,599	14.3
Benchmarking Participants						
Basque Country, Spain	330	17,203	130	2,296	15,967	14.1
British Columbia, Canada	433	51,804	150	4,256	41,735	13.9
Dubai, UAE	116	11,178	88	3,195	11,328	14.2
Massachusetts, US	468	75,805	48	1,897	67,333	14.2
Minnesota, US	656	64,566	49	1,777	55,059	14.3
Ontario, Canada	2,854	159,230	176	3,448	143,755	13.8
Quebec, Canada	605	102,112	170	3,956	85,278	14.2

Note: In Bulgaria, the sample for the eighth grade science population is 3,079 students, 139 schools, and the estimated population is 61,237.

9.2.5 Calculating Sampling Weights

The method of estimation used to produce estimates of totals from TIMSS data was through a simple weighted sum of all the responding records for the variable of interest. Estimates of percentages or means then were taken as ratios of these estimated totals. The two-stage stratified cluster PPS design used in TIMSS generally results in differential probabilities of the selection of students, requiring a unique sampling weight for each participating classroom in the study (for Australia and Thailand at grade 8 only, sampling weights varied by student's gender within classrooms—see Section 9.3.7).

The TIMSS 2007 student sampling weight comprised a series of multiplicative components. A basic weight was formed from the inverse of the probability of selecting a student from the population. This basic weight was adjusted by multiplicative factors that account for nonresponding schools, classes, and students.

Sampling weights were calculated according to a three-step procedure involving selection probabilities for schools, classrooms, and students. The first step consisted of calculating a school weight, which also incorporated weighting factors from any additional front-end sampling stages, such as regions for the Russian Federation. A school-level participation adjustment then was made to the school weight to compensate for any sampled schools that did not participate and were not replaced. This adjustment was calculated independently for each explicit stratum.

In the second step, a classroom weight reflecting the probability of the sampled classroom(s) being selected from among all the classrooms in the school at the target grade level was calculated. This classroom weight was calculated independently for each participating school. If a sampled classroom in a school did not participate or if the participation rate among students in a classroom fell below 50 percent, a classroom-level participation adjustment was made to the classroom weight. Note that a classroom participation adjustment only could occur within "participating schools" (a school was considered as a "participating school" if and only if there was at least one sampled classroom with at least 50 percent of its students participating in the study). If one (or more) selected classroom in a school did not participate, the classroom participation adjustment was computed at the explicit stratum level rather than at the school level to reduce the risk of bias.

The third and final step consisted of calculating a student weight. For most TIMSS participants, because intact classrooms were sampled, each student in the sampled classrooms was certain of selection, and, therefore, the student weight was 1.0. In Singapore however, students were further sampled within classrooms, and a student weight reflecting the probability of the sampled students being selected within the classroom was calculated. A nonparticipation adjustment then was made to adjust for sampled students who did not take part in the testing. This adjustment was calculated independently for each sampled classroom.

The basic sampling weight attached to each student record was the product of the three intermediate weights: the first stage (school) weight, the second stage (classroom) weight, and the third stage (student) weight. The overall student sampling weight was the product of the three weights including nonparticipation adjustments.

9.2.6 The First Stage (School) Weight

Essentially, the first stage weight represented the inverse of the probability of a school being sampled on the first stage. The TIMSS 2007 sample design required that school selection probabilities be proportional to school size, generally defined as enrollment in the target grade. The basic first stage weight for the i^{th} sampled school was thus defined as:

$$BW_{sc}^{i} = \frac{M}{n \cdot m_i}$$

where n was the number of sampled schools, m_i was the measure of size for the i^{th} school, and

$$M = \sum_{i=1}^{N} m_i$$

where N was the total number of schools in the explicit stratum.

For the Russian Federation that included a preliminary sampling stage, the basic first stage weight also incorporated the probability of selection in this preliminary stage. The first stage weight in such cases was simply the product of the preliminary stage weight and the first stage weight, as described earlier.

In order to avoid ending up with some basic first stage weights being less than unity, the size of large schools (schools with sizes larger than the sampling interval given by M/n), was set equal to the sampling interval. As a result, these large schools were sampled with equal probability without having to use an explicit stratification approach as in previous TIMSS cycles.

In a similar way but for different reasons, the size of small schools (see Chapter 5) was set to a constant, with the result that these small schools could be sampled with equal probability without having to use explicit stratification.

Finally, because the Basque Country, Spain had school/type entities rather than schools as its first stage sampling units, the probability of school *i* being in the sample was given as follows:

$$P_{sc}^{i} = P_{sc1}^{i} + P_{sc2}^{i} + P_{sc3}^{i} - P_{sc1}^{i}P_{sc2}^{i} - P_{sc1}^{i}P_{sc3}^{i} - P_{sc2}^{i}P_{sc3}^{i} + P_{sc1}^{i}P_{sc2}^{i}P_{sc3}^{i}$$

where $P_{sc1}^{i}, P_{sc2}^{i}, P_{sc3}^{i}$, gives the probability of school *i* being in the sample for the Castilian, mixed, and Basque types, respectively. This probability was computed as shown at the beginning of this section. The sampling school weight for the i^{th} school then becomes $1\big/P_{sc}^{i}$.

9.2.7 School Nonparticipation Adjustment

First stage weights were calculated for all sampled and replacement schools that participated (i.e., those with at least one sampled classroom having at least half of its students participating in the study). A school-level participation adjustment was required to compensate for schools that were sampled but did not participate, and were not replaced. Sampled schools that were found to be ineligible[3] were removed from the calculation of this adjustment. The school-level participation adjustment was calculated separately for each explicit stratum, as follows:

$$A_{sc} = \frac{n_s + n_{r1} + n_{r2} + n_{nr}}{n_s + n_{r1} + n_{r2}}$$

where n_s was the number of originally sampled schools that participated, n_{r1} and n_{r2} the number of first and second replacement schools, respectively, that participated, and n_{nr} was the number of schools that did not participate.

3 A sampled school was ineligible if it was found to contain no eligible students (i.e., fourth grade students). Such schools usually were in the sampling frame by mistake or were schools that had recently closed.

In Bahrain, Cyprus, Kuwait (eighth grade), Malta, and Qatar, because all schools were included in the sample (i.e., census of all schools in the target grades), the following school-level adjustment was used:

$$A_{sc} = \frac{m_s + m_{nr}}{m_s}$$

where m_s was the sum of the measures of size (number of students) from schools that participated and m_{nr} the sum of the measures of size from schools that did not participate.

The final first stage weight for the i^{th} school corrected for nonparticipating schools, thus became:

$$FW_{sc}^{i} = A_{sc} \cdot BW_{sc}^{i}$$

9.2.8 The Second Stage (Classroom) Weight

The second stage weight represented the inverse of the probability of a classroom within a sampled school being selected. All participants except Singapore sampled classrooms within schools with equal probability. In Singapore, where student subsampling was involved, classrooms were sampled using PPS techniques. Procedures for calculating sampling weights are presented below for both approaches.

Equal probability weighting: For the i^{th} school, let C^i be the total number of classrooms and c^i the number of sampled classrooms in the study. Using equal probability sampling, the basic second stage weight assigned to all sampled classrooms in the i^{th} school was:

$$BW_{cl1}^{i} = \frac{C^i}{c^i}$$

For most TIMSS participants, c^i took the values 1, 2, or 3. Some TIMSS participants sampled all classrooms in a selected school.

Probability proportional to size weighting (Singapore only): For the i^{th} school, let $k^{i,j}$ be the size of the j^{th} classroom. Using PPS sampling, the final second stage weight assigned to the j^{th} sampled classroom in the i^{th} school was

$$BW_{cl2}^{i,j} = \frac{K^i}{c^i \cdot k^{i,j}}$$

where c^i was the number of sampled classrooms in the i^{th} school, as defined earlier, and

$$K^i = \sum_{j=1}^{c^i} k^{i,j}$$

Singapore sampled two classrooms per school.

9.2.9 Classroom Nonparticipation Adjustment

Second stage weights were calculated for all sampled classrooms in the sampled and replacement schools that participated. A classroom-level participation adjustment was applied to compensate for classrooms that did not participate or where the student participation rate was below 50 percent. Sampled classrooms with student participation below 50 percent were given a weight of zero and considered to be nonparticipating. The classroom-level participation adjustment was calculated separately for each explicit stratum rather than by school to minimize the risk of bias. The adjustment was calculated as follows:

$$A_{cl} = \frac{\sum_{i}^{s+r1+r2} 1}{\sum_{i}^{s+r1+r2} \delta_i / c^i}$$

where c^i was the number of sampled classrooms in the i^{th} school, as defined earlier, and δ_i gives the number of participating classrooms in the i^{th} school.

When no subsampling of classrooms was involved, the final second stage weight assigned to all sampled classrooms in the i^{th} school became:

$$FW_{cl1}^{i,j} = A_{cl} \cdot BW_{cl1}^{i}$$

When classrooms were subsampled within schools, the final second stage weight assigned to the j^{th} sampled classroom in the i^{th} school became:

$$FW_{cl2}^{i,j} = A_{cl} \cdot BW_{cl2}^{i,j}$$

9.2.10 The Third Stage (Student) Weight

The third stage weight represented the inverse of the probability of a student in a sampled class being selected. In the usual case, when intact classrooms that included all students were sampled, as was the case for all TIMSS 2007 participants except Singapore, this probability was unity. However, countries that participated in TIMSS 2003 and participated in the bridging study assigned some portion of the tested students to the bridging sample. For these countries, the probability fell below unity. In all cases, the third stage weight was calculated independently for each sampled classroom. Procedures for calculating weights are presented below for each case.

Sampling intact classrooms (no bridging study): The basic third stage weight for the j^{th} classroom in the i^{th} school was:

$$BW_{st1}^{i,j} = 1.0$$

Subsampling students (due to bridging study but excluding Singapore): The basic third stage weight for students assigned to the regular TIMSS study for the j^{th} classroom in the i^{th} school was:

$$BW_{st2}^{ij} = \frac{n_{rg}^{i,j} + n_{bs}^{i,j}}{n_{rg}^{i,j}}$$

where $n_{rg}^{i,j}$ was the number of students assigned to the regular TIMSS study in school i and class j and $n_{bs}^{i,j}$ was the number of students assigned to the bridging study.[4] Students who tested for the bridging study were given a weight of zero.

4 Austria did not take part in the study in 2003. However, a portion of their students was assigned to a national study and, therefore, were treated the same way as the bridging study countries.

Subsampling students (Singapore only): The basic third stage weight for the j^{th} classroom in the i^{th} school was:

$$BW_{st3}^{ij} = \frac{k^{ij}}{s^{ij}} \cdot \frac{(n_{rg}^{i,j} + n_{bs}^{i,j})}{n_{rg}^{i,j}}$$

where $k^{i,j}$ was the size of the j^{th} classroom in the i^{th} school, as defined earlier, and $s^{i,j}$ was the number of sampled students per sampled classroom.

9.2.11 Adjustment for Student Nonparticipation

The student nonparticipation adjustment was calculated for each participating classroom and for each of the previously described scenarios.

First two scenarios (sampling intact classrooms or bridging study): The student nonparticipating adjustment, regardless of the participation status to the bridging study, for the j^{th} classroom in the i^{th} school was:

$$A_{st1}^{i,j} = A_{st2}^{i,j} = \frac{s_{rs}^{i,j} + s_{nr}^{i,j}}{s_{rs}^{i,j}}$$

where $s_{rs}^{i,j}$ was the number of responding students (students for which TIMSS scores were derived) in the j^{th} classroom of the i^{th} school, and $s_{nr}^{i,j}$ was the number of students from which a TIMSS score was expected but did not participate in the j^{th} classroom of the i^{th} school.

Third scenario (Singapore only): The student nonparticipating adjustment for the j^{th} classroom in the i^{th} school was:

$$A_{st3}^{i,j} = \frac{\left(s_{nl}^{i,j} + s_{rs}^{i,j} + s_{nr}^{i,j} + s_{ex}^{i,j}\right)}{\left(s_{rs}^{i,j} + s_{nr}^{i,j} + s_{ex}^{i,j}\right)} \cdot \frac{\left(s_{rs}^{i,j} + s_{nr}^{i,j}\right)}{\left(s_{rs}^{i,j}\right)}$$

where $s_{nl}^{i,j}$ was the number of students no longer at school at the time of testing in the j^{th} classroom of the i^{th} school, $s_{ex}^{i,j}$ was the number of excluded students in the j^{th} classroom of the i^{th} school and $s_{rs}^{i,j}$, $s_{nr}^{i,j}$ defined as before.

The third and final stage weight for students in the j^{th} classroom in the i^{th} school thus became

$$FW_{st}^{i,j} = A_{st\Delta}^{i,j} \cdot BW_{st\Delta}^{i,j}$$

where Δ equals 1 when there was no student subsampling, 2 for the bridging study countries except Singapore, and 3 for the Singapore data.

9.2.12 Overall Sampling Weight

The overall sampling weight was simply the product of the final first stage weight, the final second stage weight, and the final third stage weight. For example, for regular TIMSS 2007 study countries, this product is given by

$$W^{i,j} = FW_{sc}^{i} \cdot FW_{cl\Omega}^{i,j} \cdot FW_{st\Delta}^{i,j}$$

where Ω equals 1 when classes were sampled with equal probabilities and 2 otherwise, and Δ equals 1 when there was no student subsampling, 2 for the bridging study countries except Singapore, and 3 for the Singapore data.

It is important to note that with this weighting strategy, sampling weights varied by school and classroom, but participating students within the same classroom have the same sampling weights. However, this weighting strategy did not produce satisfying results for five "areas" (two states in Australia and three regions in Thailand[5]), with regard to the eighth grade student population. In these cases, the student population estimates at eighth grade by gender derived from the sample differed by roughly 10 percent from the actual population figures. A further multiplicative factor for each of these "areas" was thus added to the final weight. This factor was such that the student population estimate by gender would match the known totals for these "areas".

9.3 Calculating School and Student Participation Rates

Since nonparticipation by sampled schools, classrooms, or students can lead to bias in the study results, a variety of participation rates were computed to show the level of success each TIMSS participant achieved in securing participation from their sampled schools, classrooms, and students.

5 These are the states of Queensland and Victoria in Australia and the Bangkok, central, and northern parts of Thailand.

To monitor school participation, two school participation rates were computed: one based on originally sampled schools only and one based on sampled schools and first and second replacements. Classroom and student participation rates were also computed, as were overall participation rates.

9.3.1 Unweighted School Participation Rates

The two unweighted school participation rates that were computed were the following:

R_{unw}^{sc-s} = unweighted school participation rate for originally sampled schools only

R_{unw}^{sc-r} = unweighted school participation rate, including sampled, first, and second replacement schools.

Each unweighted school participation rate was defined as the ratio of the number of participating schools to the number of originally sampled schools, excluding any ineligible schools. A school was labeled as a "participating school" if at least one of its sampled classrooms had at least a 50 percent student participation rate. The rates were calculated as follows:

$$R_{unw}^{sc-s} = \frac{n_s}{n_s + n_{r1} + n_{r2} + n_{nr}}$$

$$R_{unw}^{sc-r} = \frac{n_s + n_{r1} + n_{r2}}{n_s + n_{r1} + n_{r2} + n_{nr}}$$

9.3.2 Unweighted Classroom Participation Rates

The unweighted classroom participation rate was computed as follows:

$$R_{unw}^{cl} = \frac{\sum_{i}^{s+r1+r2} c_*^i}{\sum_{i}^{s+r1+r2} c^i}$$

where c^i was the number of sampled classrooms in the i^{th} school, and c_*^i was the number of participating sampled classrooms in the i^{th} school. Both summations were over all participating schools.

9.3.3 Unweighted Student Participation Rates

The unweighted student participation rate was computed where summations were done over all participating schools and classrooms with at least 50 percent of its students participating in the study, as follows:

$$R_{unw}^{st} = \frac{\sum_{i,j} s_{rs}^{i,j}}{\sum_{i,j} s_{rs}^{i,j} + \sum_{i,j} s_{nr}^{i,j}}$$

9.3.4 Unweighted Overall Participation Rates

Two unweighted overall participation rates were computed for each TIMSS participant. They were as follows:

R_{unw}^{ov-s} = unweighted overall participation rate for originally sampled schools only

R_{unw}^{ov-r} = unweighted overall participation rate, including sampled, first, and second replacement schools.

For each TIMSS participant, the overall participation rate was defined as the product of the unweighted school participation rate, unweighted classroom participation rate, and the unweighted student participation rate. They were calculated as follows:

$$R_{unw}^{ov-s} = R_{unw}^{sc-s} \cdot R_{unw}^{cl} \cdot R_{unw}^{st}$$

$$R_{unw}^{ov-r} = R_{unw}^{sc-r} \cdot R_{unw}^{cl} \cdot R_{unw}^{st}$$

9.3.5 Weighted School Participation Rates

Two weighted school-level participation rates were computed for each TIMSS participant. They were as follows:

R_{wtd}^{sc-s} = weighted school participation rate for originally sampled schools only

R_{wtd}^{sc-r} = weighted school participation rate, including sampled, first, and second replacement schools.

The weighted school participation rates were calculated as follows:

$$R_{wtd}^{sc-s} = \frac{\sum_{i,j}^{s} BW_{sc}^{i} \cdot FW_{cl\Omega}^{i,j} \cdot FW_{st\Delta}^{i,j}}{\sum_{i,j}^{s+r1+r2} FW_{sc}^{i} \cdot FW_{cl\Omega}^{i,j} \cdot FW_{st\Delta}^{i,j}}$$

$$R_{wtd}^{sc-r} = \frac{\sum_{i,j}^{s+r1+r2} BW_{sc}^{i} \cdot FW_{cl\Omega}^{i,j} \cdot FW_{st\Delta}^{i,j}}{\sum_{i,j}^{s+r1+r2} FW_{sc}^{i} \cdot FW_{cl\Omega}^{i,j} \cdot FW_{st\Delta}^{i,j}}$$

where both the numerator and denominator were summations over all responding students and the appropriate classroom- and student-level sampling weights were used. Ω equals 1 when classes were sampled with equal probabilities and 2 otherwise, and Δ equals 1 when there was no student subsampling, 2 for the bridging study countries except Singapore, and 3 for the Singapore data. Note that the basic school-level weight appears in the numerator, whereas the final school-level weight appears in the denominator.

The denominator remains unchanged in all two equations and is the weighted estimate of the total enrollment in the target population. The numerator, however, changes from one equation to the next. Only students from originally sampled schools and from classrooms with at least 50 percent of their students participating in the study were included in the first equation. Students from first and second replacement schools were added in the second equation.

9.3.6 Weighted Classroom Participation Rates

The weighted classroom participation rate was computed as follows:

$$R_{wtd}^{cl} = \frac{\sum_{i,j}^{s+r1+r2} BW_{sc}^{i} \cdot BW_{cl\Omega}^{i,j} \cdot FW_{st\Delta}^{i,j}}{\sum_{i,j}^{s+r1+r2} BW_{sc}^{i} \cdot FW_{cl\Omega}^{i,j} \cdot FW_{st\Delta}^{i,j}}$$

where both the numerator and denominator were summations over all responding students from classrooms with at least 50 percent of their students participating in the study, and the appropriate student-level sampling weights were used. Note that the basic classroom-level weight appears in the numerator, whereas the final classroom-level weight appears in the denominator. Furthermore, the denominator in this formula was the same quantity that appears in the numerator of the weighted school-level participation rate for all participating schools, either sampled or replacement.

9.3.7 Weighted Student Participation Rates

The weighted student participation rate was computed as follows:

$$R_{wtd}^{st} = \frac{\sum_{i,j}^{s+r1+r2} BW_{sc}^{i} \cdot BW_{cl\Omega}^{i,j} \cdot BW_{st\Delta}^{i,j}}{\sum_{i,j}^{s+r1+r2} BW_{sc}^{i} \cdot BW_{cl\Omega}^{i,j} \cdot FW_{st\Delta}^{i,j}}$$

where both the numerator and denominator were summations over all responding students from participating schools. Note that the basic student-level weight appears in the numerator, whereas the final student-level weight appears in the denominator. Furthermore, the denominator in this formula is the same quantity that appears in the numerator of the weighted classroom-level participation rate for all participating schools, either sampled or replacement.

9.3.8 Weighted Overall Participation Rates

Three weighted overall participation rates were computed. They were as follows:

R_{wtd}^{ov-s} = weighted overall participation rate for originally sampled schools only

R_{wtd}^{ov-r} = weighted overall participation rate, including sampled, first and second replacement schools.

Each weighted overall participation rate was defined as the product of the appropriate weighted school participation rate, weighted classroom participation rate, and the weighted student participation rate. They were computed as follows:

$$R_{wtd}^{ov-s} = R_{wtd}^{sc-s} \cdot R_{wtd}^{cl} \cdot R_{wtd}^{st}$$

$$R_{wtd}^{ov-r} = R_{wtd}^{sc-r} \cdot R_{wtd}^{cl} \cdot R_{wtd}^{st}$$

Weighted school, classroom, student, and overall participation rates were computed for each TIMSS participant using these procedures.

9.3.9 Meeting TIMSS' Standards for Sampling Participation

TIMSS participants understood that the goal for sampling participation was 100 percent for all sampled schools, classrooms, and students. Guidelines for reporting achievement data for TIMSS participants securing less than full participation were modeled after IEA's previous studies for TIMSS and PIRLS. As summarized in Exhibit 9.7, countries were assigned to one of three categories on the basis of their sampling participation. Countries in Category 1 were considered to have met the TIMSS 2007 sampling requirement and to have an acceptable participation rate. Countries in Category 2 met the participation requirements only after including replacement schools. Countries that failed to meet the participation requirements even with the use of replacement schools were assigned to Category 3. One of the main goals for quality data in TIMSS 2007 was to have as many countries as possible achieve Category 1 status.

Exhibit 9.7 Categories of Sampling Participation

Category 1	Acceptable sampling **participation** rate without the use of replacement schools. In order to be placed in this category, a country had to have: • An unweighted school response rate without replacement of at least 85% (after rounding to nearest whole percent) AND an unweighted student response rate (after rounding) of at least 85% OR • A **weighted** school response rate **without** replacement of at least 85% (after rounding to nearest whole percent) AND a **weighted** student response rate (after rounding) of at least 85% OR • The product of the (unrounded) **weighted** school response rate **without** replacement and the (unrounded) **weighted** student response rate of at least 75% (after rounding to the nearest whole percent). Countries in this category would appear in the tables and figures in international reports without annotation, and will be ordered by achievement as appropriate.
Category 2	Acceptable sampling participation rate **only when replacement schools are included**. A country would be placed in this category 2 if: • It failed to meet the requirements for Category 1 but had a **weighted** school response rate without replacement of at least 50% (after rounding to the nearest percent) AND HAD EITHER • A **weighted** school response rate **with** replacement of at least 85% (after rounding to nearest whole percent) AND a **weighted** student response rate (after rounding) of at least 85% OR • The product of the (unrounded) **weighted** school response rate **with** replacement and the (unrounded) **weighted** student response rate of at least 75% (after rounding to the nearest whole percent). Countries in this category would be annotated with a "dagger" in the tables and figures in international reports, and ordered by achievement as appropriate.
Category 3	Unacceptable sampling response rate even when replacement schools are included. Countries that could provide documentation to show that they complied with TIMSS sampling procedures and requirements but did not meet the requirements for Category 1 or Category 2 would be placed in Category 3. Countries in this category would appear in a separate section of the achievement tables, below the other countries, in international reports. These countries would be presented in alphabetical order.

Exhibits 9.8 through 9.15 present the school, classroom, student, and overall participation rates and achieved sample sizes for each of the TIMSS 2007 participants. Almost all participants had excellent participation rates and belonged in Category 1. At the fourth grade however, all participants achieved the minimum acceptable participation rates, although Denmark, Scotland, the United States, along with the state of Minnesota, did so only after including replacement schools, and, therefore, their results were annotated with an obelisk in the achievement exhibits in the international reports (Category 2). Despite

efforts to secure full participation, the Netherlands' school participation at 48 percent fell below the minimum requirement of 50 percent before using replacements. However, given that this participation rate increased to 95 percent after using replacements, it was decided during the adjudication that the results for the Netherlands in the international reports would be annotated with a double-obelisk, indicating that they nearly satisfied the guidelines for sample participation rates.

At the eighth grade, England, Hong Kong SAR, Scotland, the United States, and the state of Minnesota met the sampling requirements only after including replacement schools, and, therefore, belonged in Category 2. Morocco with an overall participation rate of 55 percent belonged in Category 3. Mongolia did not provide the necessary documentation for sampling, data collection, and scoring activities. Accordingly, its achievement data were summarized in an appendix to the international reports.

Exhibit 9.8 School Participation Rates and Sample Sizes – Fourth Grade

Country	School Participation Before Replacement (Weighted Percentage)	School Participation After Replacement (Weighted Percentage)	Number of Schools in Original Sample	Number of Eligible Schools in Original Sample	Number of Schools in Original Sample That Participated	Number of Replacement Schools That Participated	Total Number of Schools That Participated
Algeria	99%	99%	150	150	149	0	149
Armenia	93%	100%	150	148	143	5	148
Australia	99%	100%	230	229	226	3	229
Austria	98%	99%	199	197	194	2	196
Chinese Taipei	100%	100%	150	150	150	0	150
Colombia	93%	99%	150	143	132	10	142
Czech Republic	89%	98%	150	147	132	12	144
Denmark	71%	91%	150	150	105	32	137
El Salvador	99%	100%	150	148	146	2	148
England	83%	90%	160	159	131	12	143
Georgia	92%	100%	152	144	131	13	144
Germany	96%	100%	250	247	239	7	246
Hong Kong SAR	81%	84%	150	150	122	4	126
Hungary	93%	99%	150	145	135	9	144
Iran, Islamic Rep. of	100%	100%	240	224	224	0	224
Italy	91%	100%	170	170	155	15	170
Japan	97%	99%	150	150	145	3	148
Kazakhstan	99%	100%	150	141	140	1	141
Kuwait	100%	100%	150	150	149	0	149
Latvia	93%	97%	150	150	140	6	146
Lithuania	99%	100%	163	156	154	2	156
Morocco	81%	81%	226	224	184	0	184
Netherlands	48%	95%	150	148	72	69	141
New Zealand	97%	100%	220	220	213	7	220
Norway	88%	97%	150	150	131	14	145
Qatar	100%	100%	114	114	114	0	114
Russian Federation	100%	100%	206	206	206	0	206
Scotland	77%	94%	150	148	114	25	139
Singapore	100%	100%	177	177	177	0	177
Slovak Republic	98%	100%	184	184	181	3	184
Slovenia	92%	99%	150	150	138	10	148
Sweden	98%	100%	160	155	151	4	155
Tunisia	100%	100%	150	150	150	0	150
Ukraine	96%	96%	150	150	144	0	144
United States	70%	89%	300	290	202	55	257
Yemen	99%	100%	150	144	143	1	144
Benchmarking Participants							
Alberta, Canada	99%	99%	150	148	146	0	146
British Columbia, Canada	98%	100%	150	150	147	3	150
Dubai, UAE	75%	75%	143	132	97	0	97
Massachusetts, US	92%	96%	50	49	45	2	47
Minnesota, US	53%	100%	50	50	30	20	50
Ontario, Canada	95%	96%	200	197	179	9	188
Quebec, Canada	97%	98%	200	192	185	1	186

Exhibit 9.9 School Participation Rates and Sample Sizes – Eighth Grade

Country	School Participation Before Replacement (Weighted Percentage)	School Participation After Replacement (Weighted Percentage)	Number of Schools in Original Sample	Number of Eligible Schools in Original Sample	Number of Schools in Original Sample That Participated	Number of Replacement Schools That Participated	Total Number of Schools That Participated
Algeria	99%	99%	150	150	149	0	149
Armenia	94%	100%	150	148	143	5	148
Australia	100%	100%	230	228	228	0	228
Bahrain	100%	100%	74	74	74	0	74
Bosnia and Herzegovina	100%	100%	150	150	150	0	150
Botswana	100%	100%	150	150	150	0	150
Bulgaria	94%	98%	170	166	158	5	163
Chinese Taipei	100%	100%	150	150	150	0	150
Colombia	96%	100%	150	148	142	6	148
Cyprus	100%	100%	67	67	67	0	67
Czech Republic	92%	100%	150	147	135	12	147
Egypt	99%	100%	237	233	231	2	233
El Salvador	99%	100%	150	145	143	2	145
England	78%	86%	160	160	126	11	137
Georgia	97%	100%	152	135	131	4	135
Ghana	100%	100%	163	163	163	0	163
Hong Kong SAR	73%	79%	152	152	112	8	120
Hungary	92%	99%	150	145	133	11	144
Indonesia	100%	100%	150	149	149	0	149
Iran, Islamic Rep. of	100%	100%	220	208	208	0	208
Israel	94%	97%	150	150	140	6	146
Italy	93%	100%	170	170	159	11	170
Japan	96%	97%	150	150	144	2	146
Jordan	100%	100%	200	200	200	0	200
Korea, Rep. of	100%	100%	150	150	150	0	150
Kuwait	97%	97%	163	163	158	0	158
Lebanon	81%	92%	150	148	120	16	136
Lithuania	98%	99%	150	144	141	1	142
Malaysia	100%	100%	150	150	150	0	150
Malta	100%	100%	60	59	59	0	59
Morocco	65%	65%	205	205	131	0	131
Norway	88%	93%	150	150	133	6	139
Oman	100%	100%	150	146	146	0	146
Palestinian Nat'l Auth.	100%	100%	155	148	147	1	148
Qatar	100%	100%	67	67	66	0	66
Romania	99%	99%	150	150	149	0	149
Russian Federation	100%	100%	210	210	210	0	210
Saudi Arabia	99%	99%	167	166	165	0	165
Scotland	74%	86%	150	150	109	20	129
Serbia	100%	100%	150	147	147	0	147
Singapore	100%	100%	164	164	164	0	164
Slovenia	92%	99%	150	150	138	10	148
Sweden	100%	100%	160	159	158	1	159
Syrian Arab Republic	100%	100%	150	150	150	0	150
Thailand	90%	100%	150	150	134	16	150
Tunisia	100%	100%	150	150	150	0	150
Turkey	100%	100%	150	146	146	0	146
Ukraine	98%	98%	150	150	146	0	146
United States	68%	83%	300	287	197	42	239
Benchmarking Participants							
Basque Country, Spain	100%	100%	130	130	130	0	130
British Columbia, Canada	98%	100%	150	150	147	3	150
Dubai, UAE	79%	79%	122	115	88	0	88
Massachusetts, US	93%	98%	50	49	45	3	48
Minnesota, US	61%	98%	50	50	32	17	49
Ontario, Canada	90%	94%	200	191	168	8	176
Quebec, Canada	93%	93%	191	183	170	0	170

Note: In Bulgaria, the figures shown above are for eighth grade mathematics. The figures for the eighth grade science population are as follows: 93%, 98%, 170, 142, 134, 5, and 139, respectively.

TIMSS & PIRLS International Study Center
Lynch School of Education, Boston College

Exhibit 9.10 Student Participation Rates and Sample Sizes – Fourth Grade

Country	Within School Student Participation (Weighted Percentage)	Number of Sampled Students in Participating Schools	Number of Students Withdrawn from Class/ School	Number of Students Excluded	Number of Students Eligible	Number of Students Absent	Number of Students Assessed
Algeria	97%	4,366	22	0	4,344	121	4,223
Armenia	96%	4,253	0	0	4,253	174	4,079
Australia	95%	4,511	78	105	4,328	220	4,108
Austria	98%	5,158	18	156	4,984	125	4,859
Chinese Taipei	100%	4,260	17	93	4,150	19	4,131
Colombia	98%	5,320	349	40	4,931	130	4,801
Czech Republic	94%	4,583	41	17	4,525	290	4,235
Denmark	94%	3,907	59	89	3,759	240	3,519
El Salvador	98%	4,467	202	0	4,265	99	4,166
England	93%	4,784	128	33	4,623	307	4,316
Georgia	98%	4,384	69	68	4,247	139	4,108
Germany	97%	5,464	78	9	5,377	177	5,200
Hong Kong SAR	96%	3,965	13	23	3,929	138	3,791
Hungary	97%	4,221	22	26	4,173	125	4,048
Iran, Islamic Rep. of	99%	3,939	53	2	3,884	51	3,833
Italy	97%	4,912	20	256	4,636	166	4,470
Japan	97%	4,677	7	20	4,650	163	4,487
Kazakhstan	100%	4,063	22	39	4,002	12	3,990
Kuwait	85%	4,468	439	0	4,029	226	3,803
Latvia	95%	4,188	2	10	4,176	268	3,908
Lithuania	94%	4,345	15	122	4,208	228	3,980
Morocco	96%	4,282	215	0	4,067	173	3,894
Netherlands	97%	3,608	152	9	3,447	98	3,349
New Zealand	96%	5,347	104	86	5,157	217	4,940
Norway	95%	4,462	21	143	4,298	190	4,108
Qatar	97%	7,411	153	18	7,240	221	7,019
Russian Federation	98%	4,659	36	42	4,581	117	4,464
Scotland	94%	4,320	92	32	4,196	267	3,929
Singapore	96%	5,235	26	1	5,208	167	5,041
Slovak Republic	97%	5,269	47	64	5,158	195	4,963
Slovenia	95%	4,664	10	57	4,597	246	4,351
Sweden	97%	4,965	60	49	4,856	180	4,676
Tunisia	99%	4,242	50	10	4,182	48	4,134
Ukraine	97%	4,459	16	0	4,443	151	4,292
United States	95%	9,000	140	543	8,317	421	7,896
Yemen	98%	6,128	180	8	5,940	129	5,811
Benchmarking Participants							
Alberta, Canada	96%	4,557	105	222	4,230	193	4,037
British Columbia, Canada	96%	4,758	67	342	4,349	196	4,153
Dubai, UAE	91%	3,421	19	4	3,398	334	3,064
Massachusetts, US	96%	1,971	11	136	1,824	77	1,747
Minnesota, US	97%	2,034	23	101	1,910	64	1,846
Ontario, Canada	95%	3,903	34	194	3,675	179	3,496
Quebec, Canada	86%	4,645	34	78	4,533	648	3,885

Exhibit 9.11 Student Participation Rates and Sample Sizes - Eighth Grade

Country	Within School Student Participation (Weighted Percentage)	Number of Sampled Students in Participating Schools	Number of Students Withdrawn from Class/ School	Number of Students Excluded	Number of Students Eligible	Number of Students Absent	Number of Students Assessed
Algeria	96%	5,793	83	0	5,710	263	5,447
Armenia	96%	4,898	0	0	4,898	209	4,689
Australia	93%	4,549	84	37	4,428	359	4,069
Bahrain	97%	4,434	61	5	4,368	138	4,230
Bosnia and Herzegovina	98%	4,373	22	44	4,307	87	4,220
Botswana	99%	4,310	63	2	4,245	37	4,208
Bulgaria	96%	4,312	87	7	4,218	199	4,019
Chinese Taipei	99%	4,164	25	53	4,086	40	4,046
Colombia	98%	5,343	368	4	4,971	98	4,873
Cyprus	96%	4,755	41	139	4,575	176	4,399
Czech Republic	95%	5,182	41	12	5,129	284	4,845
Egypt	98%	6,906	151	1	6,754	172	6,582
El Salvador	98%	4,329	191	0	4,138	75	4,063
England	88%	4,768	153	15	4,600	575	4,025
Georgia	97%	4,533	139	48	4,346	168	4,178
Ghana	98%	5,678	270	0	5,408	114	5,294
Hong Kong SAR	96%	3,657	29	2	3,626	156	3,470
Hungary	97%	4,321	21	30	4,270	159	4,111
Indonesia	97%	4,419	95	0	4,324	121	4,203
Iran, Islamic Rep. of	98%	4,140	95	0	4,045	64	3,981
Israel	94%	3,708	12	183	3,513	219	3,294
Italy	96%	4,873	40	231	4,602	194	4,408
Japan	93%	4,656	31	6	4,619	307	4,312
Jordan	96%	5,733	184	88	5,461	210	5,251
Korea, Rep. of	99%	4,358	36	19	4,303	63	4,240
Kuwait	87%	4,721	381	18	4,322	231	4,091
Lebanon	93%	4,062	0	0	4,062	276	3,786
Lithuania	91%	4,537	35	96	4,406	415	3,991
Malaysia	98%	4,589	33	0	4,556	90	4,466
Malta	95%	5,053	18	106	4,929	259	4,670
Morocco	85%	4,758	173	0	4,585	649	3,936
Norway	93%	5,085	17	78	4,990	363	4,627
Oman	99%	4,894	57	36	4,801	49	4,752
Palestinian Nat'l Auth.	98%	4,572	70	29	4,473	95	4,378
Qatar	97%	7,558	128	17	7,413	229	7,184
Romania	97%	4,447	119	12	4,316	118	4,198
Russian Federation	97%	4,706	42	51	4,613	141	4,472
Saudi Arabia	95%	4,515	1	3	4,511	268	4,243
Scotland	90%	4,700	137	19	4,544	474	4,070
Serbia	98%	4,246	16	78	4,152	107	4,045
Singapore	95%	4,828	37	0	4,791	192	4,599
Slovenia	93%	4,414	10	42	4,362	319	4,043
Sweden	94%	5,712	87	58	5,567	352	5,215
Syrian Arab Republic	96%	5,025	199	0	4,826	176	4,650
Thailand	99%	5,579	89	0	5,490	78	5,412
Tunisia	98%	4,258	84	0	4,174	94	4,080
Turkey	98%	4,682	87	19	4,576	78	4,498
Ukraine	97%	4,598	27	0	4,571	147	4,424
United States	93%	8,447	202	272	7,973	596	7,377
Benchmarking Participants							
Basque Country, Spain	98%	2,481	46	83	2,352	56	2,296
British Columbia, Canada	94%	4,836	129	146	4,561	305	4,256
Dubai, UAE	88%	3,625	17	6	3,602	407	3,195
Massachusetts, US	94%	2,093	23	56	2,014	117	1,897
Minnesota, US	95%	1,988	21	82	1,885	108	1,777
Ontario, Canada	95%	3,842	43	171	3,628	180	3,448
Quebec, Canada	85%	4,739	59	45	4,635	679	3,956

Note: In Bulgaria, the figures shown above are for eighth grade mathematics. The figures for the eighth grade science population are as follows: 96%; 3,426; 69; 124; 3,233; 154; and 3,079, respectively.

Exhibit 9.12 Unweighted school, Class, and Student Participation Rates – Fourth Grade

Country	School Participation Before Replacement	School Participation After Replacement	Class Participation	Student Participation	Overall Participation Before Replacement	Overall Participation After Replacement
Algeria	99%	99%	100%	97%	97%	97%
Armenia	97%	100%	100%	96%	93%	96%
Australia	99%	100%	100%	95%	94%	95%
Austria	98%	99%	99%	97%	95%	96%
Chinese Taipei	100%	100%	100%	100%	100%	100%
Colombia	92%	99%	100%	97%	90%	97%
Czech Republic	90%	98%	100%	94%	84%	92%
Denmark	70%	91%	99%	94%	65%	85%
El Salvador	99%	100%	100%	98%	96%	98%
England	82%	90%	100%	93%	77%	84%
Georgia	91%	100%	100%	97%	88%	97%
Germany	97%	100%	100%	97%	94%	96%
Hong Kong SAR	81%	84%	100%	96%	78%	81%
Hungary	93%	99%	100%	97%	90%	96%
Iran, Islamic Rep. of	100%	100%	100%	99%	99%	99%
Italy	91%	100%	100%	96%	88%	96%
Japan	97%	99%	100%	96%	93%	95%
Kazakhstan	99%	100%	100%	100%	99%	100%
Kuwait	99%	99%	100%	85%	85%	85%
Latvia	93%	97%	100%	94%	87%	91%
Lithuania	99%	100%	100%	95%	93%	95%
Morocco	82%	82%	100%	96%	79%	79%
Netherlands	49%	95%	97%	97%	46%	90%
New Zealand	97%	100%	100%	96%	93%	96%
Norway	87%	97%	100%	96%	83%	92%
Qatar	100%	100%	100%	97%	97%	97%
Russian Federation	100%	100%	100%	97%	97%	97%
Scotland	77%	94%	100%	94%	72%	88%
Singapore	100%	100%	100%	97%	97%	97%
Slovak Republic	98%	100%	100%	96%	95%	96%
Slovenia	92%	99%	100%	95%	87%	93%
Sweden	97%	100%	100%	96%	94%	96%
Tunisia	100%	100%	100%	99%	99%	99%
Ukraine	96%	96%	100%	97%	93%	93%
United States	70%	89%	100%	95%	66%	84%
Yemen	99%	100%	100%	98%	97%	98%
Benchmarking Participants						
Alberta, Canada	99%	99%	100%	95%	94%	94%
British Columbia, Canada	98%	100%	100%	95%	94%	95%
Dubai, UAE	73%	73%	97%	90%	64%	64%
Massachusetts, US	92%	96%	100%	96%	88%	92%
Minnesota, US	60%	100%	100%	97%	58%	97%
Ontario, Canada	91%	95%	100%	95%	86%	91%
Quebec, Canada	96%	97%	99%	86%	82%	82%

Exhibit 9.13 Unweighted School, Class, and Student Participation Rates – Eighth Grade

Country	School Participation Before Replacement	School Participation After Replacement	Class Participation	Student Participation	Overall Participation Before Replacement	Overall Participation After Replacement
Algeria	99%	99%	100%	95%	95%	95%
Armenia	97%	100%	100%	96%	92%	96%
Australia	100%	100%	100%	92%	92%	92%
Bahrain	100%	100%	100%	97%	97%	97%
Bosnia and Herzegovina	100%	100%	100%	98%	98%	98%
Botswana	100%	100%	100%	99%	99%	99%
Bulgaria	95%	98%	100%	95%	91%	94%
Chinese Taipei	100%	100%	100%	99%	99%	99%
Colombia	96%	100%	100%	98%	94%	98%
Cyprus	100%	100%	100%	96%	96%	96%
Czech Republic	92%	100%	100%	94%	87%	94%
Egypt	99%	100%	100%	97%	97%	97%
El Salvador	99%	100%	100%	98%	96%	98%
England	79%	86%	100%	88%	69%	75%
Georgia	97%	100%	100%	96%	93%	96%
Ghana	100%	100%	100%	98%	98%	98%
Hong Kong SAR	74%	79%	100%	96%	71%	76%
Hungary	92%	99%	100%	96%	88%	96%
Indonesia	100%	100%	100%	97%	97%	97%
Iran, Islamic Rep. of	100%	100%	100%	98%	98%	98%
Israel	93%	97%	100%	94%	88%	91%
Italy	94%	100%	100%	96%	89%	95%
Japan	96%	97%	100%	93%	90%	91%
Jordan	100%	100%	100%	96%	96%	96%
Korea, Rep. of	100%	100%	100%	99%	99%	99%
Kuwait	97%	97%	100%	87%	84%	84%
Lebanon	81%	92%	100%	93%	76%	86%
Lithuania	98%	99%	100%	91%	89%	89%
Malaysia	100%	100%	100%	98%	98%	98%
Malta	100%	100%	100%	95%	94%	94%
Morocco	63%	63%	100%	86%	54%	54%
Norway	89%	93%	100%	93%	82%	86%
Oman	100%	100%	100%	99%	99%	99%
Palestinian Nat'l Auth.	99%	100%	100%	98%	97%	98%
Qatar	99%	99%	100%	97%	95%	95%
Romania	99%	99%	100%	97%	97%	97%
Russian Federation	100%	100%	100%	97%	97%	97%
Saudi Arabia	99%	99%	100%	94%	93%	93%
Scotland	73%	86%	100%	90%	65%	77%
Serbia	100%	100%	100%	97%	97%	97%
Singapore	100%	100%	99%	96%	95%	95%
Slovenia	92%	99%	100%	93%	85%	91%
Sweden	99%	100%	100%	94%	93%	94%
Syrian Arab Republic	100%	100%	100%	96%	96%	96%
Thailand	89%	100%	100%	99%	88%	99%
Tunisia	100%	100%	100%	98%	98%	98%
Turkey	100%	100%	100%	98%	98%	98%
Ukraine	97%	97%	100%	97%	94%	94%
United States	69%	83%	99%	93%	63%	77%
Benchmarking Participants						
Basque Country, Spain	100%	100%	100%	98%	98%	98%
British Columbia, Canada	98%	100%	100%	93%	91%	93%
Dubai, UAE	77%	77%	99%	89%	67%	67%
Massachusetts, US	92%	98%	100%	94%	87%	92%
Minnesota, US	64%	98%	100%	94%	60%	92%
Ontario, Canada	88%	92%	100%	95%	84%	88%
Quebec, Canada	93%	93%	96%	85%	76%	76%

Note: In Bulgaria, the figures shown above are for eighth grade mathematics. The figures for the eighth grade science population are as follows: 94%, 98%, 100%, 95%, 90%, and 93%, respectively.

Exhibit 9.14 Weighted School, Class, and Student Participation Rates – Fourth Grade

Country	School Participation Before Replacement	School Participation After Replacement	Class Participation	Student Participation	Overall Participation Before Replacement	Overall Participation After Replacement
Algeria	99%	99%	100%	97%	97%	97%
Armenia	93%	100%	100%	96%	90%	96%
Australia	99%	100%	100%	95%	94%	95%
Austria	98%	99%	99%	98%	96%	97%
Chinese Taipei	100%	100%	100%	100%	100%	100%
Colombia	93%	99%	100%	98%	91%	97%
Czech Republic	89%	98%	100%	94%	83%	92%
Denmark	71%	91%	99%	94%	66%	85%
El Salvador	99%	100%	100%	98%	97%	98%
England	83%	90%	100%	93%	77%	84%
Georgia	92%	100%	100%	98%	90%	98%
Germany	96%	100%	100%	97%	93%	96%
Hong Kong SAR	81%	84%	100%	96%	78%	81%
Hungary	93%	99%	100%	97%	90%	96%
Iran, Islamic Rep. of	100%	100%	100%	99%	99%	99%
Italy	91%	100%	100%	97%	88%	97%
Japan	97%	99%	100%	97%	94%	95%
Kazakhstan	99%	100%	100%	100%	99%	100%
Kuwait	100%	100%	100%	85%	85%	85%
Latvia	93%	97%	100%	95%	89%	92%
Lithuania	99%	100%	100%	94%	93%	94%
Morocco	81%	81%	100%	96%	77%	77%
Netherlands	48%	95%	98%	97%	46%	91%
New Zealand	97%	100%	100%	96%	93%	96%
Norway	88%	97%	100%	95%	83%	92%
Qatar	100%	100%	100%	97%	97%	97%
Russian Federation	100%	100%	100%	98%	98%	98%
Scotland	77%	94%	100%	94%	72%	88%
Singapore	100%	100%	100%	96%	96%	96%
Slovak Republic	98%	100%	100%	97%	95%	97%
Slovenia	92%	99%	100%	95%	87%	93%
Sweden	98%	100%	100%	97%	94%	97%
Tunisia	100%	100%	100%	99%	99%	99%
Ukraine	96%	96%	100%	97%	93%	93%
United States	70%	89%	100%	95%	66%	84%
Yemen	99%	100%	100%	98%	97%	98%
Benchmarking Participants						
Alberta, Canada	99%	99%	100%	96%	94%	94%
British Columbia, Canada	98%	100%	100%	96%	94%	96%
Dubai, UAE	75%	75%	98%	91%	67%	67%
Massachusetts, US	92%	96%	100%	96%	88%	92%
Minnesota, US	53%	100%	100%	97%	52%	97%
Ontario, Canada	95%	96%	100%	95%	91%	92%
Quebec, Canada	97%	98%	100%	86%	83%	84%

Exhibit 9.15 Weighted School, Class, and Student Participation Rates – Eighth Grade

Country	School Participation Before Replacement	School Participation After Replacement	Class Participation	Student Participation	Overall Participation Before Replacement	Overall Participation After Replacement
Algeria	99%	99%	100%	96%	95%	95%
Armenia	94%	100%	100%	96%	90%	96%
Australia	100%	100%	100%	93%	93%	93%
Bahrain	100%	100%	100%	97%	97%	97%
Bosnia and Herzegovina	100%	100%	100%	98%	98%	98%
Botswana	100%	100%	100%	99%	99%	99%
Bulgaria	94%	98%	100%	96%	90%	94%
Chinese Taipei	100%	100%	100%	99%	99%	99%
Colombia	96%	100%	100%	98%	94%	98%
Cyprus	100%	100%	100%	96%	96%	96%
Czech Republic	92%	100%	100%	95%	87%	95%
Egypt	99%	100%	100%	98%	97%	98%
El Salvador	99%	100%	100%	98%	97%	98%
England	78%	86%	100%	88%	69%	75%
Georgia	97%	100%	100%	97%	95%	97%
Ghana	100%	100%	100%	98%	98%	98%
Hong Kong SAR	73%	79%	100%	96%	70%	75%
Hungary	92%	99%	100%	97%	89%	96%
Indonesia	100%	100%	100%	97%	97%	97%
Iran, Islamic Rep. of	100%	100%	100%	98%	98%	98%
Israel	94%	97%	100%	94%	88%	91%
Italy	93%	100%	100%	96%	89%	96%
Japan	96%	97%	100%	93%	90%	91%
Jordan	100%	100%	100%	96%	96%	96%
Korea, Rep. of	100%	100%	100%	99%	99%	99%
Kuwait	97%	97%	100%	87%	84%	84%
Lebanon	81%	92%	100%	93%	76%	85%
Lithuania	98%	99%	100%	91%	89%	90%
Malaysia	100%	100%	100%	98%	98%	98%
Malta	100%	100%	100%	95%	94%	94%
Morocco	65%	65%	100%	85%	55%	55%
Norway	88%	93%	100%	93%	82%	86%
Oman	100%	100%	100%	99%	99%	99%
Palestinian Nat'l Auth.	100%	100%	100%	98%	98%	98%
Qatar	100%	100%	100%	97%	97%	97%
Romania	99%	99%	100%	97%	97%	97%
Russian Federation	100%	100%	100%	97%	97%	97%
Saudi Arabia	99%	99%	100%	95%	94%	94%
Scotland	74%	86%	100%	90%	66%	77%
Serbia	100%	100%	100%	98%	98%	98%
Singapore	100%	100%	99%	95%	95%	95%
Slovenia	92%	99%	100%	93%	85%	92%
Sweden	100%	100%	100%	94%	93%	94%
Syrian Arab Republic	100%	100%	100%	96%	96%	96%
Thailand	90%	100%	100%	99%	88%	99%
Tunisia	100%	100%	100%	98%	98%	98%
Turkey	100%	100%	100%	98%	98%	98%
Ukraine	98%	98%	100%	97%	95%	95%
United States	68%	83%	99%	93%	63%	77%
Benchmarking Participants						
Basque Country, Spain	100%	100%	100%	98%	98%	98%
British Columbia, Canada	98%	100%	100%	94%	92%	94%
Dubai, UAE	79%	79%	99%	88%	69%	69%
Massachusetts, US	93%	98%	100%	94%	88%	92%
Minnesota, US	61%	98%	100%	95%	58%	93%
Ontario, Canada	90%	94%	100%	95%	86%	89%
Quebec, Canada	93%	93%	97%	85%	77%	77%

Note: In Bulgaria, the figures shown above are for eighth grade mathematics. The figures for the eighth grade science population are as follows: 93%, 98%, 100%, 96%, 89%, and 94%, respectively.

9.4 Trends in Student Populations

Because an important goal of the TIMSS 2007 assessment was to measure changes in students' mathematics achievement since 1995, it was important to track any changes in population composition and coverage since then that might be related to student achievement. Exhibits 9.16 and 9.17 present, for each TIMSS participant, four attributes of the fourth grade populations sampled in 2007, 2003, and 1995 and the eighth grade populations sampled in 2007, 2003, 1999, and 1995: number of years of formal schooling, average student age at time of testing, percentage of students excluded from the assessment, and overall sampling participation rate (after replacement). Most countries and provinces were very similar with regard to these attributes across the three TIMSS cycles at fourth grade and four cycles at eighth grade, although there have been changes in some countries in the age and grade structure of the assessed populations, and in the exclusion rate.

Although Australia, since 2003, has tested only fourth grade students for the fourth grade population and only eighth grade students for the eighth grade population, in 1995 the younger assessment population contained fourth grade students from some states and fifth grade students from other states, and similarly the older population contained a mixture of eighth and ninth grade students. Because of this, Australian students were somewhat older, on average, in 1995. The Russian Federation and Slovenia have undergone structural changes in the age at which children enter schools that are reflected in their samples. In 2003, the Russian fourth grade sample contained third-grade students from some regions and fourth-grade students from others, whereas all students were in fourth grade in 2007. At the eighth grade, there was still a mixture of seventh and eighth grade students in 2007, although with proportionally more eighth grade students, and correspondingly a higher average age. Slovenia is in transition towards having all children begin school at an earlier age so that they all will have four years of primary schooling at the fourth grade instead of three years, as was the case in 2003. At eighth grade, the transition was not complete in 2007.

In general, the exclusion rates do not exceed the TIMSS 2007 guidelines of 5 percent, and have not changed very much across assessments for most countries. Also, in most cases, the exclusion rates have decreased. However, the student exclusion rate was higher in 2007 than in previous assessments at fourth grade in the United States, the state of Minnesota, and the provinces of Alberta and Quebec, and at eighth grade in Serbia, the United States, and the Canadian provinces of British Columbia and Quebec.

Exhibit 9.16 Trends in Student Populations – Fourth Grade

Country	Years of Formal Schooling*			Average Age at Time of Testing			Overall Exclusion Rates			Overall Participation Rates (After Replacement)		
	2007	2003	1995	2007	2003	1995	2007	2003	1995	2007	2003	1995
Armenia	4	4		10.6	10.9		3.4%	2.9%		96%	90%	
Australia	4	4	4 or 5	9.9	9.9	10.2	4.0%	2.7%	1.8%	95%	85%	66%
Austria	4		4	10.3		10.5	5.0%		2.8%	97%		69%
Chinese Taipei	4	4		10.2	10.2		2.8%	3.1%		100%	99%	
Czech Republic	4		4	10.3		10.4	4.9%		4.1%	92%		86%
England	5	5	5	10.2	10.3	10.0	2.1%	1.9%	12.1%	84%	76%	83%
Hong Kong SAR	4	4	4	10.2	10.2	10.1	5.4%	3.8%	2.7%	81%	83%	83%
Hungary	4	4	4	10.7	10.5	10.4	4.4%	8.1%	3.8%	96%	93%	92%
Iran, Islamic Rep. of	4	4	4	10.2	10.4	10.5	3.0%	5.7%	1.3%	99%	98%	97%
Italy	4	4		9.8	9.8		5.3%	4.2%		97%	97%	
Japan	4	4	4	10.5	10.4	10.4	1.1%	0.8%	3.0%	95%	97%	92%
Latvia	4	4	4	11.0	11.1	10.5	4.6%	4.4%	2.1%	92%	88%	69%
Lithuania	4	4		10.8	10.9		5.4%	4.6%		94%	87%	
Morocco	4	4		10.6	11.0		1.4%	2.2%		77%	81%	
Netherlands	4	4	4	10.2	10.2	10.3	4.8%	5.2%	4.4%	91%	84%	59%
New Zealand	4.5 – 5.5	4.5 – 5.5	4.5 – 5.5	10.0	10.0	10.0	5.4%	4.0%	1.3%	96%	93%	95%
Norway	4	4	4	9.8	9.8	9.9	5.1%	4.4%	3.1%	92%	88%	91%
Russian Federation	4	3 or 4		10.8	10.6		3.6%	6.8%		98%	97%	
Scotland	5	5	5	9.8	9.7	9.7	4.5%	1.5%	6.7%	88%	77%	76%
Singapore	4	4	4	10.4	10.3	10.3	1.5%	0.0%	0.0%	96%	98%	98%
Slovenia	4	3 or 4	3	9.8	9.8	9.9	2.1%	1.3%	1.9%	93%	91%	77%
Tunisia	4	4		10.2	10.4		2.9%	0.9%		99%	99%	
United States	4	4	4	10.3	10.2	10.2	9.2%	5.1%	4.7%	84%	78%	80%
Benchmarking Participants												
Alberta, Canada	4		4	9.8		10.0	7.6%		–	94%		91%
Minnesota, US	4		4	10.3		10.3	8.3%		–	97%		–
Ontario, Canada	4	4	4	9.8	9.8	9.9	6.3%	4.8%	–	92%	90%	92%
Quebec, Canada	4	4	4	10.1	10.1	10.3	6.4%	3.6%	–	84%	91%	81%

* Represents years of schooling counting from the first year of ISCED Level 1.
A dash (–) indicates comparable data are not available.

Exhibit 9.17 Trends in Student Populations – Eighth Grade

Country	Years of Formal Schooling*				Average Age at Time of Testing			
	2007	2003	1999	1995	2007	2003	1999	1995
Armenia	8	8			14.9	14.9		
Australia	8	8		8 or 9	13.9	13.9		14.2
Bahrain	8	8			14.1	14.1		
Botswana	8	8			14.9	15.1		
Bulgaria	8	8	8	8	14.9	14.9	14.8	14.0
Chinese Taipei	8	8	8		14.2	14.2	14.2	
Colombia	8			8	14.5			14.5
Cyprus	8	8	8	8	13.8	13.8	13.8	13.7
Czech Republic	8		8	8	14.4		14.4	14.4
Egypt	8	8			14.1	14.4		
England	9	9	9	9	14.2	14.3	14.2	14.0
Ghana	8	8			15.8	15.5		
Hong Kong SAR	8	8	8	8	14.4	14.4	14.2	14.2
Hungary	8	8	8	8	14.6	14.5	14.4	14.3
Indonesia	8	8	8		14.3	14.5	14.6	
Iran, Islamic Rep. of	8	8	8	8	14.2	14.4	14.6	14.6
Israel	8	8	8		14.0	14.0	14.1	
Italy	8	8	8		13.9	13.9	14.0	
Japan	8	8	8	8	14.5	14.4	14.4	14.4
Jordan	8	8	8		14.0	13.9	14.0	
Korea, Rep. of**	8	8	8	8	14.3	14.6	14.4	14.2
Lebanon	8	8			14.4	14.6		
Lithuania**	8	8	8.5	8	14.9	14.9	15.2	14.3
Malaysia	8	8	8		14.3	14.3	14.4	
Norway	8	8		8	13.8	13.8		13.9
Palestinian Nat'l Auth.	8	8			14.0	14.1		
Romania	8	8	8	8	15.0	15.0	14.8	14.6
Russian Federation	7 or 8	7 or 8	7 or 8	7 or 8	14.6	14.2	14.1	14.0
Scotland	9	9		9	13.7	13.7		13.7
Serbia	8	8			14.9	14.9		
Singapore	8	8	8	8	14.4	14.3	14.4	14.5
Slovenia	7 or 8	7 or 8		7	13.8	13.8		13.8
Sweden	8	8		8	14.8	14.9		14.9
Thailand	8		8		14.3		14.5	
Tunisia	8	8	8		14.5	14.8	14.8	
United States	8	8	8	8	14.3	14.2	14.2	14.2
Benchmarking Participants								
Basque Country, Spain	8	8			14.1	14.1		
British Columbia, Canada	8		8		13.9		13.9	
Massachusetts, US	8		8		14.2		14.1	
Minnesota, US	8			8	14.3			14.3
Ontario, Canada	8	8	8	8	13.8	13.8	13.9	14.0
Quebec, Canada	8	8	8	8	14.2	14.2	14.3	14.5

Note: In Bulgaria, the figures refer to the eighth grade mathematics population. Trends are not reported for their science population.

* Represents years of schooling counting from the first year of ISCED Level 1.

** Lithuania tested the same cohort of students as other countries, but later in 1999, at the beginning of the next school year. Korea tested the same cohort of students as other countries, but later in 2003, at the beginning of the next school year.

A dash (–) indicates comparable data are not available.

Exhibit 9.17 Trends in Student Populations – Eighth Grade (Continued)

Country	Overall Exclusion Rates				Overall Participation Rates (After Replacement)			
	2007	2003	1999	1995	2007	2003	1999	1995
Armenia	3.3%	2.9%			96%	89%		
Australia	1.9%	1.3%		0.8%	93%	83%		70%
Bahrain	1.5%	0.0%			97%	98%		
Botswana	0.1%	3.0%			99%	96%		
Bulgaria	3.4%	0.5%	4.6%	0.6%	94%	92%	84%	63%
Chinese Taipei	3.3%	4.8%	1.6%		99%	99%	93%	
Colombia	1.6%			3.8%	98%			86%
Cyprus	2.5%	2.5%	0.8%	0.0%	96%	96%	97%	97%
Czech Republic	4.6%		5.2%	4.9%	95%		96%	92%
Egypt	0.5%	3.4%			98%	97%		
England	2.3%	2.1%	5.0%	11.3%	75%	46%	77%	77%
Ghana	0.9%	0.9%			98%	93%		
Hong Kong SAR	3.8%	3.4%	0.8%	2.0%	75%	80%	75%	81%
Hungary	3.9%	8.5%	4.3%	3.8%	96%	94%	93%	87%
Indonesia	3.4%	0.4%	0.0%		97%	99%	97%	
Iran, Islamic Rep. of	0.5%	6.5%	4.4%	0.3%	98%	98%	98%	98%
Israel	22.8%	22.5%	16.1%		91%	94%	94%	
Italy	5.0%	3.6%	6.7%		96%	97%	97%	
Japan	3.5%	0.6%	1.3%	0.6%	91%	93%	89%	90%
Jordan	2.0%	1.3%	3.0%		96%	96%	99%	
Korea, Rep. of**	1.6%	4.9%	4.0%	3.8%	99%	98%	100%	95%
Lebanon	1.4%	1.4%			85%	91%		
Lithuania**	4.2%	2.6%	4.5%	6.6%	90%	84%	89%	83%
Malaysia	3.3%	4.0%	4.6%		98%	98%	99%	
Norway	2.6%	2.3%		2.2%	86%	85%		93%
Palestinian Nat'l Auth.	1.0%	0.5%			98%	99%		
Romania	1.8%	0.5%	3.7%	2.8%	97%	98%	97%	89%
Russian Federation	2.3%	5.5%	1.7%	6.3%	97%	96%	97%	95%
Scotland	1.7%	0.0%		2.2%	77%	76%		73%
Serbia	6.8%	2.9%			98%	96%		
Singapore	1.8%	0.0%	0.0%	4.6%	95%	97%	98%	95%
Slovenia	1.9%	1.4%		2.6%	92%	91%		77%
Sweden	3.6%	2.8%		0.9%	94%	87%		90%
Thailand	3.4%		3.3%		99%		99%	
Tunisia	0.0%	1.8%	0.1%		98%	98%	98%	
United States	7.9%	4.9%	3.9%	2.1%	77%	73%	85%	78%
Benchmarking Participants								
Basque Country, Spain	4.2%	5.8%			98%	98%		
British Columbia, Canada	17.7%		3.6%		94%		93%	
Massachusetts, US	8.4%		5.0%		92%		93%	
Minnesota, US	7.5%			–	93%			–
Ontario, Canada	6.2%	6.0%	5.1%	–	89%	89%	93%	90%
Quebec, Canada	13.6%	4.8%	1.3%	–	77%	85%	92%	89%

References

IEA. (2006). *Windows within-school sampling software (WinW3S)* [Computer software and manual]. Hamburg: IEA Data Processing and Research Center.

Chapter 10

Reviewing the TIMSS 2007 Item Statistics

John F. Olson, Michael O. Martin, Ina V.S. Mullis, Pierre Foy, Ebru Erberber, and Corinna Preuschoff

10.1 Overview

For TIMSS 2007, similar to the process used in TIMSS 2003 and previous assessments, the TIMSS & PIRLS International Study Center conducted a review of a range of diagnostic statistics to examine and evaluate the psychometric characteristics of each achievement item in the 59 countries and 8 benchmarking participants that participated in TIMSS 2007. This review of item statistics was conducted before applying item response theory (IRT) scaling to the TIMSS 2007 achievement data to derive student mathematics and science achievement scores for analysis and reporting. The review of item statistics played a crucial role in the quality assurance of the TIMSS 2007 data, enabling the detection of unusual item properties that could signal a problem or error for a particular country. For example, an item that was uncharacteristically easy or difficult, or had an unusually low discriminating power, could indicate a potential problem with either translation or printing. Similarly, a constructed-response item with unusually low scoring reliability could indicate a problem with a scoring guide in a particular country. In the rare instances where such items were found, the country's translation verification documents and printed booklets were examined for flaws or inaccuracies and, if necessary, the item was removed from the international database for that country.

This chapter describes the basic item statistics that were calculated, the review criteria that were applied, statistics on the different types of reliability that were analyzed, and a summary of the reviews of the TIMSS 2007 item statistics. Examples from the TIMSS 2007 assessment are provided to illustrate the review process.

10.2 Statistics for Item Analysis

To begin the review process, the TIMSS & PIRLS International Study Center computed item analysis statistics for all 353 mathematics and science achievement items at the fourth grade and 429 items at the eighth grade that were administered in the TIMSS 2007 assessment. The properties of the items in each of the 59 countries and 8 benchmarking entities that participated were then carefully reviewed. Exhibits 10.1 and 10.2 show actual samples of the statistics calculated for a multiple-choice and a constructed-response item, respectively.

Exhibit 10.1 International Item Statistics for a Multiple-choice Item

Trends in International Mathematics and Science Study - TIMSS 2007 Assessment Results
International Item Statistics (Unweighted) - Review Version - 8th Grade
For Internal Review Only: DO NOT CITE OR CIRCULATE

Mathematics: Algebra / Knowing (M042077 - M10_06)
Label: Expression to equivalent to 4(3+x)
Type: MC Key: C

Country	N	Diff	Disc	Percentages Pct_A	Pct_B	Pct_C	Pct_D	Pct_E	Pct_OM	Pct_NR	Point Biserials PB_A	PB_B	PB_C	PB_D	PB_E	PB_OM	RDIFF	Flags
Algeria	752	16.2	0.11	40.2	19.8	16.2	19.3	.	4.5	2.0	0.03	-0.14	0.11	0.01	.	-0.06	0.36	DCH___
Armenia	683	73.2	0.46	15.4	5.0	73.2	3.7	.	2.8	0.4	-0.29	-0.19	0.45	-0.19	.	-0.08	-1.01	____F_
Australia	578	31.1	0.40	32.0	19.0	31.1	15.6	.	2.2	0.3	-0.11	-0.28	0.40	-0.01	.	-0.14	0.97	__H___
Bahrain	597	48.4	0.47	22.1	13.4	48.4	14.9	.	1.2	0.2	-0.15	-0.31	0.47	-0.16	.	-0.10	-1.11	__E___
Bosnia and Herzegovin	606	58.1	0.51	22.9	5.8	58.1	11.1	.	2.1	0.0	-0.30	-0.28	0.51	-0.14	.	-0.13	-0.97	____F_
Botswana	586	24.2	0.36	30.9	28.3	24.2	15.4	.	1.2	2.0	0.04	-0.30	0.36	-0.10	.	-0.05	-0.13	_CH___
Bulgaria	576	73.6	0.55	14.8	5.7	73.6	3.8	.	2.1	0.9	-0.32	-0.29	0.55	-0.21	.	-0.12	-1.44	__E_F_
Chinese Taipei	572	85.8	0.61	7.5	2.3	85.8	3.7	.	0.7	0.0	-0.42	-0.30	0.61	-0.28	.	-0.06	-1.05	____F_
Colombia	682	26.0	0.25	36.2	15.1	26.0	21.1	.	1.6	2.7	-0.14	-0.17	0.25	0.06	.	-0.06	-0.29	______
Cyprus	622	56.9	0.47	22.0	7.6	56.9	12.2	.	1.3	0.0	-0.27	-0.28	0.47	-0.11	.	-0.11	-0.73	____F_
Czech Republic	693	59.2	0.49	20.3	5.2	59.2	15.2	.	0.1	0.0	-0.33	-0.20	0.49	-0.18	.	-0.05	-0.38	____F_
Egypt	945	56.9	0.42	19.3	11.9	56.9	11.3	.	0.6	0.0	-0.16	-0.29	0.42	-0.15	.	-0.07	-1.28	__E___
El Salvador	568	24.6	0.32	34.9	19.9	24.6	19.0	.	1.6	1.4	-0.10	-0.22	0.32	0.01	.	-0.07	-0.60	_C____
England	571	46.1	0.59	20.8	18.4	46.1	13.7	.	1.1	0.0	-0.16	-0.46	0.59	-0.13	.	-0.09	0.53	__H___
Georgia	616	55.7	0.51	23.9	7.1	55.7	11.2	.	2.1	1.1	-0.30	-0.22	0.51	-0.16	.	-0.09	-1.19	__E_F_
Ghana	751	42.1	0.46	24.5	18.0	42.1	13.8	.	1.6	1.1	-0.14	-0.27	0.46	-0.12	.	-0.08	-1.45	__E___
Hong Kong SAR	502	87.5	0.51	5.6	2.8	87.5	3.8	.	0.4	0.0	-0.37	-0.32	0.51	-0.13	.	-0.10	-1.39	__E_F_
Hungary	599	63.4	0.60	15.7	6.5	63.4	13.4	.	1.0	0.2	-0.30	-0.30	0.60	-0.26	.	-0.13	-0.45	____F_
Indonesia	591	36.9	0.47	28.4	13.7	36.9	20.3	.	0.7	0.5	-0.16	-0.26	0.47	-0.14	.	0.02	-0.50	______
Iran, Islamic Rep. of	574	40.9	0.45	30.7	7.1	40.9	20.7	.	0.5	0.2	-0.18	-0.18	0.45	-0.21	.	-0.09	-0.61	____F_
Israel	453	71.3	0.52	14.8	4.0	71.3	9.3	.	0.7	1.5	-0.27	-0.26	0.52	-0.22	.	-0.14	-1.48	__E_F_
Italy	626	48.1	0.38	26.0	6.9	48.1	17.9	.	1.1	0.5	-0.19	-0.23	0.38	-0.10	.	-0.07	-0.21	__H_F_
Jordan	761	53.9	0.58	18.3	14.5	53.9	12.5	.	0.9	0.1	-0.19	-0.37	0.58	-0.23	.	-0.10	-1.02	__E___
Korea, Rep. of	603	85.6	0.58	7.5	3.6	85.6	3.3	.	0.0	0.0	-0.37	-0.32	0.58	-0.25	.	0.00	-0.89	____F_
Kuwait	571	23.5	0.32	31.2	20.8	23.5	21.9	.	2.6	0.7	0.00	-0.23	0.32	-0.07	.	-0.10	-0.36	_C____
Lebanon	538	79.0	0.38	11.7	2.0	79.0	6.5	.	0.7	0.6	-0.24	-0.13	0.38	-0.19	.	-0.13	-2.04	__E_F_
Lithuania	571	62.7	0.59	16.1	7.7	62.7	12.6	.	0.9	0.0	-0.31	-0.32	0.59	-0.24	.	-0.07	-0.52	____F_
Malaysia	634	44.6	0.50	27.3	13.1	44.6	14.4	.	0.6	0.2	-0.15	-0.35	0.50	-0.15	.	-0.12	-0.10	__H___
Malta	671	60.4	0.51	18.8	8.9	60.4	10.9	.	1.0	0.1	-0.17	-0.35	0.51	-0.22	.	-0.14	-0.65	____F_
Mongolia	610	57.4	0.49	18.7	7.9	57.4	14.4	.	1.6	0.3	-0.28	-0.25	0.49	-0.20	.	0.00	-1.38	__E_F_
Morocco	614	40.9	0.42	27.9	15.5	40.9	14.3	.	1.5	0.5	-0.17	-0.31	0.42	-0.05	.	-0.05	-0.94	______
Norway	656	15.5	0.15	36.1	29.3	15.5	15.2	.	3.8	0.3	0.10	-0.18	0.15	0.01	.	-0.14	1.54	_CH___
Oman	678	41.7	0.49	20.5	20.8	41.7	16.1	.	0.9	0.3	-0.09	-0.34	0.49	-0.17	.	-0.03	-1.08	__E___
Palestinian Nat'l Aut	653	37.1	0.44	23.6	22.4	37.1	15.3	.	1.7	0.3	-0.03	-0.34	0.44	-0.14	.	-0.08	-0.71	______
Qatar	1017	31.7	0.26	27.7	23.1	31.7	15.7	.	1.8	0.1	-0.02	-0.20	0.26	-0.06	.	-0.04	-1.07	__E___
Romania	601	65.6	0.53	20.6	5.8	65.6	7.5	.	0.5	0.3	-0.33	-0.25	0.52	-0.17	.	-0.10	-1.17	__E_F_
Russian Federation	633	77.3	0.52	12.6	2.8	77.3	6.2	.	1.1	0.8	-0.32	-0.15	0.52	-0.29	.	-0.12	-1.27	__E_F_
Saudi Arabia	612	22.2	0.31	30.2	23.0	22.2	23.4	.	1.1	0.2	0.03	-0.24	0.31	-0.07	.	-0.06	-0.47	_C____
Scotland	564	43.1	0.54	19.3	22.2	43.1	13.8	.	1.6	0.2	-0.13	-0.38	0.54	-0.13	.	-0.10	0.20	__H___
Serbia	572	69.9	0.61	16.3	4.0	69.9	8.6	.	1.2	0.2	-0.37	-0.26	0.61	-0.26	.	-0.15	-1.22	__E_F_
Singapore	655	86.3	0.58	8.1	1.8	86.3	3.5	.	0.3	0.0	-0.38	-0.28	0.58	-0.28	.	-0.12	-1.13	____F_
Slovenia	570	37.5	0.41	36.5	4.4	37.5	21.1	.	0.5	0.0	-0.25	-0.18	0.41	-0.09	.	-0.06	0.64	__H_F_
Sweden	725	24.6	0.20	36.6	24.1	24.6	11.6	.	3.2	0.5	0.08	-0.23	0.20	0.03	.	-0.18	1.36	_CH___
Syrian Arab Republic	670	46.1	0.45	23.3	13.9	46.1	15.4	.	1.3	0.3	-0.18	-0.24	0.45	-0.13	.	-0.11	-1.02	__E___
Thailand	763	37.0	0.55	27.0	18.1	37.0	16.9	.	1.0	0.4	-0.14	-0.34	0.55	-0.16	.	-0.10	0.00	__H___
Tunisia	580	34.8	0.48	14.5	31.2	34.8	18.4	.	1.0	0.3	-0.09	-0.35	0.48	-0.07	.	-0.07	-0.25	______
Turkey	663	53.5	0.60	21.0	14.9	53.5	9.8	.	0.8	0.2	-0.30	-0.34	0.60	-0.18	.	-0.04	-0.91	____F_
Ukraine	640	66.4	0.51	16.7	6.7	66.4	9.4	.	0.8	0.2	-0.26	-0.29	0.51	-0.22	.	-0.08	-1.35	__E_F_
United States	1049	50.4	0.53	28.5	4.9	50.4	15.3	.	0.9	0.2	-0.29	-0.24	0.53	-0.19	.	-0.08	0.14	__H_F_
International Avg.	.	50.5	0.46	22.6	12.5	50.5	13.1	.	1.3	0.5	-0.19	-0.27	0.46	-0.14	.	-0.09	-0.61	______
Basque Country, Spain	331	66.5	0.40	18.7	2.1	66.5	11.2	.	1.5	0.0	-0.26	-0.19	0.40	-0.12	.	-0.16	-0.83	____F_
British Columbia, Can	600	34.8	0.42	41.7	8.0	34.8	13.7	.	1.8	0.3	-0.21	-0.19	0.42	-0.10	.	-0.10	0.88	__H_F_
Massachusetts, US	282	56.0	0.58	25.2	2.8	56.0	14.5	.	1.4	0.0	-0.35	-0.20	0.58	-0.24	.	-0.14	0.39	__H_F_
Minnesota, US	247	42.1	0.52	35.6	5.7	42.1	16.2	.	0.4	0.0	-0.26	-0.23	0.52	-0.20	.	-0.11	0.79	__H_F_
Ontario, Canada	496	21.4	0.26	44.6	11.3	21.4	20.2	.	2.6	0.4	-0.02	-0.24	0.26	-0.01	.	-0.09	1.88	_CH___
Quebec, Canada	590	56.1	0.44	24.6	5.4	56.1	12.7	.	1.2	0.7	-0.20	-0.24	0.44	-0.19	.	-0.11	0.01	__H_F_

Keys: Diff: Percent correct score; Disc: Item discrimination; Pct_A...E: Percent choosing option; Pct_OM, NR: Percent Omitted and Not Reached;
PB_A...E: Point Biserial for option; PB_OM: Point Biserial for Omittted. RDIFF= Rasch difficulty.

Flags: A= Ability not ordered/Attractive distractor; C= Difficulty less than chance; D= Negative/low discrimination; E= Easier than average;
F= Distractor chosen by less than 10%; H= Harder than average; R= Scoring reliability < 80%; V= Difficulty greater than 95.

Exhibit 10.2 International Item Statistics for a Constructed-response Item

```
Trends in International Mathematics and Science Study - TIMSS 2007 Assessment Results
International Item Statistics (Unweighted) - Review Version -  8th Grade
For Internal Review Only: DO NOT CITE OR CIRCULATE

Mathematics: Number / Applying (M042304B - M04_05B)
Label: Complete the table for cycling
 Type: CR   Key: X
```

Country	N	Diff	Disc	Percentages Pct_0	Pct_1	Pct_2	Pct_OM	Pct_NR	Point Biserials PB_0	PB_1	PB_2	PB_OM	RDIFF	Reliability Cases	Score	Code	Flags
Algeria	680	12.8	0.26	57.6	22.1	1.8	18.5	1.6	-0.18	0.10	0.32	-0.02	0.49	188	60.1	57.4	____FR
Armenia	676	15.3	0.54	39.8	3.4	13.6	43.2	0.4	-0.01	0.18	0.50	-0.39	1.63	219	98.2	96.3	__H_F_
Australia	571	38.2	0.60	33.5	37.1	19.6	9.8	0.3	-0.35	0.14	0.49	-0.32	0.35	190	96.8	96.8	__E___
Bahrain	600	15.4	0.50	58.2	24.2	3.3	14.3	0.5	-0.32	0.30	0.37	-0.11	0.80	200	98.5	98.5	_____F
Bosnia and Herzegovin	600	19.8	0.48	32.2	27.7	6.0	34.2	0.0	-0.12	0.24	0.38	-0.30	1.09	210	90.5	89.5	__H_F_
Botswana	596	6.2	0.54	74.0	9.4	1.5	15.1	0.3	-0.24	0.35	0.41	-0.13	1.12	207	96.6	96.6	_____F
Bulgaria	569	22.8	0.60	36.4	21.6	12.0	30.1	0.4	-0.27	0.25	0.49	-0.29	1.25	182	80.2	79.1	__H___
Chinese Taipei	581	57.7	0.65	20.8	31.5	42.0	5.7	0.0	-0.44	0.00	0.53	-0.37	1.00	219	98.2	98.2	______
Colombia	701	24.6	0.58	46.1	32.4	8.4	13.1	2.0	-0.39	0.25	0.46	-0.13	-0.22	197	98.5	98.0	__E_F_
Cyprus	635	34.8	0.53	21.1	39.7	15.0	24.3	0.0	-0.17	0.11	0.45	-0.34	0.22	.	.	.	__E___
Czech Republic	693	43.3	0.57	23.2	42.1	22.2	12.4	0.1	-0.31	0.03	0.49	-0.27	0.10	200	96.5	96.0	__E___
Egypt	955	15.7	0.64	65.3	18.8	6.3	9.5	0.2	-0.46	0.38	0.46	-0.14	0.88	203	100.0	100.0	_____F
El Salvador	569	10.2	0.43	63.4	17.2	1.6	17.8	2.2	-0.27	0.30	0.31	-0.08	0.56	203	100.0	100.0	_____F
England	560	43.0	0.63	25.2	41.8	22.1	10.9	0.0	-0.37	0.10	0.51	-0.32	0.49	84	98.8	98.8	______
Georgia	642	15.1	0.49	32.7	19.3	5.5	42.5	0.8	-0.12	0.33	0.32	-0.31	0.87	212	91.0	89.2	_____F
Ghana	754	3.7	0.38	70.0	6.4	0.5	23.1	0.4	-0.11	0.32	0.19	-0.10	1.35	269	99.3	99.3	__H_F_
Hong Kong SAR	484	50.2	0.60	25.4	38.4	31.0	5.2	0.4	-0.39	0.05	0.48	-0.34	1.07	164	95.1	93.9	__H___
Hungary	584	49.4	0.64	17.5	44.3	27.2	11.0	0.0	-0.36	-0.05	0.56	-0.29	0.21	181	94.5	92.8	__E___
Indonesia	596	11.1	0.55	63.8	13.4	4.4	18.5	0.2	-0.27	0.34	0.40	-0.17	0.88	203	93.1	92.6	_____F
Iran, Islamic Rep. of	568	17.3	0.53	54.0	24.6	4.9	16.4	0.4	-0.32	0.29	0.40	-0.16	0.71	182	98.4	97.8	_____F
Israel	485	28.5	0.57	36.5	32.2	12.4	19.0	0.8	-0.31	0.30	0.40	-0.31	0.63	239	90.8	85.4	______
Italy	620	32.8	0.47	23.9	41.5	12.1	22.6	0.2	-0.22	0.14	0.39	-0.25	0.58	216	99.5	99.5	______
Jordan	752	14.4	0.52	68.8	18.8	5.1	7.4	0.1	-0.37	0.33	0.37	-0.14	1.10	221	98.6	98.6	__H_F_
Korea, Rep. of	601	53.7	0.71	24.5	33.6	36.9	5.0	0.0	-0.47	-0.09	0.64	-0.29	1.22	212	97.2	97.2	__H___
Kuwait	587	8.7	0.54	70.2	11.2	3.1	15.5	0.5	-0.22	0.31	0.43	-0.20	0.33	200	99.0	99.0	_____F
Lebanon	549	23.8	0.59	39.0	34.8	6.4	19.9	0.2	-0.37	0.38	0.36	-0.21	1.03	191	99.5	98.4	_____F
Lithuania	577	37.8	0.59	30.0	34.3	20.6	15.1	0.0	-0.32	0.07	0.52	-0.27	0.58	174	96.6	95.4	______
Malaysia	627	32.5	0.61	43.1	33.3	15.8	7.8	0.0	-0.46	0.21	0.48	-0.16	0.43	216	96.3	96.3	______
Malta	660	42.7	0.63	29.7	34.2	25.6	10.5	0.2	-0.39	0.10	0.52	-0.32	0.18	176	93.8	93.8	__E___
Mongolia	631	15.3	0.35	42.2	24.2	3.2	30.4	0.6	0.05	0.17	0.32	-0.34	1.18	192	81.3	72.9	__H_F_
Morocco	615	20.7	0.22	42.3	33.7	3.9	20.2	0.5	0.00	-0.05	0.37	-0.13	0.39	176	93.2	93.2	____AF_
Norway	656	34.0	0.54	28.7	39.9	14.0	17.4	0.5	-0.24	0.18	0.42	-0.33	0.07	220	94.1	92.3	__E___
Oman	675	7.0	0.53	80.0	13.0	0.4	6.5	0.9	-0.39	0.51	0.14	-0.09	1.93	226	99.1	99.1	__H_F_
Palestinian Nat'l Aut	642	9.7	0.55	75.4	13.2	3.1	8.3	0.0	-0.41	0.42	0.33	-0.09	0.83	231	94.4	94.4	_____F
Qatar	1019	5.0	0.50	71.9	8.4	0.8	18.8	0.1	-0.16	0.40	0.29	-0.16	0.98	197	98.0	97.5	_____F
Romania	599	31.1	0.60	30.4	32.2	15.0	22.4	0.3	-0.29	0.14	0.53	-0.28	0.44	209	98.6	97.6	______
Russian Federation	637	42.5	0.60	22.6	39.1	22.9	15.4	0.3	-0.32	0.02	0.54	-0.28	0.58	212	99.1	99.1	______
Saudi Arabia	600	5.0	0.36	79.0	8.3	0.8	11.8	0.7	-0.14	0.29	0.22	-0.14	0.91	211	99.1	99.1	_____F
Scotland	585	37.5	0.62	27.2	41.9	16.6	14.4	0.2	-0.32	0.14	0.50	-0.34	0.37	116	99.1	98.3	__E___
Serbia	584	33.1	0.54	19.7	34.4	15.9	30.0	0.2	-0.15	0.14	0.45	-0.38	0.65	206	97.6	97.1	______
Singapore	665	63.7	0.71	19.1	27.2	50.1	3.6	0.0	-0.50	-0.14	0.64	-0.32	0.52	220	99.1	99.1	______
Slovenia	559	41.9	0.54	24.0	44.7	19.5	11.8	0.0	-0.26	-0.04	0.52	-0.23	0.25	184	100.0	100.0	__E___
Sweden	751	41.2	0.64	22.6	34.0	24.2	19.2	0.3	-0.25	0.04	0.56	-0.40	-0.03	253	96.0	94.9	__E___
Syrian Arab Republic	668	17.0	0.38	58.1	25.6	4.2	12.1	0.0	-0.24	0.14	0.36	-0.06	0.50	195	99.5	98.5	_____F
Thailand	768	32.9	0.59	39.1	32.6	16.7	11.7	0.1	-0.37	0.16	0.49	-0.23	0.11	197	95.4	94.9	__E___
Tunisia	577	19.2	0.49	57.4	23.4	7.5	11.8	0.3	-0.38	0.24	0.38	-0.05	0.30	205	87.3	87.3	__E_F_
Turkey	654	20.0	0.55	50.2	29.8	5.0	15.0	0.0	-0.32	0.26	0.47	-0.17	1.18	184	97.8	97.8	__H_F_
Ukraine	646	18.7	0.43	46.1	20.4	8.5	24.9	0.0	-0.12	0.18	0.36	-0.26	1.13	212	85.8	84.9	__H_F_
United States	1066	41.3	0.62	31.8	40.8	20.9	6.5	0.6	-0.40	0.09	0.52	-0.26	0.50	284	93.3	92.3	______
International Avg.	.	26.9	0.54	42.7	27.7	13.1	16.5	0.4	-0.28	0.18	0.43	-0.23	0.69	.	95.1	94.3	______
Basque Country, Spain	326	38.5	0.50	23.9	45.1	16.0	15.0	0.0	-0.20	0.00	0.48	-0.26	0.28	199	99.0	99.0	__E___
British Columbia, Can	603	40.2	0.53	28.9	42.3	19.1	9.8	0.2	-0.25	0.07	0.45	-0.32	0.49	85	95.3	94.1	______
Massachusetts, US	262	51.9	0.64	23.3	41.2	31.3	4.2	0.8	-0.46	0.00	0.53	-0.24	0.54	68	95.6	95.6	______
Minnesota, US	257	42.4	0.59	31.5	45.9	19.5	3.1	0.0	-0.47	0.08	0.48	-0.08	0.69	69	94.2	92.8	______
Ontario, Canada	492	43.2	0.57	28.5	40.4	23.0	8.1	1.4	-0.33	0.07	0.47	-0.33	0.40	127	96.9	96.1	______
Quebec, Canada	569	51.0	0.61	19.2	43.9	29.0	7.9	0.9	-0.31	-0.07	0.55	-0.31	0.21	183	95.1	95.1	__E___

Keys: Diff: Percent correct score; Disc: Item discrimination; Pct_0...3: Percent obtaining score level; Pct_OM, NR: Percent Omitted and Not Reached;
PB_0...3: Point Biserial for score level; PB_OM: Point Biserial for Omittted; RDIFF= Rasch difficulty;
Reliability: (Cases) Responses double scored; (Score) Percent agreement on score; (Code) Percent agreement on code.
Flags: A= Ability not ordered/Attractive distractor; C= Difficulty less than chance; D= Negative/low discrimination; E= Easier than average;
F= Score obtained by less than 10%; H= Harder than average; R= Scoring reliability < 80%; V= Difficulty greater than 95.

For all items, regardless of item format, statistics included the number of students that responded in each country, the difficulty level (the percentage of students that answered the item correctly), and the discrimination index (the point-biserial correlation between success on the item and a total score).[1] Also provided was an estimate of the item's difficulty using a Rasch one-parameter IRT model. The international means of the item difficulties and item discriminations served as guides to the overall statistical properties of the items. Statistics for each item are displayed alphabetically by country, with the international average for each statistic in the bottom row. For those countries that tested in more than one language, statistics were calculated and examined separately by language group.

Statistics displayed for multiple-choice items included the percentage of students that chose each option, as well as the percentage of students that omitted or did not reach the item, and the point-biserial correlation between the response to each option and the total score. Statistics displayed for constructed-response items (which could have one or two score levels) included the difficulty and discrimination of each score level. Constructed-response item displays also provided information about the reliability with which the item was scored in each country, with the total number of double-scored cases and the percent exact agreement between the scorers.

10.2.1 Statistics used in Item Analysis

Definitions and detailed descriptions of the statistics that were calculated are provided below, with examples shown in Exhibits 10.1 and 10.2. The statistics were calculated separately, by grade, for mathematics and science. Statistics are listed in order of appearance in the item analysis output:

N: Number of students to whom the item was administered. If a student did not reach an item in the achievement booklet, the item was considered not administered for the purpose of the item analysis.[2]

Diff: Item difficulty is the average percent correct. For 1-point items, it is the percentage of students providing a fully correct response to the item. For the computation of this statistic, not reached items were treated as not administered.

Disc: Item discrimination was computed as the correlation between a correct response to the item and the overall score on all of the

1 For the purpose of computing the discrimination index, the total score was the percentage of mathematics or science items a student answered correctly.

2 In TIMSS, for the purposes of item analysis and item parameter estimation in scaling, items not reached by a student were treated as if they had not been administered. For purposes of estimating student proficiency, however, not reached items were treated as incorrectly answered.

mathematics or science items the student was administered.[3] Items exhibiting good measurement properties should have a moderately positive correlation.

PCT_A, **PCT_B**, **PCT_C**, **PCT_D**, and **PCT_E**: Used for multiple-choice items only (see Exhibit 10.1), each column indicates the percentage of students choosing the particular response option for the item (A, B, C, D, or E). Not reached items were excluded from the denominator for these calculations.

PCT_0, **PCT_1**, and **PCT_2**: Used for constructed-response items only (see Exhibit 10.2), each column indicates the percentage of students scoring at the particular score level, up to and including the maximum score level for the item. Not reached items were excluded from the denominator for these calculations.

PCT_OM: Percentage of students who, having reached the item, did not provide a response. Not reached items were excluded from the denominator when calculating this statistic.

PCT_NR: Percentage of students who did not reach the item. An item was coded as not reached when there was no evidence of a response to any subsequent items in the booklet and the response to the item preceding it was omitted.

PB_A, **PB_B**, **PB_C**, **PB_D**, and **PB_E**: Used for multiple-choice items only, these present the correlation between choosing each of the response options, A, B, C, D, or E, and the overall score on all of the mathematics or science items the student was administered. Items with good psychometric properties have near-zero or negative correlations for the distracter options (the incorrect options) and moderately positive correlations for the correct option.

PB_0, **PB_1**, and **PB_2**: Used for constructed-response items only, these present the correlation between the score levels on the item (0, 1, or 2) and the overall score on all of the mathematics or science items the student was administered. For items with good measurement properties, the correlation coefficients should change from negative to positive as the score on the item increases.

PB_OM: The correlation between a binary variable, indicating an omitted response to the item, and the overall score on all of the mathematics or science items the student was administered. This correlation should be negative or near zero.

3 For constructed-response items, the discrimination is the correlation between the number of score points and total score.

RDIFF: An estimate of the item's difficulty based on a Rasch one-parameter IRT model applied to each country's sample. The difficulty estimate is expressed in the logit metric (with a positive logit indicating a difficult item) and was scaled so that the average Rasch item difficulty was zero within each country.

Reliability (Cases): To provide a measure of the reliability of the scoring of the constructed-response items, those items in approximately 25 percent of the test booklets in each country were scored by two independent scorers. This column indicates the number of times each item was double-scored in a country.

Reliability (Score): This column contains the percentage of exact agreement on the scores assigned by two independent scorers.

Reliability (Code): This column contains the percentage of exact agreement on the 2-digit scoring codes.

As an aid to reviewers, the item-analysis display includes a series of "flags" signaling the presence of one or more conditions that might indicate a problem with an item. The following conditions are flagged:

- Item difficulty exceeds 95 percent in the sample as a whole.
- Item difficulty is less than 25 percent for four-option multiple-choice items in the sample as a whole.
- One or more of the distracter percentages is less than 10 percent.
- One or more of the distracter percentages is greater than the percentage for the correct answer or the point-biserial correlation for one or more of the distracters exceeds zero.
- Item discrimination (i.e., the point-biserial for the correct answer) is less than 0.2.
- Item discrimination does not increase with each score level (for constructed-response items with more than one score level).
- The Rasch difficulty estimate is easier or harder than the average across countries.
- Scoring reliability for the score points is less than 80 percent (for constructed-response items only).

Although not all of these conditions necessarily indicate a problem, the flags are a useful way to draw attention to potential sources of concern.

In order to measure trends, TIMSS 2007 included items from TIMSS 2003 at the fourth grade and from TIMSS 2003 and 1999 (those items from 1999 that were administered again in 2003) at the eighth grade.[4] For these trend items, the review included an examination of changes in item statistics between the 2003 and 2007 administrations.

An example item statistics display for a fourth grade trend item is shown in Exhibit 10.3. The information in this exhibit is different from the item statistics presented in Exhibits 10.1 and 10.2, and presents countries' statistics from the TIMSS 2007 and 2003 assessments. In reviewing these item statistics, the aim was to detect any unusual changes in item properties between assessments, which might indicate a problem in using the item to measure change.

4 For more information on trend items, see Chapter 2.

Exhibit 10.3 International Item Statistics for a Trend Item

Trends in International Mathematics and Science Study - TIMSS 2007 Bridge Assessment Results
Percent of Responses by Item Category (Science) - Trend Items - 4th Grade
For Internal Review Only: DO NOT CITE OR CIRCULATE

Science: Life Science / Factual Knowledge (S031233 - S11_03)
Label: Main features of four animals shown
Type: CR Key: X

COUNTRY	Year	N	10	70	71	79	99	V1	NOT REACHED	OMIT	1.GIRL % Right	2.BOY % Right
Armenia	2003	464	33.2	18.6	16.8	11.7	19.6	33.2	0.8	18.9	33.0	33.5
	2007	273	48.7	8.8	21.2	1.8	19.4	48.7	0.7	18.7	48.1	49.1
Australia	2003	375	64.0	17.5	11.0	5.5	2.1	64.0	0.0	2.1	65.3	62.7
	2007	291	64.6	17.2	14.1	3.4	0.7	64.6	0.0	0.7	63.5	65.6
Chinese Taipei	2003	380	76.5	3.3	18.7	0.3	1.3	76.5	0.0	1.3	76.5	76.4
	2007	300	64.0	6.7	26.0	3.0	0.3	64.0	0.0	0.3	63.3	64.6
England	2003	291	64.2	13.0	15.4	6.1	1.2	64.2	0.0	1.2	63.8	64.8
	2007	305	69.2	14.8	10.8	4.6	0.7	69.2	0.0	0.7	70.8	67.5
Hong Kong SAR	2003	373	69.8	7.9	18.9	3.0	0.4	69.8	0.0	0.4	72.5	67.7
	2007	268	69.8	5.2	22.0	1.9	1.1	69.8	0.0	1.1	69.9	69.6
Hungary	2003	268	70.5	15.4	9.9	3.6	0.6	70.5	0.0	0.6	75.9	65.1
	2007	288	72.2	11.1	11.8	3.8	1.0	72.2	0.0	1.0	73.5	71.1
Iran, Islamic Rep. of	2003	352	41.7	22.8	20.3	6.7	8.5	41.7	0.0	8.5	48.2	37.7
	2007	274	39.4	25.9	20.1	8.4	6.2	39.4	0.0	6.2	36.9	41.4
Italy	2003	353	72.4	9.2	14.3	3.7	0.5	72.4	0.0	0.5	75.4	69.8
	2007	323	66.3	10.5	15.2	3.1	5.0	66.3	0.0	5.0	65.2	67.0
Latvia	2003	295	61.0	19.5	13.5	4.8	1.1	61.0	0.0	1.1	62.0	60.1
	2007	277	68.2	15.2	11.6	3.6	1.4	68.2	0.0	1.4	71.1	65.5
Lithuania	2003	371	58.3	23.2	9.2	7.3	1.9	58.3	0.0	1.9	59.3	56.3
	2007	285	54.7	22.8	14.4	7.0	1.1	54.7	0.0	1.1	54.7	54.8
Morocco	2003	339	14.8	32.1	12.5	28.2	12.5	14.8	0.0	12.5	10.2	18.9
	2007	300	12.0	16.7	7.7	47.7	16.0	12.0	0.7	15.3	10.5	13.7
Netherlands	2003	242	54.7	15.4	19.9	10.0	0.0	54.7	0.0	0.0	57.3	51.8
	2007	237	63.3	15.2	14.8	5.5	1.3	63.3	0.0	1.3	62.5	64.1
New Zealand	2003	354	61.1	17.8	14.2	4.9	2.1	61.1	0.0	2.1	59.9	62.3
	2007	349	56.7	18.3	15.8	8.3	0.9	56.7	0.0	0.9	56.3	57.2
Norway	2003	361	58.6	20.5	12.8	6.3	1.8	58.6	0.0	1.8	61.7	55.9
	2007	290	55.9	16.9	20.7	4.1	2.4	55.9	0.0	2.4	57.4	54.5
Russian Federation	2003	325	53.9	16.6	8.3	17.1	4.2	53.9	0.0	4.2	51.7	56.1
	2007	323	62.8	15.5	11.5	7.4	2.8	62.8	0.0	2.8	60.1	65.0
Scotland	2003	330	54.5	19.1	17.3	6.8	2.3	54.5	0.4	1.9	54.2	54.9
	2007	286	59.1	21.7	10.1	7.7	1.4	59.1	0.0	1.4	50.0	66.3
Singapore	2003	562	78.1	9.7	9.1	2.4	0.7	78.1	0.0	0.7	80.7	75.5
	2007	360	78.3	10.3	9.4	1.7	0.3	78.3	0.0	0.3	81.4	75.1

V1 = Percent scoring 1 or better V2 = Percent scoring 2 or better
Percent right for boys and girls corresponds to the percent obtaining the maximum score on the item.
Because of missing gender information, some totals may appear inconsistent.

Exhibit 10.3 International Item Statistics for a Trend Item (Continued)

Trends in International Mathematics and Science Study - TIMSS 2007 Bridge Assessment Results
Percent of Responses by Item Category (Science) - Trend Items - 4th Grade
For Internal Review Only: DO NOT CITE OR CIRCULATE

Science: Life Science / Factual Knowledge (S031233 - S11_03)
Label: Main features of four animals shown
Type: CR Key: X

COUNTRY	Year	N	10	70	71	79	99	V1	NOT REACHED	OMIT	1.GIRL % Right	2.BOY % Right
Slovenia	2003	280	58.9	18.4	13.6	6.4	2.7	58.9	0.0	2.7	63.2	56.4
	2007	316	66.5	14.6	9.5	8.5	0.9	66.5	0.0	0.9	69.7	63.4
Tunisia	2003	354	18.5	26.2	20.3	27.2	7.8	18.5	0.0	7.8	16.4	20.6
	2007	290	18.6	29.3	13.1	30.0	9.0	18.6	0.0	9.0	20.7	16.6
United States	2003	809	60.7	20.0	11.5	6.7	1.0	60.7	0.0	1.0	62.9	58.6
	2007	566	61.3	19.3	12.0	6.0	1.4	61.3	0.2	1.2	59.2	63.3
International Avg.	2003	.	56.3	17.3	14.4	8.4	3.6	56.3	0.1	3.6	57.5	55.2
	2007	.	57.6	15.8	14.6	8.4	3.7	57.6	0.1	3.6	57.2	57.8
Ontario, Canada	2003	359	55.1	21.3	14.6	7.1	2.0	55.1	0.0	2.0	49.1	61.9
	2007	255	50.6	22.7	12.9	8.6	5.1	50.6	0.0	5.1	49.3	52.1
Quebec, Canada	2003	373	51.9	25.2	15.6	7.0	0.3	51.9	0.0	0.3	47.7	55.5
	2007	276	61.2	15.2	13.4	6.9	3.3	61.2	0.0	3.3	64.5	58.0

V1 = Percent scoring 1 or better V2 = Percent scoring 2 or better
Percent right for boys and girls corresponds to the percent obtaining the maximum score on the item.
Because of missing gender information, some totals may appear inconsistent.

10.2.2 Item-by-Country Interaction

Although countries are expected to exhibit some variation in performance across items, in general, as a whole, countries with high average performance on the assessment should perform relatively well on each of the items, and low-scoring countries should do less well on each of items. When this does not occur (i.e., when a high-scoring country has a low performance on an item on which other countries are doing well), there is said to be an item-by-country interaction. When large, such item-by-country interactions may be a sign of an item that is flawed in some way, and steps should be taken to address the problem.

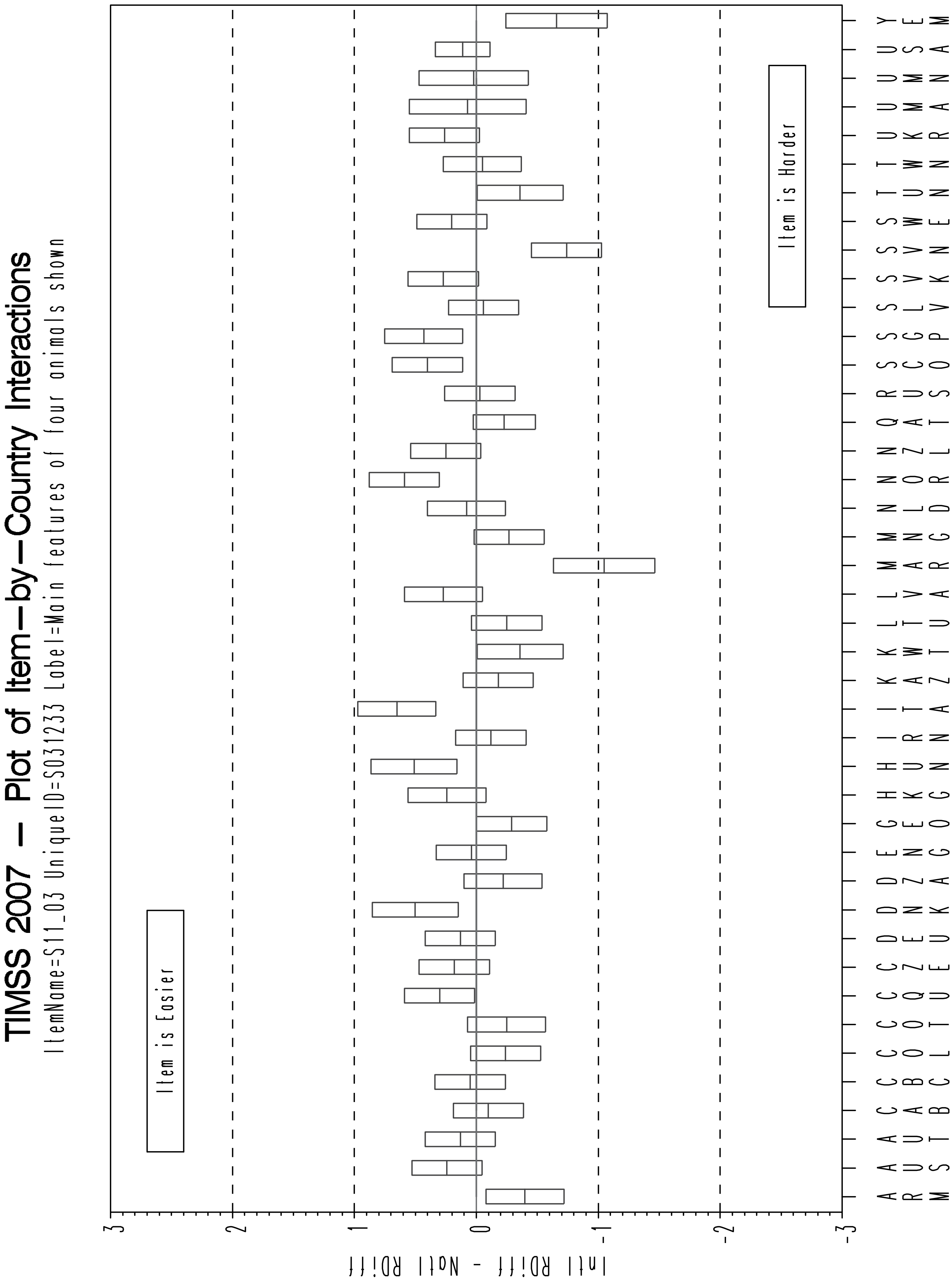
To assist in detecting sizeable item-by-country interactions, the TIMSS & PIRLS International Study Center produced a graphical display for each item showing the average probability across all countries of a correct response for a student of average international proficiency, compared with the probability of a correct response by a student of average proficiency in each country. Exhibit 10.4 provides an example of a TIMSS item-by-country interaction display. The probability for each country is presented as a 95 percent confidence interval, which includes a built-in Bonferroni correction for multiple comparisons. The limits for the confidence interval are computed as follows:

$$\text{Upper Limit} = 1 - \frac{e^{RDIFF_{ik} + SE_{RDIFF_{ik}} \text{x} Z_b}}{1 + e^{RDIFF_{ik} + SE_{RDIFF_{ik}} \text{x} Z_b}}$$

$$\text{Lower Limit} = 1 - \frac{e^{RDIFF_{ik} - SE_{RDIFF_{ik}} \text{x} Z_b}}{1 + e^{RDIFF_{ik} - SE_{RDIFF_{ik}} \text{x} Z_b}}$$

where $RDIFF_{ik}$ is the Rasch difficulty of item k within country i, $SE_{RDIFFik}$ is the standard error of the difficulty of item k in country i and Z_b is the critical value from the Z distribution, corrected for multiple comparisons using the Bonferroni procedure.

Exhibit 10.4 Sample Plot of Item-by-Country Interaction for a TIMSS 2007 Item



10.2.3 Trend Item Analysis

Because an important part of the TIMSS 2007 assessment was the measuring of trends across cycles, there was an additional stage of the review process to ensure that the trend items had similar characteristics in both cycles (i.e., an item that was relatively easy in 2003 should be relatively easy in 2007). The comparison between cycles was made in a number of ways. For each trend country, almanacs of item statistics displayed the percentage of students within each score category (or response option, for multiple-choice items) for each cycle, as well as the difficulty of the item and the percent correct by gender. While some changes were anticipated as countries' overall achievement may have improved or declined, items were noted if the trend difference was greater than 2 logits for a particular country.

In addition, TIMSS 2007 included a bridge study to examine the effect of changes to the assessment design and booklets. Countries measuring trend were required to participate in a bridge study where they administered a subset of TIMSS 2003 booklets in TIMSS 2007 under the TIMSS 2003 conditions. During the trend item analysis and review stage, comparisons then were made for these items to examine for any differences.

The TIMSS & PIRLS International Study Center used two different graphical displays to examine the differences between item difficulties in 2003 to 2007. The first of these, shown in Exhibit 10.5, displays the difference in Rasch difficulty estimates (in logits) for trend comparisons between 2003 and the 2007 bridge data. A positive difference indicates that the item was relatively easier in a country in 2007, and a negative difference indicates that an item was relatively more difficult. The second, Exhibit 10.6, shows a country's performance on all trend items simultaneously. Individually for each country, a scatterplot graphed the Rasch difficulty of each item in 2003 against the difficulty for that item in 2007. Where there are no differences between the difficulties in the 2003 and 2007 bridge data, the data points will align on or near the diagonal indicating a one-to-one correlation between cycles.

These graphs were used in conjunction with one another to detect items that performed differently in the two cycles. When such items were found, the source of the difference was investigated using booklets from both cycles, translation verifier's comments, national adaptation forms, and trend scoring reliability data.

Exhibit 10.5 Sample Plot of Difference in Rasch Difficulties for a TIMSS 2007 Item

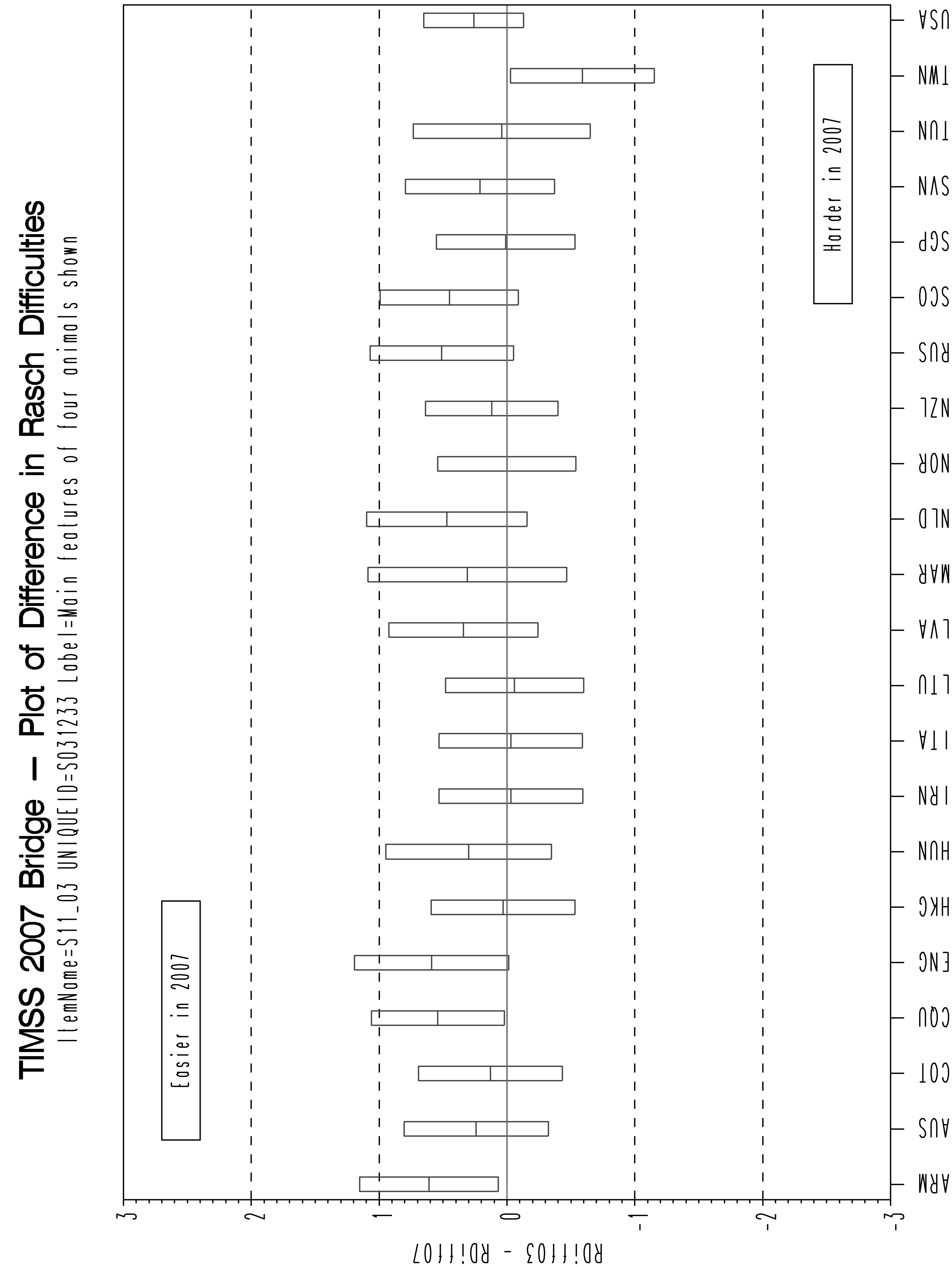
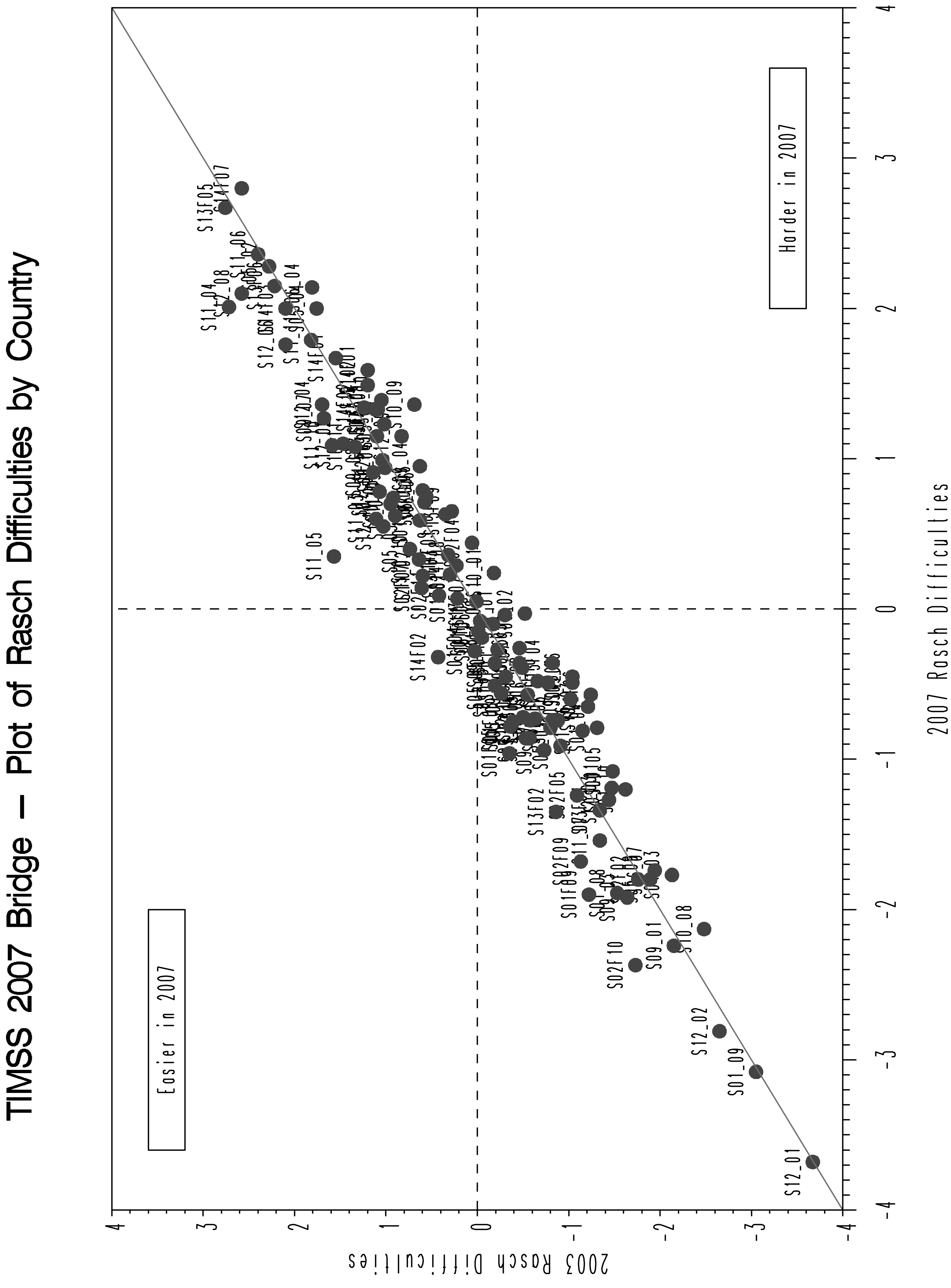


Exhibit 10.6 Sample Plot of Rasch Difficulties by Country



10.3 Reliability

10.3.1 Test Reliability

Exhibits 10.7 and 10.8 display the mathematics and science test reliability coefficients for each country. This coefficient is the median Cronbach's alpha reliability across the 14 test booklets. In general, at both grade levels in mathematics, median reliabilities were relatively high, with an international median (the median of the reliability coefficients for all countries) of 0.83 for fourth grade and 0.88 for eighth grade. In science, median reliabilities were 0.80 for fourth grade and 0.84 for eighth grade. Despite the generally high reliabilities, there were some countries with median reliabilities below 0.70 at one or both grades in mathematics, namely Algeria, Botswana, El Salvador, Ghana, Kuwait, Qatar, Saudi Arabia, and Yemen. Countries with median reliabilities below 0.70 at one or both grades in science were Algeria and Yemen.

Exhibit 10.7 Cronbach's Alpha Reliability Coefficient – TIMSS 2007 Mathematics Test

Country	Reliability Coefficient	
	4th Grade	8th Grade
Algeria	0.77	0.66
Armenia	0.87	0.88
Australia	0.86	0.89
Austria	0.82	
Bahrain		0.80
Bosnia and Herzegovina		0.84
Botswana		0.69
Bulgaria		0.90
Chinese Taipei	0.83	0.93
Colombia	0.77	0.77
Cyprus		0.88
Czech Republic	0.83	0.88
Denmark	0.84	
Egypt		0.84
El Salvador	0.70	0.63
England	0.88	0.90
Georgia	0.83	0.84
Germany	0.83	
Ghana		0.68
Hong Kong SAR	0.81	0.92
Hungary	0.88	0.90
Indonesia		0.83
Iran, Islamic Rep. of	0.81	0.84
Israel		0.90
Italy	0.85	0.87
Japan	0.85	0.91
Jordan		0.88
Kazakhstan	0.87	
Korea, Rep. of		0.92
Kuwait	0.69	0.69
Latvia	0.83	
Lebanon		0.84
Lithuania	0.85	0.89
Malaysia		0.88
Malta		0.89
Morocco	0.78	0.76
Netherlands	0.79	
New Zealand	0.87	
Norway	0.82	0.84
Oman		0.80
Palestinian Nat'l Auth.		0.83
Qatar	0.58	0.64
Romania		0.90
Russian Federation	0.86	0.90
Saudi Arabia		0.62
Scotland	0.85	0.89
Serbia		0.89
Singapore	0.87	0.92
Slovak Republic	0.86	
Slovenia	0.84	0.88
Sweden	0.82	0.87
Syrian Arab Republic		0.79
Thailand		0.88
Tunisia	0.78	0.78
Turkey		0.91
Ukraine	0.84	0.88
United States	0.85	0.89
Yemen	0.55	
International Median	0.83	0.88
Benchmarking Participants		
Alberta, Canada	0.82	
Basque Country, Spain		0.85
British Columbia, Canada	0.84	0.87
Dubai, UAE	0.84	0.89
Massachusetts, US	0.82	0.89
Minnesota, US	0.86	0.87
Ontario, Canada	0.82	0.87
Quebec, Canada	0.82	0.87

Exhibit 10.8 Cronbach's Alpha Reliability Coefficient - TIMSS 2007 Science Test

Country	Reliability Coefficient	
	4th Grade	8th Grade
Algeria	0.76	0.65
Armenia	0.88	0.88
Australia	0.81	0.85
Austria	0.81	
Bahrain		0.84
Bosnia and Herzegovina		0.82
Botswana		0.75
Bulgaria		0.87
Chinese Taipei	0.80	0.88
Colombia	0.81	0.78
Cyprus		0.82
Czech Republic	0.80	0.83
Denmark	0.80	
Egypt		0.82
El Salvador	0.79	0.71
England	0.82	0.87
Georgia	0.77	0.79
Germany	0.80	
Ghana		0.72
Hong Kong SAR	0.78	0.86
Hungary	0.83	0.84
Indonesia		0.76
Iran, Islamic Rep. of	0.83	0.83
Israel		0.88
Italy	0.82	0.83
Japan	0.78	0.85
Jordan		0.87
Kazakhstan	0.80	
Korea, Rep. of		0.85
Kuwait	0.82	0.82
Latvia	0.76	
Lebanon		0.83
Lithuania	0.76	0.85
Malaysia		0.85
Malta		0.88
Morocco	0.79	0.73
Netherlands	0.73	
New Zealand	0.83	
Norway	0.79	0.82
Oman		0.82
Palestinian Nat'l Auth.		0.85
Qatar	0.77	0.78
Romania		0.84
Russian Federation	0.82	0.85
Saudi Arabia		0.73
Scotland	0.80	0.85
Serbia		0.83
Singapore	0.86	0.91
Slovak Republic	0.82	
Slovenia	0.79	0.83
Sweden	0.79	0.85
Syrian Arab Republic		0.80
Thailand		0.84
Tunisia	0.85	0.73
Turkey		0.85
Ukraine	0.80	0.84
United States	0.82	0.86
Yemen	0.69	
International Median	0.80	0.84
Benchmark Participants		
Alberta, Canada	0.79	
Basque Country, Spain		0.81
British Columbia, Canada	0.79	0.83
Dubai, UAE	0.85	0.86
Massachusetts, US	0.78	0.86
Minnesota, US	0.81	0.85
Ontario, Canada	0.81	0.82
Quebec, Canada	0.76	0.81

10.3.2 Scoring Reliability for Constructed-response Items

About one-third of the items in the TIMSS 2007 assessment were constructed-response items, comprising nearly half of the score points for the assessment.[5] An essential requirement for use of such items is that they be reliably scored by all participants. That is, a particular student response should receive the same score, regardless of the scorer. In conducting TIMSS 2007, measures taken to ensure that the constructed-response items were scored reliably in all countries included developing scoring guides for each constructed-response question (which provided descriptions of acceptable responses for each score point value)[6] and providing extensive training in the application of the scoring guides. Procedures for organizing and monitoring the scoring sessions are outlined in the *TIMSS 2007 Survey Operations Procedures Unit 5: Scoring the TIMSS 2007 Assessment* (TIMSS, 2006).

10.3.2.1 Within-Country Scoring Reliability

To gather and document information about the within-country agreement among scorers, a random sample of at least 200 student responses to each item was selected to be scored independently by two scorers.[7] The inter-rater agreement for each item in each country was examined as part of the item review process. The average and range of the within-country exact percent of agreement across all items for both grades is presented in Exhibit 10.9 for mathematics and Exhibit 10.10 for science.

Agreement across items was high on average across countries. The exact percent agreement was 98 percent at both grades in mathematics and 96 percent at both grades in science. All countries had an average exact percent agreement above 92 percent at the fourth grade and 95 percent at the eighth grade in mathematics and above 85 percent at the fourth grade and 90 at the eighth grade in science.

5 For details on the development of the TIMSS 2007 assessment items, see Chapter 2.

6 A discussion of the development of the scoring guides for constructed-response items is provided in Chapter 2.

7 Since individual items appear in two booklets, 100 of each of the 14 booklets were chosen randomly for double-scoring. For a sample of 4,500, this amounts to almost 25 percent of the total sample.

Exhibit 10.9 TIMSS 2007 Within-country Scoring Reliability for the Fourth Grade Constructed-response Mathematics Items

Countries	Correctness Score Agreement			Diagnostic Score Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement	
		Min	Max		Min	Max
Algeria	92	58	99	85	54	98
Armenia	99	94	100	97	91	100
Australia	100	98	100	99	95	100
Austria	99	95	100	99	94	100
Chinese Taipei	98	84	100	97	83	100
Colombia	99	93	100	97	89	100
Czech Republic	98	90	100	96	77	100
Denmark	97	83	100	93	74	99
El Salvador	99	96	100	98	85	100
England	99	91	100	98	89	100
Georgia	97	88	100	94	68	100
Germany	97	75	100	95	71	100
Hong Kong SAR	100	98	100	100	98	100
Hungary	100	97	100	99	95	100
Iran, Islamic Rep. of	99	96	100	96	84	100
Italy	99	94	100	99	79	100
Japan	99	94	100	98	84	100
Kazakhstan	99	96	100	99	94	100
Kuwait	100	98	100	98	95	100
Latvia	95	41	100	92	39	100
Lithuania	98	88	100	97	50	100
Morocco	95	33	100	88	29	98
Netherlands	97	86	100	95	72	100
New Zealand	99	95	100	97	90	100
Norway	99	92	100	97	88	100
Qatar	99	91	100	95	78	100
Russian Federation	100	98	100	99	96	100
Scotland	99	91	100	98	87	100
Singapore	99	93	100	97	90	100
Slovak Republic	99	92	100	98	90	100
Slovenia	100	99	100	99	94	100
Sweden	98	89	100	97	87	100
Tunisia	98	86	100	93	77	99
Ukraine	100	98	100	100	98	100
United States	98	83	100	96	72	100
Yemen	98	83	100	93	80	99
International Avg.	98	88	100	96	81	100
Benchmark Participants						
Alberta, Canada	99	93	100	98	90	100
British Columbia, Canada	99	96	100	99	91	100
Dubai, UAE	97	87	100	94	78	100
Massachusetts, US	98	82	100	96	72	100
Minnesota, US	98	79	100	96	68	100
Ontario, Canada	99	88	100	98	88	100
Quebec, Canada	98	90	100	97	86	100

Exhibit 10.9 TIMSS 2007 Within-country Scoring Reliability for the Eighth Grade Constructed-response Mathematics Items (Continued)

Countries	Correctness Score Agreement			Diagnostic Score Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement	
		Min	Max		Min	Max
Algeria	95	60	100	90	57	97
Armenia	99	94	100	97	75	100
Australia	99	93	100	97	86	100
Bahrain	100	97	100	99	96	100
Bosnia and Herzegovina	98	90	100	96	83	100
Botswana	98	84	100	96	76	100
Bulgaria	96	70	100	94	68	100
Chinese Taipei	98	47	100	97	43	100
Colombia	99	92	100	97	89	100
Czech Republic	98	86	100	96	81	100
Egypt	99	94	100	97	89	100
El Salvador	100	98	100	100	96	100
England	99	94	100	98	85	100
Georgia	97	76	100	95	75	100
Ghana	100	98	100	99	92	100
Hong Kong SAR	99	95	100	99	94	100
Hungary	98	84	100	97	80	100
Indonesia	98	90	100	95	82	100
Iran, Islamic Rep. of	99	93	100	97	86	100
Israel	96	82	100	92	69	99
Italy	99	85	100	98	68	100
Japan	97	84	100	94	71	100
Jordan	100	97	100	98	93	100
Korea, Rep. of	99	96	100	99	93	100
Kuwait	99	96	100	98	93	100
Lebanon	100	97	100	98	94	100
Lithuania	98	94	100	97	91	100
Malaysia	99	96	100	99	96	100
Malta	97	81	100	95	73	100
Norway	99	94	100	97	86	100
Oman	99	95	100	97	93	100
Palestinian Nat'l Auth.	98	89	100	96	83	100
Qatar	99	91	100	98	86	100
Romania	99	96	100	99	95	100
Russian Federation	100	98	100	99	96	100
Saudi Arabia	100	97	100	99	92	100
Scotland	99	95	100	98	89	100
Serbia	99	94	100	98	93	100
Singapore	98	93	100	97	91	100
Slovenia	100	98	100	99	96	100
Sweden	98	86	100	96	84	100
Syrian Arab Republic	99	95	100	98	93	100
Thailand	98	89	100	97	82	100
Tunisia	97	87	100	95	74	100
Turkey	100	95	100	99	92	100
Ukraine	98	80	100	97	79	100
United States	97	86	100	95	77	100
Morocco	95	75	100	89	57	99
International Avg.	98	89	100	97	83	100
Benchmark Participants						
Basque Country, Spain	99	89	100	98	85	100
British Columbia, Canada	98	89	100	97	85	100
Dubai, UAE	97	87	100	95	83	100
Massachusetts, US	97	78	100	95	74	100
Minnesota, US	97	81	100	95	76	100
Ontario, Canada	98	87	100	97	85	100
Quebec, Canada	99	94	100	98	91	100

Exhibit 10.10 TIMSS 2007 Within-country Scoring Reliability for the Fourth Grade Constructed-response Science Items

Countries	Correctness Score Agreement			Diagnostic Score Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement	
		Min	Max		Min	Max
Algeria	88	69	98	78	50	96
Armenia	98	93	100	95	77	100
Australia	99	95	100	98	92	100
Austria	98	90	100	96	89	100
Chinese Taipei	97	74	100	96	74	100
Colombia	98	92	100	97	89	100
Czech Republic	94	78	100	91	74	100
Denmark	91	72	100	86	68	99
El Salvador	99	78	100	98	72	100
England	98	88	100	95	84	100
Georgia	92	68	100	86	68	98
Germany	93	73	100	91	69	100
Hong Kong SAR	99	98	100	99	97	100
Hungary	99	96	100	99	94	100
Iran, Islamic Rep. of	97	83	100	96	78	100
Italy	98	85	100	97	82	100
Japan	97	88	100	95	82	100
Kazakhstan	99	97	100	99	97	100
Kuwait	99	94	100	96	89	99
Latvia	85	42	100	80	36	99
Lithuania	95	80	100	92	78	100
Morocco	93	75	100	85	43	98
Netherlands	92	71	100	88	61	99
New Zealand	97	90	100	95	86	100
Norway	97	88	100	95	87	99
Qatar	99	94	100	96	88	100
Russian Federation	100	99	100	100	98	100
Scotland	97	87	100	95	80	100
Singapore	96	90	100	95	90	100
Slovak Republic	99	97	100	98	93	100
Slovenia	99	93	100	99	93	100
Sweden	93	65	100	89	62	100
Tunisia	93	77	100	88	67	99
Ukraine	100	98	100	100	98	100
United States	94	68	100	90	66	100
Yemen	96	85	100	89	67	98
International Avg.	96	83	100	93	78	100
Benchmark Participants						
Alberta, Canada	98	86	100	97	86	100
British Columbia, Canada	99	89	100	96	84	100
Dubai, UAE	93	73	100	89	71	99
Massachusetts, US	94	72	100	91	65	100
Minnesota, US	94	74	100	91	55	100
Ontario, Canada	98	90	100	97	88	100
Quebec, Canada	99	91	100	97	88	100

Exhibit 10.10 TIMSS 2007 Within-country Scoring Reliability for the Eighth Grade Constructed-response Science Items (Continued)

Countries	Correctness Score Agreement			Diagnostic Score Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement	
		Min	Max		Min	Max
Algeria	94	75	100	89	70	99
Armenia	98	89	100	95	75	100
Australia	97	88	100	95	86	100
Bahrain	94	78	100	90	67	100
Bosnia and Herzegovina	95	74	100	91	72	99
Botswana	95	79	100	89	73	100
Bulgaria	91	69	100	86	59	100
Chinese Taipei	94	66	100	90	63	100
Colombia	98	88	100	96	84	100
Czech Republic	93	75	100	90	64	100
Egypt	97	88	100	94	80	99
El Salvador	100	98	100	99	92	100
England	97	88	100	95	80	100
Georgia	92	67	100	85	53	100
Ghana	99	96	100	98	94	100
Hong Kong SAR	99	96	100	98	94	100
Hungary	95	86	100	92	80	100
Indonesia	97	81	100	92	75	100
Iran, Islamic Rep. of	97	86	100	95	79	100
Israel	92	73	100	84	66	99
Italy	96	63	100	94	60	100
Japan	91	54	100	85	54	100
Jordan	99	93	100	96	74	100
Korea, Rep. of	99	95	100	98	87	100
Kuwait	99	88	100	97	87	100
Lebanon	100	97	100	98	95	100
Lithuania	97	90	100	96	84	100
Malaysia	99	96	100	98	93	100
Malta	93	81	100	89	75	99
Norway	97	88	100	95	85	100
Oman	99	95	100	94	81	100
Palestinian Nat'l Auth.	94	82	100	88	69	99
Qatar	99	95	100	98	91	100
Romania	99	89	100	98	89	100
Russian Federation	99	93	100	98	92	100
Saudi Arabia	99	90	100	98	88	100
Scotland	97	84	100	95	77	100
Serbia	97	74	100	94	74	100
Singapore	96	90	100	94	90	100
Slovenia	100	95	100	99	93	100
Sweden	92	70	100	88	64	100
Syrian Arab Republic	99	92	100	98	91	100
Thailand	90	73	100	83	63	100
Tunisia	91	61	100	85	61	100
Turkey	97	81	100	94	63	100
Ukraine	92	68	100	86	52	100
United States	93	73	100	88	61	100
Morocco	90	58	99	81	49	98
International Avg.	96	82	100	93	76	100
Benchmark Participants						
Basque Country, Spain	97	86	100	96	77	100
British Columbia, Canada	96	81	100	92	77	100
Dubai, UAE	96	88	100	94	83	100
Massachusetts, US	92	76	100	88	62	100
Minnesota, US	93	77	100	89	61	100
Ontario, Canada	96	84	100	93	82	100
Quebec, Canada	95	84	100	92	82	100

10.3.2.2 Trend Item Scoring Reliability

TIMSS 2007 also took steps to show that the constructed-response items from 2003 that were used in 2007 were scored in the same way in both assessments. In anticipation of this, countries that participated in TIMSS 2003 sent samples of scored student booklets from the 2003 data collection to the IEA Data Processing and Research Center, where they were digitally scanned and stored in presentation software for later use. As a check on scoring consistency from 2003 to 2007, staff members working in each country on scoring the 2007 fourth- and eighth-grade data were asked also to score these 2003 responses using the DPC software.

As shown in Exhibit 10.11 for mathematics and Exhibit 10.12 for science, there was a very high degree of scoring consistency, with 97 percent exact agreement for both grades in mathematics, on average internationally, between the scores awarded in 2003 and those given by the 2007 scorers. The average exact percent agreement in science was 93 percent for fourth grade and 94 percent for eighth grade. There also was high agreement in mathematics at the diagnostic score level, with 96 and 94 percent exact agreement, on average, for grades four and eight, respectively. It was somewhat less in science, with 86 percent at grade four and 88 percent at grade eight, on average.

Exhibit 10.11 TIMSS 2007 Trend Scoring Reliability (2003–2007) for the Fourth Grade Constructed-response Mathematics Items

Countries	Correctness Score Agreement			Diagnostic Score Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement	
		Min	Max		Min	Max
Armenia	96	90	100	93	65	99
Australia	97	84	100	96	83	100
Chinese Taipei	97	93	100	96	88	100
England	98	92	100	97	87	100
Hong Kong SAR	99	93	100	98	87	100
Hungary	99	96	100	97	92	100
Iran, Islamic Rep. of	98	95	100	96	86	100
Japan	98	93	100	96	88	100
Lithuania	97	88	100	94	74	100
Netherlands	97	90	99	95	88	99
New Zealand	98	95	100	97	90	100
Norway	98	96	100	97	93	100
Russian Federation	99	95	100	98	92	100
Scotland	96	91	100	95	90	100
Singapore	95	86	100	93	83	100
Slovenia	96	68	99	93	47	99
Tunisia	98	97	100	95	81	100
United States	98	92	100	96	88	100
International Avg.	97	91	100	96	83	100
Benchmark Participants						
Alberta, Canada	98	91	99	96	85	99
British Columbia, Canada	98	91	99	96	85	99
Ontario, Canada	98	91	99	96	85	99
Quebec, Canada	98	91	99	96	85	99

Exhibit 10.11 TIMSS 2007 Trend Scoring Reliability (2003–2007) for the Eighth Grade Constructed-response Mathematics Items (Continued)

Countries	Correctness Score Agreement			Diagnostic Score Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement	
		Min	Max		Min	Max
Armenia	96	80	100	94	74	100
Bahrain	98	79	100	96	77	100
Botswana	95	87	99	93	81	98
Bulgaria	95	80	100	92	76	100
Chinese Taipei	96	83	100	94	70	100
Egypt	97	82	100	92	75	100
England	97	92	100	95	83	100
Ghana	99	96	100	97	93	100
Hong Kong SAR	98	94	100	97	90	100
Hungary	96	88	100	94	80	100
Indonesia	98	88	100	95	88	100
Iran, Islamic Rep. of	98	92	100	95	88	99
Israel	95	86	99	91	75	98
Japan	97	91	100	95	80	100
Jordan	97	63	100	96	45	100
Korea, Rep. of	96	86	100	94	81	100
Lithuania	97	82	100	93	70	100
Malaysia	97	89	100	95	82	99
Norway	97	87	100	94	79	100
Palestinian Nat'l Auth.	95	83	100	93	80	100
Russian Federation	98	94	100	95	84	100
Scotland	94	84	100	92	77	100
Serbia	96	87	100	94	85	99
Singapore	96	80	100	94	78	100
Slovenia	96	86	100	94	75	100
Sweden	97	89	100	94	82	100
Tunisia	98	90	100	95	82	100
United States	97	88	100	94	74	100
International Avg.	97	86	100	94	79	100
Benchmark Participants						
Basque Country, Spain	97	89	100	95	80	100
British Columbia, Canada	96	83	100	92	68	99
Ontario, Canada	96	83	100	92	68	99
Quebec, Canada	96	83	100	92	68	99

Exhibit 10.12 TIMSS 2007 Trend Scoring Reliability (2003–2007) for the Fourth Grade Constructed-response Science Items

Countries	Correctness Score Agreement			Diagnostic Score Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement	
		Min	Max		Min	Max
Armenia	91	75	99	80	57	91
Australia	93	88	100	88	77	99
Chinese Taipei	91	33	99	85	33	97
England	95	86	99	90	79	99
Hong Kong SAR	93	86	100	89	73	99
Hungary	94	85	100	88	72	100
Iran, Islamic Rep. of	92	80	99	84	77	98
Japan	92	85	99	87	70	98
Lithuania	94	87	100	85	71	99
Netherlands	92	84	97	85	75	97
New Zealand	94	85	100	87	67	100
Norway	95	88	99	91	81	99
Russian Federation	95	85	100	91	72	97
Scotland	92	80	100	88	69	100
Singapore	92	84	99	88	77	95
Slovenia	89	75	100	65	40	88
Tunisia	94	76	99	86	74	97
United States	92	84	99	84	64	98
International Avg.	93	80	99	86	68	97
Benchmark Participants						
Alberta, Canada	91	80	100	84	65	99
British Columbia, Canada	91	80	100	84	65	99
Ontario, Canada	91	80	100	84	65	99
Quebec, Canada	91	80	100	84	65	99

Exhibit 10.12 TIMSS 2007 Trend Scoring Reliability (2003–2007) for the Eighth Grade Constructed-response Science Items (Continued)

Countries	Correctness Score Agreement			Diagnostic Score Agreement		
	Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement		Average of Exact Percent Agreement Across Items	Range of Exact Percent Agreement	
		Min	Max		Min	Max
Armenia	93	75	99	87	56	99
Bahrain	96	91	99	90	81	97
Botswana	92	79	99	86	67	98
Bulgaria	94	85	100	88	70	100
Chinese Taipei	91	67	100	81	36	100
Egypt	91	74	98	82	65	98
England	91	67	100	87	59	100
Ghana	99	95	100	96	87	99
Hong Kong SAR	95	87	100	91	74	100
Hungary	94	88	99	89	73	98
Indonesia	96	91	100	91	80	100
Iran, Islamic Rep. of	94	86	100	87	72	100
Israel	94	85	100	86	61	100
Japan	94	78	100	85	57	100
Jordan	99	96	100	98	85	100
Korea, Rep. of	94	80	100	88	68	99
Lithuania	94	82	100	87	74	100
Malaysia	95	86	100	91	75	99
Norway	93	84	100	87	72	100
Palestinian Nat'l Auth.	94	87	100	87	76	99
Russian Federation	97	92	100	93	86	99
Scotland	94	83	100	90	68	100
Serbia	95	86	99	90	74	99
Singapore	93	80	100	87	69	100
Slovenia	91	77	99	83	64	99
Sweden	93	83	100	87	76	99
Tunisia	97	84	100	90	75	100
United States	92	79	99	85	71	99
International Avg.	94	83	100	88	70	99
Benchmark Participants						
Basque Country, Spain	95	87	100	90	75	99
British Columbia, Canada	91	79	99	84	65	98
Ontario, Canada	91	79	99	84	65	98
Quebec, Canada	91	79	99	84	65	98

10.3.2.3 Cross-Country Scoring Reliability Study

Because of the many different languages in use in TIMSS 2007, establishing the reliability of constructed-response scoring across all countries was not feasible. However, TIMSS 2007 did conduct a cross-country study of scoring reliability among Northern Hemisphere countries that had scorers who were proficient in English.[8] A sample of student responses was provided by the English-speaking Southern Hemisphere countries. It included 200 student responses for each of 18 fourth-grade and 20 eighth-grade mathematics items and 23 fourth-grade and 20 eighth-grade science items (81 in total, representing about one-quarter of constructed-response items at the two grades) collected from Australia, Botswana, New Zealand, and Singapore. This set of 16,200 student responses in English was then scored independently in each country that had two scorers proficient in English. In all, 52 scorers from 30 countries at fourth grade and 67 scorers from 38 countries at eighth grade participated in the study. Scoring for this study took place shortly after the other scoring reliability activities were completed. Making all possible comparisons among scorers gave 1,225 comparisons at fourth grade and 2,211 comparisons at eighth grade for each student response to each item. This resulted in 265,200 total comparisons at fourth grade and 442,200 total comparisons at eighth grade when aggregated across all 200 student responses to that item. Agreement across countries was defined in terms of the percentage of these comparisons that were in exact agreement.

Exhibits 10.13 and 10.14 show that scorer reliability across countries was high for mathematics, with the percent exact agreement averaging 95 percent across the 18 items for the correctness score and 93 percent for the diagnostic score at fourth grade, and 91 percent across the 20 mathematics items for the correctness score and 90 percent for the diagnostic score at eighth grade. For science, the percent exact agreement averaged 91 percent across the 23 items for the correctness score and 86 percent for the diagnostic score at fourth grade, and 83 percent across the 20 science items for the correctness score and 76 percent for the diagnostic score at eighth grade.

8 See Chapter 6 for further details.

Exhibit 10.13 TIMSS 2007 Cross-country Scoring Reliability for Constructed-response Mathematics Items – Fourth Grade

Item Label	Total Valid Comparisons	Exact Percent Agreement	
		Correctness Score Agreement	Diagnostic Score Agreement
M04_02 - M041056	265200	98	96
M04_04 - M041076	265200	99	98
M04_07 - M041146	265200	92	92
M04_09 - M041258A	265200	96	94
M04_09 - M041258B	265200	86	74
M04_11 - M041275	265200	85	85
M05_02 - M031309	265200	99	99
M05_04 - M031242A	265200	98	97
M05_04 - M031242B	265200	97	96
M05_05 - M031247	265200	94	91
M11_02 - M031009	265200	100	99
M11_04 - M031316	265200	99	99
M11_06 - M031079B	261579	99	99
M11_06 - M031079C	261579	97	97
M11_09 - M031325	265200	97	92
M12_04 - M041059	265200	99	95
M12_13 - M041276A	265200	98	98
M12_13 - M041276B	265200	83	79
Average Percent Agreement		95	93

TIMSS 2007 Cross-country Scoring Reliability for Constructed-response Mathematics Items – Eighth Grade

Item Label	Total Valid Comparisons	Exact Percent Agreement	
		Correctness Score Agreement	Diagnostic Score Agreement
M04_05 - M042304A	442200	94	93
M04_05 - M042304B	442200	86	85
M04_05 - M042304C	442200	93	93
M04_05 - M042304D	442200	79	77
M04_11 - M042130	442200	92	87
M04_12 - M042303A	442200	93	90
M04_12 - M042303B	442200	88	88
M05_03 - M032640	442200	91	91
M05_04 - M032344	442002	94	94
M05_05 - M032754	442200	92	92
M05_06 - M032755	442200	89	84
M11_02 - M032725	442200	94	92
M11_03 - M032683	442200	89	83
M11_13 - M032681A	442200	93	91
M11_13 - M032681B	442200	93	92
M11_13 - M032681C	442200	94	94
M12_03 - M042194	442200	95	95
M12_04 - M042114A	442200	93	91
M12_04 - M042114B	442200	94	94
M12_07 - M042050	442200	95	95
Average Percent Agreement		91	90

Exhibit 10.14 TIMSS 2007 Cross-country Scoring Reliability for Constructed-response Science Items – Fourth Grade

Item Label	Total Valid Comparisons	Exact Percent Agreement	
		Correctness Score Agreement	Diagnostic Score Agreement
S04_02 - S041023	265200	90	88
S04_04 - S041001	265200	74	74
S04_05 - S041029	265200	91	86
S04_08 - S041179	265200	99	99
S04_11 - S041216	265200	94	93
S04_12 - S041061	265200	99	99
S04_13 - S041202	265200	84	81
S05_02 - S031240A	265200	83	74
S05_02 - S031240B	265200	84	76
S05_04 - S031235A	265200	95	87
S05_04 - S031235B	265200	92	83
S05_06 - S031399A	265200	92	85
S05_06 - S031399B	265200	97	86
S05_07 - S031393	265200	93	84
S05_08 - S031278	265200	93	85
S11_03 - S031233	265200	99	97
S11_04 - S031204	265200	88	87
S11_06 - S031299	265200	92	90
S11_10 - S031088A	265200	95	93
S11_10 - S031088B	265200	83	68
S12_01 - S041027	265200	99	99
S12_02 - S041043	265200	89	85
S12_05 - S041006	265200	79	79
Average Percent Agreement		91	86

TIMSS 2007 Cross-country Scoring Reliability for Constructed-response Science Items – Eighth Grade

Item Label	Total Valid Comparisons	Exact Percent Agreement	
		Correctness Score Agreement	Diagnostic Score Agreement
S04_04 - S042052	442200	65	65
S04_06 - S042043	442200	89	66
S04_07 - S042196	416000	90	90
S04_09 - S042292	415874	78	71
S04_11 - S042232A	409600	86	86
S04_11 - S042232B	409600	86	86
S04_13 - S042149	442200	75	75
S04_14 - S042155	442200	83	83
S05_02 - S022292	416000	89	74
S05_06 - S022078	416000	89	79
S05_08 - S022281	416000	88	82
S05_11 - S032519	442200	75	63
S05_14 - S032120A	442200	77	63
S05_14 - S032120B	442200	86	69
S11_03 - S032306	442200	82	78
S11_04 - S032640	442200	81	68
S11_06 - S032570	416000	80	75
S11_08 - S032272	416000	92	82
S11_10 - S032060	442200	93	93
Average Percent Agreement		83	76

10.4 Summary of Review of TIMSS 2007 Item Statistics

Based on the information from the comprehensive collection of item analyses and reliability data that were computed and summarized for TIMSS 2007, as described in this chapter, the TIMSS & PIRLS International Study Center thoroughly reviewed all item statistics for every participating country to ensure that the items were performing comparably across countries. In particular, items with the following problems were considered for possible deletion from the international database:

- An error was detected during TIMSS 2007 translation verification but was not corrected before test administration.
- Data checking revealed a multiple-choice item with more or fewer options than in the international version.
- The item analysis showed the item to have a negative biserial, or, for an item with more than 1 score point, a nonmonotonic relationship between score level and total score.
- The item-by-country interaction results showed a very large negative interaction for a particular country.
- For constructed-response items, the within-country scoring reliability data showed an agreement of less than 70 percent.
- For trend items, an item performed substantially differently in 2007 compared to 2003, or an item was not included in the 2003 assessment for a particular country.

When the item statistics indicated a problem with an item, the documentation from the translation verification[9] was used as an aid in checking the test booklets. If a question remained about potential translation or cultural issues, however, then the National Research Coordinator was consulted before deciding how the item should be treated. If a problem could be detected by the TIMSS & PIRLS International Study Center (such as a negative point-biserial for a correct answer or too few options for a multiple-choice item), the item was deleted from the international scaling.

The checking of the TIMSS 2007 achievement data involved review of 782 items for 59 countries and 8 benchmarking participants at both grades (total of more than 52,000 item-country combinations), and resulted in the detection of very few items that were inappropriate for international comparisons. Among the few items singled out in the review process

9 See Chapter 4 for a description of the process for translating and verifying the TIMSS 2007 data-collection instruments.

were mostly items with differences attributable to either translation or printing problems. Appendix C, Country Adaptations to Items and Item Scoring, provides a list of deleted items, as well as a list of recodes made to constructed-response item codes.

References

TIMSS & PIRLS International Study Center. (2006). *TIMSS 2007 survey operations procedures unit 5: Scoring the TIMSS 2007 assessment*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Chapter 11

Scaling the Data from the TIMSS 2007 Mathematics and Science Assessments

Pierre Foy, Joseph Galia, and Isaac Li

11.1 Overview

The TIMSS 2007 goals of broad coverage of the mathematics and science curriculum and of measuring trends across assessments necessitated a complex matrix-sampling booklet design,[1] with individual students responding to just a subset of the mathematics and science items in the assessment, and not the entire assessment item pool. Given the complexities of the data collection and the need to have student scores on the entire assessment for analysis and reporting purposes, TIMSS 2007 relied on Item Response Theory (IRT) scaling to describe student achievement on the assessment and to provide accurate measures of trends from previous assessments. The TIMSS IRT scaling approach used multiple imputation—or "plausible values"—methodology to obtain proficiency scores in mathematics and science for all students, even though each student responded to only a part of the assessment item pool. To enhance the reliability of the student scores, the TIMSS scaling combined student responses to the items they were administered with information about students' backgrounds, a process known as "conditioning."

This chapter first reviews the psychometric models and the conditioning and plausible values methodology used in scaling the TIMSS 2007 data, and then describes how this approach was applied to the TIMSS 2007 data and to the data from the previous TIMSS 2003 study, in order to measure trends in achievement. It also describes how "bridging" data, specifically collected in TIMSS 2007 to examine for any possible differences between the booklet designs from 2003 and 2007, were used in the scaling to preserve the TIMSS trend measures. The TIMSS scaling was conducted jointly by the TIMSS & PIRLS International Study Center

1 The TIMSS 2007 assessment design is described in Chapter 2.

at Boston College and Educational Testing Service, using software from Educational Testing Service.[2]

11.2 TIMSS 2007 Scaling Methodology[3]

The IRT scaling approach used by TIMSS was developed originally by Educational Testing Service for use in the U.S. National Assessment of Educational Progress. It is based on psychometric models that were first used in the field of educational measurement in the 1950s and have become popular since the 1970s for use in large-scale surveys, test construction, and computer adaptive testing.[4] This approach also has been used to scale IEA's PIRLS data to measure progress in reading literacy.

Three distinct IRT models, depending on item type and scoring procedure, were used in the analysis of the TIMSS 2007 assessment data. Each is a "latent variable" model that describes the probability that a student will respond in a specific way to an item in terms of the student's proficiency, which is an unobserved, or "latent", trait, and various characteristics (or "parameters") of the item. A three-parameter model was used with multiple-choice items, which were scored as correct or incorrect, and a two-parameter model for constructed-response items with just two response options, which also were scored as correct or incorrect. Since each of these item types has just two response categories, they are known as dichotomous items. A partial credit model was used with polytomous constructed-response items, i.e., those with more than two response options.

11.2.1 Two- and Three-Parameter IRT Models for Dichotomous Items

The fundamental equation of the three-parameter (3PL) model gives the probability that a student whose proficiency on a scale k is characterized by the unobservable variable θ_k will respond correctly to item i as:

$$(1) \qquad P\left(x_i = 1 \mid \theta_k, a_i, b_i, c_i\right) = c_i + \frac{1 - c_i}{1 + \exp\left(-1.7 \cdot a_i \cdot (\theta_k - b_i)\right)} \equiv P_{i,1}\left(\theta_k\right)$$

2 TIMSS is indebted to Matthias Von Davier, Ed Kulick, Scott Davis, and John Barone of Educational Testing Service for their advice and support.

3 This section describing the TIMSS scaling methodology has been adapted with permission from Chapter 14 of the *TIMSS 1999 Technical Report* (Yamamoto and Kulick, 2000).

4 For a description of IRT scaling see Birnbaum (1968); Lord and Novick (1968); Lord (1980); Van Der Linden and Hambleton (1996). The theoretical underpinning of the multiple imputation methodology was developed by Rubin (1987), applied to large-scale assessment by Mislevy (1991), and studied further by Mislevy, Johnson and Muraki (1992), and Beaton and Johnson (1992). The procedures used in TIMSS have been used in several other large-scale surveys, including Progress in International Reading Literacy Study (PIRLS), the U.S. National Assessment of Educational Progress (NAEP), the U.S. National Adult Literacy Survey (NALS), the International Adult Literacy Survey (IALS), and the International Adult Literacy and Life Skills Survey (IALLS).

where

x_i is the response to item i, 1 if correct and 0 if incorrect;

θ_k is the proficiency of a student on a scale k (note that a student with higher proficiency has a greater probability of responding correctly);

a_i is the slope parameter of item i, characterizing its discriminating power;

b_i is the location parameter of item i, characterizing its difficulty;

c_i is the lower asymptote parameter of item i, reflecting the chances of students with very low proficiency selecting the correct answer.

The probability of an incorrect response to the item is defined as:

$$(2) \qquad P_{i,0} = P\left(x_i = 0 \mid \theta_k, a_i, b_i, c_i\right) = 1 - P_{i,1}\left(\theta_k\right)$$

The two-parameter (2PL) model was used for the constructed-response items that were scored as either correct or incorrect. The form of the 2PL model is the same as Equations (1) and (2) with the c_i parameter fixed at zero.

11.2.2 IRT Model for Polytomous Items

In TIMSS 2007, as in previous study cycles, constructed-response items requiring an extended response were scored for partial credit, with 0, 1, and 2 as the possible score levels. These polytomous items were scaled using a generalized partial credit model (Muraki, 1992). The fundamental equation of this model gives the probability that a student with proficiency θ_k on scale k will have, for the i^{th} item, a response x_i that is scored in the l^{th} of m_i ordered score categories as:

$$(3) \qquad P\left(x_i = l \mid \theta_k, a_i, b_i, d_{i,1}, \cdots, d_{i,m_i-1}\right) = \frac{\exp\left(\sum_{v=0}^{l} 1.7 \cdot a_i \cdot \left(\theta_k - b_i + d_{i,v}\right)\right)}{\sum_{g=0}^{m_i-1} \exp\left(\sum_{v=0}^{g} 1.7 \cdot a_i \cdot \left(\theta_k - b_i + d_{i,v}\right)\right)} \equiv P_{i,l}\left(\theta_k\right)$$

where

m_i is the number of response categories for item i, usually 3;

x_i is the response to item i, ranging between 0 and m_i –1;

θ_k is the proficiency of a student on a scale k;

a_i is the slope parameter of item i;

b_i is its location parameter, characterizing its difficulty;

$d_{i,l}$ is the category l threshold parameter (l = 0, ..., m_i –1).

The indeterminacy of model parameters in the polytomous model is resolved by setting $d_{i,0} = 0$ and $\sum_{j=1}^{m_i - 1} d_{i,j} = 0$.

For all of the IRT models there is a linear indeterminacy between the values of item parameters and proficiency parameters, i.e., mathematically equivalent but different values of item parameters can be estimated on an arbitrarily linearly transformed proficiency scale. This linear indeterminacy can be resolved by setting the origin and unit size of the proficiency scale to arbitrary constants, such as a mean of 500 and a standard deviation of 100, as was done for TIMSS back in 1995. The indeterminacy is most apparent when the scale is set for the first time.

IRT modeling relies on a number of assumptions, the most important being conditional independence. Under this assumption, item response probabilities depend only on θ_k (a measure of a student's proficiency) and the specified parameters of the item, and are unaffected by the demographic characteristics or unique experiences of the students, the data collection conditions, or the other items presented in the test. Under this assumption, the joint probability of a particular response pattern x across a set of n items is given by:

(4) $$P\left(x \mid \theta_k, \textit{item parameters}\right) = \prod_{i=1}^{n} \prod_{l=0}^{m_i - 1} P_{i,l}\left(\theta_k\right)^{u_{i,l}}$$

where $P_{i,l}(\theta_k)$ is of the form appropriate to the type of item (dichotomous or polytomous), m_i is equal to 2 for dichotomously scored items, and $u_{i,l}$ is an indicator variable defined as:

(5) $$u_{i,l} = \begin{cases} 1 & \textit{if response } x_i \textit{ is in category } l; \\ 0 & \textit{otherwise.} \end{cases}$$

Replacing the hypothetical response pattern with the real scored data, the above function can be viewed as a likelihood function to be maximized by a given set of item parameters. In TIMSS 2007, the item parameters for each scale were estimated independently of the parameters of other scales. Once items were calibrated in this manner, a likelihood function for the proficiency θ_k was induced from student responses to the calibrated items. This likelihood function for the proficiency θ_k is called the posterior distribution of the θ's for each student.

11.2.3 Proficiency Estimation Using Plausible Values

Most cognitive skills testing is concerned with accurately assessing the performance of individual students for the purposes of diagnosis, selection, or placement. Regardless of the measurement model used, whether classical test theory or item response theory, the accuracy of these measurements can be improved—that is, the amount of measurement error can be reduced—by increasing the number of items given to the individual. Thus, it is common to see achievement tests designed to provide information on individual students that contain more than 70 items. Since the uncertainty associated with each θ in such tests is negligible, the distribution of θ, or the joint distribution of θ with other variables, can be approximated using each individual's estimated θ.

For the distribution of proficiencies in large populations, however, more efficient estimates can be obtained from a matrix-sampling design like that used in TIMSS. This design solicits relatively few responses from each sampled student while maintaining a wide range of content representation when responses are aggregated across all students. With this approach, however, the advantage of estimating population characteristics more efficiently is offset by the inability to make precise statements about individuals. The uncertainty associated with individual θ estimates becomes too large to be ignored. In this situation, aggregations of individual student scores can lead to seriously biased estimates of population characteristics (Wingersky, Kaplan, & Beaton, 1987).

Plausible values methodology was developed as a way to address this issue. Instead of first computing estimates of individual θ's and then aggregating these to estimate population parameters, the plausible values approach uses all available data, students' responses to the items they were administered together with all background data, to estimate directly the characteristics of student populations and subpopulations. Although these directly estimated population characteristics could be used for reporting purposes, instead the usual plausible values approach is to generate multiple imputed scores, called plausible values, from the estimated ability distributions and to use these in analyses and reporting, making use of standard statistical software. By including all available background data in the model, a process known as "conditioning", relationships between these background variables and the estimated proficiencies will be appropriately accounted for in the plausible values. Because of this, analyses conducted using plausible values will provide an accurate representation of these underlying relationships. A detailed review of the plausible values methodology is given in Mislevy (1991).[5]

The following is a brief overview of the plausible values approach. Let y represent the responses of all sampled students to background questions or background data of sampled students collected from other sources, and let θ represent the proficiency of interest. If θ were known for all sampled students, it would be possible to compute a statistic $t(\theta,y)$, such as a sample mean or sample percentile point, to estimate a corresponding population quantity T.

Because of the latent nature of the proficiency, however, θ values are not known even for sampled students. The solution to this problem is to follow Rubin (1987) by considering θ as "missing data" and approximate $t(\theta,y)$ by its expectation given (x,y), the data that actually were observed, as follows:

$$
\begin{aligned}
t^*(x,y) &= E\left|\, t(\underline{\theta},\underline{y}) \mid \underline{x},\underline{y} \,\right| \\
&= \int t(\underline{\theta},\underline{y})\, p\left(\underline{\theta} \,\middle|\, \underline{x},\underline{y}\right) d\underline{\theta}
\end{aligned}
\tag{6}
$$

It is possible to approximate t^* using random draws from the conditional distribution of the scale proficiencies given the student's item responses x_j, the student's background variables y_j, and model parameters for the items. These values are referred to as imputations in the sampling literature, and as plausible values in large-scale surveys such as PIRLS, TIMSS, NAEP, NALS, and IALLS. The value of θ for any student that would enter into the computation of t is thus replaced by a randomly selected value from his or her conditional distribution. Rubin (1987) proposed repeating this process several times so that the uncertainly associated with imputation can be quantified. For example, the average of multiple estimates of t, each computed

5 Along with theoretical justifications, Mislevy presents comparisons with standard procedures; discusses biases that arise in some secondary analyses; and offers numerical examples.

from a different set of plausible values, is a numerical approximation of t^* of the above equation; the variance among them reflects the uncertainty due to not observing $\underline{\theta}$. It should be noted that this variance does not include the variability of sampling from the population. That variability is estimated separately by a jackknife variance estimation procedure, which is presented later in this chapter.

Plausible values are not intended to be estimates of individual student scores, but rather are imputed scores for like students—students with similar response patterns and background characteristics in the sampled population—that may be used to estimate population characteristics correctly. When the underlying model is correctly specified, plausible values will provide consistent estimates of population characteristics, even though they are not generally unbiased estimates of the proficiencies of the individuals with whom they are associated. Taking the average of the plausible values still will not yield suitable estimates of individual student scores.[6]

Plausible values for each student j are drawn from the conditional distribution $P\left(\theta_j \middle| x_j, y_j, \Gamma, \Sigma\right)$, where Γ is a matrix of regression coefficients for the background variables, and Σ is a common variance matrix of residuals. Using standard rules of probability, the conditional probability of proficiency can be represented as:

$$(7) \quad P\left(\theta_j \middle| x_j, y_j, \Gamma, \Sigma\right) \propto P\left(x_j \middle| \theta_j, y_j, \Gamma, \Sigma\right) P\left(\theta_j \middle| y_j, \Gamma, \Sigma\right) = P\left(x_j \middle| \theta_j\right) P\left(\theta_j \middle| y_j, \Gamma, \Sigma\right)$$

where θ_j is a vector of scale values, $P\left(x_j \middle| \theta_j\right)$ is the product over the scales of the independent likelihoods induced by responses to items within each scale, and $P\left(\theta_j \middle| y_j, \Gamma, \Sigma\right)$ is the multivariate joint density of proficiencies for the scales, conditional on the observed values y_j of background responses and parameters Γ and Σ. Item parameter estimates are fixed and regarded as population values in the computations described in this section.

11.2.4 Conditioning

A multivariate normal distribution was assumed for $P\left(\theta_j \middle| y_j, \Gamma, \Sigma\right)$, with a common variance Σ, and with a mean given by a linear model with regression parameters Γ. Since in large-scale studies like TIMSS there are many hundreds of background variables, it is customary to conduct a principal components analysis to reduce the number of variables to be

6 For further discussion, see Mislevy, Beaton, Kaplan, and Sheehan (1992).

used in Γ. Typically, components accounting for 90 percent of the variance in the data are selected. These principal components are referred to as the conditioning variables and denoted as y^c. The following model is then fit to the data:

(8) $$\theta = \Gamma' y^c + \varepsilon$$

where ε is normally distributed with mean zero and variance Σ. As in a regression analysis, Γ is a matrix each of whose columns is the effects for each scale and Σ is the matrix of residual variance between scales.

Note that in order to be strictly correct for all functions Γ of θ, it is necessary that $P(\theta \mid y)$ be correctly specified for all background variables in the survey. Estimates of functions Γ involving background variables not conditioned in this manner are subject to estimation error due to misspecification. The nature of these errors is discussed in detail in Mislevy (1991). In TIMSS 2007, however, principal component scores based on nearly all background variables were used. Those selected variables were chosen to reflect high relevance to policy and to education practices. The computation of marginal means and percentile points of θ for these variables is nearly optimal.

The basic method for estimating Γ and Σ with the Expectation and Maximization (EM) procedure is described in Mislevy (1985) for a single scale case. The EM algorithm requires the computation of the mean θ, and variance Σ, of the posterior distribution in equation (7).

11.2.5 Generating Proficiency Scores

After completing the EM algorithm, plausible values for all sampled students are drawn from the joint distribution of the values of Γ in a three-step process. First, a value of Γ is drawn from a normal approximation $P(\Gamma, \Sigma \mid x_j, y_j)$ to that fixes Σ at the value $\hat{\Sigma}$ (Thomas, 1993). Second, conditional on the generated value of Γ (and the fixed value of $\Sigma = \hat{\Sigma}$), the mean θ_j and variance Σ_j^p of the posterior distribution in equation (7), where p is the number of scales, are computed using the methods applied in the EM algorithm. In the third step, the proficiency values are drawn independently from a multivariate normal distribution with mean θ_j and variance Σ_j^p

These three steps are repeated five times, producing five imputations of θ_j for each sampled student.

For students with an insufficient number of responses, the Γ's and Σ's described in the previous paragraph are fixed. Hence, all students—regardless of the number of items attempted—are assigned a set of plausible values.

The plausible values can then be employed to evaluate equation (6) for an arbitrary function T as follows:

- Using the first vector of plausible values for each student, evaluate T as if the plausible values were the true values of θ. Denote the result as T_1.
- Evaluate the sampling variance of T_1, or Var_1, with respect to students' first vector of plausible values.
- Carry out steps 1 and 2 for the second through fifth vectors of plausible values, thus obtaining T_u and Var_u for $u = 2, \ldots, 5$.
- The best estimate of T obtainable from the plausible values is the average of the five values obtained from the different sets of plausible values:

$$\hat{T} = \frac{\sum_u T_u}{5}$$

- An estimate of the variance of $\hat{T}$ is the sum of two components: an estimate of Var_u obtained by averaging as in the previous step, and the variance among the T_u's.

  Let $\bar{U} = \frac{\sum_u Var_u}{M}$, and let $B_M = \frac{\sum_u \left(T_u - \hat{T}\right)^2}{M-1}$ be the variance among the M plausible values. Then the estimate of the total variance of $\hat{T}$ is:

(9) $$Var\left(\hat{T}\right) = \bar{U} + \left(1 + M^{-1}\right) B_M$$

The first component in $Var\left(\hat{T}\right)$ reflects the uncertainty due to sampling students from the population; the second reflects the uncertainty due to the fact that sampled students' θ's are not known precisely, but only indirectly through x and y.

11.2.6 Working with Plausible Values

The plausible values methodology was used in TIMSS 2007 to ensure the accuracy of estimates of the proficiency distributions for the TIMSS population as a whole and particularly for comparisons between subpopulations. A further advantage of this method is that the variation between the five plausible values generated for each student reflects the uncertainty associated with proficiency estimates for individual students. However, retaining this component of uncertainty requires that additional analytical procedures be used to estimate students' proficiencies.

If the θ values were observed for all sampled students, the statistic $\left(t-T\right)/U^{1/2}$ would follow a t-distribution with d degrees of freedom. Then the incomplete-data statistic $\left(T-\hat{T}\right) / \left[Var\left(\hat{T}\right)\right]^{1/2}$ is approximately t-distributed, with degrees of freedom (Johnson & Rust, 1992) given by:

$$\nu = \frac{1}{\dfrac{f_M{}^2}{M-1} + \dfrac{\left(1-f_M\right)^2}{d}} \tag{10}$$

where d is the degrees of freedom for the complete-data statistic, and f_M is the proportion of total variance due to not observing the θ values:

$$f_M = \frac{\left(1+M^{-1}\right) B_M}{Var\left(\hat{T}\right)} \tag{11}$$

When B_M is small relative to $\bar{U}$, the reference distribution for the incomplete-data statistic differs little from the reference distribution for the corresponding complete-data statistic. If, in addition, d is large, the normal approximation can be used instead of the t-distribution.

For a k-dimensional function T, such as the k coefficients in a multiple regression analysis, each U and $\bar{U}$ is a covariance matrix, and B_M is an average of squares and cross-products rather than simply an average of squares. In this case, the quantity $\left(\underline{T}-\hat{\underline{T}}\right) Var^{-1}\left(\hat{\underline{T}}\right)\left(\underline{T}-\hat{\underline{T}}\right)'$ is approximately

F-distributed with degrees of freedom equal to k and ν, with ν defined as above but with a matrix generalization of f_M:

$$f_M = \left(1+M^{-1}\right) Trace\left[B_M Var^{-1}\left(\hat{T}\right)\right] \Big/ k \tag{12}$$

For the same reason that the normal distribution can approximate the *t*-distribution, a chi-square distribution with k degrees of freedom can be used in place of the *F*-distribution for evaluating the significance of the above quantity $\left(\underline{T}-\hat{\underline{T}}\right)' Var^{-1}\left(\hat{\underline{T}}\right)\left(\underline{T}-\hat{\underline{T}}\right)$.

Statistics $\hat{T}$, the estimates of proficiency conditional on responses to cognitive items and background variables, are consistent estimates of the corresponding population values T, as long as background variables are included in the conditioning variables. The consequences of violating this restriction are described by Beaton & Johnson (1990), Mislevy (1991), and Mislevy & Sheehan (1987). To avoid such biases, the TIMSS 2007 analyses included all student background variables, as well as the class means to preserve between-class differences—the between- and within-classroom variance structure essential for hierarchical modeling.

11.3 Implementing the Scaling Procedures for the TIMSS 2007 Assessment Data

The application of IRT scaling and plausible values methodology to the TIMSS 2007 assessment data involved four major tasks: calibrating the achievement test items (estimating model parameters for each item), creating principal components from the student questionnaire data for use in conditioning; generating IRT scale scores (proficiency scores) for overall mathematics and science and for each of the mathematics and science content and cognitive domains; and placing the proficiency scores on the metric used to report the results from previous assessments.

The TIMSS eighth-grade reporting metric was established in 1995 by setting the average of the mean scores of the countries that participated in TIMSS 1995 at the eighth grade to 500 and the standard deviation to 100. To enable comparisons between 2007, 2003, 1999 and 1995, the TIMSS 2007, TIMSS 2003, and TIMSS 1999 eighth-grade data also were placed on this metric. This was done by concurrently scaling the assessment data from each successive TIMSS cycle with the assessment data from the previous cycle and applying linear transformations to set the scores from each successive cycle on the same metric as the scores from the previous cycle. Placing the TIMSS 2007 eighth-grade results on this common metric permitted trend results from four points in time: 1995, 1999, 2003, and 2007.

The TIMSS fourth-grade reporting metric was set in much the same way as was done for the eighth grade, with the notable exception that TIMSS 1999 did not have a fourth-grade assessment. The TIMSS 2003 fourth-grade data were placed directly on the 1995 fourth-grade scale, which also had a mean of 500 and standard deviation of 100 based on the countries that participated in TIMSS 1995 at the fourth grade. This enabled comparisons between results from 1995 and 2003. Subsequently, the TIMSS 2007 fourth-grade data were put on the 1995 metric to produce trend results from all three survey cycles: 1995, 2003, and 2007. In 2007, as in previous TIMSS cycles, scale metrics were aligned for trend reporting only for overall mathematics and overall science; there were insufficient trend items from previous survey cycles to reliably measure trends in the content and cognitive domains.

11.3.1 The Bridging Study

In 2003, TIMSS introduced a new assessment design, consisting of a series of interlinked student booklets, each containing six blocks of assessment items.[7] From examination of the TIMSS 2003 data, it was apparent that not all students had sufficient time to complete their 2003 assessment booklets. This led to a "position effect",[8] whereby items positioned later in a booklet appeared to be more difficult than the same items positioned earlier in the booklet. The position effect was detectable because of the counterbalanced design of the 2003 assessment booklets. A new booklet design was introduced in TIMSS 2007, providing more time for students to respond to the items. Unlike the TIMSS 2003 booklets, which each contained six blocks of items, the TIMSS 2007 booklets each comprised just four of these blocks, to be completed in the same amount of time (i.e., 72 minutes at the fourth grade and 90 minutes at the eighth grade). Concerned that the 2007 assessment booklets might appear easier because students had more time, TIMSS implemented a "bridging study" to see if this was indeed the case. The bridging study involved the administration of a subset of the TIMSS 2003 assessment booklets at both grades in 2007 to establish a bridge between the 2003 and 2007 assessments. The data from the bridging study would

7 The TIMSS 2003 assessment design is described in the *TIMSS Assessment Frameworks and Specifications 2003 – 2nd Edition* (Mullis, et al., 2003).

8 The TIMSS 2003 position effect is described in the *TIMSS 2003 Technical Report* (Martin, et al., 2004, p. 264).

reveal if the change in booklet design from 2003 to 2007 had any effect on the difficulty of the achievement items, and if so, would provide a basis for maintaining the measurement of trends by adjusting for this effect.

It was important to establish that a subset of 2003 booklets could be a suitable representation of the TIMSS 2003 assessment as a whole. This evaluation was done by re-scaling the 2003 data using items only from four selected 2003 booklets: booklets 5, 6, 11, and 12. These were selected to maximize the number of common item blocks between the 2003 and 2007 assessments. A comparison of the resulting national average scale scores to the ones published in the 2003 international reports, showed that virtually all differences were well within sampling error. As well, an examination of Cronbach's alpha reliability coefficients across the set of items in these four booklets revealed that they remained as high, or nearly so, when compared to the reliability coefficients across all TIMSS 2003 items.

By inserting them into the rotation of the fourteen 2007 assessment booklets, the four bridge booklets were administered alongside the TIMSS 2007 assessment booklets to randomly equivalent samples of students in all trend countries (countries that participated in both TIMSS 2003 and TIMSS 2007).[9] All item blocks in the bridge booklets also were part of the TIMSS 2003 assessment, and four mathematics and four science blocks in the bridge booklets (at each grade level) also were included in the TIMSS 2007 assessment booklets. Presenting the same items using the 2007 bridge booklets and the 2007 assessment booklets allowed TIMSS to isolate the effect of changing the booklet design, and to provide enough data to adjust for this effect, as necessary.

A comparison of the average percent correct statistics of the common items in the 2007 bridge booklets and 2007 assessment booklets confirmed that the items were easier, on average, in the TIMSS 2007 assessment booklets, particularly at the eighth grade, as shown in Exhibit 11.1. The percent correct averaged across all fourth-grade mathematics items were 0.3% higher in the 2007 assessment booklets; the fourth-grade science items were 0.9% higher. The percent correct averaged across the eighth-grade mathematics items were 1.2% higher; the eighth-grade science items were 1.1% higher. Thus, because of the change in booklet design, the trend items in the TIMSS 2007 assessment booklets could not be assumed to have behaved as they had in the TIMSS 2003 booklets. The bridging data

9 The assignment of TIMSS 2007 bridge booklets and TIMSS 2007 assessment booklets was done automatically by the WinW3S software, as described in Chapter 6.

show what could have been expected if the booklet design had not been changed. Consequently, it was necessary to incorporate this effect into the trend scaling. The trend scaling of overall mathematics and overall science was performed by combining the assessment data from the TIMSS 2003 assessment booklets, the TIMSS 2007 bridge booklets, and the TIMSS 2007 assessment booklets using all items from the bridge booklets as trend items from the 2003 assessment and freeing all items in the 2007 assessment booklets to have their own IRT model parameters.

Exhibit 11.1 Overall Percent Correct and Percent Not Reached for Common Items in TIMSS 2007 Bridge Booklets and Assessment Booklets

Grade and Subject		Number of Common Items	TIMSS 2007 Bridge Booklets		TIMSS 2007 Assessment Booklets	
			Overall Percent Correct	Overall Percent Not Reached	Overall Percent Correct	Overall Percent Not Reached
Fourth Grade (19 Countries)	Mathematics	47	53.4	1.2	53.7	2.1
	Science	47	58.1	0.4	59.0	1.9
Eighth Grade (32 Countries)	Mathematics	52	44.6	0.2	45.8	1.3
	Science	57	43.6	0.1	44.7	1.2

11.3.2 Calibrating the TIMSS 2007 Assessment Data

As described in the TIMSS 2007 Assessment Frameworks (Mullis, Martin, Ruddock, O'Sullivan, Arora, & Erberber, 2005), the TIMSS 2007 achievement test design consisted of a total of 14 mathematics blocks and 14 science blocks at each grade, distributed across 14 assessment booklets. Each block contained either mathematics or science items, drawn from a range of content and cognitive domains. The 14 mathematics blocks were designated M01 through M14, and the 14 science blocks S01 through S14. All odd-numbered item blocks were previously used in the 2003 assessment and all even-numbered blocks consisted of newly-developed items for the 2007 assessment. Each assessment booklet contained four blocks—two mathematics and two science blocks. Two of the blocks (one mathematics and one science) were new in 2007 and two had previously been used in 2003.

The TIMSS 2007 test administration also included the four bridge booklets for trend countries, i.e., countries that also had participated in the 2003 assessment. Thus each sampled student in a trend country completed either one of the fourteen 2007 assessment booklets, or one of the four 2007 bridge booklets. Students in "non-trend" countries completed one of the fourteen 2007 assessment booklets. The booklets were distributed among the students in each sampled class according to a scheme that ensured comparable random samples of students responded to each booklet, including the bridge booklets in trend countries.

In line with the TIMSS assessment framework, IRT scales were constructed for reporting overall student achievement in mathematics and science, as well as for reporting separately for each of the mathematics and science content and cognitive domains. Item calibration for the content and cognitive domains was conducted by the TIMSS & PIRLS International Study Center using the commercially-available Parscale software (Muraki & Bock, 1991). Item calibration for the overall mathematics and science scales was performed by ETS using their in-house version of Parscale and included data from the TIMSS 2003 assessment, the TIMSS 2007 assessment and the 2007 bridging study. The calibration was conducted using all available data from each country's TIMSS student samples and from all three assessments. All student samples were weighted so that each country contributed equally to the item calibration.

The first step in constructing the scales for TIMSS 2007 was to estimate the IRT model item parameters for each item on each of the scales. The trend scales for overall mathematics and science typically are based on a concurrent item calibration approach. The general concurrent calibration approach consists of three steps that look to build a linkage between the item calibration that was done in the previous assessment—called the previous calibration—and the current assessment. The first step consists of establishing a common set of item parameters for the two assessments through a concurrent calibration of both sets of assessment data, and setting common items to have the same item parameter estimates. It is then possible to obtain the mean and standard deviation of the latent ability distribution of students in both assessments under the concurrent calibration. The difference between these two distributions is the change in achievement from the previous to the current assessment. However, this difference is in the logit metric, and not the metric of the previous assessment, which would be necessary to measure growth.

The second step is to find the linear transformation that transforms the distribution of the previous assessment data under the concurrent calibration to match the distribution of these data under the previous calibration. The third step is to apply this same transformation to the current assessment data scaled using the concurrent calibration. This places the current assessment data on the metric of the previous assessment.

Exhibit 11.2 illustrates how the concurrent calibration approach customarily has been applied in the context of TIMSS trend scaling. The observed gap between both calibrations on the previous assessment data is generally small and arises from slight differences in the item parameter estimations, which in turn are due mostly to the previous assessment data being calibrated with other assessment data in the two calibrations. The linear transformation removes this gap by shifting the two distributions from the concurrent calibration, such that the distribution of the previous assessment from the concurrent calibration aligns with the distribution of the previous assessment from the previous calibration, while preserving the gap between the previous and current assessment data under the concurrent calibration. This latter gap is the change in achievement between the previous and current assessments that TIMSS seeks to measure as its trend.

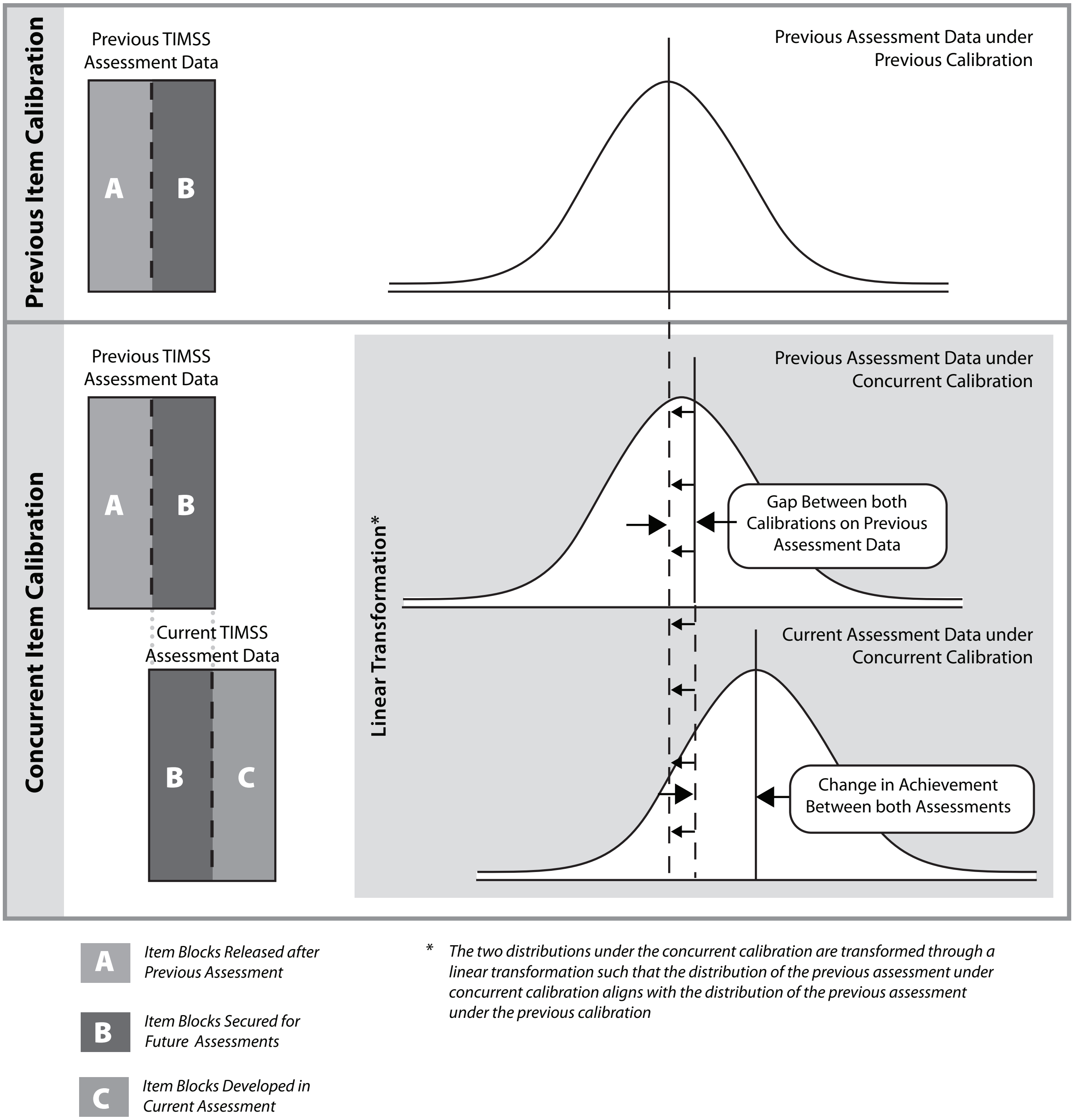
Because the bridging study demonstrated that the common items did not behave similarly across the 2003 and 2007 assessment booklets, it was necessary to adapt the concurrent calibration approach to include the 2007 bridging data. Accordingly, the 2007 concurrent calibration included the original 2003 data, the 2007 bridging data, and the 2007 data. Only countries that participated in both 2003 and 2007 were included in this concurrent calibration. All of the items contained in the 2007 bridge booklets also were contained in the 2003 booklets, so that these received the same item parameters in the concurrent scaling. This constituted the link between the 2003 assessment and the 2007 bridging data. The 2007 bridge booklets and the 2007 assessment booklets were administered to randomly equivalent samples of the 2007 assessment populations, which constituted the link between the 2007 bridging data and the 2007 assessment data.

Having estimated the item parameters from the concurrent calibration, new achievement distributions were generated by applying these item parameters to the 2003 assessment data, the 2007 bridging data, and the 2007 assessment data. Following the procedure outlined above, the next step was to identify the linear transformation that transformed the 2003 assessment distribution generated by the concurrent calibration item parameters to match the 2003 assessment distribution generated by the item parameters from the original 2003 calibration, and to apply this same transformation to the 2007 bridging data distribution (also generated by the concurrent calibration item parameters). An additional step, however, was required to establish a second linear transformation to make the distribution of the

2007 assessment data match the now-transformed distribution of the 2007 bridging data. This was done on the basis that both the 2007 assessment data and the 2007 bridging data came from randomly equivalent samples of the same 2007 assessment population.

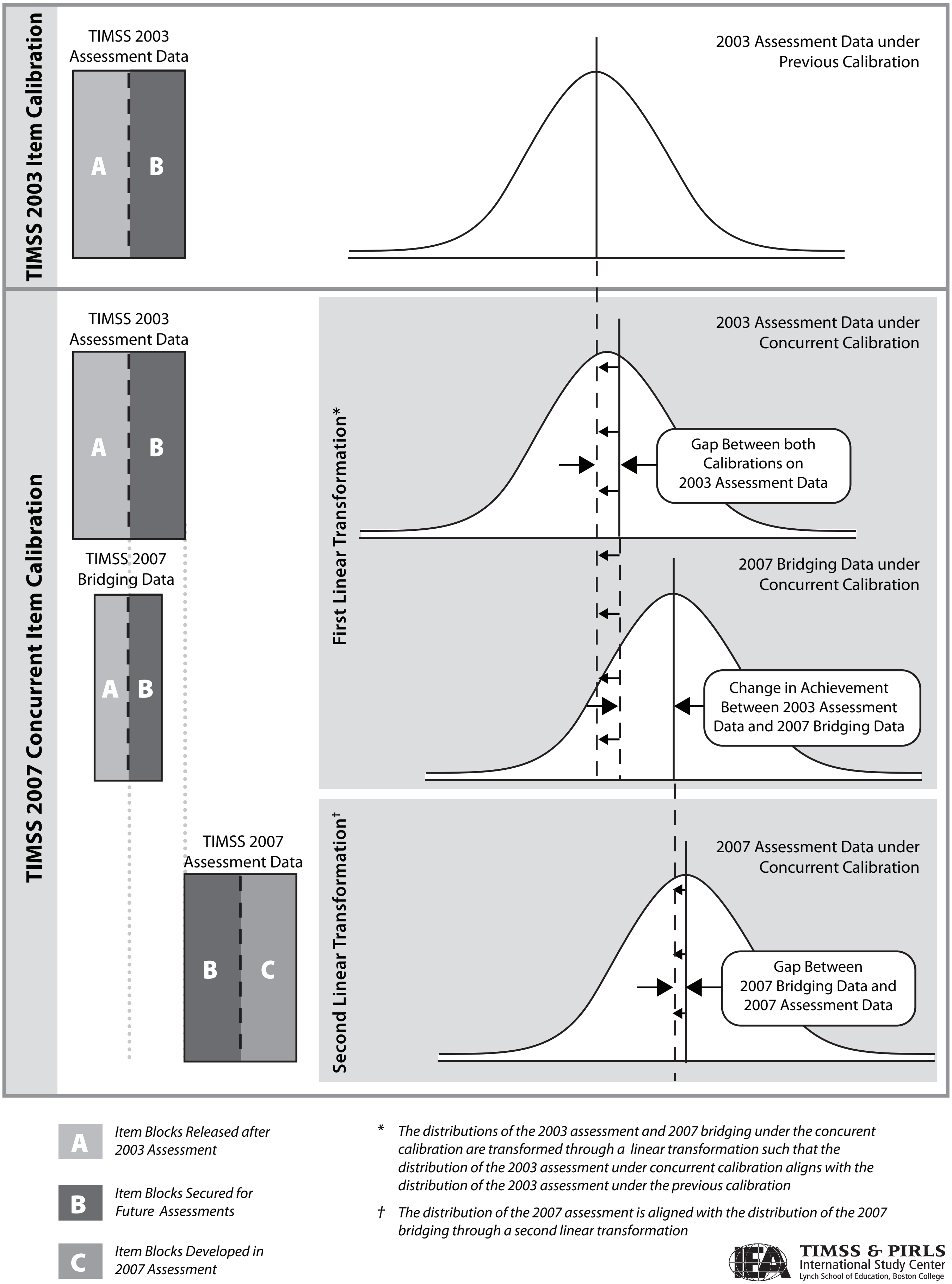
Exhibit 11.3 demonstrates how this modified concurrent calibration approach was implemented in TIMSS 2007. As was explained in Exhibit 11.2, the gap between both calibrations on the 2003 assessment data was due largely to minor differences in the estimated item parameters arising from the fact that the 2003 assessment data were combined with the 1999 assessment data (the 1995 assessment data at the fourth grade) in the 2003 calibration and combined with the 2007 bridging data and 2007 assessment data in the 2007 calibration. The first linear transformation served to remove this gap while preserving the gap between the 2003 assessment data and the 2007 bridging data under the 2007 concurrent calibration, which was the change in achievement used to determine the TIMSS measure of trend. Finally, the gap between the 2007 bridging data and 2007 assessment data was primarily the result of minor sampling differences across the national samples of students between the two sets of data and was removed by the second linear transformation, which aligned the distribution of the 2007 assessment data with the distribution of the 2007 bridging data.

Exhibit 11.2 Concurrent Calibration Model Used Traditionally for TIMSS



* *The two distributions under the concurrent calibration are transformed through a linear transformation such that the distribution of the previous assessment under concurrent calibration aligns with the distribution of the previous assessment under the previous calibration*

Exhibit 11.3 Concurrent Calibration Model Used for TIMSS 2007



* *The distributions of the 2003 assessment and 2007 bridging under the concurent calibration are transformed through a linear transformation such that the distribution of the 2003 assessment under concurrent calibration aligns with the distribution of the 2003 assessment under the previous calibration*

† *The distribution of the 2007 assessment is aligned with the distribution of the 2007 bridging through a second linear transformation*

Exhibit 11.4 shows the distribution of items included in the TIMSS 2007 concurrent calibrations for reporting trends in overall mathematics and science at both grades. All data were included from the 2003 and 2007 assessments, as well as the data from the 2007 bridge booklets to account for the modified TIMSS 2007 assessment design. Items were categorized as items unique to the TIMSS 2003 assessment, items in the TIMSS 2007 bridge booklets—which by design also were included in the TIMSS 2003 assessment and constituted the set of common items—and items in the TIMSS 2007 assessment booklets. Taking eighth-grade mathematics as an example, the TIMSS 2007 assessment booklets contributed 214 items worth 236 points, the TIMSS 2007 bridge booklets contributed 151 items worth 165 points (these same items were also in the TIMSS 2003 assessment booklets), and there were 216 items worth 237 points unique to the TIMSS 2003 assessment booklets.

Exhibit 11.4 Items Included in the TIMSS 2007 Concurrent Item Calibrations of Overall Mathematics and Science

TIMSS 2007 Trend Scales		Items in TIMSS 2007 Assessment Booklets		Items in TIMSS 2007 Bridge Booklets		Items Unique to TIMSS 2003 Assessment Booklets		TOTAL	
		Number	Points	Number	Points	Number	Points	Number	Points
Fourth Grade	Mathematics	177	188	125	130	171	179	473	497
	Science	170	189	119	130	159	175	448	494
Eighth Grade	Mathematics	214	236	151	165	216	237	581	638
	Science	210	231	151	163	202	220	563	614

At the fourth grade, to construct separate overall mathematics and science scales for reporting trends, as well as performance generally in 2007, concurrent item calibrations were conducted using data from the 21 countries that participated in both 2003 and 2007 assessments. These calibrations included 93,863 student records from the 2003 assessment, 25,952 records from the 2007 bridging study, and 91,204 records from the 2007 assessment, for a total of 211,019 student records. The item parameters established in these calibrations were used subsequently for estimating student scores for all 37 countries and 7 benchmarking entities that participated in 2007.

At the eighth grade, concurrent item calibrations for the overall mathematics and science scales were conducted using data from the 33 countries that participated in both 2003 and 2007 assessments. They included 158,477 student records from the 2003 assessment, 41,377 records from the 2007 bridging study, and 145,349 records from the 2007 assessment, for a total of 345,203 student records. The item parameters established in these calibrations were used subsequently for estimating student scores for all 50 countries and 7 benchmarking entities that participated in 2007. All countries and their samples included in these calibrations for reporting trends are presented in Exhibit 11.5.

Because there were insufficient items to construct reliable scales for measuring trends in each of the content and cognitive domains, scales for these domains were constructed using 2007 data only. At the fourth grade, separate calibrations were conducted for each of the three mathematics and three science content domains and the three mathematics and three science cognitive domains. These calibrations were based on 160,922 student records from the 36 countries that participated in the 2007 assessment.[10] Similarly at the eighth grade, separate calibrations were conducted for each of the four mathematics and four science content domains and the three mathematics and three science cognitive domains. These calibrations were based on 220,788 student records from the 49 countries that participated in the 2007 assessment at the eighth grade.[10] All countries and their samples included in the item calibrations for the content and cognitive domains are presented in Exhibit 11.6.

Item calibrations for the content and cognitive domains included only the items from the TIMSS 2007 assessment booklets. Exhibit 11.7 and Exhibit 11.8 show the number of items and score points included in each content and cognitive domain at the fourth and eighth grades, respectively.

Exhibits D.1 through D.30 in Appendix D present the item parameters generated from all item calibrations. In Exhibits D.1 through D.4, items where the parameters were freed in 2003, to address the position effect in 2003, have an “F” in the second character position of the item label. All items from the TIMSS 2007 assessment booklets have the letter “Z” in the second character position of the item label. As a by-product of the calibrations, interim scores in mathematics, science, and all content and cognitive domains were produced for use in constructing conditioning variables.

10 Data from Mongolia and the seven benchmarking participants were not included in these item calibrations.

Exhibit 11.5 Sample Sizes for Item Calibrations of Overall Mathematics and Science for Countries Participating in both TIMSS 2003 and TIMSS 2007

Country	Fourth Grade			Eighth Grade		
	TIMSS 2003 Assessment Booklets	TIMSS 2007 Bridge Booklets	TIMSS 2007 Assessment Booklets	TIMSS 2003 Assessment Booklets	TIMSS 2007 Bridge Booklets	TIMSS 2007 Assessment Booklets
Armenia	5,674	1,139	4,079	5,726	1,307	4,689
Australia	4,321	1,186	4,108	4,791	1,164	4,069
Bahrain	—	—	—	4,199	1,210	4,230
Botswana	—	—	—	5,150	1,197	4,208
Bulgaria	—	—	—	4,117	1,141	4,019
Chinese Taipei	4,661	1,192	4,131	5,379	1,155	4,046
Cyprus	—	—	—	4,002	1,255	4,399
Egypt	—	—	—	7,095	1,871	6,582
England	3,585	1,208	4,316	2,830	1,159	4,025
Ghana	—	—	—	5,100	1,498	5,294
Hong Kong SAR	4,608	1,072	3,791	4,972	986	3,470
Hungary	3,319	1,155	4,048	3,302	1,183	4,111
Indonesia	—	—	—	5,762	967	3,374
Iran, Islamic Rep. of	4,352	1,087	3,833	4,942	1,115	3,981
Israel	—	—	—	4,318	926	3,294
Italy	4,282	1,277	4,470	4,278	1,242	4,408
Japan	4,535	1,274	4,487	4,856	1,221	4,312
Jordan	—	—	—	4,489	1,492	5,251
Korea, Rep. of	—	—	—	5,309	1,208	4,240
Latvia	2,451	1,101	3,908	—	—	—
Lebanon	—	—	—	3,814	1,073	3,786
Lithuania	4,422	1,134	3,980	4,964	1,141	3,991
Malaysia	—	—	—	5,314	1,285	4,466
Morocco	4,264	1,090	3,894	—	—	—
Netherlands	2,937	962	3,349	—	—	—
New Zealand	4,254	1,405	4,940	—	—	—
Norway	4,342	1,165	4,108	4,133	1,317	4,627
Palestinian Nat'l Auth.	—	—	—	5,357	1,253	4,378
Romania	—	—	—	4,104	1,201	4,198
Russian Federation	3,963	1,277	4,464	4,667	1,277	4,472
Scotland	3,936	1,123	3,929	3,516	1,156	4,070
Serbia	—	—	—	4,296	1,153	4,045
Singapore	6,668	1,440	5,041	6,018	1,329	4,599
Slovenia	3,126	1,244	4,351	3,578	1,150	4,043
Sweden	—	—	—	4,256	1,473	5,215
Tunisia	4,334	1,160	4,081	4,931	1,175	4,080
United States	9,829	2,261	7,896	8,912	2,097	7,377
Total	**93,863**	**25,952**	**91,204**	**158,477**	**41,377**	**145,349**

Exhibit 11.6 Sample Sizes for Scaling the Content and Cognitive Domains for All Countries Participating in TIMSS 2007

Country	Fourth Grade		Eighth Grade	
	Item Calibration	Proficiency Estimation	Item Calibration	Proficiency Estimation
Algeria	4,223	4,223	5,447	5,447
Armenia	4,079	4,079	4,689	4,689
Australia	4,108	4,108	4,069	4,069
Austria	4,859	4,859	—	—
Bahrain	—	—	4,230	4,230
Bosnia and Herzegovina	—	—	4,220	4,220
Botswana	—	—	4,208	4,208
Bulgaria	—	—	4,019	4,019
Chinese Taipei	4,131	4,131	4,046	4,046
Colombia	4,801	4,801	4,873	4,873
Cyprus	—	—	4,399	4,399
Czech Republic	4,235	4,235	4,845	4,845
Denmark	3,519	3,519	—	—
Egypt	—	—	6,582	6,582
El Salvador	4,166	4,166	4,063	4,063
England	4,316	4,316	4,025	4,025
Georgia	4,108	4,108	4,178	4,178
Germany	5,200	5,200	—	—
Ghana	—	—	5,294	5,294
Hong Kong SAR	3,791	3,791	3,470	3,470
Hungary	4,048	4,048	4,111	4,111
Indonesia	—	—	4,203	4,203
Iran, Islamic Rep. of	3,833	3,833	3,981	3,981
Israel	—	—	3,294	3,294
Italy	4,470	4,470	4,408	4,408
Japan	4,487	4,487	4,312	4,312
Jordan	—	—	5,251	5,251
Korea, Rep. of	—	—	4,240	4,240
Kazakhstan	3,990	3,990	—	—
Kuwait	3,803	3,803	4,091	4,091
Latvia	3,908	3,908	—	—
Lebanon	—	—	3,786	3,786
Lithuania	3,980	3,980	3,991	3,991
Malaysia	—	—	4,466	4,466
Malta	—	—	4,670	4,670
Mongolia	—	4,523	—	4,499
Morocco	3,894	3,894	3,060	3,060
Netherlands	3,349	3,349	—	—
New Zealand	4,940	4,940	—	—
Norway	4,108	4,108	4,627	4,627
Oman	—	—	4,752	4,752
Palestinian Nat'l Auth.	—	—	4,378	4,378
Qatar	7,019	7,019	7,184	7,184
Romania	—	—	4,198	4,198
Russian Federation	4,464	4,464	4,472	4,472
Saudi Arabia	—	—	4,243	4,243
Scotland	3,929	3,929	4,070	4,070
Serbia	—	—	4,045	4,045
Singapore	5,041	5,041	4,599	4,599
Slovak Republic	4,963	4,963	—	—
Slovenia	4,351	4,351	4,043	4,043
Sweden	4,676	4,676	5,215	5,215
Syrian Arab Republic	—	—	4,650	4,650
Thailand	—	—	5,412	5,412
Tunisia	4,134	4,134	4,080	4,080
Turkey	—	—	4,498	4,498
Ukraine	4,292	4,292	4,424	4,424
United States	7,896	7,896	7,377	7,377
Yemen	5,811	5,811	—	—
Benchmarking Participants				
Alberta, Canada	—	4,037	—	—
Basque Country, Spain	—	—	—	2,296
British Columbia, Canada	—	4,153	—	4,256
Dubai, UAE	—	3,064	—	3,195
Massachusetts, US	—	1,747	—	1,897
Minnesota, US	—	1,846	—	1,777
Ontario, Canada	—	3,496	—	3,448
Quebec, Canada	—	3,885	—	3,956
Total	**160,922**	**187,673**	**220,788**	**246,112**

Exhibit 11.7 TIMSS 2007 Items by Content and Cognitive Domains at the Fourth Grade

TIMSS 2007 Scales in the Content and Cognitive Domains at the Fourth Grade			Items in TIMSS 2007 Assessment Booklets: Number	Items in TIMSS 2007 Assessment Booklets: Points
Mathematics	Overall		177	188
	Content Domains	Number	91	96
		Geometric Shapes and Measures	60	64
		Data Display	26	28
	Cognitive Domains	Knowing	68	71
		Applying	70	74
		Reasoning	39	43
Science	Overall		170	189
	Content Domains	Life Science	71	81
		Physical Science	64	66
		Earth Science	35	42
	Cognitive Domains	Knowing	74	84
		Applying	63	68
		Reasoning	33	37

Exhibit 11.8 TIMSS 2007 Items by Content and Cognitive Domains at the Eighth Grade

TIMSS 2007 Scales in the Content and Cognitive Domains at the Eighth Grade			Items in TIMSS 2007 Assessment Booklets: Number	Items in TIMSS 2007 Assessment Booklets: Points
Mathematics	Overall		214	236
	Content Domains	Number	63	72
		Algebra	64	69
		Geometry	47	49
		Data and chance	40	46
	Cognitive Domains	Knowing	81	83
		Applying	88	97
		Reasoning	45	56
Science	Overall		210	231
	Content Domains	Biology	75	86
		Chemistry	41	45
		Physics	54	57
		Earth Science	40	43
	Cognitive Domains	Knowing	83	87
		Applying	84	95
		Reasoning	43	49

11.3.3 Omitted and Not-Reached Responses

Apart from missing data on items that by design were not administered to a student, missing data could also occur because a student did not answer an item—whether because the student did not know the answer, omitted it by mistake, or did not have time to attempt the item. An item was considered not reached when—within part 1 or part 2 of the booklet—the item itself and the item immediately preceding it were not answered, and there were no other items completed in the remainder of that part of the booklet.

In TIMSS 2007, as in previous TIMSS assessments, not-reached items were treated differently in estimating item parameters and in generating student proficiency scores. In estimating the values of the item parameters, items in the TIMSS 2007 assessment booklets that were considered not to have been reached by students were treated as if they had not been administered. This approach was considered optimal for parameter estimation. Because of the position effect described earlier, items located in positions 3 and 6 of the test booklets in the TIMSS 2003 assessment data and TIMSS 2007 bridging data that were considered not to have been reached by the students were treated as incorrect. However, not-reached items were always considered as incorrect responses when student proficiency scores were generated.

11.3.4 Evaluating Fit of IRT Models to the TIMSS 2007 Data

After the item calibrations were completed, checks were performed to verify that the item parameters obtained from Parscale adequately reproduced the observed distribution of student responses across the proficiency continuum. The fit of the IRT models to the TIMSS 2007 data was examined by comparing the item response function curves generated using the item parameters estimated from the data with the empirical item response functions calculated from the posterior distributions of the θ's for each student that responded to the item. When the empirical results fall near the fitted curves for any given item, the IRT model fits the data well and leads to more accurate and reliable measurement of the underlying proficiency scale. Graphical plots of these response function curves are called item characteristic curves (ICC).

Exhibit 11.9 shows an ICC plot of the empirical and fitted item response functions for a dichotomous item. In the plot, the horizontal axis represents the proficiency scale, and the vertical axis represents the probability of a correct response. The fitted curve based on the estimated item parameters is shown as a solid line. Empirical results are represented by triangles. The empirical results were obtained by first dividing the proficiency scale into intervals of equal size and then counting the number of students responding to the item whose EAP scores from Parscale fell in each interval. Then the proportion of students in each interval that responded correctly to the item was calculated. In the exhibit, the center of each triangle represents this empirical proportion of correct responses. The size of each triangle is proportional to the number of students contributing to the estimation of its empirical proportion correct.

Exhibit 11.9 TIMSS 2007 Mathematics Assessment Example Item Response Function for a Dichotomous Item

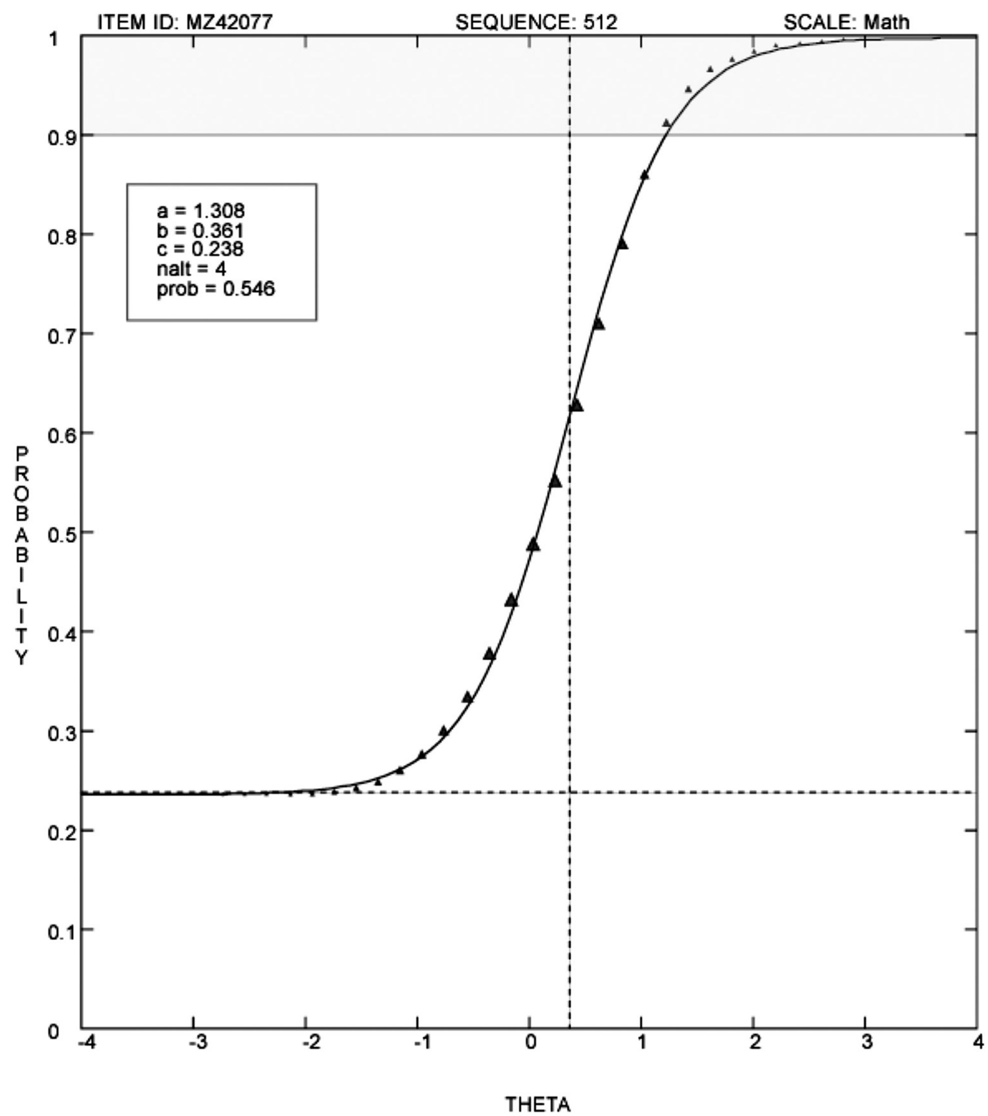


Exhibit 11.10 TIMSS 2007 Mathematics Assessment Example Item Response Function for a Polytomous Item

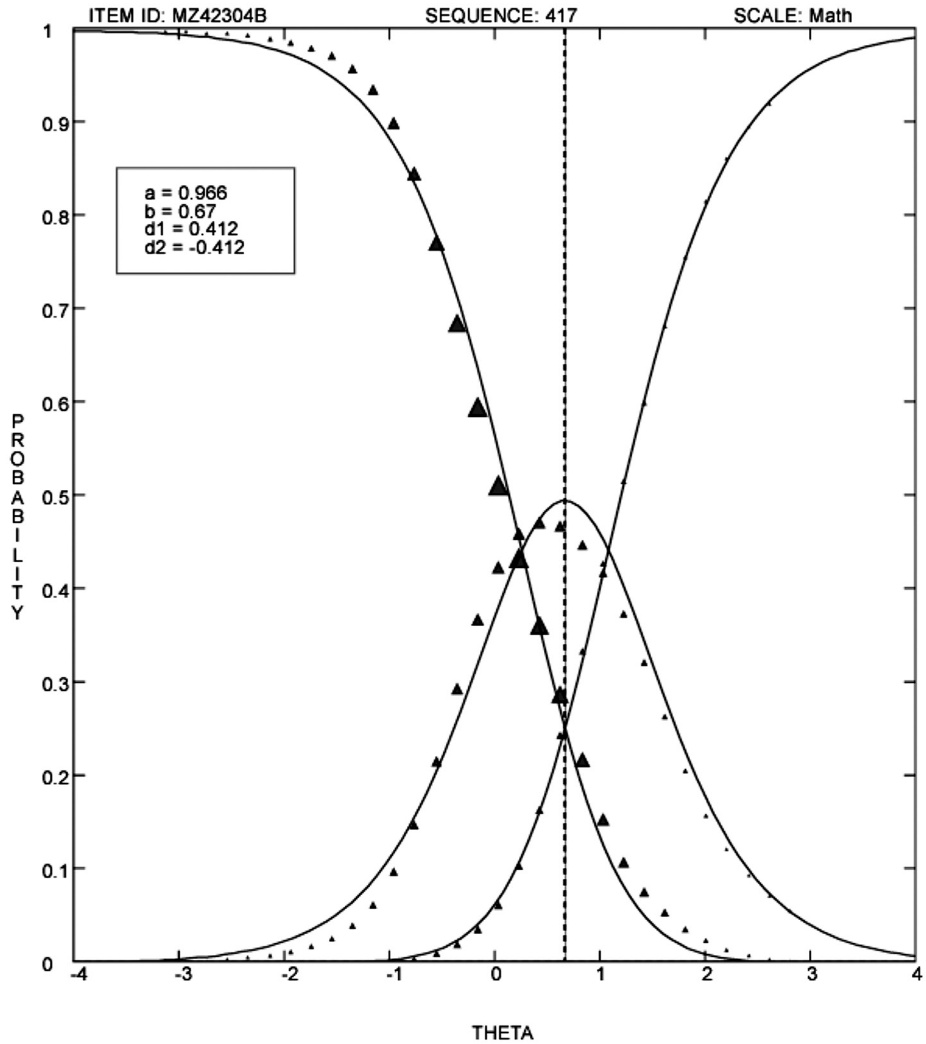


Exhibit 11.10 contains an ICC plot of the empirical and fitted item response functions for a polytomous item. As for the dichotomous item plot, the horizontal axis represents the proficiency scale, but the vertical axis represents the probability of having a response in a given response category. The fitted curves based on the estimated item parameters are shown as solid lines. Empirical results are represented by triangles. The interpretation of the triangles is the same as in Exhibit 11.9. The curve starting at the top left of the chart plots the probability of a score of zero on the item, which decreases as θ increases. The bell-shaped curve shows the probability of a score of one point—starting low for low-ability students, reaching a maximum for

medium-ability students, and decreasing for high-ability students. The curve ending at the top right corner of the chart shows the probability of a score of two points—full credit, starting low for low-ability students and increasing as θ increases.

11.3.5 Variables for Conditioning the TIMSS 2007 Data

Because there were so many background variables that could be used in conditioning, TIMSS followed the practice established by NAEP and followed by other large-scale studies of using principal components analysis to reduce the number of variables while explaining most of their common variance. Principal components for the TIMSS 2007 background data were constructed as follows:

- For categorical variables (questions with a small number of fixed response options), a "dummy coded" variable was created for each response option, with a value of one if the option was chosen and zero otherwise. If a student omitted or was not administered a particular question, all dummy coded variables associated with that question were assigned the value zero.
- Background variables with numerous response options (such as year of birth or number of people who live in the home) were recoded using criterion scaling.[11] This was done by replacing each response option with an interim achievement score. For the overall mathematics and science scales, the interim achievement scores were the average across the interim mathematics and science scores produced from the item calibration. For the content domain scales, the interim achievement scores from the calibration in each subject were averaged to form a composite mathematics and a composite science score, and the average of these composite scores was used as the interim achievement score.
- Separately for each TIMSS country, all the dummy-coded and criterion-scaled variables were included in a principal components analysis. Those principal components accounting for 90 percent of the variance of the background variables were retained for use as conditioning variables. Because the principal components analysis was performed separately for each country, different numbers of principal components were required to account for 90% of the common variance in each country's background variables.[12]

11 The process of generating criterion-scaled variables is described in Beaton (1969).

12 The criterion was reduced to 80% when applied to the TIMSS 2007 bridging data because of the smaller student sample sizes.

In addition to the principal components, student gender (dummy coded), the language of the test (dummy coded), an indicator of the classroom in the school to which the student belonged (criterion scaled), and an optional country-specific variable (dummy coded) were included as primary conditioning variables, thereby accounting for most of the variance between students and preserving the between- and within-classrooms variance structure in the scaling model. Exhibit 11.11 and Exhibit 11.12 show the total number of variables that were used in the principal component analysis and the number of principal components selected within each country. Conditioning variables were needed for the TIMSS 2007 assessment data of all participants, as well as for the TIMSS 2007 bridging data and the TIMSS 2003 assessment data of all trend countries.

Exhibit 11.11 Number of Variables and Principal Components for Conditioning in TIMSS 2007 at the Fourth Grade

Country	TIMSS 2003 Assessment Booklets			TIMSS 2007 Bridge Booklets			TIMSS 2007 Assessment Booklets		
	Number of Primary Conditioning Variables	Total Number of Principal Components	Number of Principal Components Retained	Number of Primary Conditioning Variables	Total Number of Principal Components	Number of Principal Components Retained	Number of Primary Conditioning Variables	Total Number of Principal Components	Number of Principal Components Retained
Algeria	—	—	—	—	—	—	2	285	172
Armenia	2	291	178	2	287	114	2	287	172
Australia	2	301	166	2	293	110	2	293	163
Austria	—	—	—	—	—	—	2	293	168
Chinese Taipei	2	313	172	2	293	116	2	293	165
Colombia	—	—	—	—	—	—	2	285	168
Czech Republic	—	—	—	—	—	—	2	293	168
Denmark	—	—	—	—	—	—	2	285	159
El Salvador	—	—	—	—	—	—	2	293	173
England	2	295	165	2	291	115	2	291	165
Georgia	—	—	—	—	—	—	2	289	171
Germany	—	—	—	—	—	—	2	293	163
Hong Kong SAR	2	313	171	3	291	110	3	293	160
Hungary	2	307	172	2	291	115	2	291	166
Iran, Islamic Rep. of	2	305	172	2	293	115	2	293	170
Italy	2	311	173	2	236	110	2	237	152
Japan	2	313	175	2	293	116	2	293	165
Kazakhstan	—	—	—	—	—	—	3	291	158
Kuwait	—	—	—	—	—	—	2	285	171
Latvia	3	313	173	2	293	110	2	293	164
Lithuania	2	290	163	2	293	114	2	293	166
Moldova, Rep. of	—	—	—	—	—	—	3	291	145
Mongolia	—	—	—	—	—	—	3	277	165
Morocco	2	297	177	2	291	118	2	291	174
Netherlands	2	289	164	2	285	108	2	285	160
New Zealand	8	311	174	7	293	120	7	293	168
Norway	2	313	177	2	293	114	2	293	165
Qatar	—	—	—	—	—	—	3	291	176
Russian Federation	2	241	134	2	293	114	2	293	167
Scotland	2	295	168	2	291	115	2	291	166
Singapore	2	301	170	2	293	118	2	293	164
Slovak Republic	—	—	—	—	—	—	3	293	169
Slovenia	2	313	172	2	293	119	2	293	168
Sweden	—	—	—	—	—	—	2	293	166
Tunisia	2	311	184	2	293	123	2	293	176
Ukraine	—	—	—	—	—	—	3	291	169
United States	8	287	168	7	283	125	7	283	166
Yemen	—	—	—	—	—	—	2	285	180
Benchmarking Participants									
Alberta, Canada	—	—	—	—	—	—	3	287	162
British Columbia, Canada	—	—	—	—	—	—	3	287	162
Dubai, UAE	—	—	—	—	—	—	3	291	163
Massachusetts, US	—	—	—	—	—	—	2	281	155
Minnesota, US	—	—	—	—	—	—	2	283	156
Ontario, Canada	3	291	160	3	287	103	3	287	159
Quebec, Canada	3	291	165	3	287	108	3	287	162

Exhibit 11.12 Number of Variables and Principal Components for Conditioning in TIMSS 2007 at the Eighth Grade

Country	TIMSS 2003 Assessment Booklets			TIMSS 2007 Bridge Booklets			TIMSS 2007 Assessment Booklets		
	Number of Primary Conditioning Variables	Total Number of Principal Components	Number of Principal Components Retained	Number of Primary Conditioning Variables	Total Number of Principal Components	Number of Principal Components Retained	Number of Primary Conditioning Variables	Total Number of Principal Components	Number of Principal Components Retained
Algeria	—	—	—	—	—	—	3	811	391
Armenia	2	891	430	3	892	233	3	892	445
Australia	2	417	225	3	399	139	3	399	217
Bahrain	3	429	242	4	396	152	4	396	226
Bosnia and Herzegovina	—	—	—	—	—	—	5	895	453
Botswana	2	424	248	3	399	162	3	399	237
Bulgaria	2	913	409	3	899	179	3	899	375
Chinese Taipei	2	432	231	3	396	139	3	396	208
Colombia	—	—	—	—	—	—	3	388	225
Cyprus	2	897	420	3	897	218	3	898	407
Czech Republic	—	—	—	—	—	—	3	900	460
Egypt	4	418	249	4	396	167	4	396	237
El Salvador	—	—	—	—	—	—	3	399	230
England	2	410	216	3	375	135	3	381	207
Georgia	—	—	—	—	—	—	3	895	416
Ghana	2	410	245	3	399	163	3	399	236
Hong Kong SAR	2	432	233	3	399	135	3	399	211
Hungary	2	907	437	3	898	241	3	899	445
Indonesia	2	633	336	3	899	231	3	901	421
Iran, Islamic Rep. of	2	424	243	3	399	151	3	399	228
Israel	3	432	241	4	396	145	4	396	222
Italy	2	430	234	3	325	137	3	326	198
Japan	2	425	231	3	394	139	3	395	212
Jordan	2	432	247	3	396	154	3	396	229
Korea, Rep. of	2	432	234	3	377	141	3	396	214
Kuwait	—	—	—	—	—	—	3	386	221
Lebanon	2	745	376	4	734	194	4	734	361
Lithuania	3	811	392	3	900	233	3	900	442
Malaysia	2	412	231	3	396	150	3	397	220
Malta	—	—	—	—	—	—	3	897	409
Moldova, Rep. of	—	—	—	—	—	—	4	867	319
Mongolia	—	—	—	—	—	—	4	897	425
Morocco	—	—	—	—	—	—	3	891	403
Norway	2	429	236	3	396	146	3	396	217
Oman	—	—	—	—	—	—	4	396	231
Palestinian Nat'l Auth.	3	432	252	3	392	157	3	392	231
Qatar	—	—	—	—	—	—	4	394	227
Romania	3	919	453	4	899	231	4	901	438
Russian Federation	2	915	446	3	898	225	3	897	431
Saudi Arabia	—	—	—	—	—	—	3	387	226
Scotland	2	410	224	3	381	141	3	381	210
Serbia	2	919	444	3	837	226	3	894	435
Singapore	2	420	233	3	398	145	3	398	214
Slovenia	2	766	372	3	786	223	3	786	395
Sweden	2	916	398	3	901	218	3	901	396
Syrian Arab Republic	—	—	—	—	—	—	3	901	464
Thailand	—	—	—	—	—	—	3	399	224
Tunisia	2	410	242	3	399	159	3	399	234
Turkey	—	—	—	—	—	—	3	396	227
Ukraine	—	—	—	—	—	—	4	901	439
United States	8	404	229	8	389	160	8	389	222
Benchmarking Participants									
Basque Country, Spain	3	429	230	4	377	122	4	377	202
British Columbia, Canada	—	—	—	—	—	—	4	388	215
Dubai, UAE	—	—	—	—	—	—	3	397	217
Massachusetts, US	—	—	—	—	—	—	3	389	209
Minnesota, US	—	—	—	—	—	—	3	389	204
Ontario, Canada	3	410	219	4	388	128	4	388	209
Quebec, Canada	3	410	223	4	388	136	4	388	212

11.3.6 Generating IRT Proficiency Scores for the TIMSS 2007 Data

Educational Testing Service's MGROUP program (Sheehan, 1985)[13] was used to generate the IRT proficiency scores. This program takes as input the students' responses to the items they were given, the item parameters estimated at the calibration stage, and the conditioning variables, and generates as output the plausible values that represent student proficiency. A useful feature of MGROUP is its ability to perform multi-dimensional scaling using the responses to all items across the scales and the correlations among the scales to improve the reliability of each individual scale. Because the redesigned TIMSS 2007 assessment booklets were balanced in terms of their mathematics and science content, TIMSS was able to capitalize on this feature for the first time in 2007. In this way, the overall mathematics and science scales were established simultaneously using a two-dimensional MGROUP run. This feature of MGROUP also was used to generate multi-dimensional scales across the mathematics content domains, the mathematics cognitive domains, the science content domains, and the science cognitive domains.

In addition to generating plausible values for the TIMSS 2007 assessment data, the parameters estimated at the calibration stage also were used to generate plausible values on the overall mathematics and science scales using the fourth-grade 2003 assessment data and 2007 bridging data for the 21 trend countries that also participated in the TIMSS 2003 fourth-grade assessment, and the eighth-grade 2003 assessment data and 2007 bridging data for the 33 countries that also participated in the 2003 eighth-grade assessment. These additional plausible values were then used to establish the two successive linear transformations necessary to place the TIMSS 2007 assessment on the TIMSS trend scale.

In all, a total of 209 (86 at the fourth grade and 123 at the eighth grade) two-dimensional MGROUP runs were required for the overall mathematics and science scales, and 404 (176 at the fourth grade and 228 at the eighth grade) multidimensional MGROUP runs for the content and cognitive scales. Exhibit 11.13 shows the sizes of the student samples—2003 assessment data, 2007 bridging data, and 2007 assessment data—for which proficiency scores using the 2007 item parameters were generated on the overall mathematics and science scales. At the fourth grade, scores on the 2003 assessment data were generated for 103,865 students, scores on the 2007 bridging data were generated for 28,098 students, and scores on the

13 The MGROUP program was provided by ETS under contract to the TIMSS & PIRLS International Study Center at Boston College. It is now commercially available as DESI.

2007 assessment data for 187,673 students. At the eighth grade, scores on the 2003 assessment data were generated for 169,619 students, scores on the 2007 bridging data for 44,350 students, and scores on the 2007 assessment data for 246,112 students. Exhibit 11.6, presented previously, shows that a total of 187,673 students received proficiency scores on the 2007 assessment data in the content and cognitive domains at the fourth grade and 246,112 students at the eighth grade.

Exhibit 11.13 Sample Sizes for TIMSS 2007 Proficiency Estimation of Overall Mathematics and Science

Country	Fourth Grade			Eighth Grade		
	TIMSS 2003 Assessment Booklets	TIMSS 2007 Bridge Booklets	TIMSS 2007 Assessment Booklets	TIMSS 2003 Assessment Booklets	TIMSS 2007 Bridge Booklets	TIMSS 2007 Assessment Booklets
Algeria	—	—	4,223	—	—	5,447
Armenia	5,674	1,139	4,079	5,726	1,307	4,689
Australia	4,321	1,186	4,108	4,791	1,164	4,069
Austria	—	—	4,859	—	—	—
Bahrain	—	—	—	4,199	1,210	4,230
Bosnia and Herzegovina	—	—	—	—	—	4,220
Botswana	—	—	—	5,150	1,197	4,208
Bulgaria	—	—	—	4,117	1,141	4,019
Chinese Taipei	4,661	1,192	4,131	5,379	1,155	4,046
Colombia	—	—	4,801	—	—	4,873
Cyprus	—	—	—	4,002	1,255	4,399
Czech Republic	—	—	4,235	—	—	4,845
Denmark	—	—	3,519	—	—	—
Egypt	—	—	—	7,095	1,871	6,582
El Salvador	—	—	4,166	—	—	4,063
England	3,585	1,208	4,316	2,830	1,159	4,025
Georgia	—	—	4,108	—	—	4,178
Germany	—	—	5,200	—	—	—
Ghana	—	—	—	5,100	1,498	5,294
Hong Kong SAR	4,608	1,072	3,791	4,972	986	3,470
Hungary	3,319	1,155	4,048	3,302	1,183	4,111
Indonesia	—	—	—	5,762	1,202	4,203
Iran, Islamic Rep. of	4,352	1,087	3,833	4,942	1,115	3,981
Israel	—	—	—	4,318	926	3,294
Italy	4,282	1,277	4,470	4,278	1,242	4,408
Japan	4,535	1,274	4,487	4,856	1,221	4,312
Jordan	—	—	—	4,489	1,492	5,251
Kazakhstan	—	—	3,990	—	—	—
Korea, Rep. of	—	—	—	5,309	1,208	4,240
Kuwait	—	—	3,803	—	—	4,091
Latvia	3,687	1,101	3,908	—	—	—
Lebanon	—	—	—	3,814	1,073	3,786
Lithuania	4,422	1,134	3,980	4,964	1,141	3,991
Malaysia	—	—	—	5,314	1,285	4,466
Malta	—	—	—	—	—	4,670
Mongolia	—	—	4,523	—	—	4,499
Morocco	4,264	1,090	3,894	—	—	3,060
Netherlands	2,937	962	3,349	—	—	—
New Zealand	4,308	1,405	4,940	—	—	—
Norway	4,342	1,165	4,108	4,133	1,317	4,627
Oman	—	—	—	—	—	4,752
Palestinian Nat'l Auth.	—	—	—	5,357	1,253	4,378
Qatar	—	—	7,019	—	—	7,184
Romania	—	—	—	4,104	1,201	4,198
Russian Federation	3,963	1,277	4,464	4,667	1,277	4,472
Saudi Arabia	—	—	—	—	—	4,243
Scotland	3,936	1,123	3,929	3,516	1,156	4,070
Serbia	—	—	—	4,296	1,153	4,045
Singapore	6,668	1,440	5,041	6,018	1,329	4,599
Slovak Republic	—	—	4,963	—	—	—
Slovenia	3,126	1,244	4,351	3,578	1,150	4,043
Sweden	—	—	4,676	4,256	1,473	5,215
Syrian Arab Republic	—	—	—	—	—	4,650
Thailand	—	—	—	—	—	5,412
Tunisia	4,334	1,174	4,134	4,931	1,175	4,080
Turkey	—	—	—	—	—	4,498
Ukraine	—	—	4,292	—	—	4,424
United States	9,829	2,261	7,896	8,912	2,097	7,377
Yemen	—	—	5,811	—	—	—
Benchmarking Participants						
Alberta, Canada	—	—	4,037	—	—	—
Basque Country, Spain	—	—	—	2,514	645	2,296
British Columbia, Canada	—	—	4,153	—	—	4,256
Dubai, UAE	—	—	3,064	—	—	3,195
Massachusetts, US	—	—	1,747	—	—	1,897
Minnesota, US	—	—	1,846	—	—	1,777
Ontario, Canada	4,362	1,021	3,496	4,217	989	3,448
Quebec, Canada	4,350	1,111	3,885	4,411	1,104	3,956
Total	**103,865**	**28,098**	**187,673**	**169,619**	**44,350**	**246,112**

11.3.7 Transforming the Mathematics and Science Scores to Measure Trends

To provide results for TIMSS 2007 that would be comparable to results from previous TIMSS assessments, the 2007 proficiency scores (plausible values) for overall mathematics and science had to be transformed to the metric used in 1995, 1999, and 2003. This was accomplished through two successive linear transformations as part of the concurrent calibration approach.

First, the means and standard deviations of the mathematics and science 2003 scores produced in 2007—the plausible values from the TIMSS 2003 assessment data based on the 2007 concurrent item calibrations—were made to match the means and standard deviations of the scores reported in the TIMSS 2003 assessment—the plausible values produced in 2003 using the 2003 item calibrations—by applying the appropriate linear transformations. These linear transformations were given by:

$$PV_{k,i}^{*} = A_{k,i} + B_{k,i} \cdot PV_{k,i} \tag{13}$$

where

$PV_{k,i}$ was the plausible value i of scale k prior to transformation;

$PV_{k,i}^{*}$ was the plausible value i of scale k after transformation;

and $A_{k,i}$ and $B_{k,i}$ were the linear transformation constants.

The linear transformation constants were obtained by first computing the international means and standard deviations of the proficiency scores for the overall mathematics and science scales using the plausible values produced in 2003 based on the 2003 item calibrations for the trend countries. Next, the same calculations were done using the plausible values from the TIMSS 2003 assessment data based on the 2007 item calibrations for the same set of countries. The linear transformation constants were defined as:

$$\begin{aligned} B_{k,i} &= \sigma_{k,i} / \sigma_{k,i}^{*} \\ A_{k,i} &= \mu_{k,i} - B_{k,i}\, \mu_{k,i}^{*} \end{aligned} \tag{14}$$

where

$\mu_{k,i}$ was the international mean of scale k based on plausible value i released in 2003;

$\mu_{k,i}^*$ was the international mean of scale k based on plausible value i from the TIMSS 2003 assessment data based on the 2007 concurrent item calibrations;

$\sigma_{k,i}$ was the international standard deviation of scale k based on plausible value i released in 2003;

$\sigma_{k,i}^*$ was the international standard deviation of scale k based on plausible value i from the TIMSS 2003 assessment data based on the 2007 concurrent item calibrations.

Exhibit 11.14 shows the linear transformation constants that were computed in this first step. Once the linear transformation constants had been established, all of the mathematics and science plausible values generated on the TIMSS 2007 bridging data were transformed by applying the linear transformations.

Exhibit 11.14 Linear Transformation Constants Applied to the TIMSS 2007 Bridge Scores

Scale		Plausible Value	TIMSS 2003 Scores Using 2003 Item Calibrations		TIMSS 2003 Scores Using 2007 Item Calibrations		$A_{k,i}$	$B_{k,i}$
			Mean	Standard Deviation	Mean	Standard Deviation		
Fourth Grade	Mathematics	PV1	498.12622	104.81269	–0.06579	0.99477	505.05840	105.36413
		PV2	498.31619	103.90056	–0.06546	0.99426	505.15723	104.50041
		PV3	498.14926	104.01856	–0.06582	0.99533	505.02747	104.50692
		PV4	498.51640	104.36297	–0.06712	0.99476	505.55795	104.91235
		PV5	498.33038	103.88447	–0.06510	0.99498	505.12714	104.40824
	Science	PV1	495.05010	109.62454	–0.05554	0.98941	501.20328	110.79794
		PV2	494.22197	109.40731	–0.05360	0.98730	500.16177	110.81421
		PV3	494.23251	110.17620	–0.05360	0.98717	500.21478	111.60831
		PV4	494.34316	109.52188	–0.05348	0.98990	500.26064	110.63879
		PV5	495.13090	109.68009	–0.05185	0.98629	500.89740	111.20455
Eighth Grade	Mathematics	PV1	476.14829	105.92163	0.00510	0.98871	475.60194	107.13090
		PV2	476.39770	107.36384	0.00539	0.99167	475.81398	108.26543
		PV3	476.33494	107.48064	0.00480	0.99012	475.81336	108.55323
		PV4	475.96981	107.31753	0.00481	0.98907	475.44768	108.50350
		PV5	476.42089	107.00376	0.00551	0.99005	475.82554	108.07918
	Science	PV1	481.84829	105.24281	0.00707	0.98023	481.08890	107.36518
		PV2	481.99746	105.50264	0.00785	0.98128	481.15317	107.51570
		PV3	482.40244	104.91097	0.00804	0.97856	481.54006	107.20928
		PV4	482.08413	105.81120	0.00856	0.97901	481.15912	108.08008
		PV5	482.51302	104.94370	0.00939	0.97924	481.50676	107.16884

Next, the means and standard deviations of the mathematics and science proficiency scores on the TIMSS 2007 assessment data were made to match the means and standard deviations of the now-transformed scores on the TIMSS 2007 bridging data by applying appropriate linear transformations. These linear transformations were derived using the same equations given above, with the linear transformation constants obtained by first computing the international means and standard deviations of the now-transformed scores on the TIMSS 2007 bridging data for the overall mathematics and science scales across the trend countries, and then the same calculations using the plausible values generated on the TIMSS 2007 assessment data across the trend countries.

Exhibit 11.15 shows the linear transformation constants that were computed in this second step. Once these linear transformation constants had been established, all of the 2007 mathematics and science proficiency scores—the plausible values generated on the TIMSS 2007 assessment data—for all participating countries and benchmarking participants were transformed by applying the linear transformations. This provided mathematics and science student achievement scores for the TIMSS 2007 assessment that were directly comparable to the scores from the 1995, 1999 (only at the eighth grade), and 2003 assessments.

Exhibit 11.15 Linear Transformation Constants Applied to the TIMSS 2007 Proficiency Scores

Scale		Plausible Value	Transformed TIMSS 2007 Bridge Scores		TIMSS 2007 Proficiency Scores		$A_{k,i}$	$B_{k,i}$
			Mean	Standard Deviation	Mean	Standard Deviation		
Fourth Grade	Mathematics	PV1	506.17533	108.02573	–0.01243	1.04972	507.45462	102.90944
		PV2	506.23088	107.63611	–0.01115	1.04540	507.37904	102.96198
		PV3	506.62376	107.29968	–0.01037	1.04678	507.68705	102.50484
		PV4	506.15659	108.10783	–0.01021	1.04853	507.20928	103.10455
		PV5	506.19823	107.37574	–0.01337	1.04727	507.56872	102.52942
	Science	PV1	504.92173	112.88966	0.01118	1.01466	503.67776	111.25838
		PV2	503.55827	112.77187	0.01470	1.00669	501.91179	112.02242
		PV3	503.42470	113.64933	0.01197	1.00968	502.07753	112.55966
		PV4	503.36473	112.95516	0.01129	1.01015	502.10236	111.82060
		PV5	504.79464	112.70603	0.01263	1.01355	503.38990	111.19905
Eighth Grade	Mathematics	PV1	474.29429	109.44201	–0.01422	1.01544	475.82719	107.77822
		PV2	474.61572	110.62798	–0.01264	1.01579	475.99222	108.90822
		PV3	474.52757	111.06244	–0.01359	1.01350	476.01716	109.58307
		PV4	474.22239	110.91719	–0.01266	1.01656	475.60358	109.11081
		PV5	475.17257	110.29007	–0.01343	1.01490	476.63216	108.67084
	Science	PV1	481.92084	105.72417	0.00330	0.97876	481.56437	108.01886
		PV2	482.06417	105.48861	0.00376	0.97833	481.65864	107.82554
		PV3	482.56974	104.81989	0.00504	0.97830	482.03002	107.14473
		PV4	481.56147	106.10752	0.00105	0.98092	481.44803	108.17180
		PV5	482.65436	105.06218	0.00228	0.97759	482.40927	107.47102

11.3.8 Setting the Metric for the Mathematics and Science Content and Cognitive Domain Scales

As described earlier, the IRT scales for the mathematics and science content and cognitive domains had no provision for measuring trends, so there was no need to establish links to previous assessment metrics. Instead, the plausible values for each content and cognitive domain scale were transformed to the same metric as its respective overall subject scale in 2007. For example, in eighth-grade mathematics, the mean and standard deviation for the number, algebra, geometry, and data and chance scales were set to have the same mean and standard deviation as the 2007 eighth-grade mathematics scale. Setting linear transformation constants was done in the same manner as described in the previous section, with the exception that the means and standard deviations of the overall subject scales were averaged across the five plausible values. Exhibits 11.16 through 11.19 show the transformations that were applied to all the content and cognitive domains. Taking fourth-grade mathematics as an example, the plausible values of all fourth-grade mathematics content and cognitive domains were transformed to have a mean of 472.9372 and a standard deviation of 123.6880, the international mean and standard deviation for overall mathematics across the 36 fourth-grade countries.

Exhibit 11.16 Linear Transformation Constants for the TIMSS 2007 Fourth-Grade Mathematics Content and Cognitive Domains

Scale		Plausible Values	Mean	Standard Deviation		
Mathematics		PV1	472.7558	123.9167		
		PV2	472.8534	123.9992		
		PV3	473.3264	123.3602		
		PV4	472.7947	123.8875		
		PV5	472.9556	123.2766		
		Overall	**472.9372**	**123.6880**	$A_{k,i}$	$B_{k,i}$
Content Domains	Number	PV1	–0.1044	1.1129	484.5396	111.1409
		PV2	–0.1036	1.1094	484.4879	111.4913
		PV3	–0.1052	1.1138	484.6169	111.0519
		PV4	–0.1049	1.1126	484.6034	111.1682
		PV5	–0.1059	1.1145	484.6843	110.9775
	Geometric Shapes and Mesures	PV1	–0.1654	1.1350	490.9571	108.9716
		PV2	–0.1661	1.1340	491.0578	109.0680
		PV3	–0.1654	1.1366	490.9363	108.8250
		PV4	–0.1635	1.1351	490.7560	108.9620
		PV5	–0.1663	1.1366	491.0385	108.8265
	Data Display	PV1	–0.2348	1.2274	496.5946	100.7747
		PV2	–0.2298	1.2283	496.0757	100.7013
		PV3	–0.2376	1.2257	496.9106	100.9138
		PV4	–0.2318	1.2256	496.3298	100.9167
		PV5	–0.2263	1.2204	495.8715	101.3540
Cognitive Domains	Knowing	PV1	–0.1233	1.0819	487.0345	114.3231
		PV2	–0.1207	1.0824	486.7352	114.2708
		PV3	–0.1198	1.0801	486.6588	114.5102
		PV4	–0.1212	1.0777	486.8516	114.7689
		PV5	–0.1217	1.0788	486.8850	114.6544
	Applying	PV1	–0.1649	1.1292	491.0036	109.5366
		PV2	–0.1661	1.1313	491.0923	109.3315
		PV3	–0.1648	1.1281	491.0032	109.6410
		PV4	–0.1635	1.1301	490.8277	109.4483
		PV5	–0.1656	1.1278	491.0955	109.6730
	Reasoning	PV1	–0.1643	1.1908	490.0010	103.8737
		PV2	–0.1640	1.1930	489.9417	103.6768
		PV3	–0.1653	1.1929	490.0784	103.6869
		PV4	–0.1627	1.1931	489.8072	103.6689
		PV5	–0.1591	1.1895	489.4796	103.9819

Exhibit 11.17 Linear Transformation Constants for the TIMSS 2007 Fourth-Grade Science Content and Cognitive Domains

Scale		Plausible Values	Mean	Standard Deviation		
Science		PV1	476.8554	127.8734		
		PV2	475.2254	128.4317		
		PV3	475.0733	128.9199		
		PV4	475.1666	128.1879		
		PV5	476.6620	128.0548		
		Overall	**475.7965**	**128.2935**	$A_{k,i}$	$B_{k,i}$
Content Domains	Life Science	PV1	–0.0991	1.0267	488.1824	124.9529
		PV2	–0.0989	1.0222	488.2120	125.5117
		PV3	–0.1012	1.0243	488.4686	125.2512
		PV4	–0.0995	1.0261	488.2427	125.0362
		PV5	–0.1015	1.0258	488.4969	125.0695
	Physical Science	PV1	–0.1244	1.0591	490.8606	121.1338
		PV2	–0.1270	1.0588	491.1865	121.1670
		PV3	–0.1236	1.0580	490.7812	121.2581
		PV4	–0.1268	1.0616	491.1226	120.8508
		PV5	–0.1250	1.0602	490.9197	121.0130
	Earth Science	PV1	–0.1738	1.1588	495.0349	110.7096
		PV2	–0.1759	1.1559	495.3152	110.9871
		PV3	–0.1729	1.1604	494.9164	110.5598
		PV4	–0.1759	1.1589	495.2658	110.7030
		PV5	–0.1727	1.1595	494.9014	110.6414
Cognitive Domains	Knowing	PV1	–0.0979	1.0077	488.2655	127.3159
		PV2	–0.1015	1.0130	488.6458	126.6496
		PV3	–0.1000	1.0098	488.4998	127.0429
		PV4	–0.0996	1.0124	488.4181	126.7196
		PV5	–0.1000	1.0101	488.4992	127.0106
	Applying	PV1	–0.1053	1.0206	489.0330	125.7006
		PV2	–0.1064	1.0213	489.1652	125.6169
		PV3	–0.1074	1.0243	489.2459	125.2444
		PV4	–0.1070	1.0193	489.2690	125.8602
		PV5	–0.1051	1.0216	488.9907	125.5752
	Reasoning	PV1	–0.1044	1.1160	487.7931	114.9562
		PV2	–0.1061	1.1128	488.0338	115.2891
		PV3	–0.1028	1.1156	487.6219	114.9956
		PV4	–0.1075	1.1136	488.1796	115.2033
		PV5	–0.1054	1.1165	487.9106	114.9097

Exhibit 11.18 Linear Transformation Constants for the TIMSS 2007 Eighth-Grade Mathematics Content and Cognitive Domains

Scale		Plausible Values	Mean	Standard Deviation		
Mathematics		PV1	450.7160	111.5804		
		PV2	450.8086	112.6485		
		PV3	450.5763	113.2416		
		PV4	450.2712	113.1116		
		PV5	451.3883	112.4931		
		Overall	**450.7521**	**112.6151**	$A_{k,i}$	$B_{k,i}$
Content Domains	Number	PV1	–0.0300	1.0349	454.0165	108.8143
		PV2	–0.0335	1.0350	454.4005	108.8096
		PV3	–0.0323	1.0346	454.2695	108.8481
		PV4	–0.0309	1.0338	454.1148	108.9336
		PV5	–0.0344	1.0346	454.4955	108.8470
	Algebra	PV1	–0.0044	1.0900	451.2025	103.3148
		PV2	–0.0044	1.0906	451.2070	103.2605
		PV3	–0.0038	1.0905	451.1481	103.2675
		PV4	–0.0056	1.0910	451.3284	103.2175
		PV5	–0.0098	1.0935	451.7566	102.9858
	Geometry	PV1	–0.0828	1.0827	459.3668	104.0144
		PV2	–0.0802	1.0803	459.1119	104.2434
		PV3	–0.0824	1.0820	459.3264	104.0787
		PV4	–0.0814	1.0795	459.2466	104.3209
		PV5	–0.0798	1.0808	459.0657	104.1960
	Data and Chance	PV1	–0.0674	1.0645	457.8778	105.7897
		PV2	–0.0717	1.0606	458.3616	106.1821
		PV3	–0.0697	1.0633	458.1391	105.9131
		PV4	–0.0716	1.0597	458.3578	106.2665
		PV5	–0.0706	1.0603	458.2455	106.2111
Cognitive Domains	Knowing	PV1	–0.0671	1.0395	458.0263	108.3317
		PV2	–0.0717	1.0415	458.5033	108.1275
		PV3	–0.0670	1.0393	458.0155	108.3590
		PV4	–0.0656	1.0409	457.8484	108.1890
		PV5	–0.0672	1.0402	458.0238	108.2634
	Applying	PV1	–0.0495	1.0458	456.0790	107.6794
		PV2	–0.0516	1.0464	456.3011	107.6224
		PV3	–0.0517	1.0472	456.3152	107.5379
		PV4	–0.0508	1.0483	456.2053	107.4299
		PV5	–0.0519	1.0449	456.3465	107.7801
	Reasoning	PV1	–0.0441	1.0749	455.3733	104.7632
		PV2	–0.0414	1.0752	455.0850	104.7371
		PV3	–0.0474	1.0745	455.7219	104.8029
		PV4	–0.0463	1.0752	455.6066	104.7384
		PV5	–0.0469	1.0723	455.6766	105.0259

Exhibit 11.19 Linear Transformation Constants for the TIMSS 2007 Eighth-Grade Science Content and Cognitive Domains

Scale		Plausible Values	Mean	Standard Deviation		
Science		PV1	465.4845	106.0061		
		PV2	465.6370	105.7173		
		PV3	466.0839	105.1221		
		PV4	465.1039	106.3709		
		PV5	466.2519	105.4609		
		Overall	**465.7122**	**105.7354**	$A_{k,i}$	$B_{k,i}$
Content Domains	Biology	PV1	–0.0398	0.8496	470.6701	124.4517
		PV2	–0.0401	0.8504	470.7007	124.3389
		PV3	–0.0415	0.8476	470.8869	124.7403
		PV4	–0.0413	0.8480	470.8605	124.6949
		PV5	–0.0422	0.8514	470.9481	124.1906
	Chemistry	PV1	–0.0654	1.0273	472.4460	102.9242
		PV2	–0.0656	1.0270	472.4647	102.9571
		PV3	–0.0652	1.0240	472.4397	103.2609
		PV4	–0.0649	1.0288	472.3856	102.7784
		PV5	–0.0650	1.0311	472.3790	102.5499
	Physics	PV1	–0.0827	0.9906	474.5414	106.7348
		PV2	–0.0842	0.9905	474.7044	106.7466
		PV3	–0.0805	0.9882	474.3278	107.0004
		PV4	–0.0774	0.9886	473.9906	106.9579
		PV5	–0.0821	0.9865	474.5151	107.1846
	Earth Science	PV1	–0.0951	1.0407	475.3735	101.6041
		PV2	–0.0920	1.0419	475.0517	101.4861
		PV3	–0.0922	1.0393	475.0911	101.7377
		PV4	–0.0962	1.0372	475.5150	101.9418
		PV5	–0.0939	1.0436	475.2263	101.3174
Cognitive Domains	Knowing	PV1	–0.0454	0.8542	471.3322	123.7832
		PV2	–0.0443	0.8545	471.1986	123.7342
		PV3	–0.0428	0.8545	471.0059	123.7376
		PV4	–0.0435	0.8535	471.1056	123.8873
		PV5	–0.0448	0.8553	471.2555	123.6291
	Applying	PV1	–0.0596	0.8704	472.9576	121.4767
		PV2	–0.0606	0.8681	473.0985	121.7949
		PV3	–0.0600	0.8684	473.0205	121.7655
		PV4	–0.0594	0.8702	472.9248	121.5055
		PV5	–0.0585	0.8696	472.8282	121.5880
	Reasoning	PV1	–0.0815	1.0554	473.8798	100.1850
		PV2	–0.0838	1.0618	474.0618	99.5821
		PV3	–0.0822	1.0580	473.9259	99.9417
		PV4	–0.0794	1.0586	473.6424	99.8829
		PV5	–0.0801	1.0576	473.7238	99.9776

11.4 Capturing the Uncertainty in the TIMSS Student Achievement Scores

To obtain estimates of students' proficiency in mathematics and science that were both accurate and cost-effective, TIMSS 2007 made extensive use of probability sampling techniques to sample students from national eighth- and fourth-grade student populations, and applied matrix sampling methods to target individual students with a subset of the entire set of assessment materials. Statistics computed from these student samples were used to estimate population parameters. This approach made an efficient use of resources, in particular keeping student response burden to a minimum, but at a cost of some variance or uncertainty in the statistics. To quantify this uncertainty, each statistic in the TIMSS 2007 international reports (Martin et al., 2008; Mullis et al., 2008) is accompanied by an estimate of its standard error. These standard errors incorporate components reflecting the uncertainty due to generalizing from student samples to the entire eighth- or fourth-grade student populations (sampling variance), and to inferring students' performance on the entire assessment from their performance on the subset of items that they took (imputation variance).

11.4.1 Estimating Sampling Variance

The TIMSS 2007 sampling design applied a stratified multistage cluster-sampling technique to the problem of selecting efficient and accurate samples of students while working with schools and classes. This design capitalized on the structure of the student population (i.e., students grouped in classrooms within schools) to derive student samples that permitted efficient and economical data collection. Unfortunately, however, such a complex sampling design complicates the task of computing standard errors to quantify sampling variability.

When, as in TIMSS, the sampling design involves multistage cluster sampling, there are several options for estimating sampling errors that avoid the assumption of simple random sampling (Wolter, 1985). The jackknife repeated replication technique (JRR) was chosen by TIMSS because it is computationally straightforward and provides approximately unbiased estimates of the sampling errors of means, totals, and percentages.

The variation on the JRR technique used in TIMSS 2007 is described in Johnson and Rust (1992). It assumes that the primary sampling units (PSUs) can be paired in a manner consistent with the sampling design, with each pair regarded as members of a pseudo-stratum for variance estimation purposes. When used in this way, the JRR technique appropriately accounts for the combined effect of the between- and within-PSU contributions to the sampling variance. The general use of JRR entails systematically assigning pairs of schools to sampling zones, and randomly selecting one of these schools to have its contribution doubled and the other to have its contribution zeroed, so as to construct a number of "pseudo-replicates" of the original sample. The statistic of interest is computed once for the entire original sample, and once again for each jackknife pseudo-replicate sample. The variation between the estimates for each of the jackknife replicate samples and the original sample estimate is the jackknife estimate of the sampling error of the statistic.

11.4.2 Constructing Sampling Zones for Sampling Variance Estimation

To apply the JRR technique used in TIMSS 2007, the sampled schools were paired and assigned to a series of groups known as sampling zones. This was done at Statistics Canada by working through the list of sampled schools in the order in which they were selected and assigning the first and second participating schools to the first sampling zone, the third and fourth participating schools to the second zone, and so on. In total, 75 zones were used, allowing for 150 schools per country. When more than 75 zones were constructed, they were collapsed to keep the total number to 75.

Sampling zones were constructed within design domains, or explicit strata. When there was an odd number of schools in an explicit stratum, either by design or because of school non-response, the students in the remaining school were randomly divided to make up two "quasi" schools for the purposes of calculating the jackknife standard error.[14] Each sampling zone then consisted of a pair of schools or "quasi" schools. Exhibit 11.20 shows the number of sampling zones in each country.

Within each sampling zone, both schools were assigned an indicator (u_j), coded randomly to 0 or 1, such that one school had a value of 0, and the other a value of 1. This indicator determined whether the weights for the sampled students in the school in this zone were to be doubled ($u_j = 1$) or zeroed ($u_j = 0$) for the purposes of creating the pseudo-replicate samples.

14 If the remaining school consisted of 2 sampled classrooms, each classroom became a "quasi" school.

Exhibit 11.20 Number of Sampling Zones Used in All TIMSS 2007 Countries

Country	TIMSS 2007 Sampling Zones	
	Fourth Grade	Eighth Grade
Algeria	75	75
Armenia	74	74
Australia	75	75
Austria	75	—
Bahrain	—	75
Bosnia and Herzegovina	—	75
Botswana	—	75
Bulgaria	—	75
Chinese Taipei	75	75
Colombia	72	75
Cyprus	—	75
Czech Republic	72	74
Denmark	69	—
Egypt	—	75
El Salvador	75	73
England	72	69
Georgia	75	71
Germany	75	—
Ghana	—	75
Hong Kong SAR	64	61
Hungary	73	72
Indonesia	—	75
Iran, Islamic Rep. of	75	75
Israel	—	74
Italy	75	75
Japan	75	74
Jordan	—	75
Kazakhstan	71	—
Korea, Rep. of	—	75
Kuwait	75	75
Latvia	74	—
Lebanon	—	68
Lithuania	75	72
Malaysia	—	75
Malta	—	75
Mongolia	75	75
Morocco	75	68
Netherlands	71	—
New Zealand	75	—
Norway	73	70
Oman	—	75
Palestinian Nat'l Auth.	—	75
Qatar	75	75
Romania	—	75
Russian Federation	61	63
Saudi Arabia	—	75
Scotland	70	65
Serbia	—	74
Singapore	75	75
Slovak Republic	75	—
Slovenia	74	74
Sweden	75	75
Syrian Arab Republic	—	75
Thailand	—	75
Tunisia	75	75
Turkey	—	74
Ukraine	73	74
United States	75	75
Yemen	73	—
Benchmark Participants		
Alberta, Canada	73	—
Basque Country, Spain	—	65
British Columbia, Canada	75	75
Dubai, UAE	75	75
Massachusetts, US	24	24
Minnesota, US	25	25
Ontario, Canada	75	75
Quebec, Canada	75	75

11.4.3 Computing Sampling Variance Using the JRR Method

To compute a statistic t from the sample of a country, the formula for the sampling variance estimate of the statistic t, based on the JRR algorithm used in TIMSS 2007, is given by the following equation:

$$Var_{jrr}\left(t\right)=\sum_{h=1}^{H}\left[t\left(J_h\right)-t\left(S\right)\right]^2$$

where H is the total number of sampling zones in the sample of the country under consideration. The term $t(S)$ corresponds to the statistic of interest for the whole sample computed with the overall sampling weights (as described in Chapter 9). The term $t(J_h)$ denotes the same statistic using the h^{th} jackknife replicate sample J_h and its set of replicate sampling weights, which are identical to the overall sampling weights, except for the students in the h^{th} sampling zone. For the students in the h^{th} zone, all students belonging to one of the randomly selected schools of the pair were removed, and the students belonging to the other school in the zone were included twice. In practice, this was accomplished by recoding to zero the weights for the students in the school to be excluded from the replication, and multiplying by two the weights of the remaining students within the h^{th} pair. Each sampled student was assigned a vector of 75 replicate sampling weights W_{hi}, where h took values from 1 to 75. If W_{0i} was the overall sampling weight of student i, the h replicate weights for that student were computed as

$$W_{hi}=W_{0i}\cdot k_{hi}$$

where

$$k_{hi}=\begin{cases} 2\cdot u_j & \text{if student } i \text{ is in school } j \text{ of sampling zone } h \\ 1 & \text{otherwise} \end{cases}$$

The school-level indicators u_j determined which students in a sampling zone would get zero weights and which ones would get double weights, on the basis of the school within the pair from which the students were sampled. The process of setting the k_{hi} values for all sampled students and across all sampling zones is illustrated in Exhibit 11.21. Thus, the computation of the JRR variance estimate for any statistic in TIMSS 2007 required the computation of the statistic up to 76 times for any given country: once to obtain the statistic for the full sample based on the overall weights W_{0i}, and up to 75 times to obtain the statistics for each of the jackknife replicate samples J_h using a set of replicate weights W_{hi}.

Exhibit 11.21 Construction of Replicate Weights Across Sampling Zones in TIMSS 2007

Sampling Zone	School Replicate Indicator (u_i)	Replicate Factors for Computing JRR Replicate Sampling Weights (k_{hi})						
		1	2	3	⋯	h	⋯	75
1	0	0	1	1	⋯	1	⋯	1
	1	2						
2	0	1	0	1	⋯	1	⋯	1
	1		2					
3	0	1	1	0	⋯	1	⋯	1
	1			2				
⋮	⋮	⋮	⋮	⋮	⋱	⋮	⋯	⋮
h	0	1	1	1	⋯	0	⋯	1
	1					2		
⋮	⋮	⋮	⋮	⋮	⋯	⋮	⋱	⋮
75	0	1	1	1	⋯	1	⋯	0
	1							2

In the TIMSS 2007 analyses, 75 replicate weights were computed for each country regardless of the number of actual zones within the country. If a country had fewer than 75 zones, then the additional replicate weights where h was greater than the number of zones within the country were all made equal to the overall sampling weight. Although this involved some redundant computations, having 75 replicate weights for each country had no effect on the magnitude of the error variance computed using the jackknife formula and it simplified the computation of standard errors for numerous countries at a time. All standard errors presented in the TIMSS 2007 international reports were computed using SAS programs developed at the TIMSS & PIRLS International Study Center.

11.4.4 Estimating Imputation Variance

The TIMSS 2007 item pool was far too extensive to be administered in its entirety to any one student, and so a matrix-sampling test design was developed whereby each student was given a single test booklet containing only a part of the entire assessment.[15] The results for all of the booklets were then aggregated using item response theory to provide results for the entire assessment. Since each student responded to just a subset of the assessment items, multiple imputation (the generation of plausible values) was used to derive reliable estimates of student performance on the assessment as a whole. Since every student proficiency estimate incorporates some uncertainty arising from the use of IRT models, TIMSS followed the customary procedure of generating five estimates for each student and using the variability among them as a measure of this imputation uncertainty, or error. In the TIMSS 2007 international reports, the imputation error for each variable has been combined with the sampling error for that variable to provide a standard error that incorporates both.

The general procedure for estimating the imputation variance using plausible values is described in Mislevy, Beaton, Kaplan, and Sheenan (1992). First, compute the statistic t for each set of M plausible values. The statistics t_m, where $m = 1, 2, \ldots, 5$, can be anything estimable from the data, such as a mean, the difference between means, percentiles, and so forth.

Once the statistics t_m are computed, the imputation variance is then calculated as:

$$Var_{imp} = \left(1 + \frac{1}{M}\right) Var\left(t_1, \cdots, t_m\right)$$

where M is the number of plausible values used in the calculation, and $Var\left(t_1, \ldots, t_M\right)$ is the usual variance of the M estimates computed using each plausible value.

11.4.5 Combining Sampling and Imputation Variance

The standard errors of all proficiency statistics reported by TIMSS include both sampling and imputation variance components. These standard errors were computed using the following formula:

15 The TIMSS 2007 assessment design is described in Chapter 2.

$$Var\left(t_{pv}\right) = Var_{jrr}\left(t_1\right) + Var_{imp}$$

where $Var_{jrr}(t_1)$ is the sampling variance computed for the first plausible value[16] and Var_{imp} is the imputation variance. The *TIMSS 2007 User Guide for the International Database* (Foy & Olson, 2009) contains programs in SAS and SPSS that compute each of these variance components for the TIMSS 2007 data. Furthermore, the IDB Analyzer—software provided with the international database—automatically computes standard errors as described in this section.

Exhibits 11.22 through 11.25 show basic summary statistics for overall mathematics and science achievement in the TIMSS 2007 assessment for the fourth and eighth grades. Each exhibit presents the student sample size, the mean and standard deviation averaged across the five plausible values, the jackknife sampling error for the mean, and the overall standard error for the mean, which includes the imputation error. Appendix E contains tables showing the same summary statistics for the mathematics and science content and cognitive domains at the fourth and eighth grades.

16 Under ideal circumstances and with unlimited computing resources, the JRR sampling variance would be computed for each of the plausible values and the imputation variance as described here. This would require computing the same statistic up to 380 times (once overall for each of the five plausible values using the overall sampling weights, and then 75 times more for each plausible value using the complete set of replicate weights). An acceptable shortcut, however, is to compute the JRR sampling variance component using only one plausible value (the first one), and then the imputation variance using the five plausible values. Using this approach, a statistic needs to be computed only 80 times.

Exhibit 11.22 Summary Statistics and Standard Errors for Proficiency in Mathematics at the Fourth Grade

Country	Sample Size	Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	377.645	89.560	5.058	5.176
Armenia	4,079	499.513	89.523	4.245	4.286
Australia	4,108	516.062	83.306	3.468	3.509
Austria	4,859	505.389	67.937	1.905	2.005
Chinese Taipei	4,131	575.819	69.225	1.633	1.733
Colombia	4,801	355.450	90.178	4.794	4.974
Czech Republic	4,235	486.399	71.458	2.665	2.781
Denmark	3,519	523.106	70.835	2.335	2.403
El Salvador	4,166	329.906	90.819	3.463	4.104
England	4,316	541.465	86.044	2.856	2.882
Georgia	4,108	438.458	88.430	4.180	4.207
Germany	5,200	525.155	68.149	2.224	2.254
Hong Kong SAR	3,791	606.802	67.126	3.429	3.584
Hungary	4,048	509.720	91.160	3.505	3.547
Iran, Islamic Rep. of	3,833	402.422	83.522	3.617	4.054
Italy	4,470	506.750	77.025	3.132	3.135
Japan	4,487	568.157	76.075	2.093	2.121
Kazakhstan	3,990	549.348	83.807	7.117	7.146
Kuwait	3,803	315.535	99.299	3.412	3.646
Latvia	3,908	537.200	71.904	2.089	2.306
Lithuania	3,980	529.799	75.761	2.288	2.372
Morocco	3,894	341.305	95.265	4.509	4.668
Netherlands	3,349	534.952	61.346	2.130	2.145
New Zealand	4,940	492.475	86.135	2.216	2.313
Norway	4,108	473.216	76.222	2.430	2.543
Qatar	7,019	296.268	90.067	0.974	1.043
Russian Federation	4,464	544.045	83.370	4.909	4.911
Scotland	3,929	494.449	78.926	2.182	2.214
Singapore	5,041	599.406	84.146	3.716	3.744
Slovak Republic	4,963	495.975	84.937	4.428	4.468
Slovenia	4,351	501.843	71.399	1.628	1.811
Sweden	4,676	502.574	66.482	2.385	2.527
Tunisia	4,134	327.435	110.809	4.406	4.469
Ukraine	4,292	469.003	84.479	2.893	2.912
United States	7,896	529.009	75.329	2.395	2.448
Yemen	5,811	223.683	110.136	5.637	5.968
Benchmarking Participants					
Alberta, Canada	4,037	505.320	66.059	2.938	2.952
British Columbia, Canada	4,153	505.219	71.314	2.543	2.749
Dubai, UAE	3,064	444.334	89.598	1.896	2.141
Massachusetts, US	1,747	572.484	69.772	3.468	3.513
Minnesota, US	1,846	554.117	77.714	5.823	5.863
Ontario, Canada	3,496	511.614	68.001	3.008	3.100
Quebec, Canada	3,885	519.103	67.347	2.999	3.028

Exhibit 11.23 Summary Statistics and Standard Errors for Proficiency in Science at the Fourth Grade

Country	Sample Size	Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	353.819	101.883	5.810	6.024
Armenia	4,079	484.387	118.784	5.529	5.684
Australia	4,108	527.397	80.497	3.149	3.341
Austria	4,859	525.627	77.410	2.182	2.520
Chinese Taipei	4,131	556.696	77.353	1.911	2.002
Colombia	4,801	400.305	97.459	5.397	5.446
Czech Republic	4,235	515.052	75.607	2.895	3.124
Denmark	3,519	516.917	76.937	2.709	2.854
El Salvador	4,166	389.583	93.202	3.191	3.368
England	4,316	541.527	80.219	2.790	2.852
Georgia	4,108	417.637	84.662	4.094	4.556
Germany	5,200	527.554	79.119	2.283	2.403
Hong Kong SAR	3,791	554.181	67.885	3.460	3.498
Hungary	4,048	536.226	84.807	3.113	3.346
Iran, Islamic Rep. of	3,833	435.639	97.424	4.071	4.275
Italy	4,470	535.217	81.368	3.090	3.172
Japan	4,487	547.780	69.631	1.672	2.066
Kazakhstan	3,990	532.830	74.326	5.481	5.631
Kuwait	3,803	348.151	123.080	4.096	4.367
Latvia	3,908	541.895	66.857	2.142	2.288
Lithuania	3,980	514.205	65.196	1.807	2.366
Morocco	3,894	297.447	123.744	5.580	5.864
Netherlands	3,349	523.176	59.870	2.209	2.610
New Zealand	4,940	504.066	90.091	2.369	2.626
Norway	4,108	476.551	76.659	2.488	3.484
Qatar	7,019	294.396	129.491	1.240	2.559
Russian Federation	4,464	546.231	80.524	4.636	4.781
Scotland	3,929	500.409	76.241	2.002	2.275
Singapore	5,041	586.654	93.044	3.905	4.091
Slovak Republic	4,963	525.691	87.247	4.634	4.765
Slovenia	4,351	518.393	76.172	1.887	1.936
Sweden	4,676	524.810	73.575	2.763	2.876
Tunisia	4,134	318.474	141.383	5.524	5.907
Ukraine	4,292	473.814	82.912	2.605	3.085
United States	7,896	538.574	83.990	2.579	2.714
Yemen	5,811	197.365	130.062	6.650	7.188
Benchmarking Participants					
Alberta, Canada	4,037	542.588	73.632	3.655	3.828
British Columbia, Canada	4,153	536.690	72.661	2.476	2.691
Dubai, UAE	3,064	459.648	107.310	2.601	2.752
Massachusetts, US	1,747	570.894	69.230	3.845	4.253
Minnesota, US	1,846	551.478	79.542	6.056	6.089
Ontario, Canada	3,496	535.869	78.245	3.289	3.722
Quebec, Canada	3,885	517.122	66.651	2.415	2.664

Exhibit 11.24 Summary Statistics and Standard Errors for Proficiency in Mathematics at the Eighth Grade

Country	Sample Size	Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	386.752	59.250	1.493	2.142
Armenia	4,689	498.680	84.735	3.438	3.505
Australia	4,069	496.232	79.426	3.874	3.934
Bahrain	4,230	398.071	83.601	1.320	1.567
Bosnia and Herzegovina	4,220	455.863	77.801	2.678	2.697
Botswana	4,208	363.539	76.579	1.981	2.268
Bulgaria	4,019	463.630	101.605	4.857	4.965
Chinese Taipei	4,046	598.301	105.505	4.337	4.533
Colombia	4,873	379.636	78.935	3.600	3.632
Cyprus	4,399	465.477	89.319	1.569	1.648
Czech Republic	4,845	503.807	73.686	2.313	2.392
Egypt	6,582	390.557	100.247	3.409	3.571
El Salvador	4,063	340.441	72.822	2.664	2.756
England	4,025	513.404	83.579	4.790	4.816
Georgia	4,178	409.617	96.464	5.889	5.950
Ghana	5,294	309.370	91.597	4.150	4.364
Hong Kong SAR	3,470	572.487	93.734	5.675	5.793
Hungary	4,111	516.895	84.678	3.417	3.474
Indonesia	4,203	397.110	87.341	3.692	3.808
Iran, Islamic Rep. of	3,981	403.380	86.095	3.968	4.116
Israel	3,294	463.251	98.873	3.866	3.949
Italy	4,408	479.626	76.231	2.925	3.037
Japan	4,312	569.810	85.416	2.063	2.407
Jordan	5,251	426.893	102.208	4.037	4.117
Korea, Rep. of	4,240	597.266	92.069	2.471	2.707
Kuwait	4,091	353.670	78.636	2.196	2.316
Lebanon	3,786	449.061	74.637	3.827	3.984
Lithuania	3,991	505.818	79.744	2.218	2.324
Malaysia	4,466	473.886	79.248	5.005	5.029
Malta	4,670	487.752	91.772	0.868	1.210
Morocco	3,060	380.784	80.326	2.753	2.970
Norway	4,627	469.216	65.665	1.918	1.976
Oman	4,752	372.434	94.944	2.848	3.370
Palestinian Nat'l Auth.	4,378	367.155	102.436	3.399	3.549
Qatar	7,184	306.791	93.360	0.727	1.374
Romania	4,198	461.318	99.748	4.038	4.099
Russian Federation	4,472	511.734	83.079	4.045	4.101
Saudi Arabia	4,243	329.337	76.433	2.174	2.852
Scotland	4,070	487.406	79.727	3.606	3.705
Serbia	4,045	485.767	89.451	3.077	3.316
Singapore	4,599	592.785	92.958	3.732	3.814
Slovenia	4,043	501.476	71.618	1.996	2.110
Sweden	5,215	491.300	70.052	2.093	2.260
Syrian Arab Republic	4,650	394.838	82.402	3.407	3.765
Thailand	5,412	441.390	91.617	4.897	4.951
Tunisia	4,080	420.413	66.519	2.343	2.433
Turkey	4,498	431.810	108.742	4.680	4.753
Ukraine	4,424	462.162	89.231	3.600	3.621
United States	7,377	508.454	76.736	2.773	2.830
Benchmarking Participants					
Basque Country, Spain	2,296	498.559	68.590	2.723	2.990
British Columbia, Canada	4,256	509.449	72.443	3.016	3.032
Dubai, UAE	3,195	460.616	96.176	2.257	2.370
Massachusetts, US	1,897	547.130	79.234	4.510	4.559
Minnesota, US	1,777	532.450	67.764	4.299	4.411
Ontario, Canada	3,448	517.232	70.214	3.485	3.518
Quebec, Canada	3,956	528.110	68.410	3.221	3.512

Exhibit 11.25 Summary Statistics and Standard Errors for Proficiency in Science at the Eighth Grade

Country	Sample Size	Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	408.060	62.603	1.488	1.738
Armenia	4,689	487.960	101.142	5.511	5.755
Australia	4,069	514.788	80.324	3.610	3.648
Bahrain	4,230	467.448	86.027	1.411	1.718
Bosnia and Herzegovina	4,220	465.745	79.444	2.772	2.815
Botswana	4,208	354.534	99.425	2.537	3.054
Bulgaria	3,079	470.284	102.622	5.676	5.892
Chinese Taipei	4,046	561.003	89.274	3.603	3.686
Colombia	4,873	417.182	76.652	3.466	3.515
Cyprus	4,399	451.624	85.319	1.655	2.044
Czech Republic	4,845	538.878	71.394	1.892	1.919
Egypt	6,582	408.242	99.381	3.356	3.563
El Salvador	4,063	387.274	69.770	2.745	2.926
England	4,025	541.505	85.398	4.458	4.479
Georgia	4,178	420.902	83.326	4.603	4.768
Ghana	5,294	303.272	108.360	5.006	5.356
Hong Kong SAR	3,470	530.209	80.969	4.847	4.919
Hungary	4,111	539.034	76.583	2.840	2.919
Indonesia	4,203	426.990	74.181	3.168	3.366
Iran, Islamic Rep. of	3,981	458.929	81.340	3.484	3.594
Israel	3,294	467.922	100.906	4.304	4.338
Italy	4,408	495.147	77.517	2.773	2.818
Japan	4,312	553.815	77.108	1.852	1.897
Jordan	5,251	481.721	97.720	3.945	3.962
Korea, Rep. of	4,240	553.139	75.862	1.939	2.034
Kuwait	4,091	417.956	89.241	2.552	2.818
Lebanon	3,786	413.611	96.812	5.808	5.932
Lithuania	3,991	518.559	78.205	2.266	2.550
Malaysia	4,466	470.801	88.199	5.981	6.027
Malta	4,670	457.167	113.859	1.238	1.365
Morocco	3,060	401.831	78.550	2.597	2.898
Norway	4,627	486.758	73.272	2.059	2.187
Oman	4,752	422.502	95.744	2.911	2.964
Palestinian Nat'l Auth.	4,378	404.126	110.930	3.456	3.504
Qatar	7,184	318.854	125.866	0.927	1.734
Romania	4,198	461.900	87.893	3.672	3.850
Russian Federation	4,472	529.570	77.651	3.819	3.883
Saudi Arabia	4,243	403.245	77.978	2.213	2.448
Scotland	4,070	495.732	81.116	3.319	3.397
Serbia	4,045	470.307	84.720	3.007	3.151
Singapore	4,599	567.250	103.889	4.373	4.448
Slovenia	4,043	537.544	72.017	2.133	2.213
Sweden	5,215	510.690	78.033	2.477	2.557
Syrian Arab Republic	4,650	451.976	74.713	2.678	2.885
Thailand	5,412	470.614	82.735	4.268	4.297
Tunisia	4,080	444.898	60.475	1.921	2.117
Turkey	4,498	454.159	91.892	3.648	3.711
Ukraine	4,424	485.063	83.992	3.418	3.459
United States	7,377	519.989	82.274	2.832	2.857
Benchmarking Participants					
Basque Country, Spain	2,296	497.706	72.028	2.746	2.956
British Columbia, Canada	4,256	525.717	70.793	2.660	2.685
Dubai, UAE	3,195	488.865	94.001	2.601	2.762
Massachusetts, US	1,897	556.041	79.367	4.354	4.554
Minnesota, US	1,777	538.510	71.850	4.716	4.762
Ontario, Canada	3,448	526.128	69.455	3.574	3.648
Quebec, Canada	3,956	506.589	68.973	2.897	3.054

References

Beaton, A.E. (1969). Criterion scaling of questionnaire items. *Socio-Economic Planning Sciences*, 2, 355–362.

Beaton, A.E., & Johnson, E.G. (1990). The average response method of scaling. *Journal of Educational Statistics*, 15, 9–38.

Beaton, A.E., & Johnson, E.G. (1992). Overview of the scaling methodology used in the national assessment. *Journal of Educational Measurement*, 26(2), 163–175.

Birnbaum, A. (1968). Some latent trait models and their use in inferring an examinee's ability. In F. M. Lord & M. R. Novick, *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley Publishing.

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

Johnson, E.G., & Rust, K.F. (1992). Population inferences and variance estimation for NAEP data. *Journal of Educational Statistics*, 17(2), 175–190.

Lord, F.M. (1980). *Applications of items response theory to practical testing problems*. Hillsdales, NJ: Lawrence Erlbaum Associates.

Lord, F.M., & Novick, M.R. (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

Martin, M.O., Mullis, I.V.S., & Chrostowski, S.J. (2004). *TIMSS 2003 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Martin, M.O., Mullis, I.V.S., & Foy, P. (with Olson, J.F., Erberber, E., Preuschoff, C., & Galia, J.). (2008). *TIMSS 2007 international science report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mislevy, R.J. (1985). Estimation of latent group effects. *Journal of the American Statistical Association*, 80, 993–97.

Mislevy, R.J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56, 177–196.

Mislevy, R.J., Beaton, A., Kaplan, B.A., & Sheehan, K. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, 29(2), 133–161.

Mislevy, R.J., Johnson, E.G. & Muraki, E. (1992). Scaling procedures in NAEP. *Journal of Educational Statistics*, 17(2), 131–154.

Mislevy, R.J. & Sheehan, K. (1987). Marginal estimation procedures. In A. E. Beaton (Ed.), *Implementing the new design*: *The NAEP 1983–84 technical report* (pp. 293–360). (no. 15-TR-20) Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Mullis, I.V.S., Martin, M.O., & Foy, P. (with Olson, J.F., Preuschoff, C., Erberber, E., Arora, A., & Galia, J.). (2008). *TIMSS 2007 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

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

Mullis, I.V.S., Martin, M.O., Smith, T.A., Garden, R.A., Gregory, K.D., Gonzalez, E.J., Chrostowki, S.J., & O'Connor, K.M. (2003). *TIMSS assessment frameworks and specifications 2003* (2nd ed.). Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Muraki, E. (1992). A generalized partial credit model: Application of an EM algorithm. *Applied Psychological Measurement*, 16(2), 159–176.

Muraki, E., & Bock, R.D. (1991). *PARSCALE: Parameter scaling of rating data* [Software Version 4.1]. Chicago, IL: Scientific Software, Inc.

Rubin, D.B. (1987). *Multiple imputation for nonresponse in surveys*. New York: John Wiley & Sons.

Sheehan, K. M. (1985). *M-GROUP*: *Estimation of group effects in multivariate models* [Software Version 3.2]. Princeton, NJ: Educational Testing Service.

Thomas, N. (1993). Asymptotic corrections for multivariate posterior moments with factored likelihood functions. *Journal of Computational and Graphical Statistics*, 2, 309–22.

Van Der Linden, W.J. & Hambleton, R. (1996). *Handbook of modern item response theory*. New York. Springer-Verlag.

Wingersky, M., Kaplan, B.A., & Beaton, A.E. (1987). Joint estimation procedures. In A. E. Beaton (Ed.), *Implementing the new design: The NAEP 1983–84 technical report* (pp.285–92) (No. 15-TR-20). Princeton, NJ: Educational Testing Service, National Assessment of Educational Progress.

Wolter, K. M. (1985). *Introduction to variance estimation*. New York: Springer-Verlag.

Yamamoto, K., & Kulick, E. (2000). Scaling methodology and procedures for the TIMSS mathematics and science scales. In M.O. Martin, K.D. Gregory, & S.E. Stemler (Eds.), *TIMSS 1999 technical report*. Chestnut Hill, MA: Boston College.

Chapter 12

Creating the TIMSS 2007 Background Indices

Michael O. Martin and Corinna Preuschoff

12.1 Overview

The TIMSS 2007 international reports (Martin, Mullis, & Foy, 2008; Mullis, Martin, & Foy, 2008) presented factors related to teaching and learning mathematics and science helpful in understanding the achievement results. To describe the educational context for mathematics and science achievement and to provide useful information to policy-makers, curriculum specialists, and researchers, data on hundreds of background variables were collected from students, teachers, schools, and National Research Coordinators (NRCs). These questionnaire data were summarized in a concise manner in the exhibits (pictures and tables) of the international reports to make them as accessible and useful as possible. One of the principal ways of doing this was through the computation of index variables, multiple-item indicators that combined data from several questions in the TIMSS 2007 questionnaires.

As described in Chapter 3, TIMSS contextual data were collected through four sets of questionnaires: student, teacher, school, and curriculum. The present chapter describes the TIMSS 2007 background indices used to summarize and report these data, and provides information on the reliability and validity of the scales underlying these indices.

12.2 Computing Background Indices

In the TIMSS reports, an index is a composite variable that assigns students to one of three levels—high, medium, and low—on the basis of responses to a series of component variables. The high category of an index is defined in terms of the student responses (or those responses of teachers or school principals) that are expected to be most characteristic of a supportive learning environment, whereas the low category is defined in terms of the responses expected to characterize the least supportive learning environment. The medium level is somewhere in between. The TIMSS indices are intended to describe factors fostering mathematics and science achievement in terms of responses to the questions that were actually asked, thereby preserving a high degree of direct interpretability.

As an example, the Index of Students' Perception of Being Safe in School (SPBSS) (described later in this chapter) groups students according to their reports of the frequency of incidents affecting their safety: 1) Something of mine was stolen; 2) I was hit or hurt by other student(s) (for example, shoving, hitting, kicking); 3) I was made to do things that I didn't want to do by other students; 4) I was made fun of or called names; and 5) I was left out of activities by other students. Students at the high level of the index (i.e., those that perceived school to be very safe) reported that no such incidents happened to them during the past month. In contrast, students at the low level of this index reported three or more such incidents.

TIMSS used two different methods to create composite scales: the combined response method and the scale method. The combined response method was used to directly classify cases into the high, medium, or low level of an index, depending on the combination of responses to the source questions. For example, the Index of Good Attendance at School (GAS) (described later in this chapter) classified students into the three index levels based on principals' reports on frequency of occurrence and seriousness of three aspects of attendance problems: 1) Arriving late at school; 2) Absenteeism (i.e., unjustified absences); and 3) Skipping class. Responses were assigned to the high level of the index if the school principal reported that all three behaviors either never occur or that they are not a serious problem. Responses were assigned to the low level if the principal indicated that two or more of the behaviors were a serious problem, or two behaviors were minor problems and a third behavior a serious problem. All other response combinations were assigned to the medium category. The scale method was used when the construct of interest had an underlying quantitative continuum. The index scores were computed by averaging the numerical values associated with each response option. Following this, students were assigned to the three levels based on cutoff points. This method often was employed for items that made use of Likert scale format (e.g., response options are *agree a lot* coded 1, *agree a little* coded 2, *disagree a little* coded 3, and *disagree a lot* coded 4). Examples of this type of index are the measures of students' attitudes toward mathematics and science presented in Chapter 4 of the international reports.

Underlying each TIMSS background index was a scale made up of the component variables of the index. In constructing an index, it was important that the component variables of the underlying scale were intercorrelated so that together they formed a reliable scale and also that they were correlated to some extent with students' mathematics and science achievement. The process of identifying the response combinations that defined the high, medium, and low level of the index was informed by the relationship with achievement, but where possible these combinations were chosen based on a judgment of which responses could be expected to most effectively capture constructs describing environments supportive for learning mathematics and science.

12.3 Developing the Background Indices

Planning for reporting the questionnaire data and creating the TIMSS 2007 background indices began with a review of the questionnaires that had been administered in TIMSS 2007 and in previous TIMSS cycles. Staff at the TIMSS & PIRLS International Study Center identified TIMSS 2007 variables that also had been used in 2003, 1999, and 1995 to determine if they could be used to measure trends. They also checked to see if improvements could be made to indices developed in previous cycles by adding new items from the TIMSS 2007 questionnaires. Newly developed questions were reviewed in the context of the TIMSS 2007 framework to identify variables for creating new indices.

Countries following a Southern Hemisphere school year administered the TIMSS 2007 assessment at the end of 2006 (the end of their school year), and so data from some of these—Australia, Botswana, El Salvador, New Zealand, Malaysia, and Singapore— were available for use in exploratory analyses before the data from Northern Hemisphere countries became available. These exploratory analyses had three primary purposes: identifying new indices that could be created from variables added in the 2007 cycle, ensuring that indices used in previous cycles still performed similarly in 2007, and exploring the impact of improving indices created in previous cycles by adding extra component variables. These analyses used principal component analysis to explore the dimensionality of proposed indices using different combinations of variables, and also examined the reliability of each underlying scale and the relationship between its component variables and mathematics and science achievement.

Based on the exploratory analyses, specifications were developed for the construction of all indices. These described the source variables to be used, how they should be recoded and combined, and how the resulting indices should be presented in the international reports. The analysis specifications guided the programmers and production staff who implemented these analyses and created exhibits for the international reports, and were made available to NRCs to aid their reviews of the exhibits. The final report exhibits were produced using custom-designed SAS programs that calculated student achievement averages using all five imputed scores (plausible values) for each student, including standard errors calculated using the jackknife procedure (see Chapter 11).

Representatives from participating countries had several opportunities to review proposed exhibits and make suggestions for additions and modifications. The draft exhibits first were reviewed in conjunction with the TIMSS 2007 international reports outline, background data almanacs, and analysis notes, at the seventh NRC meeting in Salzburg, Austria in December 2007. Based on NRCs' comments, the exhibits and data were further refined for a second review at the eighth NRC meeting in Gaborone, Botswana in June 2008. At this meeting, NRCs were provided with a draft of the TIMSS 2007 international reports containing complete versions of the report exhibits. NRCs approved these final exhibits, including index definitions.

As a final step, all indices were made available for secondary analysis as part of the TIMSS 2007 International Database. Supplement 3 of the *TIMSS 2007 User Guide for the International Database* (Foy & Olson, 2009) provides a detailed description of all indices included in the international database.

Background indices were presented throughout Chapters 4–8 of the TIMSS 2007 international reports. In all these exhibits, the student was the unit of analysis even if the information had been supplied by teachers or principals. Results always were presented in terms of the percentage of students possessing a particular characteristic. This approach presents the data from the perspective of students' educational experiences and is consistent with the TIMSS sampling and assessment design. In many exhibits, the average mathematics or science achievement of the students at each index level also was presented.

Since one of the major benefits of TIMSS is the ability to measure trends over time, background indices, which spanned across assessment cycles (1995, 1999, 2003, and 2007), were included whenever possible. In these exhibits, for example, the change from 2003 in the percentage of students at each index level was displayed for countries that participated in the 2003 assessment, with an arrow indicating if the percent in 2007 was significantly higher or lower.

12.4 Reliability and Validity of Background Indices

In this section, the composition of each index variable reported in the TIMSS 2007 international reports is briefly described and indicators of reliability and validity for the component variables of these indices are presented. The reliability of the underlying scales is assessed using Cronbach's alpha, and the relationship with achievement is summarized by the multiple correlation between the component variables of the scales underlying the indices and achievement (multiple R), and the percent of variance in achievement accounted for by the component variables (R-square). These statistics provide a sense of how well the component variables are related to mathematics and science achievement, which is an aspect of the validity of the index. In addition, confirmatory factor analysis was used to examine the dimensionality of the scales underlying the indices and to present a latent trait measurement model of each scale and its component variables.

In the exhibits in this chapter, reliability and validity indicators are presented for each TIMSS 2007 participant, together with the median indicator across countries. Indicators are presented separately for mathematics and science at fourth and eighth grades. For countries teaching science as a single integrated subject, a single index was created for each science exhibit. For countries where the sciences are taught as separate subjects (biology, earth science, chemistry, and physics) at the eighth grade, students were asked separately about each subject. Thus, separate indices were created for each science subject, and the reliability and validity indicators for separate science countries are presented in a separate panel (e.g., Exhibit 12.1).

The factor analyses were conducted using the Mplus software package (Muthén & Muthén, 2007). Mplus was chosen because of its ability to model complex survey data and use information efficiently in the presence of missing data. The Mplus analyses reported in this chapter were conducted using a variation of the TIMSS sampling weight (SENWGT; see Foy & Olson, 2009) that weights each country equally, while taking into account the complex TIMSS sampling design and correcting for unequal selection probabilities as necessary. The analyses were conducted using data from 49 countries at the eighth grade and 36 countries at the fourth grade. The benchmarking participants were not included in the analyses.

12.4.1 Student-level Indices

In the *TIMSS 2007 Student Questionnaire*, students were asked about their home environments and school experiences, and their attitudes toward mathematics and science. At the fourth grade, two indices were constructed representing different aspects of students' attitudes toward mathematics and science: positive affect and self-confidence. An index of time students spend on homework in mathematics and science and an index of students' perceptions of being safe in school also were constructed at the fourth grade. At the eighth grade, three indices were constructed representing three aspects of students' attitudes toward mathematics and science: positive affect, self-confidence, and valuing the subject. The eighth grade also included an index of time students spend on homework in mathematics and science and an index of students' perceptions of being safe in school. Reliability and validity indicators for the attitudinal indices are presented in Exhibits 12.1 to 12.3. The results from confirmatory factor analysis, representing further evidence of the validity of the TIMSS attitude scales, are presented in Exhibit 12.4.

The Index of Students' Positive Affect Toward Mathematics (PATM) and the Index of Students' Positive Affect Toward Science (PATS) examined students' general affect toward mathematics and science. The index was presented in Exhibit 4.8 of the TIMSS 2007 international reports. The exhibit shows trends from 1995 at the fourth grade, and from 1995 and 1999 at the eighth grade (comparable data were not available from 2003).

For mathematics the index is based on students' responses to three statements about mathematics: 1) I enjoy learning mathematics; 2) Mathematics is boring; and 3) I like mathematics. For science the index is based on students' responses to three statements about science: 1) I enjoy learning science; 2) Science is boring; and 3) I like science. The negatively worded statements "mathematics is boring" and "science is boring" were reverse coded. An average was computed across the three items based on a 4-point scale: *agree a lot* = 1, *agree a little* = 2, *disagree a little* = 3, and *disagree a lot* = 4. A high level indicates an average score of less than or equal to 2, corresponding to students agreeing a little or a lot, on average. A low level indicates an average score equal to or greater than 3, corresponding to students disagreeing a little or a lot, on average. A medium level indicates an average score of greater than 2 but less than 3. For countries that taught biology, earth science, chemistry, and physics as separate subjects at the eighth grade, the questions were asked about each individual science subject, and students responded with respect to each science course they were taking. Thus, separate indices were created for each science subject and the reliability and validity indicators for separate science countries are presented in a separate panel for eighth grade in Exhibit 12.1.

A similar index of students' general attitudes toward mathematics and science was presented in the TIMSS 1999 international reports (Martin, M.O., Mullis, I.V.S., Gonzales, E.J., Gregory, K.D., Smith, T.A., Chrostowski, S.J., Garden, R.A., & O'Connor, K.M., 2000; Mullis, I.V.S., Martin, M.O., Gonzales, E.J., Gregory, K.D., Garden, R.A., O'Connor, K.M., Chrostowski, S.J., & Smith, T.A., 2000), including two more variables. For mathematics these were "mathematics is important to everyone's life" and "I would like a job that involved using mathematics", which were not part of the *TIMSS 2007 Student Questionnaire*. Thus, the percentage of students at each index level in 1999 was recomputed based on the TIMSS 2007 index definition.

The three index components also were part of the *TIMSS 1995 Student Questionnaire*. At the eighth grade the *TIMSS 1995 Student Questionnaire*, however, asked about physical science and not about chemistry and physics. Thus, the same data were presented in the "difference in percent from 1995" column of the physics and chemistry panels in the *TIMSS 2007 International Science Report*.

As shown in Exhibit 12.1, the three component variables (statements) form a fairly reliable scale, with median reliability coefficients (Cronbach's alpha) across countries of 0.82 and 0.81 for mathematics and science, respectively, at the fourth grade, and 0.81 and 0.78, respectively, at the eighth grade. At the fourth grade, the median multiple correlation between the three component variables and student achievement was 0.18 for mathematics and 0.16 for science, corresponding to an R-square of 0.03 in each case, after rounding. At the eighth grade, the median multiple correlation between the three component variables and student achievement was 0.28 for mathematics and 0.24 for general science, corresponding to R-squares of 0.08 and 0.06, respectively. For the separate sciences, the reliabilities were similar to those for general science, although the correlations were somewhat lower, with the median multiple correlations ranging between 0.12 and 0.15, corresponding to R-squares between 0.01 and 0.02.

Exhibit 12.1 Index of Students' Positive Affect Toward Mathematics (PATM) / Science (PATS)—Reliability and Validity Indicators

Countries	Grade 4						Grade 8					
	Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables		Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science
Algeria	0.42	0.42	0.32	0.30	0.10	0.09	0.66	–	0.29	–	0.09	–
Armenia	0.61	0.65	0.11	0.16	0.01	0.02	0.73	–	0.14	–	0.02	–
Australia	0.85	0.86	0.18	0.15	0.03	0.02	0.85	0.88	0.27	0.24	0.07	0.06
Austria	0.85	0.82	0.17	0.16	0.03	0.03	◊	◊	◊	◊	◊	◊
Bahrain	◊	◊	◊	◊	◊	◊	0.81	0.78	0.21	0.18	0.05	0.03
Bosnia and Herzegovina	◊	◊	◊	◊	◊	◊	0.85	–	0.23	–	0.05	–
Botswana	◊	◊	◊	◊	◊	◊	0.69	0.65	0.32	0.45	0.10	0.20
Bulgaria	◊	◊	◊	◊	◊	◊	0.83	–	0.20	–	0.04	–
Chinese Taipei	0.83	0.78	0.28	0.22	0.08	0.05	0.89	0.88	0.49	0.39	0.24	0.15
Colombia	0.50	0.47	0.31	0.32	0.10	0.10	0.71	0.69	0.13	0.17	0.02	0.03
Cyprus	◊	◊	◊	◊	◊	◊	0.83	–	0.34	–	0.11	–
Czech Republic	0.84	0.85	0.16	0.08	0.03	0.01	0.84	–	0.30	–	0.09	–
Denmark	0.85	0.88	0.05	0.13	0.00	0.02	◊	◊	◊	◊	◊	◊
Egypt	◊	◊	◊	◊	◊	◊	0.62	0.60	0.24	0.30	0.06	0.09
El Salvador	0.47	0.41	0.31	0.33	0.10	0.11	0.69	0.64	0.25	0.29	0.06	0.08
England	0.87	0.88	0.12	0.09	0.01	0.01	0.86	0.88	0.22	0.27	0.05	0.07
Georgia	0.57	0.64	0.26	0.19	0.07	0.04	0.73	–	0.24	–	0.06	–
Germany	0.83	0.82	0.14	0.15	0.02	0.02	◊	◊	◊	◊	◊	◊
Ghana	◊	◊	◊	◊	◊	◊	0.45	0.39	0.34	0.40	0.12	0.16
Hong Kong SAR	0.87	0.83	0.28	0.24	0.08	0.06	0.86	0.85	0.36	0.28	0.13	0.08
Hungary	0.86	0.86	0.16	0.17	0.03	0.03	0.84	–	0.33	–	0.11	–
Indonesia	◊	◊	◊	◊	◊	◊	0.66	–	0.22	–	0.05	–
Iran, Islamic Rep. of	0.74	0.76	0.35	0.35	0.12	0.12	0.80	0.80	0.29	0.17	0.09	0.03
Israel	◊	◊	◊	◊	◊	◊	0.82	0.82	0.10	0.20	0.01	0.04
Italy	0.82	0.81	0.16	0.15	0.02	0.02	0.86	0.85	0.30	0.20	0.09	0.04
Japan	0.84	0.83	0.29	0.16	0.08	0.03	0.84	0.85	0.39	0.30	0.15	0.09
Jordan	◊	◊	◊	◊	◊	◊	0.75	0.75	0.29	0.21	0.08	0.04
Kazakhstan	0.55	0.56	0.25	0.19	0.06	0.04	◊	◊	◊	◊	◊	◊
Korea, Rep. of	◊	◊	◊	◊	◊	◊	0.89	0.88	0.47	0.39	0.22	0.15
Kuwait	0.57	0.54	0.28	0.35	0.08	0.12	0.81	0.77	0.20	0.15	0.04	0.02
Latvia	0.81	0.81	0.12	0.05	0.02	0.00	◊	◊	◊	◊	◊	◊
Lebanon	◊	◊	◊	◊	◊	◊	0.69	–	0.27	–	0.07	–
Lithuania	0.81	0.80	0.24	0.12	0.06	0.02	0.77	–	0.36	–	0.13	–
Malaysia	◊	◊	◊	◊	◊	◊	0.82	0.81	0.30	0.26	0.09	0.07
Malta	◊	◊	◊	◊	◊	◊	0.87	–	0.27	–	0.07	–
Morocco	0.47	0.44	0.28	0.35	0.08	0.13	0.64	–	0.29	–	0.09	–
Netherlands	0.89	0.92	0.09	0.10	0.01	0.01	◊	◊	◊	◊	◊	◊
New Zealand	0.82	0.82	0.12	0.23	0.01	0.05	◊	◊	◊	◊	◊	◊
Norway	0.88	0.89	0.10	0.14	0.01	0.02	0.88	0.90	0.30	0.17	0.09	0.03
Oman	◊	◊	◊	◊	◊	◊	0.66	0.59	0.37	0.35	0.14	0.12
Palestinian Nat'l Auth.	◊	◊	◊	◊	◊	◊	0.69	0.67	0.27	0.25	0.07	0.06
Qatar	0.62	0.62	0.32	0.28	0.10	0.08	0.81	0.74	0.15	0.14	0.02	0.02
Romania	◊	◊	◊	◊	◊	◊	0.81	–	0.23	–	0.05	–
Russian Federation	0.73	0.75	0.24	0.16	0.06	0.03	0.81	–	0.28	–	0.08	–
Saudi Arabia	◊	◊	◊	◊	◊	◊	0.72	0.70	0.15	0.21	0.02	0.05
Scotland	0.85	0.86	0.06	0.11	0.00	0.01	0.86	0.87	0.18	0.33	0.03	0.11
Serbia	◊	◊	◊	◊	◊	◊	0.86	–	0.31	–	0.10	–
Singapore	0.87	0.84	0.21	0.22	0.04	0.05	0.88	0.86	0.33	0.30	0.11	0.09
Slovak Republic	0.80	0.78	0.17	0.16	0.03	0.03	◊	◊	◊	◊	◊	◊
Slovenia	0.83	0.83	0.17	0.14	0.03	0.02	0.84	–	0.25	–	0.06	–
Sweden	0.88	0.87	0.05	0.13	0.00	0.02	0.88	–	0.34	–	0.11	–
Syrian Arab Republic	◊	◊	◊	◊	◊	◊	0.71	–	0.28	–	0.08	–
Thailand	◊	◊	◊	◊	◊	◊	0.73	0.69	0.25	0.18	0.06	0.03
Tunisia	0.37	0.41	0.44	0.44	0.20	0.20	0.76	0.71	0.30	0.18	0.09	0.03
Turkey	◊	◊	◊	◊	◊	◊	0.76	0.74	0.31	0.18	0.10	0.03
Ukraine	0.75	0.77	0.24	0.19	0.06	0.03	0.82	–	0.23	–	0.05	–
United States	0.85	0.85	0.14	0.14	0.02	0.02	0.86	0.86	0.24	0.20	0.06	0.04
Yemen	0.27	0.37	0.18	0.28	0.03	0.08	◊	◊	◊	◊	◊	◊
International Median	0.82	0.81	0.18	0.16	0.03	0.03	0.81	0.78	0.28	0.24	0.08	0.06
Benchmarking Participants												
Alberta, Canada	0.85	0.85	0.18	0.11	0.03	0.01	◊	◊	◊	◊	◊	◊
Basque Country, Spain	◊	◊	◊	◊	◊	◊	0.86	0.85	0.34	0.28	0.12	0.08
British Columbia, Canada	0.85	0.87	0.16	0.10	0.03	0.01	0.85	0.88	0.31	0.25	0.10	0.06
Dubai, UAE	0.75	0.76	0.16	0.27	0.03	0.07	0.84	0.82	0.22	0.22	0.05	0.05
Massachusetts, US	0.87	0.87	0.18	0.12	0.03	0.01	0.87	0.87	0.26	0.20	0.07	0.04
Minnesota, US	0.85	0.85	0.16	0.10	0.02	0.01	0.86	0.87	0.30	0.27	0.09	0.07
Ontario, Canada	0.87	0.87	0.17	0.13	0.03	0.02	0.86	0.87	0.32	0.29	0.10	0.08
Quebec, Canada	0.85	0.86	0.20	0.16	0.04	0.03	0.88	0.90	0.27	0.18	0.07	0.03

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

Exhibit 12.1 Index of Students' Positive Affect Toward Mathematics (PATM) / Science (PATS)—Reliability and Validity Indicators (Continued)

Countries	Grade 8 Separate Science											
	Cronbach's Alpha Between the Component Variables				Multiple R Between Student Achievement and Component Variables				Percent of Variance in Student Achievement Accounted for by the Component Variables			
	Biology	Earth Science	Chemistry	Physics	Biology	Earth Science	Chemistry	Physics	Biology	Earth Science	Chemistry	Physics
Algeria	0.64	0.61	0.64	0.66	0.17	0.13	0.15	0.14	0.03	0.02	0.02	0.02
Armenia	0.70	0.67	0.68	0.68	0.08	0.07	0.07	0.12	0.01	0.01	0.01	0.01
Bosnia and Herzegovina	0.84	0.83	0.83	0.80	0.05	0.04	0.05	0.08	0.00	0.00	0.00	0.01
Bulgaria	0.78	0.75	0.78	0.74	0.19	0.19	0.18	0.17	0.04	0.04	0.03	0.03
Cyprus	0.19	0.82	0.81	0.79	0.26	0.11	0.19	0.25	0.07	0.01	0.04	0.06
Czech Republic	0.85	0.85	0.86	0.84	0.06	0.08	0.05	0.12	0.00	0.01	0.00	0.02
Georgia	0.73	0.64	0.71	0.71	0.19	0.13	0.14	0.18	0.04	0.02	0.02	0.03
Hungary	0.87	0.87	0.85	0.83	0.03	0.07	0.06	0.15	0.00	0.00	0.00	0.02
Indonesia	0.62	–	–	0.65	0.22	–	–	0.24	0.05	–	–	0.06
Lebanon	0.68	–	0.67	0.67	0.27	–	0.19	0.16	0.08	–	0.04	0.03
Lithuania	0.81	0.79	0.78	0.76	0.06	0.08	0.12	0.14	0.00	0.01	0.01	0.02
Malta	0.88	0.88	0.88	0.88	0.34	0.16	0.32	0.21	0.12	0.03	0.10	0.04
Morocco	0.59	0.62	0.64	0.64	0.27	0.19	0.28	0.27	0.07	0.04	0.08	0.07
Romania	0.80	0.81	0.79	0.75	0.09	0.14	0.09	0.07	0.01	0.02	0.01	0.00
Russian Federation	0.82	0.81	0.81	0.79	0.03	0.09	0.10	0.15	0.00	0.01	0.01	0.02
Serbia	0.85	0.85	0.85	0.81	0.08	0.04	0.09	0.08	0.01	0.00	0.01	0.01
Slovenia	0.87	0.87	0.87	0.83	0.09	0.12	0.21	0.15	0.01	0.01	0.05	0.02
Sweden	0.88	0.87	0.88	0.86	0.17	0.12	0.22	0.22	0.03	0.02	0.05	0.05
Syrian Arab Republic	0.62	0.63	0.67	0.65	0.23	0.19	0.17	0.19	0.05	0.03	0.03	0.04
Ukraine	0.80	0.79	0.82	0.79	0.10	0.09	0.12	0.12	0.01	0.01	0.01	0.01
International Median	0.80	0.81	0.81	0.78	0.14	0.12	0.14	0.15	0.02	0.01	0.02	0.02

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

A diamond (◊) indicates the country did not participate in the assessment.

A dash (–) indicates comparable data are not available or country administered separate science version of the student questionnaire.

The Index of Students' Self-Confidence in Learning Mathematics (SCM) and the Index of Students' Self-Confidence in Learning Science (SCS) examined how students think about their abilities in mathematics and science. The index, first developed in 2003, is presented with trends in Exhibit 4.10 of the TIMSS 2007 international reports. In addition, Exhibit 4.11 reports the percentage of students at each index level by gender.

For mathematics, the index is based on students' responses to four statements about mathematics: 1) I usually do well in mathematics; 2) I learn things quickly in mathematics; 3) Mathematics is more difficult for me than for many of my classmates (eighth grade version) and mathematics is harder for me than for many of my classmates (fourth grade version); and 4) Mathematics is not one of my strengths (eighth grade version) and I'm just not good at mathematics (fourth grade version). For science the index is based on students' responses to four statements about science: 1) I usually do well in science; 2) I learn things quickly in science; 3) Science is more difficult for me than for many of my classmates (eighth grade version) and science is harder for me than for many of my classmates (fourth grade version); and 4) Science is not one of my strengths (eighth grade version) and I'm just not good at science (fourth grade version). The two negatively worded statements were reverse coded.

An average was computed across the four items based on a 4-point scale: *agree a lot* = 1, *agree a little* = 2, *disagree a little* = 3, and *disagree a lot* = 4. A high level indicates an average score of less than or equal to 2, corresponding to students agreeing a little or a lot, on average. A low level indicates an average score equal to or greater than 3, corresponding to students disagreeing a little or a lot, on average. A medium level indicates an average score of greater than 2 but less than 3. For countries that taught biology, earth science, chemistry, and physics as separate subjects at the eighth grade, the questions were asked about each individual science subject, and students responded with respect to each science course they were taking. Thus, separate indices were created for each science subject, and the reliability and validity indicators for separate science countries are presented in a separate panel for eighth grade in Exhibit 12.2.

As shown in Exhibit 12.2, the four component variables (statements) form a fairly reliable scale, with median reliability coefficients (Cronbach's alpha) across countries of 0.72 for both mathematics and science at the fourth grade and 0.73 and 0.66, respectively, at the eighth grade. At the fourth grade, the median multiple correlation between the four component variables and student achievement was 0.43 for mathematics and 0.31 for science, corresponding to R-squares of 0.18 and 0.10, respectively. At the eighth grade, the median multiple correlation between the four component variables and student achievement was 0.46 for mathematics and 0.37 for general science, corresponding to R-squares of 0.21 and 0.14, respectively. For the separate sciences, the reliabilities were similar, but the correlations were somewhat lower than for general science, with the median multiple correlations ranging between 0.28 and 0.33, corresponding to R-squares between 0.08 and 0.11.

The Index of Students' Valuing Mathematics (SVM) and the Index of Students' Valuing Science (SVS) summarize eighth grade students' reports of their motivation to learn and their perception of mathematics and science as advantageous for their future lives. There was not a comparable index at fourth grade. The index, modified from the 2003 index, is presented in Exhibit 4.9 of the TIMSS 2007 international reports, including trends from 2003.

For mathematics, the index is based on eighth grade students' responses to four statements about mathematics: 1) I think learning mathematics will help me in my daily life; 2) I need mathematics to learn other school subjects; 3) I need to do well in mathematics to get into the university of my choice; and 4) I would like to do well in mathematics to get the job I want. For science the index is based on students' responses to four similar statements about science: 1) I think learning science will help me in my daily life; 2) I need science to learn other school subjects; 3) I need to do well in science to get into the university of my choice; and 4) I would like to do well in science to get the job I want. An average was computed across the four items based on a 4-point scale: *agree a lot* = 1, *agree a little* = 2, *disagree a little* = 3, and *disagree a lot* = 4. A high level indicates an average score of less than or equal to 2, corresponding to students agreeing with the statements a little or a lot, on average. A low level indicates an average score equal to or greater than 3, corresponding to students disagreeing a little or a lot, on average. A medium level indicates an average score of greater than 2 but less than 3. For countries that taught biology, earth science, chemistry, and physics as separate subjects at the eighth grade, the questions were asked about each individual science

Exhibit 12.2 Index of Students' Self-Confidence in Learning Mathematics (SCM) / Science (SCS)—Reliability and Validity Indicators

Countries	Grade 4						Grade 8					
	Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables		Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science
Algeria	0.36	0.41	0.28	0.27	0.08	0.08	0.54	–	0.44	–	0.20	–
Armenia	0.60	0.61	0.17	0.17	0.03	0.03	0.66	–	0.21	–	0.04	–
Australia	0.75	0.74	0.46	0.27	0.22	0.07	0.81	0.81	0.55	0.37	0.30	0.14
Austria	0.78	0.75	0.47	0.34	0.22	0.12	◊	◊	◊	◊	◊	◊
Bahrain	◊	◊	◊	◊	◊	◊	0.67	0.58	0.51	0.46	0.26	0.21
Bosnia and Herzegovina	◊	◊	◊	◊	◊	◊	0.78	–	0.51	–	0.27	–
Botswana	◊	◊	◊	◊	◊	◊	0.46	0.43	0.29	0.36	0.09	0.13
Bulgaria	◊	◊	◊	◊	◊	◊	0.70	–	0.42	–	0.18	–
Chinese Taipei	0.73	0.73	0.47	0.26	0.22	0.07	0.84	0.81	0.55	0.41	0.31	0.17
Colombia	0.43	0.46	0.36	0.35	0.13	0.12	0.68	0.63	0.37	0.30	0.14	0.09
Cyprus	◊	◊	◊	◊	◊	◊	0.79	–	0.52	–	0.28	–
Czech Republic	0.75	0.77	0.43	0.31	0.18	0.10	0.85	–	0.53	–	0.28	–
Denmark	0.78	0.76	0.43	0.26	0.18	0.07	◊	◊	◊	◊	◊	◊
Egypt	◊	◊	◊	◊	◊	◊	0.46	0.53	0.35	0.40	0.12	0.16
El Salvador	0.35	0.33	0.33	0.36	0.11	0.13	0.57	0.37	0.36	0.34	0.13	0.12
England	0.75	0.79	0.44	0.29	0.20	0.08	0.79	0.84	0.46	0.37	0.21	0.14
Georgia	0.51	0.56	0.34	0.25	0.12	0.06	0.66	–	0.38	–	0.14	–
Germany	0.81	0.76	0.49	0.35	0.24	0.13	◊	◊	◊	◊	◊	◊
Ghana	◊	◊	◊	◊	◊	◊	0.51	0.52	0.33	0.35	0.11	0.12
Hong Kong SAR	0.72	0.68	0.40	0.29	0.16	0.09	0.80	0.75	0.38	0.26	0.15	0.07
Hungary	0.78	0.79	0.51	0.39	0.26	0.15	0.84	–	0.56	–	0.31	–
Indonesia	◊	◊	◊	◊	◊	◊	0.43	–	0.30	–	0.09	–
Iran, Islamic Rep. of	0.73	0.78	0.48	0.44	0.23	0.19	0.74	0.73	0.46	0.35	0.21	0.13
Israel	◊	◊	◊	◊	◊	◊	0.73	0.74	0.41	0.44	0.17	0.20
Italy	0.69	0.68	0.35	0.24	0.12	0.06	0.84	0.81	0.48	0.31	0.23	0.10
Japan	0.76	0.75	0.47	0.28	0.22	0.08	0.78	0.79	0.50	0.40	0.25	0.16
Jordan	◊	◊	◊	◊	◊	◊	0.65	0.62	0.52	0.42	0.27	0.18
Kazakhstan	0.79	0.79	0.28	0.21	0.08	0.04	◊	◊	◊	◊	◊	◊
Korea, Rep. of	◊	◊	◊	◊	◊	◊	0.86	0.86	0.64	0.48	0.40	0.23
Kuwait	0.35	0.42	0.38	0.39	0.14	0.15	0.59	0.53	0.43	0.34	0.18	0.11
Latvia	0.72	0.71	0.50	0.32	0.25	0.10	◊	◊	◊	◊	◊	◊
Lebanon	◊	◊	◊	◊	◊	◊	0.65	–	0.46	–	0.21	–
Lithuania	0.71	0.70	0.54	0.34	0.29	0.12	0.79	–	0.58	–	0.33	–
Malaysia	◊	◊	◊	◊	◊	◊	0.64	0.66	0.40	0.28	0.16	0.08
Malta	◊	◊	◊	◊	◊	◊	0.78	–	0.47	–	0.22	–
Morocco	0.44	0.42	0.28	0.28	0.08	0.08	0.63	–	0.45	–	0.20	–
Netherlands	0.82	0.78	0.43	0.29	0.18	0.08	◊	◊	◊	◊	◊	◊
New Zealand	0.69	0.68	0.48	0.35	0.23	0.12	◊	◊	◊	◊	◊	◊
Norway	0.68	0.72	0.39	0.32	0.15	0.10	0.80	0.79	0.61	0.37	0.38	0.13
Oman	◊	◊	◊	◊	◊	◊	0.49	0.49	0.46	0.43	0.21	0.19
Palestinian Nat'l Auth.	◊	◊	◊	◊	◊	◊	0.54	0.53	0.46	0.45	0.21	0.20
Qatar	0.41	0.47	0.36	0.36	0.13	0.13	0.61	0.53	0.40	0.34	0.16	0.12
Romania	◊	◊	◊	◊	◊	◊	0.63	–	0.46	–	0.21	–
Russian Federation	0.74	0.75	0.38	0.28	0.14	0.08	0.84	–	0.52	–	0.27	–
Saudi Arabia	◊	◊	◊	◊	◊	◊	0.49	0.48	0.44	0.44	0.19	0.19
Scotland	0.72	0.74	0.32	0.25	0.10	0.06	0.77	0.83	0.45	0.48	0.20	0.23
Serbia	◊	◊	◊	◊	◊	◊	0.82	–	0.64	–	0.41	–
Singapore	0.76	0.75	0.50	0.31	0.25	0.10	0.82	0.82	0.45	0.26	0.20	0.07
Slovak Republic	0.73	0.73	0.45	0.36	0.21	0.13	◊	◊	◊	◊	◊	◊
Slovenia	0.66	0.65	0.50	0.32	0.25	0.10	0.76	–	0.54	–	0.30	–
Sweden	0.72	0.73	0.38	0.29	0.15	0.09	0.82	–	0.58	–	0.33	–
Syrian Arab Republic	◊	◊	◊	◊	◊	◊	0.57	–	0.42	–	0.17	–
Thailand	◊	◊	◊	◊	◊	◊	0.58	0.61	0.28	0.24	0.08	0.06
Tunisia	0.45	0.49	0.47	0.43	0.22	0.19	0.73	0.62	0.48	0.39	0.23	0.16
Turkey	◊	◊	◊	◊	◊	◊	0.76	0.71	0.50	0.38	0.25	0.14
Ukraine	0.69	0.68	0.46	0.33	0.21	0.11	0.79	–	0.53	–	0.28	–
United States	0.76	0.78	0.46	0.34	0.21	0.12	0.84	0.82	0.46	0.34	0.21	0.12
Yemen	0.09	0.31	0.22	0.23	0.05	0.05	◊	◊	◊	◊	◊	◊
International Median	0.72	0.72	0.43	0.31	0.18	0.10	0.73	0.66	0.46	0.37	0.21	0.14
Benchmarking Participants												
Alberta, Canada	0.77	0.77	0.46	0.33	0.21	0.11	◊	◊	◊	◊	◊	◊
Basque Country, Spain	◊	◊	◊	◊	◊	◊	0.80	0.75	0.56	0.45	0.31	0.21
British Columbia, Canada	0.77	0.76	0.46	0.32	0.21	0.10	0.86	0.84	0.56	0.40	0.32	0.16
Dubai, UAE	0.62	0.64	0.38	0.36	0.14	0.13	0.69	0.68	0.46	0.38	0.21	0.14
Massachusetts, US	0.78	0.80	0.46	0.29	0.21	0.09	0.84	0.85	0.53	0.40	0.28	0.16
Minnesota, US	0.76	0.77	0.51	0.30	0.26	0.09	0.85	0.85	0.56	0.46	0.31	0.21
Ontario, Canada	0.76	0.78	0.48	0.31	0.23	0.10	0.87	0.85	0.61	0.45	0.37	0.20
Quebec, Canada	0.78	0.77	0.55	0.30	0.30	0.09	0.87	0.85	0.56	0.32	0.31	0.10

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

Exhibit 12.2 Index of Students' Self-Confidence in Learning Mathematics (SCM) / Science (SCS)—Reliability and Validity Indicators (Continued)

Countries	Grade 8 Separate Science											
	Cronbach's Alpha Between the Component Variables				Multiple R Between Student Achievement and Component Variables				Percent of Variance in Student Achievement Accounted for by the Component Variables			
	Biology	Earth Science	Chemistry	Physics	Biology	Earth Science	Chemistry	Physics	Biology	Earth Science	Chemistry	Physics
Algeria	0.56	0.33	0.48	0.50	0.27	0.22	0.21	0.24	0.07	0.05	0.04	0.06
Armenia	0.53	0.55	0.48	0.56	0.14	0.15	0.08	0.18	0.02	0.02	0.01	0.03
Bosnia and Herzegovina	0.72	0.70	0.73	0.70	0.30	0.28	0.27	0.29	0.09	0.08	0.07	0.09
Bulgaria	0.66	0.66	0.65	0.62	0.26	0.29	0.28	0.25	0.07	0.08	0.08	0.06
Cyprus	–0.46	0.73	0.75	0.74	0.30	0.37	0.37	0.43	0.09	0.13	0.14	0.18
Czech Republic	0.81	0.83	0.85	0.84	0.22	0.20	0.21	0.28	0.05	0.04	0.04	0.08
Georgia	0.66	0.49	0.60	0.61	0.37	0.34	0.29	0.30	0.14	0.12	0.09	0.09
Hungary	0.81	0.83	0.82	0.83	0.26	0.27	0.24	0.35	0.07	0.08	0.06	0.12
Indonesia	0.43	–	–	0.42	0.34	–	–	0.34	0.12	–	–	0.12
Lebanon	0.56	–	0.57	0.54	0.42	–	0.35	0.35	0.17	–	0.12	0.12
Lithuania	0.75	0.75	0.79	0.77	0.28	0.29	0.25	0.29	0.08	0.09	0.06	0.09
Malta	0.79	0.75	0.80	0.77	0.45	0.33	0.37	0.32	0.20	0.11	0.14	0.10
Morocco	0.55	0.46	0.52	0.52	0.35	0.31	0.35	0.34	0.12	0.10	0.12	0.12
Romania	0.56	0.65	0.58	0.53	0.30	0.29	0.21	0.16	0.09	0.09	0.04	0.02
Russian Federation	0.81	0.80	0.83	0.81	0.30	0.35	0.28	0.37	0.09	0.12	0.08	0.14
Serbia	0.72	0.72	0.76	0.72	0.33	0.32	0.30	0.34	0.11	0.11	0.09	0.11
Slovenia	0.79	0.78	0.80	0.77	0.35	0.34	0.41	0.36	0.12	0.11	0.16	0.13
Sweden	0.78	0.79	0.79	0.78	0.34	0.25	0.36	0.38	0.12	0.06	0.13	0.14
Syrian Arab Republic	0.50	0.43	0.49	0.46	0.37	0.32	0.24	0.28	0.14	0.10	0.06	0.08
Ukraine	0.77	0.77	0.79	0.77	0.35	0.35	0.28	0.35	0.12	0.12	0.08	0.12
International Median	0.69	0.73	0.75	0.71	0.32	0.30	0.28	0.33	0.10	0.09	0.08	0.11

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

A diamond (◊) indicates the country did not participate in the assessment.

A dash (–) indicates comparable data are not available or country administered separate science version of the student questionnaire.

subject, and students responded with respect to each science course they were taking. Thus, separate indices were created for each science subject, and the reliability and validity indicators for separate science countries are presented in a separate panel for eighth grade in Exhibit 12.3.

A similar index of students' valuing mathematics and science was presented in the TIMSS 2003 international reports (Martin, M.O., Mullis I.V.S., Gonzales, E.J., & Chrostowski, S.J., 2004; Mullis I.V.S., Martin, M.O., Gonzales, E.J., & Chrostowski, S.J., 2004) that included three more variables for both subjects. "I would like to take more mathematics in school", "I enjoy learning mathematics", and "I would like a job that involved using mathematics" were included in the TIMSS 2003 index calculations for mathematics but not in the TIMSS 2007 index calculations. The percentage of students at each index level in 2003 was recomputed based on the TIMSS 2007 index definition.

As shown in Exhibit 12.3, the four components form a fairly reliable scale, with a median reliability coefficient (Cronbach's alpha) of 0.70 for mathematics and 0.78 for general science. For the separate sciences, reliabilities ranged from 0.76 to 0.83. The median multiple correlation between the four statements and student achievement was 0.19 for mathematics and 0.21 for general science, corresponding to an R-square of 0.04, after rounding. For the separate sciences, the median multiple correlations ranged from 0.15 to 0.20, corresponding to R-squares of 0.02 to 0.04.

Exhibit 12.3 Index of Students' Valuing Mathematics (SVM) / Science (SVS)—Reliability and Validity Indicators

Countries	Grade 8					
	Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
	Mathematics	Science	Mathematics	Science	Mathematics	Science
Algeria	0.64	–	0.18	–	0.03	–
Armenia	0.66	–	0.07	–	0.00	–
Australia	0.79	0.88	0.15	0.23	0.02	0.05
Bahrain	0.73	0.78	0.16	0.15	0.02	0.02
Bosnia and Herzegovina	0.70	–	0.09	–	0.01	–
Botswana	0.58	0.64	0.31	0.39	0.10	0.15
Bulgaria	0.73	–	0.16	–	0.02	–
Chinese Taipei	0.76	0.83	0.35	0.40	0.12	0.16
Colombia	0.66	0.76	0.04	0.14	0.00	0.02
Cyprus	0.72	–	0.21	–	0.04	–
Czech Republic	0.66	–	0.13	–	0.02	–
Egypt	0.58	0.64	0.21	0.20	0.04	0.04
El Salvador	0.64	0.76	0.12	0.18	0.02	0.03
England	0.72	0.83	0.12	0.20	0.01	0.04
Georgia	0.60	–	0.14	–	0.02	–
Ghana	0.63	0.69	0.24	0.27	0.06	0.07
Hong Kong SAR	0.82	0.84	0.28	0.32	0.08	0.10
Hungary	0.64	–	0.22	–	0.05	–
Indonesia	0.63	–	0.07	–	0.00	–
Iran, Islamic Rep. of	0.65	0.73	0.18	0.14	0.03	0.02
Israel	0.73	0.85	0.25	0.19	0.06	0.04
Italy	0.68	0.76	0.17	0.17	0.03	0.03
Japan	0.70	0.79	0.23	0.32	0.05	0.10
Jordan	0.70	0.74	0.21	0.21	0.04	0.04
Korea, Rep. of	0.74	0.80	0.33	0.34	0.11	0.12
Kuwait	0.80	0.83	0.20	0.16	0.04	0.03
Lebanon	0.68	–	0.20	–	0.04	–
Lithuania	0.72	–	0.20	–	0.04	–
Malaysia	0.75	0.80	0.26	0.39	0.07	0.15
Malta	0.69	–	0.26	–	0.07	–
Morocco	0.62	–	0.22	–	0.05	–
Norway	0.77	0.84	0.17	0.15	0.03	0.02
Oman	0.69	0.69	0.29	0.23	0.08	0.05
Palestinian Nat'l Auth.	0.73	0.74	0.28	0.27	0.08	0.07
Qatar	0.82	0.85	0.19	0.14	0.04	0.02
Romania	0.72	–	0.12	–	0.02	–
Russian Federation	0.71	–	0.21	–	0.05	–
Saudi Arabia	0.69	0.74	0.14	0.13	0.02	0.02
Scotland	0.74	0.85	0.15	0.25	0.02	0.06
Serbia	0.70	–	0.10	–	0.01	–
Singapore	0.76	0.83	0.22	0.38	0.05	0.14
Slovenia	0.69	–	0.18	–	0.03	–
Sweden	0.74	–	0.17	–	0.03	–
Syrian Arab Republic	0.65	–	0.17	–	0.03	–
Thailand	0.69	0.75	0.22	0.21	0.05	0.05
Tunisia	0.67	0.72	0.19	0.08	0.04	0.01
Turkey	0.60	0.72	0.19	0.17	0.04	0.03
Ukraine	0.70	–	0.15	–	0.02	–
United States	0.73	0.82	0.18	0.22	0.03	0.05
International Median	0.70	0.78	0.19	0.21	0.04	0.04
Benchmarking Participants						
Basque Country, Spain	0.75	0.85	0.25	0.17	0.06	0.03
British Columbia, Canada	0.75	0.83	0.23	0.26	0.05	0.07
Dubai, UAE	0.70	0.80	0.17	0.16	0.03	0.02
Massachusetts, US	0.73	0.82	0.21	0.21	0.04	0.05
Minnesota, US	0.73	0.84	0.24	0.22	0.06	0.05
Ontario, Canada	0.72	0.83	0.20	0.28	0.04	0.08
Quebec, Canada	0.71	0.82	0.15	0.25	0.02	0.06

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

A dash (–) indicates comparable data are not available or country administered separate science version of the student questionnaire.

Exhibit 12.3 Index of Students' Valuing Mathematics (SVM) / Science (SVS)—Reliability and Validity Indicators (Continued)

Countries	Grade 8 Separate Science											
	Cronbach's Alpha Between the Component Variables				Multiple R Between Student Achievement and Component Variables				Percent of Variance in Student Achievement Accounted for by the Component Variables			
	Biology	Earth Science	Chemistry	Physics	Biology	Earth Science	Chemistry	Physics	Biology	Earth Science	Chemistry	Physics
Algeria	0.67	0.78	0.78	0.76	0.10	0.09	0.05	0.03	0.01	0.01	0.00	0.00
Armenia	0.69	0.72	0.79	0.78	0.06	0.06	0.06	0.10	0.00	0.00	0.00	0.01
Bosnia and Herzegovina	0.76	0.81	0.85	0.86	0.19	0.24	0.19	0.15	0.04	0.06	0.03	0.02
Bulgaria	0.76	0.81	0.83	0.84	0.20	0.16	0.15	0.10	0.04	0.03	0.02	0.01
Cyprus	0.88	0.81	0.84	0.86	0.20	0.23	0.13	0.18	0.04	0.05	0.02	0.03
Czech Republic	0.77	0.78	0.81	0.82	0.08	0.18	0.15	0.18	0.01	0.03	0.02	0.03
Georgia	0.75	0.80	0.83	0.82	0.22	0.23	0.20	0.19	0.05	0.05	0.04	0.04
Hungary	0.78	0.79	0.83	0.83	0.17	0.20	0.19	0.13	0.03	0.04	0.03	0.02
Indonesia	0.68	–	–	0.79	0.17	–	–	0.07	0.03	–	–	0.01
Lebanon	0.74	–	0.79	0.81	0.20	–	0.11	0.10	0.04	–	0.01	0.01
Lithuania	0.82	0.83	0.85	0.86	0.17	0.23	0.16	0.15	0.03	0.05	0.02	0.02
Malta	0.80	0.81	0.84	0.83	0.27	0.20	0.33	0.25	0.07	0.04	0.11	0.06
Morocco	0.73	0.81	0.80	0.81	0.11	0.17	0.11	0.09	0.01	0.03	0.01	0.01
Romania	0.78	0.80	0.84	0.86	0.28	0.31	0.23	0.26	0.08	0.09	0.05	0.07
Russian Federation	0.81	0.81	0.84	0.83	0.19	0.17	0.13	0.14	0.03	0.03	0.02	0.02
Serbia	0.76	0.80	0.85	0.86	0.16	0.24	0.16	0.15	0.03	0.06	0.02	0.02
Slovenia	0.78	0.80	0.84	0.86	0.13	0.20	0.18	0.24	0.02	0.04	0.03	0.06
Sweden	0.82	0.82	0.87	0.87	0.22	0.22	0.19	0.22	0.05	0.05	0.04	0.05
Syrian Arab Republic	0.67	0.75	0.76	0.77	0.11	0.08	0.08	0.08	0.01	0.01	0.01	0.01
Ukraine	0.78	0.81	0.83	0.84	0.21	0.22	0.17	0.15	0.04	0.05	0.03	0.02
International Median	0.76	0.80	0.83	0.83	0.18	0.20	0.16	0.15	0.03	0.04	0.02	0.02

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

Exhibit 12.4 presents latent factor models for the mathematics and science attitudinal indices at fourth and eighth grades. At each grade level, the mathematics model is presented graphically, while the corresponding models for the sciences are presented in tabular form to conserve space. The latent factors corresponding to the TIMSS 2007 indices are represented graphically by large darkened ovals, with correlations between the latent constructs represented by curved double-headed arrows. The fourth grade section of Exhibit 12.4 has two latent factors: *Positive Affect Toward Mathematics* and *Self-Confidence in Learning Mathematics*, and the estimated correlation between them is 0.662. Each latent factor is shown with arrows pointing to its observed component variables. For example, *Positive Affect Toward Mathematics* has three observed component variables, "I enjoy learning mathematics", "Mathematics is boring (reversed)", and "I like mathematics." The figure next to each arrow is the estimated factor loading, or the correlation between the latent factor and the component variable. The greater the loading, the stronger is the relationship between the observed variable and the latent factor. The loadings of the three component variables of *Positive Affect Toward Mathematics* were 0.864, 0.664, and 0.943, respectively. Also shown in the small ovals on the right hand side are the standardized residuals corresponding to each observed variable. The residuals are a function of the factor loadings; the greater the loading, the smaller the residual.

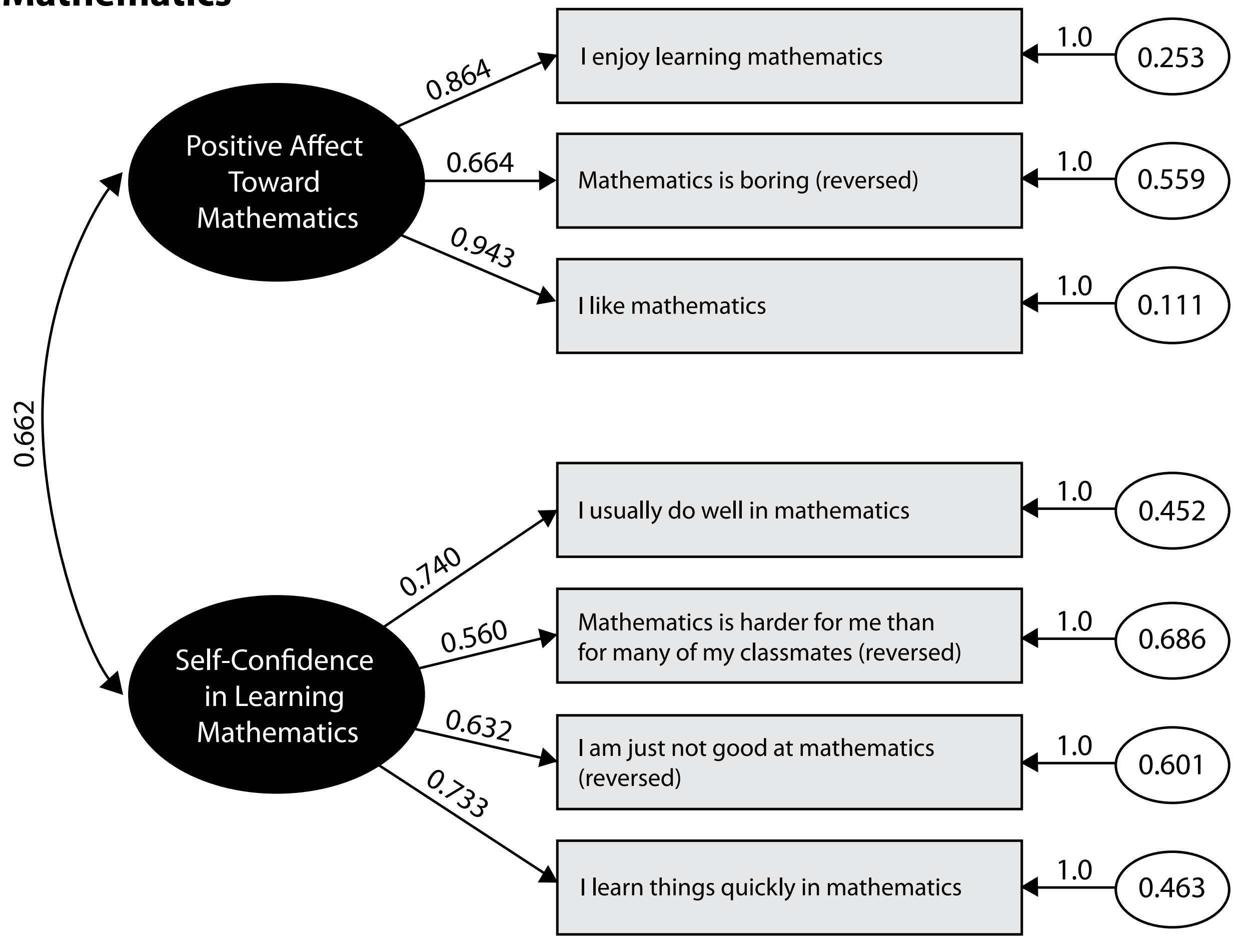
The confirmatory factor analyses reported in this chapter provide two commonly-used indicators of how well the factor models account for the TIMSS data: the Chi-square and the Root Mean Square Error Approximation (RMSEA). The Chi-square is not very useful for large sample-studies such as TIMSS, as it is sensitive to large sample size. However, the Root Mean Square Error Approximation is a more informative criterion, with values up to 0.10 indicating reasonable fit (Byrne, 2001).

Exhibit 12.4 also shows the measurement model for the eighth grade attitudinal indices. There were three latent factors for mathematics at the eighth grade: *Positive Affect Toward Mathematics, Self-Confidence in Learning Mathematics,* and *Valuing Mathematics,* based on 11 observed component variables. There were three corresponding factors in science, although these were fitted separately for countries teaching general science as well as individually for each science for countries teaching the sciences as separate subjects. In all instances, the correlations between these latent factors were strongly positive. For example, the correlation between *Positive Affect Toward Mathematics* and *Self-Confidence in Learning Mathematics* was 0.724; the correlation between *Positive Affect Toward Mathematics* and *Valuing Mathematics* was 0.589; and the correlation between *Self-Confidence in Learning Mathematics* and *Valuing Mathematics* was 0.421. Correlations among the latent factors for science were of similar magnitude.

The RMSEA value indicated quite good model fit for mathematics and general science (0.087 and 0.049, respectively) at the eighth grade, but somewhat less good at fourth grade and for the separate science subjects.

Exhibit 12.4 Latent Variable Model of Students' Attitudes Toward Mathematics/Science, Grade 4

Mathematics



Chi-square = 28161.726 ; Df = 9 ; RMSEA = 0.141

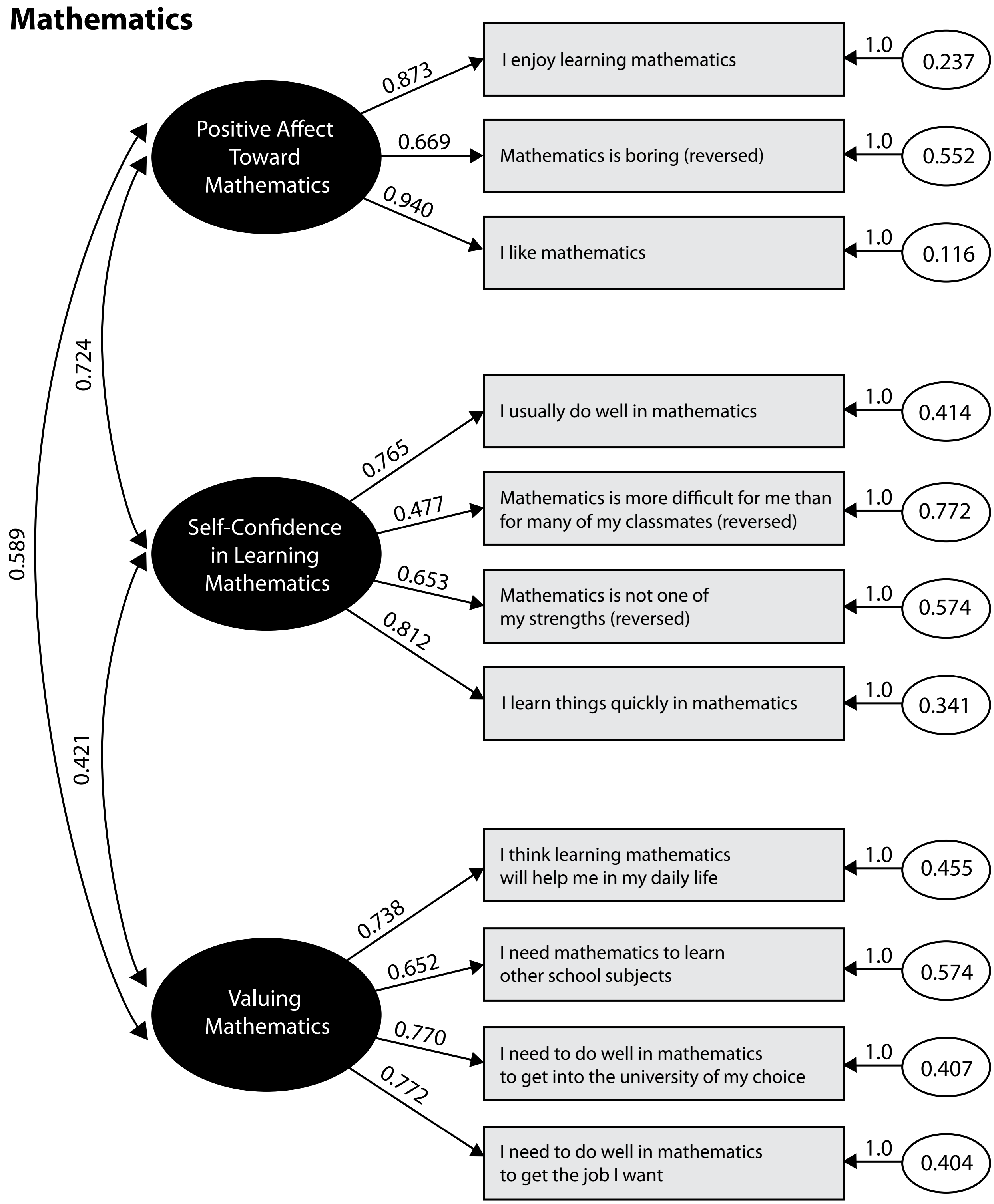
Science

Factors: Positive Affect Toward Science, Self-confidence in Learning Science

	Positive Affect Toward Science	Self-Confidence in Learning Science
Observed Variable	*Factor Loadings*	
I enjoy learning science	0.883	—
Science is boring (reversed)	0.685	—
I like science	0.921	—
I usually do well in science	—	0.719
Science is harder for me than for many of my classmates (reversed)	—	0.580
I am just not good at science (reversed)	—	0.640
I learn things quickly in science	—	0.766
	Positive Affect Toward Science	**Self-Confidence in Learning Science**
Factors	*Factor Intercorrelations*	
Positive Affect Toward Science	1.0	0.776
Self-Confidence in Learning Science	0.776	1.0

Chi-square= 37051.229; Df= 8; RMSEA= 0.172

Exhibit 12.4 Latent Variable Model of Students' Attitudes Toward Mathematics/Science, Grade 8 (Continued)



Chi-square = 45116.031 ; Df = 27 ; RMSEA = 0.087

Exhibit 12.4 Latent Variable Model of Students' Attitudes Toward Mathematics/Science, Grade 8 (Continued)

General Science

Factors: Positive Affect Toward Science, Self-Confidence in Learning Science, Valuing Science

	Positive Affect Toward Science	Self-Confidence in Learning Science	Valuing Science
Observed Variable	*Factor Loadings*		
I enjoy learning science	0.826	—	—
Science is boring (reversed)	0.663	—	—
I like science	0.875	—	—
I usually do well in science	—	0.692	—
Science is more difficult for me than for many of my classmates (reversed)	—	0.515	—
Science is not one of my strengths (reversed)	—	0.663	—
I learn things quickly in science	—	0.757	—
I think learning science will help me in my daily life	—	—	0.705
I need science to learn other school subjects	—	—	0.660
I need to do well in science to get into the university of my choice	—	—	0.787
I need to do well in science to get the job I want	—	—	0.762

	Positive Affect Toward Science	Self-Confidence in Learning Science	Valuing Science
Factors	*Factor Intercorrelations*		
Positive Affect Toward Science	1.0	0.883	0.625
Self-Confidence in Learning Science	0.883	1.0	0.497
Valuing Science	0.625	0.497	1.0

Chi-square= 7149.560; Df= 26; RMSEA= 0.049

Biology

Factors: Positive Affect Toward Biology, Self-Confidence in Learning Biology, Valuing Biology

	Positive Affect Toward Biology	Self-Confidence in Learning Biology	Valuing Biology
Observed Variable	*Factor Loadings*		
I enjoy learning biology	0.878	—	—
Biology is boring (reversed)	0.659	—	—
I like biology	0.926	—	—
I usually do well in biology	—	0.755	—
Biology is more difficult for me than for many of my classmates (reversed)	—	0.452	—
Biology is not one of my strengths (reversed)	—	0.535	—
I learn things quickly in biology	—	0.822	—
I think learning biology will help me in my daily life	—	—	0.700
I need biology to learn other school subjects	—	—	0.687
I need to do well in biology to get into the university of my choice	—	—	0.860
I need to do well in biology to get the job I want	—	—	0.862

	Positive Affect Toward Biology	Self-Confidence in Learning Biology	Valuing Biology
Factors	*Factor Intercorrelations*		
Positive Affect Toward Biology	1.0	0.742	0.622
Self-Confidence in Learning Biology	0.742	1.0	0.323
Valuing Biology	0.622	0.323	1.0

Chi-square= 30009.949; Df=23; RMSEA= 0.131

Exhibit 12.4 Latent Variable Model of Students' Attitudes Toward Mathematics/Science, Grade 8 (Continued)

Earth Science

Factors: Positive Affect Toward Earth Science, Self-Confidence in Learning Earth Science, Valuing Earth Science

	Positive Affect Toward Earth Science	Self-Confidence in Learning Earth Science	Valuing Earth Science
Observed Variable		*Factor Loadings*	
I enjoy learning earth science	0.889	—	—
Earth science is boring (reversed)	0.637	—	—
I like earth science	0.931	—	—
I usually do well in earth science	—	0.773	—
Earth science is more difficult for me than for many of my classmates (reversed)	—	0.474	—
Earth science is not one of my strengths (reversed)	—	0.529	—
I learn things quickly in earth science	—	0.864	—
I think learning earth science will help me in my daily life	—	—	0.705
I need earth science to learn other school subjects	—	—	0.751
I need to do well in earth science to get into the university of my choice	—	—	0.878
I need to do well in earth science to get the job I want	—	—	0.864
	Positive Affect Toward Earth Science	Self-Confidence in Learning Earth Science	Valuing Earth Science
Factors		*Factor Intercorrelations*	
Positive Affect Toward Earth Science	1.0	0.752	0.557
Self-Confidence in Learning Earth Science	0.752	1.0	0.265
Valuing Earth Science	0.557	0.265	1.0

Chi-square= 34479.811; Df= 18; RMSEA= 0.162

Chemistry

Factors: Positive Affect Toward Chemistry, Self-Confidence in Learning Chemistry, Valuing Chemistry

	Positive Affect Toward Chemistry	Self-Confidence in Learning Chemistry	Valuing Chemistry
Observed Variable		*Factor Loadings*	
I enjoy learning chemistry	0.918	—	—
Chemistry is boring (reversed)	0.595	—	—
I like chemistry	0.928	—	—
I usually do well in chemistry	—	0.828	—
Chemistry is more difficult for me than for many of my classmates (reversed)	—	0.405	—
Chemistry is not one of my strengths (reversed)	—	0.498	—
I learn things quickly in chemistry	—	0.874	—
I think learning chemistry will help me in my daily life	—	—	0.781
I need chemistry to learn other school subjects	—	—	0.775
I need to do well in chemistry to get into the university of my choice	—	—	0.890
I need to do well in chemistry to get the job I want	—	—	0.883
	Positive Affect Toward Chemistry	Self-Confidence in Learning Chemistry	Valuing Chemistry
Factors		*Factor Intercorrelations*	
Positive Affect Toward Chemistry	1.0	0.828	0.631
Self-Confidence in Learning Chemistry	0.828	1.0	0.445
Valuing Chemistry	0.631	0.445	1.0

Chi-square= 42363.636; Df= 18; RMSEA= 0.175

Exhibit 12.4 Latent Variable Model of Students' Attitudes Toward Mathematics/Science, Grade 8 (Continued)

Physics

Factors: Positive Affect Toward Physics, Self-Confidence in Learning Physics, Valuing Physics

	Positive Affect Toward Physics	Self-Confidence in Learning Physics	Valuing Physics
Observed Variable	*Factor Loadings*		
I enjoy learning physics	0.918	—	—
Physics is boring (reversed)	0.588	—	—
I like physics	0.933	—	—
I usually do well in physics	—	0.813	—
Physics is more difficult for me than for many of my classmates (reversed)	—	0.368	—
Physics is not one of my strengths (reversed)	—	0.459	—
I learn things quickly in physics	—	0.877	—
I think learning physics will help me in my daily life	—	—	0.807
I need physics to learn other school subjects	—	—	0.796
I need to do well in physics to get into the university of my choice	—	—	0.880
I need to do well in physics to get the job I want	—	—	0.884
	Positive Affect Toward Physics	**Self-Confidence in Learning Physics**	**Valuing Physics**
Factors	*Factor Intercorrelations*		
Positive Affect Toward Physics	1.0	0.834	0.631
Self-Confidence in Learning Physics	0.834	1.0	0.460
Valuing Physics	0.631	0.460	1.0

Chi-square= 51693.532; Df= 18 RMSEA= 0.187

In constructing the Index of Time Students Spend on Doing Mathematics Homework (TMH) and the Index of Time Students Spend on Doing Science Homework (TSH), students were categorized according to their responses to two questions on the frequency of homework they are given and the amount of time they spend on that homework. A high level indicates homework in mathematics or science assigned at least 3 or 4 times a week and students spend more than 30 minutes on that homework. A low level indicates homework in these subjects assigned no more than twice a week, and students spend no more than 30 minutes on that homework. A medium level indicates all other combinations of frequencies.

These TIMSS indices are unique for two reasons: they are comprised of only two variables, and the way the categories of the two variables are combined does not lend itself well to the Cronbach Alpha measure of reliability. Also, the categories for grouping students are sensitive to differences across countries in the role of homework in mathematics and science instruction. The index is presented in Exhibit 4.7 of the TIMSS 2007 international reports. Similar indices were reported in previous TIMSS cycles, but the questions and the index definition have been refined over time. Thus, no trends were reported for this index.

As shown in Exhibit 12.5, the variables comprising this index have relatively low reliability (international median Cronbach's alpha ranging between 0.05 and 0.14 for mathematics and general science) and only a weak relationship with achievement (international median multiple-R of 0.16 or less, corresponding to R-squares less than 0.02) as compared to other indices. These statistics suggest that while homework may be an important part of instruction in many countries, there is great variation across countries in how homework is used, and often students receiving the greatest amounts of homework or spending most time on it may not be the high performers.

The Index of Students' Perception of Being Safe in School (SPBSS) summarizes students' reports of how safe and secure they feel in their schools. The index, developed in 2003, is presented in Exhibit 8.14 of the *TIMSS 2007 International Mathematics Report* and Exhibit 8.15 of the *TIMSS 2007 International Science Report*. The index groups students according to their reports about the frequency of incidents affecting their safety: 1) Something of mine was stolen; 2) I was hit or hurt by other student(s) (for example, shoving, hitting, kicking); 3) I was made to do things that I didn't want to do by other students; 4) I was made fun of or called names; and 5) I was left out of activities by other students. Because this index had no components specific to particular branches of science, it was not necessary to construct separate indices for separate sciences countries. Students at the high level of the index reported that no such incidents occurred during the past month. Students at the low level reported three or more incidents during this period. Students at the medium level reported at least one but no more than two such incidents.

As shown in Exhibit 12.6 the five component variables form a fairly reliable scale, with a median reliability coefficient across countries of 0.61 at fourth grade and 0.62 at eighth grade. The median multiple correlation between the component variables and student achievement was 0.20 for both mathematics and science at fourth grade and 0.16 and 0.18 for mathematics and science, respectively, at eighth grade.

As shown in Exhibit 12.7, factor loadings ranged from 0.551 for "something of mine was stolen" to 0.737 for "I was made fun of or called names" at the fourth grade. At the eighth grade, the factor structure was similar, with factor loadings ranging from 0.550 for "something of mine was stolen" to 0.754 for "I was made to do things I didn't want to do by other students." With an RMSEA value of less than 0.05 the model fits the data well at both grades.

Exhibit 12.5 Index of Time Students Spend on Doing Mathematics Homework (TMH) / Science Homework (TSH) in a Normal School Week—Reliability and Validity Indicators

Countries	Grade 4						Grade 8					
	Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables		Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science
Algeria	0.19	0.08	0.09	0.08	0.01	0.01	◊	–	◊	–	◊	–
Armenia	–0.03	0.02	0.05	0.08	0.00	0.01	–0.01	–	0.03	–	0.00	–
Australia	–0.16	–0.09	0.09	0.28	0.01	0.08	0.28	0.20	0.23	0.10	0.05	0.01
Austria	–0.03	0.28	0.11	0.27	0.01	0.07	◊	◊	◊	◊	◊	◊
Bahrain	◊	◊	◊	◊	◊	◊	0.14	0.22	0.09	0.07	0.01	0.00
Bosnia and Herzegovina	◊	◊	◊	◊	◊	◊	0.27	–	0.08	–	0.01	–
Botswana	◊	◊	◊	◊	◊	◊	0.02	0.02	0.09	0.09	0.01	0.01
Bulgaria	◊	◊	◊	◊	◊	◊	0.20	–	0.01	–	0.00	–
Chinese Taipei	0.07	0.21	0.13	0.24	0.02	0.06	0.21	0.18	0.19	0.18	0.04	0.03
Colombia	0.05	–0.01	0.11	0.19	0.01	0.04	0.03	0.10	0.06	0.12	0.00	0.02
Cyprus	◊	◊	◊	◊	◊	◊	0.07	–	0.13	–	0.02	–
Czech Republic	0.11	0.13	0.12	0.13	0.01	0.02	0.00	–	0.19	–	0.04	–
Denmark	0.15	0.35	0.15	0.18	0.02	0.03	◊	◊	◊	◊	◊	◊
Egypt	◊	◊	◊	◊	◊	◊	0.02	–0.10	0.15	0.23	0.02	0.05
El Salvador	–0.07	–0.21	0.13	0.11	0.02	0.01	–0.01	–0.05	0.09	0.14	0.01	0.02
England	0.08	0.02	0.07	0.19	0.01	0.04	0.23	0.29	0.16	0.22	0.02	0.05
Georgia	–0.03	–0.02	0.09	0.15	0.01	0.02	0.05	–	0.15	–	0.02	–
Germany	0.05	0.20	0.14	0.15	0.02	0.02	◊	◊	◊	◊	◊	◊
Ghana	◊	◊	◊	◊	◊	◊	–0.15	–0.09	0.16	0.15	0.02	0.02
Hong Kong SAR	0.01	0.15	0.19	0.13	0.04	0.02	0.29	–0.08	0.11	0.17	0.01	0.03
Hungary	0.07	0.20	0.01	0.10	0.00	0.01	0.13	–	0.06	–	0.00	–
Indonesia	◊	◊	◊	◊	◊	◊	–0.09	–	0.13	–	0.02	–
Iran, Islamic Rep. of	0.07	0.03	0.14	0.12	0.02	0.01	0.09	0.24	0.20	0.08	0.04	0.01
Israel	◊	◊	◊	◊	◊	◊	0.14	0.30	0.07	0.17	0.01	0.03
Italy	0.09	0.03	0.11	0.14	0.01	0.02	0.05	0.22	0.12	0.09	0.02	0.01
Japan	0.09	0.02	0.21	0.18	0.04	0.03	0.07	0.05	0.14	0.14	0.02	0.02
Jordan	◊	◊	◊	◊	◊	◊	0.11	0.11	0.16	0.14	0.02	0.02
Kazakhstan	0.06	0.18	0.02	0.19	0.00	0.04	◊	◊	◊	◊	◊	◊
Korea, Rep. of	◊	◊	◊	◊	◊	◊	0.12	0.05	0.10	0.08	0.01	0.01
Kuwait	0.12	0.17	0.16	0.18	0.02	0.03	0.21	0.17	0.22	0.10	0.05	0.01
Latvia	0.03	0.11	0.13	0.15	0.02	0.02	◊	◊	◊	◊	◊	◊
Lebanon	◊	◊	◊	◊	◊	◊	0.05	–	0.15	–	0.02	–
Lithuania	0.08	0.21	0.12	0.16	0.01	0.03	0.15	–	0.10	–	0.01	–
Malaysia	◊	◊	◊	◊	◊	◊	0.06	0.14	0.10	0.04	0.01	0.00
Malta	◊	◊	◊	◊	◊	◊	0.15	–	0.16	–	0.03	–
Morocco	0.09	–0.03	0.05	0.10	0.00	0.01	0.04	–	0.09	–	0.01	–
Netherlands	0.14	0.07	0.23	0.27	0.05	0.07	◊	◊	◊	◊	◊	◊
New Zealand	0.05	0.03	0.16	0.27	0.03	0.07	◊	◊	◊	◊	◊	◊
Norway	–0.03	0.16	0.10	0.19	0.01	0.04	–0.01	0.11	0.07	0.07	0.01	0.00
Oman	◊	◊	◊	◊	◊	◊	0.03	0.05	0.08	0.13	0.01	0.02
Palestinian Nat'l Auth.	◊	◊	◊	◊	◊	◊	0.10	0.14	0.11	0.10	0.01	0.01
Qatar	0.05	0.08	0.11	0.14	0.01	0.02	0.12	0.14	0.09	0.08	0.01	0.01
Romania	◊	◊	◊	◊	◊	◊	0.21	–	0.28	–	0.08	–
Russian Federation	0.04	0.09	0.06	0.15	0.00	0.02	–0.01	–	0.11	–	0.01	–
Saudi Arabia	◊	◊	◊	◊	◊	◊	0.21	0.14	0.10	0.18	0.01	0.03
Scotland	–0.02	0.11	0.12	0.24	0.02	0.06	0.02	0.14	0.21	0.13	0.05	0.02
Serbia	◊	◊	◊	◊	◊	◊	0.34	–	0.01	–	0.00	–
Singapore	0.05	0.07	0.10	0.05	0.01	0.00	0.20	0.23	0.18	0.12	0.03	0.01
Slovak Republic	0.14	0.26	0.13	0.20	0.02	0.04	◊	◊	◊	◊	◊	◊
Slovenia	–0.08	0.04	0.15	0.10	0.02	0.01	0.24	–	0.07	–	0.00	–
Sweden	0.08	0.19	0.15	0.20	0.02	0.04	0.17	–	0.11	–	0.01	–
Syrian Arab Republic	◊	◊	◊	◊	◊	◊	–0.02	–	0.16	–	0.02	–
Thailand	◊	◊	◊	◊	◊	◊	0.15	0.05	0.18	0.20	0.03	0.04
Tunisia	0.01	0.01	0.06	0.05	0.00	0.00	0.22	0.25	0.04	0.13	0.00	0.02
Turkey	◊	◊	◊	◊	◊	◊	0.09	0.18	0.17	0.14	0.03	0.02
Ukraine	–0.01	0.01	0.07	0.20	0.01	0.04	0.02	–	0.10	–	0.01	–
United States	0.05	0.03	0.05	0.18	0.00	0.03	0.29	0.26	0.15	0.08	0.02	0.01
Yemen	0.01	0.14	0.10	0.09	0.01	0.01	◊	◊	◊	◊	◊	◊
International Median	0.05	0.08	0.11	0.16	0.01	0.02	0.10	0.14	0.11	0.13	0.01	0.02
Benchmarking Participants												
Alberta, Canada	0.18	0.18	0.09	0.19	0.01	0.04	◊	◊	◊	◊	◊	◊
Basque Country, Spain	◊	◊	◊	◊	◊	◊	0.13	0.18	0.10	0.06	0.01	0.00
British Columbia, Canada	0.22	0.15	0.13	0.21	0.02	0.04	0.26	0.34	0.07	0.08	0.01	0.01
Dubai, UAE	0.11	0.24	0.12	0.24	0.02	0.06	0.24	0.38	0.16	0.13	0.03	0.02
Massachusetts, US	–0.11	0.12	0.02	0.14	0.00	0.02	0.35	0.32	0.18	0.04	0.03	0.00
Minnesota, US	0.12	0.17	0.10	0.24	0.01	0.06	0.21	0.15	0.17	0.09	0.03	0.01
Ontario, Canada	0.21	–0.03	0.07	0.18	0.00	0.03	0.26	0.15	0.14	0.05	0.02	0.00
Quebec, Canada	0.05	0.08	0.21	0.21	0.04	0.05	0.32	0.30	0.14	0.06	0.02	0.00

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

Exhibit 12.5 Index of Time Students Spend on Doing Mathematics Homework (TMH) / Science Homework (TSH) in a Normal School Week—Reliability and Validity Indicators (Continued)

Countries	Grade 8 Separate Science											
	Cronbach's Alpha Between the Component Variables				Multiple R Between Student Achievement and Component Variables				Percent of Variance in Student Achievement Accounted for by the Component Variables			
	Biology	Earth Science	Chemistry	Physics	Biology	Earth Science	Chemistry	Physics	Biology	Earth Science	Chemistry	Physics
Algeria	–	–	–	–	–	–	–	–	–	–	–	–
Armenia	0.18	0.17	0.09	0.06	0.05	0.04	0.04	0.07	0.00	0.00	0.00	0.00
Bosnia and Herzegovina	0.24	0.25	0.29	0.28	0.21	0.24	0.21	0.18	0.04	0.06	0.04	0.03
Bulgaria	0.29	0.29	0.31	0.28	0.16	0.11	0.13	0.13	0.03	0.01	0.02	0.02
Cyprus	0.10	0.20	0.17	0.17	0.22	0.17	0.19	0.20	0.05	0.03	0.04	0.04
Czech Republic	0.09	0.17	0.21	0.24	0.18	0.15	0.17	0.19	0.03	0.02	0.03	0.04
Georgia	–0.02	–0.04	0.00	0.04	0.09	0.11	0.02	0.04	0.01	0.01	0.00	0.00
Hungary	0.26	0.17	0.24	0.22	0.16	0.17	0.15	0.17	0.03	0.03	0.02	0.03
Indonesia	–0.08	–	–	–0.07	0.07	–	–	0.07	0.00	–	–	0.00
Lebanon	0.14	–	0.15	0.05	0.14	–	0.14	0.14	0.02	–	0.02	0.02
Lithuania	0.23	0.25	0.27	0.23	0.15	0.16	0.18	0.16	0.02	0.03	0.03	0.03
Malta	0.09	0.10	0.12	0.18	0.18	0.21	0.15	0.16	0.03	0.04	0.02	0.03
Morocco	–0.07	0.00	–0.08	0.05	0.13	0.11	0.14	0.15	0.02	0.01	0.02	0.02
Romania	0.38	0.41	0.37	0.23	0.21	0.17	0.17	0.21	0.04	0.03	0.03	0.04
Russian Federation	0.13	0.12	0.15	0.14	0.14	0.16	0.13	0.13	0.02	0.02	0.02	0.02
Serbia	0.32	0.31	0.33	0.33	0.18	0.24	0.20	0.21	0.03	0.06	0.04	0.04
Slovenia	0.14	0.19	0.19	0.20	0.18	0.16	0.19	0.17	0.03	0.02	0.04	0.03
Sweden	0.03	0.07	0.08	0.09	0.12	0.13	0.12	0.13	0.01	0.02	0.01	0.02
Syrian Arab Republic	0.09	0.08	0.03	0.02	0.04	0.07	0.02	0.08	0.00	0.01	0.00	0.01
Ukraine	0.08	0.08	0.13	0.04	0.10	0.10	0.10	0.12	0.01	0.01	0.01	0.01
International Median	0.13	0.17	0.16	0.17	0.15	0.16	0.14	0.15	0.02	0.02	0.02	0.02

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

A diamond (◊) indicates the country did not participate in the assessment.

A dash (–) indicates comparable data are not available or country administered separate science version of the student questionnaire.

Exhibit 12.6 Index of Students' Perception of Being Safe in School (SPBSS)—Reliability and Validity Indicators

Country	Grade 4					Grade 8				
	Cronbach's Alpha Between the Component Variables	Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables		Cronbach's Alpha Between the Component Variables	Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
		Mathematics	Science	Mathematics	Science		Mathematics	Science	Mathematics	Science
Algeria	0.48	0.24	0.24	0.06	0.06	0.56	0.14	0.13	0.02	0.02
Armenia	0.60	0.06	0.11	0.00	0.01	0.68	0.08	0.07	0.01	0.01
Australia	0.64	0.20	0.18	0.04	0.03	0.65	0.10	0.10	0.01	0.01
Austria	0.68	0.18	0.20	0.03	0.04	◊	◊	◊	◊	◊
Bahrain	◊	◊	◊	◊	◊	0.61	0.20	0.26	0.04	0.07
Bosnia and Herzegovina	◊	◊	◊	◊	◊	0.65	0.17	0.17	0.03	0.03
Botswana	◊	◊	◊	◊	◊	0.20	0.41	0.49	0.17	0.24
Bulgaria	◊	◊	◊	◊	◊	0.64	0.19	0.20	0.04	0.04
Chinese Taipei	0.70	0.18	0.17	0.03	0.03	0.70	0.07	0.09	0.01	0.01
Colombia	0.59	0.21	0.24	0.04	0.06	0.52	0.11	0.10	0.01	0.01
Cyprus	◊	◊	◊	◊	◊	0.66	0.20	0.20	0.04	0.04
Czech Republic	0.58	0.19	0.16	0.03	0.03	0.59	0.13	0.11	0.02	0.01
Denmark	0.59	0.14	0.14	0.02	0.02	◊	◊	◊	◊	◊
Egypt	◊	◊	◊	◊	◊	0.62	0.27	0.30	0.08	0.09
El Salvador	0.59	0.20	0.22	0.04	0.05	0.54	0.08	0.10	0.01	0.01
England	0.63	0.20	0.22	0.04	0.05	0.62	0.12	0.11	0.01	0.01
Georgia	0.53	0.23	0.24	0.06	0.06	0.70	0.12	0.20	0.02	0.04
Germany	0.65	0.21	0.20	0.04	0.04	◊	◊	◊	◊	◊
Ghana	◊	◊	◊	◊	◊	0.41	0.24	0.30	0.06	0.09
Hong Kong SAR	0.65	0.16	0.16	0.03	0.03	0.66	0.16	0.15	0.03	0.02
Hungary	0.64	0.26	0.26	0.07	0.07	0.64	0.15	0.14	0.02	0.02
Indonesia	◊	◊	◊	◊	◊	0.58	0.19	0.20	0.04	0.04
Iran, Islamic Rep. of	0.54	0.14	0.14	0.02	0.02	0.54	0.22	0.21	0.05	0.04
Israel	◊	◊	◊	◊	◊	0.70	0.26	0.27	0.07	0.07
Italy	0.63	0.14	0.17	0.02	0.03	0.58	0.11	0.09	0.01	0.01
Japan	0.67	0.14	0.14	0.02	0.02	0.66	0.10	0.11	0.01	0.01
Jordan	◊	◊	◊	◊	◊	0.62	0.21	0.25	0.05	0.06
Kazakhstan	0.60	0.06	0.08	0.00	0.01	◊	◊	◊	◊	◊
Korea, Rep. of	◊	◊	◊	◊	◊	0.55	0.09	0.11	0.01	0.01
Kuwait	0.66	0.31	0.33	0.09	0.11	0.70	0.22	0.24	0.05	0.06
Latvia	0.55	0.21	0.22	0.04	0.05	◊	◊	◊	◊	◊
Lebanon	◊	◊	◊	◊	◊	0.65	0.28	0.32	0.08	0.10
Lithuania	0.60	0.24	0.23	0.06	0.06	0.59	0.14	0.14	0.02	0.02
Malaysia	◊	◊	◊	◊	◊	0.55	0.17	0.24	0.03	0.06
Malta	◊	◊	◊	◊	◊	0.63	0.23	0.22	0.05	0.05
Morocco	0.40	0.22	0.23	0.05	0.05	0.52	0.19	0.20	0.04	0.04
Netherlands	0.62	0.22	0.20	0.05	0.04	◊	◊	◊	◊	◊
New Zealand	0.65	0.23	0.26	0.05	0.07	◊	◊	◊	◊	◊
Norway	0.66	0.23	0.22	0.05	0.05	0.58	0.10	0.09	0.01	0.01
Oman	◊	◊	◊	◊	◊	0.63	0.24	0.27	0.06	0.07
Palestinian Nat'l Auth.	◊	◊	◊	◊	◊	0.60	0.23	0.27	0.05	0.07
Qatar	0.66	0.25	0.24	0.06	0.06	0.68	0.23	0.22	0.05	0.05
Romania	◊	◊	◊	◊	◊	0.63	0.20	0.20	0.04	0.04
Russian Federation	0.52	0.15	0.14	0.02	0.02	0.54	0.14	0.12	0.02	0.01
Saudi Arabia	◊	◊	◊	◊	◊	0.58	0.17	0.20	0.03	0.04
Scotland	0.67	0.15	0.17	0.02	0.03	0.64	0.14	0.15	0.02	0.02
Serbia	◊	◊	◊	◊	◊	0.64	0.13	0.15	0.02	0.02
Singapore	0.59	0.25	0.26	0.06	0.07	0.63	0.20	0.19	0.04	0.03
Slovak Republic	0.61	0.26	0.27	0.07	0.08	◊	◊	◊	◊	◊
Slovenia	0.64	0.19	0.19	0.04	0.04	0.65	0.13	0.16	0.02	0.03
Sweden	0.62	0.21	0.18	0.04	0.03	0.72	0.17	0.16	0.03	0.03
Syrian Arab Republic	◊	◊	◊	◊	◊	0.69	0.16	0.20	0.03	0.04
Thailand	◊	◊	◊	◊	◊	0.62	0.19	0.19	0.04	0.04
Tunisia	0.49	0.25	0.22	0.06	0.05	0.59	0.12	0.15	0.01	0.02
Turkey	◊	◊	◊	◊	◊	0.58	0.16	0.18	0.02	0.03
Ukraine	0.60	0.19	0.18	0.04	0.03	0.60	0.15	0.16	0.02	0.03
United States	–	–	–	–	–	–	–	–	–	–
Yemen	0.58	0.15	0.13	0.02	0.02	◊	◊	◊	◊	◊
International Median	0.61	0.20	0.20	0.04	0.04	0.62	0.16	0.18	0.03	0.03
Benchmarking Participants										
Alberta, Canada	0.65	0.18	0.21	0.03	0.05	◊	◊	◊	◊	◊
Basque Country, Spain	◊	◊	◊	◊	◊	0.62	0.20	0.14	0.04	0.02
British Columbia, Canada	0.65	0.19	0.20	0.04	0.04	0.65	0.12	0.15	0.02	0.02
Dubai, UAE	0.55	0.23	0.27	0.05	0.08	0.62	0.18	0.18	0.03	0.03
Massachusetts, US	–	–	–	–	–	–	–	–	–	–
Minnesota, US	–	–	–	–	–	–	–	–	–	–
Ontario, Canada	0.65	0.19	0.20	0.04	0.04	0.63	0.07	0.07	0.00	0.01
Quebec, Canada	0.64	0.19	0.18	0.04	0.03	0.59	0.12	0.10	0.01	0.01

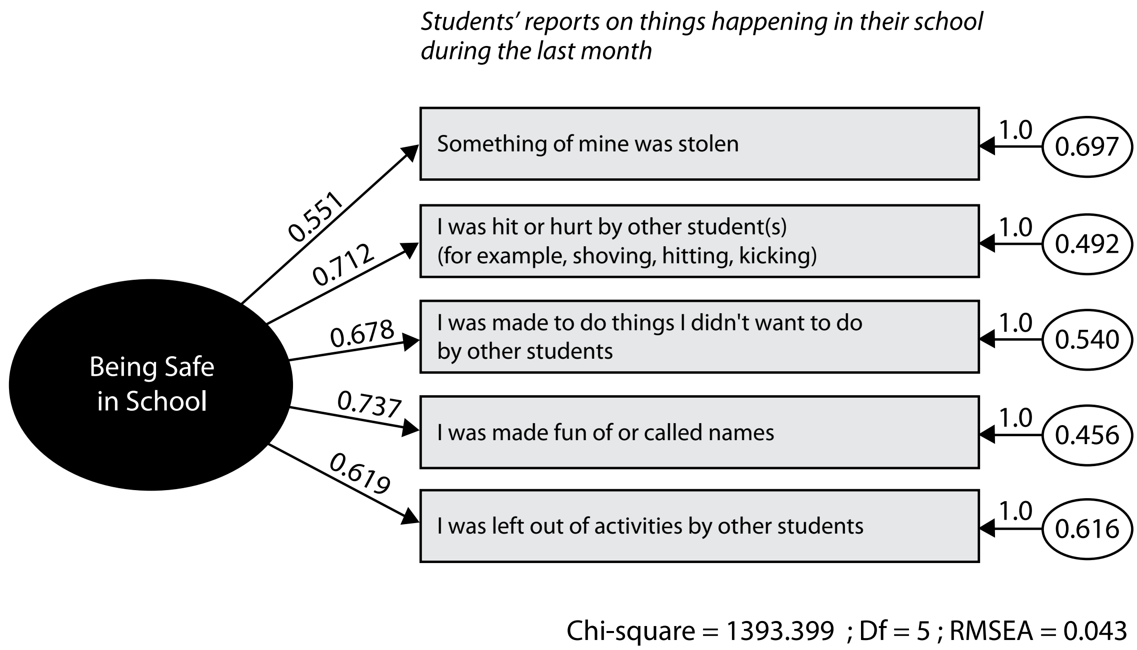
SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

A diamond (◊) indicates the country did not participate in the assessment.

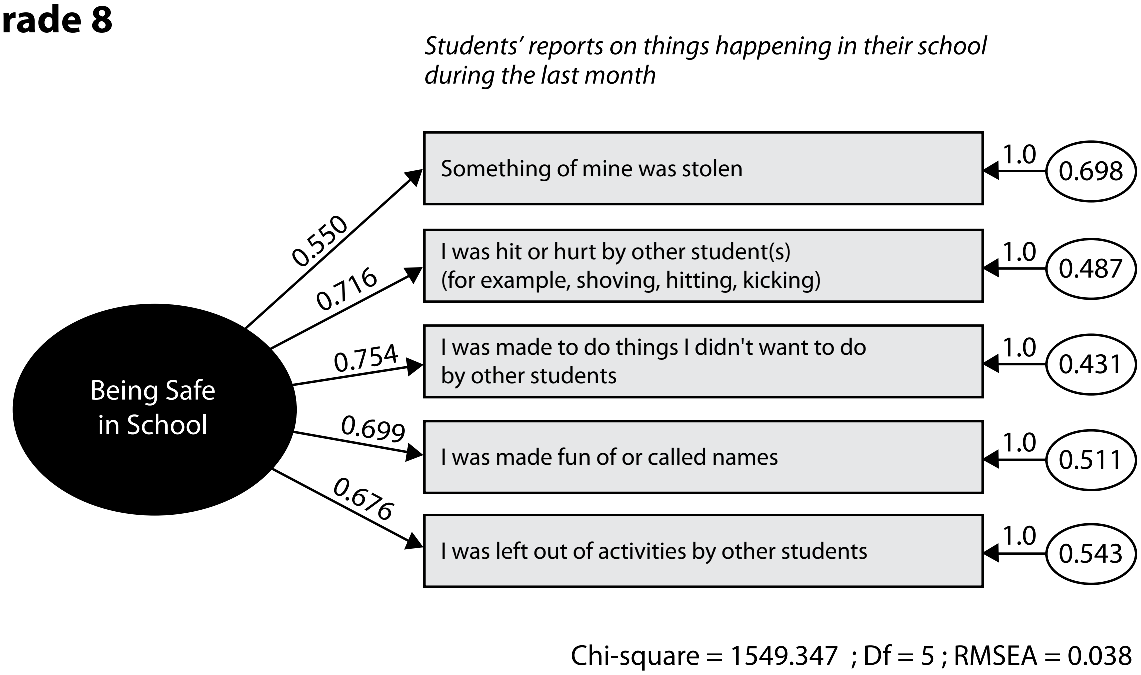
A dash (–) indicates comparable data are not available.

Exhibit 12.7 Latent Variable Model of Students' Perception of Being Safe in School

Grade 4



Grade 8



12.4.2 Teacher-level Indices

The *TIMSS 2007 Teacher Questionnaires* collected information about teachers' education and training, instructional practices, and the implemented curriculum in mathematics and science. At the fourth grade, a single questionnaire addressed both subjects, whereas there were separate versions for mathematics and science teachers at the eighth grade. Five indices presented in the TIMSS 2007 international reports were based on questions in the teacher questionnaires.

The Index of Teachers' Reports on Teaching Mathematics Classes with Few or No Limitations (MCFL) and the Index of Teachers' Reports on Teaching Science Classes with Few or No Limitations (SCFL) group students according to their teachers' reports on the instructional impact of five characteristics of their students: 1) Students with different academic abilities; 2) Students who come from a wide range of backgrounds; 3) Students with special needs; 4) Uninterested students; and 5) Disruptive students. The index, modified from an earlier version from 2003, is presented in Exhibit 7.3 of the TIMSS 2007 international reports, including trends from 2003 at the eighth grade. The item "low morale among students" was included in the TIMSS 2003 index calculations but not in the 2007 index calculations. Thus, the percentage of students at each index level in 2003 was recomputed excluding this item. Trends were not reported at the fourth grade because the component variables were not part of the fourth grade teacher questionnaire in 2003.

Teachers rated the impact of the five statements about student factors limiting mathematics and science instruction on a 4-point scale: *not at all/not applicable* = 1; *a little* = 2; *some* = 3; and *a lot* = 4. An average was computed across the five items. Students were placed in the high category, if the average was less than or equal to 2. If the average across the five items was greater than 3, students were placed in the low category. A medium level indicates averages greater than 2 but less than 3.

As shown in Exhibit 12.8, the five components form a fairly reliable scale, with median reliability coefficients (Cronbach's alpha) across countries of 0.71 and 0.73 for mathematics and science, respectively, at the fourth grade, and 0.69 and 0.68 at the eighth grade. The median multiple correlation between the five statements and student achievement was 0.15 and 0.12 for mathematics and science at the fourth grade, and 0.19 and 0.14, respectively, at the eighth grade.

From the latent factor measurement model shown in Exhibit 12.9, it appears that, for both mathematics and science at both grades, “uninterested students” and “disruptive students” are the dominant student characteristics, having the highest factor loadings on all four scales. It may be that for teachers everywhere, such students pose a challenge for instruction in the classroom. By comparison, “students with different academic abilities”, “students who come from a wide range of backgrounds”, and “students with special needs” had somewhat lower loadings, implying that the challenge posed by such students is of a different nature, and may vary more from classroom to classroom and country to country.

The Index of Teachers’ Emphasis on Mathematics Homework (EMH) and the Index of Teachers’ Emphasis on Science Homework (ESH) categorize fourth and eighth grade students according to their teachers’ responses to two questions about the frequency of assigning homework and the amount of homework assigned. By describing teachers’ practices in assigning mathematics and science homework, these indices complement the indices on students’ reports of the time they actually spend on homework (Exhibit 12.5). Students at the high level of the teacher indices had teachers who reported assigning more than 30 minutes of homework in half of the lessons or more. Students at the low level had teachers who reported assigning less than 30 minutes of homework in fewer than half of the lessons. A medium level indicates all other combinations of amount and frequency of homework assignments. Like the student indices described earlier, the teacher indices are sensitive to differences across countries in the role of homework in mathematics and science instruction. These indices were presented in Exhibit 7.12 of the *TIMSS 2007 International Mathematics Report* and Exhibit 7.11 of the *TIMSS 2007 International Science Report.*

As shown in Exhibit 12.10, the variables comprising this index have relatively low reliability (international median Cronbach’s alpha of 0.08 or less) and show no substantive relationship with achievement. This underlines the different purpose homework serves in instructional contexts and particularly its use for remedial instruction.

The Index of Teachers’ Adequate Working Conditions (TAWC) summarizes teachers’ perspectives on the availability of school resources and how these affect their capacity to provide effective mathematics and science instruction. Teachers were asked to rate problems in their school by severity on a 3-point scale: *not a problem* = 1; *minor problem* = 2; and

Exhibit 12.8 Index of Teachers' Reports on Teaching Mathematics (MFCL) / Science (SFCL) Classes with Few or No Limitations on Instruction Due to Student Factors—Reliability and Validity Indicators

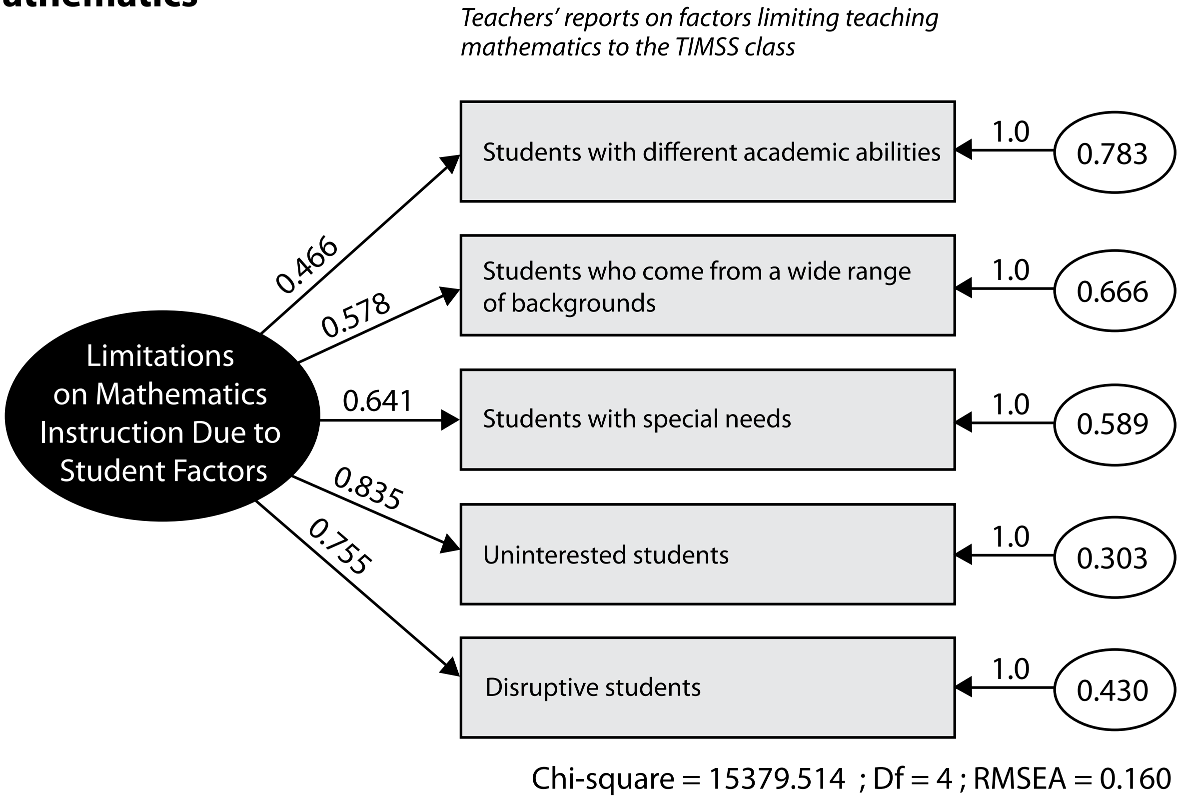
Country	Grade 4						Grade 8					
	Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables		Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science
Algeria	0.71	0.75	0.18	0.09	0.03	0.01	0.63	0.59	0.05	0.08	0.00	0.01
Armenia	0.76	0.81	0.12	0.09	0.01	0.01	0.61	0.63	0.13	0.06	0.02	0.00
Australia	0.76	0.83	0.19	0.14	0.03	0.02	0.81	0.77	0.42	0.22	0.17	0.05
Austria	0.71	0.76	0.15	0.19	0.02	0.04	◊	◊	◊	◊	◊	◊
Bahrain	◊	◊	◊	◊	◊	◊	0.63	0.64	0.13	0.14	0.02	0.02
Bosnia and Herzegovina	◊	◊	◊	◊	◊	◊	0.68	0.68	0.06	0.04	0.00	0.00
Botswana	◊	◊	◊	◊	◊	◊	0.64	0.62	0.11	0.12	0.01	0.01
Bulgaria	◊	◊	◊	◊	◊	◊	0.55	0.67	0.27	0.20	0.07	0.04
Chinese Taipei	0.72	0.72	0.08	0.07	0.01	0.00	0.63	0.66	0.21	0.13	0.04	0.02
Colombia	0.69	0.69	0.12	0.11	0.02	0.01	0.69	0.69	0.19	0.17	0.04	0.03
Cyprus	◊	◊	◊	◊	◊	◊	0.70	0.73	0.06	0.03	0.00	0.00
Czech Republic	0.72	0.74	0.18	0.08	0.03	0.01	0.74	0.69	0.28	0.16	0.08	0.03
Denmark	0.73	0.76	0.11	0.12	0.01	0.01	◊	◊	◊	◊	◊	◊
Egypt	◊	◊	◊	◊	◊	◊	0.69	0.60	0.19	0.13	0.04	0.02
El Salvador	0.66	0.69	0.13	0.13	0.02	0.02	0.71	0.70	0.14	0.13	0.02	0.02
England	0.79	0.79	0.18	0.18	0.03	0.03	0.67	0.76	0.51	0.37	0.26	0.14
Georgia	0.71	0.68	0.12	0.13	0.01	0.02	0.70	0.73	0.14	0.09	0.02	0.01
Germany	0.68	0.74	0.24	0.24	0.06	0.06	◊	◊	◊	◊	◊	◊
Ghana	◊	◊	◊	◊	◊	◊	0.57	0.68	0.25	0.21	0.06	0.04
Hong Kong SAR	0.72	0.72	0.34	0.12	0.12	0.01	0.79	0.74	0.47	0.38	0.23	0.15
Hungary	0.76	0.77	0.28	0.32	0.08	0.10	0.70	0.71	0.24	0.21	0.06	0.04
Indonesia	◊	◊	◊	◊	◊	◊	0.71	0.61	0.17	0.11	0.03	0.01
Iran, Islamic Rep. of	0.72	0.75	0.18	0.14	0.03	0.02	0.65	0.69	0.19	0.14	0.04	0.02
Israel	◊	◊	◊	◊	◊	◊	0.71	0.77	0.23	0.19	0.05	0.04
Italy	0.64	0.72	0.15	0.12	0.02	0.01	0.65	0.67	0.07	0.09	0.00	0.01
Japan	0.67	0.68	0.05	0.05	0.00	0.00	0.75	0.74	0.18	0.22	0.03	0.05
Jordan	◊	◊	◊	◊	◊	◊	0.75	0.68	0.21	0.13	0.04	0.02
Kazakhstan	0.67	0.68	0.18	0.11	0.03	0.01	◊	◊	◊	◊	◊	◊
Korea, Rep. of	◊	◊	◊	◊	◊	◊	0.66	0.66	0.10	0.03	0.01	0.00
Kuwait	0.77	0.73	0.13	0.12	0.02	0.01	0.71	0.61	0.06	0.13	0.00	0.02
Latvia	0.59	0.65	0.10	0.08	0.01	0.01	◊	◊	◊	◊	◊	◊
Lebanon	◊	◊	◊	◊	◊	◊	0.71	0.65	0.23	0.18	0.05	0.03
Lithuania	0.66	0.74	0.14	0.10	0.02	0.01	0.67	0.72	0.20	0.14	0.04	0.02
Malaysia	◊	◊	◊	◊	◊	◊	0.76	0.67	0.36	0.37	0.13	0.13
Malta	◊	◊	◊	◊	◊	◊	0.83	0.75	0.37	0.34	0.13	0.11
Morocco	0.62	0.72	0.17	0.11	0.03	0.01	0.51	0.59	0.09	0.17	0.01	0.03
Netherlands	0.75	0.81	0.21	0.19	0.04	0.04	◊	◊	◊	◊	◊	◊
New Zealand	0.76	0.79	0.21	0.20	0.05	0.04	◊	◊	◊	◊	◊	◊
Norway	0.74	0.78	0.16	0.11	0.03	0.01	0.52	0.68	0.07	0.10	0.00	0.01
Oman	◊	◊	◊	◊	◊	◊	0.55	0.54	0.18	0.14	0.03	0.02
Palestinian Nat'l Auth.	◊	◊	◊	◊	◊	◊	0.74	0.60	0.07	0.09	0.00	0.01
Qatar	0.68	0.67	0.15	0.16	0.02	0.03	0.73	0.59	0.12	0.24	0.01	0.06
Romania	◊	◊	◊	◊	◊	◊	0.68	0.72	0.09	0.08	0.01	0.01
Russian Federation	0.69	0.71	0.12	0.16	0.01	0.02	0.72	0.72	0.22	0.16	0.05	0.03
Saudi Arabia	◊	◊	◊	◊	◊	◊	0.69	0.71	0.09	0.08	0.01	0.01
Scotland	0.75	0.71	0.18	0.19	0.03	0.04	0.80	0.80	0.48	0.25	0.23	0.06
Serbia	◊	◊	◊	◊	◊	◊	0.52	0.73	0.13	0.07	0.02	0.00
Singapore	0.81	0.82	0.30	0.23	0.09	0.05	0.76	0.77	0.44	0.40	0.19	0.16
Slovak Republic	0.64	0.73	0.18	0.19	0.03	0.04	◊	◊	◊	◊	◊	◊
Slovenia	0.68	0.72	0.09	0.07	0.01	0.00	0.70	0.65	0.14	0.04	0.02	0.00
Sweden	0.75	0.83	0.17	0.18	0.03	0.03	0.76	0.74	0.19	0.14	0.03	0.02
Syrian Arab Republic	◊	◊	◊	◊	◊	◊	0.69	0.68	0.12	0.03	0.01	0.00
Thailand	◊	◊	◊	◊	◊	◊	0.58	0.65	0.33	0.27	0.11	0.07
Tunisia	0.65	0.69	0.15	0.06	0.02	0.00	0.55	0.67	0.08	0.04	0.01	0.00
Turkey	◊	◊	◊	◊	◊	◊	0.64	0.72	0.25	0.28	0.06	0.08
Ukraine	0.61	0.72	0.10	0.07	0.01	0.01	0.74	0.72	0.22	0.10	0.05	0.01
United States	0.81	0.83	0.20	0.17	0.04	0.03	0.78	0.81	0.36	0.25	0.13	0.06
Yemen	0.46	0.53	0.12	0.11	0.02	0.01	◊	◊	◊	◊	◊	◊
International Median	0.71	0.73	0.15	0.12	0.02	0.01	0.69	0.68	0.19	0.14	0.03	0.02
Benchmarking Participants												
Alberta, Canada	0.79	0.83	0.22	0.21	0.05	0.04	◊	◊	◊	◊	◊	◊
Basque Country, Spain	◊	◊	◊	◊	◊	◊	0.76	0.65	0.21	0.18	0.04	0.03
British Columbia, Canada	0.69	0.75	0.10	0.09	0.01	0.01	0.77	0.83	0.30	0.12	0.09	0.01
Dubai, UAE	0.60	0.73	0.14	0.30	0.02	0.09	0.74	0.61	0.34	0.17	0.12	0.03
Massachusetts, US	0.77	0.80	0.14	0.19	0.02	0.04	0.83	0.84	0.42	0.33	0.18	0.11
Minnesota, US	0.74	0.87	0.27	0.20	0.07	0.04	0.74	0.71	0.44	0.14	0.19	0.02
Ontario, Canada	0.77	0.77	0.17	0.17	0.03	0.03	0.76	0.82	0.12	0.16	0.02	0.03
Quebec, Canada	0.75	0.80	0.18	0.12	0.03	0.01	0.78	0.71	0.39	0.29	0.15	0.09

A diamond (◊) indicates the country did not participate in the assessment.

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

Exhibit 12.9 Latent Variable Model of Teachers' Reports on Limitations on Instruction Due to Student Factors, Grade 4

Mathematics



Science

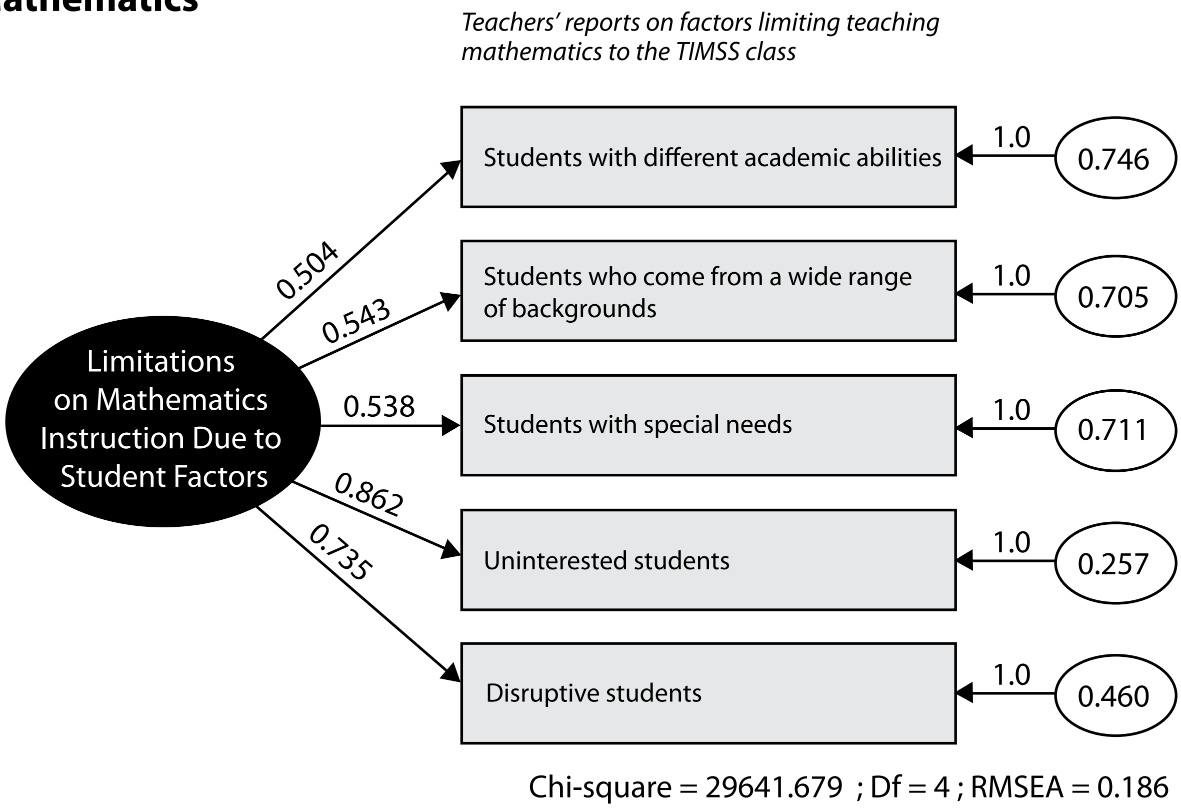
Factor: Limitations on Science Instruction Due to Student Factors

		Limitations on Science Instruction Due to Student Factors
	Observed Variable	*Factor Loadings*
Teachers' reports on factors limiting teaching science to the TIMSS class	Students with different academic abilities	0.603
	Students who come from a wide range of backgrounds	0.675
	Students with special needs	0.686
	Uninterested students	0.851
	Disruptive students	0.785

Chi-square=23354.600; Df=4; RMSEA=0.200

Exhibit 12.9 Latent Variable Model of Teachers' Reports on Limitations on Instruction Due to Student Factors, Grade 8 (Continued)

Mathematics



Science

Factor: Limitations on Science Instruction Due to Student Factors

		Limitations on Science Instruction Due to Student Factors
Teachers' reports on factors limiting teaching science to the TIMSS class	*Observed Variable*	*Factor Loadings*
	Students with different academic abilities	0.520
	Students who come from a wide range of backgrounds	0.555
	Students with special needs	0.521
	Uninterested students	0.830
	Disruptive students	0.754

Chi-square= 56470.929; Df=4; RMSEA=0.190

Exhibit 12.10 Index of Teachers' Emphasis on Mathematics (EMH) / Science (ESH) Homework—Reliability and Validity Indicators

Country	Grade 4						Grade 8					
	Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables		Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science
Algeria	0.13	0.07	0.04	0.06	0.00	0.00	–0.41	–0.37	0.02	0.04	0.00	0.00
Armenia	0.08	–0.10	0.04	0.02	0.00	0.00	0.02	0.22	0.06	0.03	0.00	0.00
Australia	0.02	0.01	0.13	0.18	0.02	0.03	–0.27	–0.29	0.28	0.16	0.08	0.03
Austria	–0.11	–0.05	0.05	0.06	0.00	0.00	◊	◊	◊	◊	◊	◊
Bahrain	◊	◊	◊	◊	◊	◊	–0.03	–0.31	0.13	0.12	0.02	0.01
Bosnia and Herzegovina	◊	◊	◊	◊	◊	◊	–0.07	0.04	0.16	0.04	0.02	0.00
Botswana	◊	◊	◊	◊	◊	◊	0.11	–0.20	0.12	0.07	0.01	0.01
Bulgaria	◊	◊	◊	◊	◊	◊	0.16	–0.22	0.19	0.04	0.04	0.00
Chinese Taipei	0.04	–0.18	0.06	0.04	0.00	0.00	–0.02	–0.01	0.20	0.13	0.04	0.02
Colombia	–0.46	–0.07	0.04	0.13	0.00	0.02	–0.05	0.21	0.12	0.05	0.01	0.00
Cyprus	◊	◊	◊	◊	◊	◊	0.02	–0.27	0.05	0.02	0.00	0.00
Czech Republic	0.04	0.21	0.04	0.07	0.00	0.00	–0.02	0.06	0.14	0.05	0.02	0.00
Denmark	–0.33	0.26	0.04	0.10	0.00	0.01	◊	◊	◊	◊	◊	◊
Egypt	◊	◊	◊	◊	◊	◊	–0.21	–0.03	0.04	0.06	0.00	0.00
El Salvador	0.14	0.09	0.02	0.06	0.00	0.00	0.09	–0.39	0.05	0.13	0.00	0.02
England	–0.03	–0.07	0.10	0.07	0.01	0.01	–0.26	0.01	0.31	0.28	0.09	0.08
Georgia	–0.04	0.08	0.05	0.11	0.00	0.01	0.26	0.32	0.07	0.05	0.00	0.00
Germany	0.19	0.06	0.06	0.04	0.00	0.00	◊	◊	◊	◊	◊	◊
Ghana	◊	◊	◊	◊	◊	◊	–0.27	–0.53	0.11	0.12	0.01	0.02
Hong Kong SAR	0.41	0.07	0.15	0.02	0.02	0.00	–0.34	–0.36	0.29	0.09	0.08	0.01
Hungary	0.14	0.32	0.03	0.07	0.00	0.01	0.15	0.39	0.14	0.05	0.02	0.00
Indonesia	◊	◊	◊	◊	◊	◊	–0.04	–0.30	0.09	0.06	0.01	0.00
Iran, Islamic Rep. of	0.06	0.25	0.06	0.13	0.00	0.02	0.29	0.27	0.02	0.07	0.00	0.00
Israel	◊	◊	◊	◊	◊	◊	0.26	0.11	0.22	0.06	0.05	0.00
Italy	0.25	0.20	0.08	0.10	0.01	0.01	–0.07	0.30	0.06	0.02	0.00	0.00
Japan	0.08	0.06	0.03	0.06	0.00	0.00	–0.19	–0.02	0.09	0.04	0.01	0.00
Jordan	◊	◊	◊	◊	◊	◊	–0.16	–0.03	0.08	0.11	0.01	0.01
Kazakhstan	0.07	–0.14	0.10	0.06	0.01	0.00	◊	◊	◊	◊	◊	◊
Korea, Rep. of	◊	◊	◊	◊	◊	◊	–0.43	–0.42	0.05	0.05	0.00	0.00
Kuwait	–0.24	0.16	0.08	0.08	0.01	0.01	0.22	–0.04	0.05	0.09	0.00	0.01
Latvia	–0.16	0.15	0.03	0.03	0.00	0.00	◊	◊	◊	◊	◊	◊
Lebanon	◊	◊	◊	◊	◊	◊	0.17	0.25	0.10	0.07	0.01	0.01
Lithuania	0.25	0.09	0.05	0.02	0.00	0.00	0.24	0.27	0.10	0.05	0.01	0.00
Malaysia	◊	◊	◊	◊	◊	◊	0.00	0.21	0.09	0.18	0.01	0.03
Malta	◊	◊	◊	◊	◊	◊	0.32	–0.50	0.24	0.26	0.06	0.07
Morocco	0.12	0.18	0.11	0.10	0.01	0.01	–0.85	–0.19	0.13	0.03	0.02	0.00
Netherlands	0.11	0.20	0.15	0.05	0.02	0.00	◊	◊	◊	◊	◊	◊
New Zealand	0.12	0.10	0.06	0.02	0.00	0.00	◊	◊	◊	◊	◊	◊
Norway	–0.34	0.17	0.03	0.06	0.00	0.00	0.08	–0.20	0.06	0.04	0.00	0.00
Oman	◊	◊	◊	◊	◊	◊	0.23	–0.33	0.08	0.04	0.01	0.00
Palestinian Nat'l Auth.	◊	◊	◊	◊	◊	◊	0.24	–0.42	0.10	0.04	0.01	0.00
Qatar	0.09	–0.35	0.06	0.08	0.00	0.01	–0.64	–0.03	0.09	0.14	0.01	0.02
Romania	◊	◊	◊	◊	◊	◊	0.02	0.28	0.12	0.07	0.01	0.00
Russian Federation	0.23	–0.09	0.07	0.14	0.00	0.02	–0.01	0.10	0.07	0.03	0.00	0.00
Saudi Arabia	◊	◊	◊	◊	◊	◊	0.00	–0.76	0.08	0.17	0.01	0.03
Scotland	–0.04	0.16	0.08	0.16	0.01	0.03	–0.45	–0.15	0.42	0.19	0.18	0.04
Serbia	◊	◊	◊	◊	◊	◊	0.09	–0.13	0.03	0.03	0.00	0.00
Singapore	–0.05	–0.08	0.07	0.01	0.01	0.00	0.05	–0.25	0.25	0.11	0.06	0.01
Slovak Republic	0.09	–0.03	0.09	0.06	0.01	0.00	◊	◊	◊	◊	◊	◊
Slovenia	0.16	0.18	0.02	0.03	0.00	0.00	0.01	–0.08	0.10	0.03	0.01	0.00
Sweden	0.06	0.11	0.06	0.10	0.00	0.01	0.04	–0.17	0.05	0.06	0.00	0.00
Syrian Arab Republic	◊	◊	◊	◊	◊	◊	0.07	0.17	0.06	0.09	0.00	0.01
Thailand	◊	◊	◊	◊	◊	◊	0.09	–0.02	0.11	0.05	0.01	0.00
Tunisia	–0.18	0.10	0.04	0.04	0.00	0.00	–0.23	0.02	0.02	0.03	0.00	0.00
Turkey	◊	◊	◊	◊	◊	◊	0.20	0.05	0.07	0.05	0.00	0.00
Ukraine	–0.19	0.17	0.09	0.04	0.01	0.00	–0.02	0.06	0.04	0.02	0.00	0.00
United States	0.23	0.05	0.04	0.02	0.00	0.00	0.30	0.04	0.22	0.05	0.05	0.00
Yemen	–0.08	–0.36	0.13	0.05	0.02	0.00	◊	◊	◊	◊	◊	◊
International Median	0.07	0.08	0.06	0.06	0.00	0.00	0.01	–0.03	0.09	0.05	0.01	0.00
Benchmarking Participants												
Alberta, Canada	–0.40	–0.04	0.02	0.03	0.00	0.00	◊	◊	◊	◊	◊	◊
Basque Country, Spain	◊	◊	◊	◊	◊	◊	0.10	0.20	0.07	0.06	0.01	0.00
British Columbia, Canada	0.15	–0.09	0.05	0.10	0.00	0.01	0.34	0.29	0.15	0.05	0.02	0.00
Dubai, UAE	0.06	–0.16	0.34	0.11	0.11	0.01	0.25	0.05	0.10	0.16	0.01	0.03
Massachusetts, US	–0.15	–0.09	0.11	0.08	0.01	0.01	0.15	–0.72	0.30	0.18	0.09	0.03
Minnesota, US	0.15	0.26	0.11	0.18	0.01	0.03	0.16	–0.37	0.24	0.10	0.06	0.01
Ontario, Canada	0.17	0.22	0.05	0.10	0.00	0.01	0.33	–0.11	0.11	0.11	0.01	0.01
Quebec, Canada	0.24	–0.09	0.12	0.07	0.01	0.01	0.05	0.02	0.24	0.21	0.06	0.04

A diamond (◊) indicates the country did not participate in the assessment.

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

serious problems = 3. For mathematics, an average was computed across three statements: 1) The school building needs significant repair; 2) Classrooms are overcrowded; and 3) Teachers do not have adequate workspace outside of their classroom. For science an additional statement about the "availability of materials to conduct science experiments or investigations" was included in the index computation. Students at the high level of the index had teachers with an average score equal to 1, i.e., their teachers reported that none of the issues presented above constituted a problem. Students at the medium level had teachers with an average response value greater than 1 but less than or equal to 2. Students at the low level had teachers with an average score greater than 2.

Developed in 2007, the index is presented in Exhibit 8.9 of the *TIMSS 2007 International Mathematics Report* and Exhibit 8.10 of the *TIMSS 2007 International Science Report.* The median reliability coefficients for fourth grade mathematics and science were 0.58 and 0.62, respectively, and 0.60 and 0.66 for mathematics and science at the eighth grade. The relationship to mathematics and science achievement varied considerably across countries, perhaps reflecting the status of the teaching profession and the resources available for support. In some countries, such as El Salvador and Morocco, where teaching conditions may not be optimal, the index was positively related to achievement whereas in others (e.g., Chinese Taipei and Japan) there was no relationship. This is reflected in a relatively low international median multiple correlation between the component variables (ranging between 0.08 and 0.13) and R-square values less than 0.02, as shown in Exhibit 12.11.

Exhibit 12.12 presents the latent factor models corresponding to these indices. The models are similar for mathematics and science, except that science includes an extra statement about the availability of materials for conducting science experiments or investigations. In all models, factor loadings were strongly positive, 0.5 or greater, with the highest loading associated with the statement "teachers do not have adequate workspace outside of their classroom". For science, the RMSEA value of 0.065 indicates reasonable fit at the fourth grade but somewhat less fit at the eighth grade (0.105). For mathematics no fit statistics could be computed because the model was just identified yielding trivially perfect fit. [1]

1 A model is just-identified if all the parameters are uniquely determined because there is just enough information in the sample variance-covariance matrix (Schumacker & Lonax, 2004)

Exhibit 12.11 Index of Teachers' Adequate Working Conditions (TAWC)—Reliability and Validity Indicators

Country	Grade 4						Grade 8					
	Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables		Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science
Algeria	0.57	0.63	0.05	0.05	0.00	0.00	0.62	0.56	0.05	0.05	0.00	0.00
Armenia	0.20	0.43	0.10	0.14	0.01	0.02	0.70	0.56	0.06	0.05	0.00	0.00
Australia	0.56	0.67	0.09	0.13	0.01	0.02	0.71	0.72	0.18	0.19	0.03	0.04
Austria	0.63	0.56	0.08	0.12	0.01	0.01	◊	◊	◊	◊	◊	◊
Bahrain	◊	◊	◊	◊	◊	◊	0.60	0.70	0.13	0.18	0.02	0.03
Bosnia and Herzegovina	◊	◊	◊	◊	◊	◊	0.54	0.64	0.07	0.07	0.01	0.01
Botswana	◊	◊	◊	◊	◊	◊	0.60	0.66	0.11	0.12	0.01	0.01
Bulgaria	◊	◊	◊	◊	◊	◊	0.49	0.60	0.07	0.06	0.01	0.00
Chinese Taipei	0.72	0.78	0.03	0.06	0.00	0.00	0.78	0.71	0.11	0.12	0.01	0.01
Colombia	0.70	0.76	0.25	0.20	0.06	0.04	0.66	0.72	0.25	0.28	0.06	0.08
Cyprus	◊	◊	◊	◊	◊	◊	0.55	0.66	0.04	0.02	0.00	0.00
Czech Republic	0.22	0.43	0.07	0.08	0.01	0.01	0.29	0.49	0.11	0.14	0.01	0.02
Denmark	0.56	0.60	0.07	0.09	0.00	0.01	◊	◊	◊	◊	◊	◊
Egypt	◊	◊	◊	◊	◊	◊	0.57	0.67	0.13	0.08	0.02	0.01
El Salvador	0.58	0.63	0.21	0.31	0.04	0.10	0.51	0.65	0.17	0.20	0.03	0.04
England	0.58	0.64	0.12	0.12	0.01	0.01	0.64	0.72	0.16	0.18	0.03	0.03
Georgia	0.63	0.60	0.08	0.06	0.01	0.00	0.64	0.54	0.06	0.05	0.00	0.00
Germany	0.59	0.56	0.08	0.14	0.01	0.02	◊	◊	◊	◊	◊	◊
Ghana	◊	◊	◊	◊	◊	◊	0.65	0.53	0.08	0.17	0.01	0.03
Hong Kong SAR	0.73	0.78	0.09	0.12	0.01	0.01	0.74	0.72	0.19	0.12	0.04	0.01
Hungary	0.60	0.66	0.06	0.06	0.00	0.00	0.65	0.68	0.16	0.13	0.03	0.02
Indonesia	◊	◊	◊	◊	◊	◊	0.67	0.68	0.12	0.15	0.01	0.02
Iran, Islamic Rep. of	0.56	0.59	0.13	0.20	0.02	0.04	0.60	0.61	0.13	0.23	0.02	0.05
Israel	◊	◊	◊	◊	◊	◊	0.59	0.63	0.20	0.17	0.04	0.03
Italy	0.59	0.62	0.11	0.13	0.01	0.02	0.52	0.51	0.11	0.10	0.01	0.01
Japan	0.69	0.72	0.03	0.04	0.00	0.00	0.67	0.70	0.06	0.08	0.00	0.01
Jordan	◊	◊	◊	◊	◊	◊	0.60	0.67	0.16	0.14	0.03	0.02
Kazakhstan	0.76	0.79	0.06	0.07	0.00	0.01	◊	◊	◊	◊	◊	◊
Korea, Rep. of	◊	◊	◊	◊	◊	◊	0.52	0.60	0.11	0.09	0.01	0.01
Kuwait	0.59	0.77	0.13	0.14	0.02	0.02	0.62	0.75	0.05	0.09	0.00	0.01
Latvia	0.38	0.47	0.14	0.15	0.02	0.02	◊	◊	◊	◊	◊	◊
Lebanon	◊	◊	◊	◊	◊	◊	0.61	0.71	0.07	0.32	0.01	0.11
Lithuania	0.42	0.51	0.08	0.07	0.01	0.00	0.51	0.52	0.13	0.14	0.02	0.02
Malaysia	◊	◊	◊	◊	◊	◊	0.67	0.60	0.09	0.10	0.01	0.01
Malta	◊	◊	◊	◊	◊	◊	0.58	0.71	0.37	0.37	0.14	0.13
Morocco	0.51	0.52	0.31	0.27	0.09	0.07	0.72	0.72	0.13	0.13	0.02	0.02
Netherlands	0.64	0.64	0.08	0.11	0.01	0.01	◊	◊	◊	◊	◊	◊
New Zealand	0.47	0.56	0.07	0.10	0.01	0.01	◊	◊	◊	◊	◊	◊
Norway	0.66	0.65	0.10	0.08	0.01	0.01	0.62	0.68	0.10	0.07	0.01	0.00
Oman	◊	◊	◊	◊	◊	◊	0.72	0.75	0.10	0.13	0.01	0.02
Palestinian Nat'l Auth.	◊	◊	◊	◊	◊	◊	0.52	0.64	0.11	0.17	0.01	0.03
Qatar	0.74	0.74	0.05	0.31	0.00	0.09	0.66	0.66	0.11	0.28	0.01	0.08
Romania	◊	◊	◊	◊	◊	◊	0.52	0.55	0.09	0.13	0.01	0.02
Russian Federation	0.54	0.62	0.14	0.12	0.02	0.01	0.42	0.55	0.04	0.15	0.00	0.02
Saudi Arabia	◊	◊	◊	◊	◊	◊	–	0.70	–	0.11	–	0.01
Scotland	0.58	0.55	0.07	0.10	0.01	0.01	0.57	0.64	0.08	0.17	0.01	0.03
Serbia	◊	◊	◊	◊	◊	◊	0.59	0.69	0.09	0.10	0.01	0.01
Singapore	0.72	0.74	0.01	0.08	0.00	0.01	0.73	0.72	0.18	0.10	0.03	0.01
Slovak Republic	0.48	0.58	0.10	0.13	0.01	0.02	◊	◊	◊	◊	◊	◊
Slovenia	0.58	0.63	0.08	0.08	0.01	0.01	0.54	◊	0.06	◊	0.00	◊
Sweden	0.54	0.60	0.06	0.06	0.00	0.00	0.60	0.64	0.05	0.08	0.00	0.01
Syrian Arab Republic	◊	◊	◊	◊	◊	◊	0.54	0.59	0.12	0.10	0.02	0.01
Thailand	◊	◊	◊	◊	◊	◊	0.69	0.66	0.32	0.33	0.10	0.11
Tunisia	0.54	0.57	0.07	0.10	0.01	0.01	0.68	0.66	0.09	0.11	0.01	0.01
Turkey	◊	◊	◊	◊	◊	◊	0.51	0.64	0.15	0.26	0.02	0.07
Ukraine	0.50	0.54	0.13	0.12	0.02	0.01	0.48	0.52	0.10	0.09	0.01	0.01
United States	0.62	0.65	0.17	0.18	0.03	0.03	0.65	0.70	0.16	0.16	0.03	0.03
Yemen	0.62	0.52	0.16	0.18	0.03	0.03	◊	◊	◊	◊	◊	◊
International Median	0.58	0.62	0.08	0.12	0.01	0.01	0.60	0.66	0.11	0.13	0.01	0.02
Benchmarking Participants												
Alberta, Canada	0.54	0.64	0.09	0.14	0.01	0.02	◊	◊	◊	◊	◊	◊
Basque Country, Spain	◊	◊	◊	◊	◊	◊	0.47	0.53	0.13	0.14	0.02	0.02
British Columbia, Canada	0.66	0.67	0.04	0.05	0.00	0.00	0.59	0.61	0.08	0.07	0.01	0.01
Dubai, UAE	0.51	0.85	0.19	0.29	0.04	0.08	0.65	0.60	0.19	0.23	0.04	0.05
Massachusetts, US	0.56	0.57	0.11	0.08	0.01	0.01	0.66	0.55	0.21	0.28	0.04	0.08
Minnesota, US	0.58	0.53	0.10	0.25	0.01	0.06	0.61	0.69	0.18	0.28	0.03	0.08
Ontario, Canada	0.39	0.54	0.15	0.11	0.02	0.01	0.63	0.69	0.12	0.12	0.01	0.01
Quebec, Canada	0.63	0.59	0.05	0.05	0.00	0.00	0.67	0.57	0.17	0.22	0.03	0.05

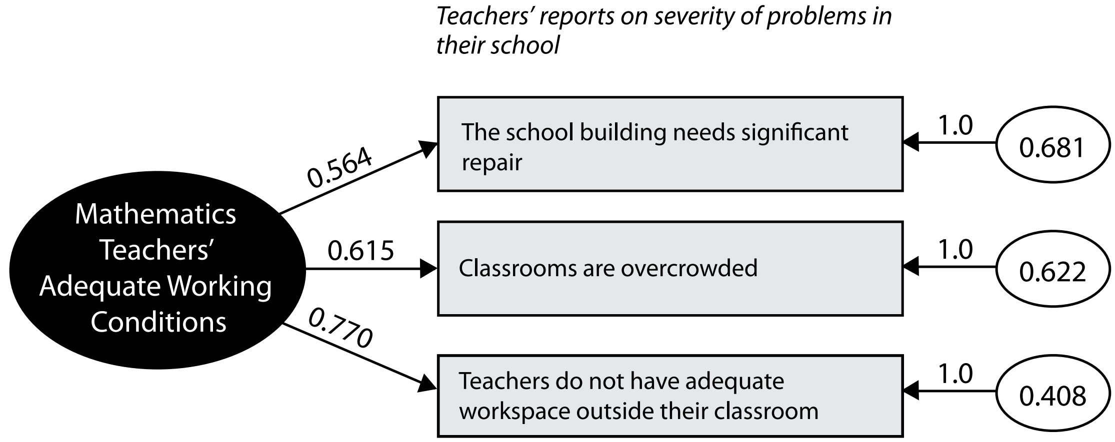
A diamond (◊) indicates the country did not participate in the assessment.

A dash (–) indicates comparable data are not available.

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

Exhibit 12.12 Latent Variable Model of Teachers' Adequate Working Conditions, Grade 4

Mathematics



Science

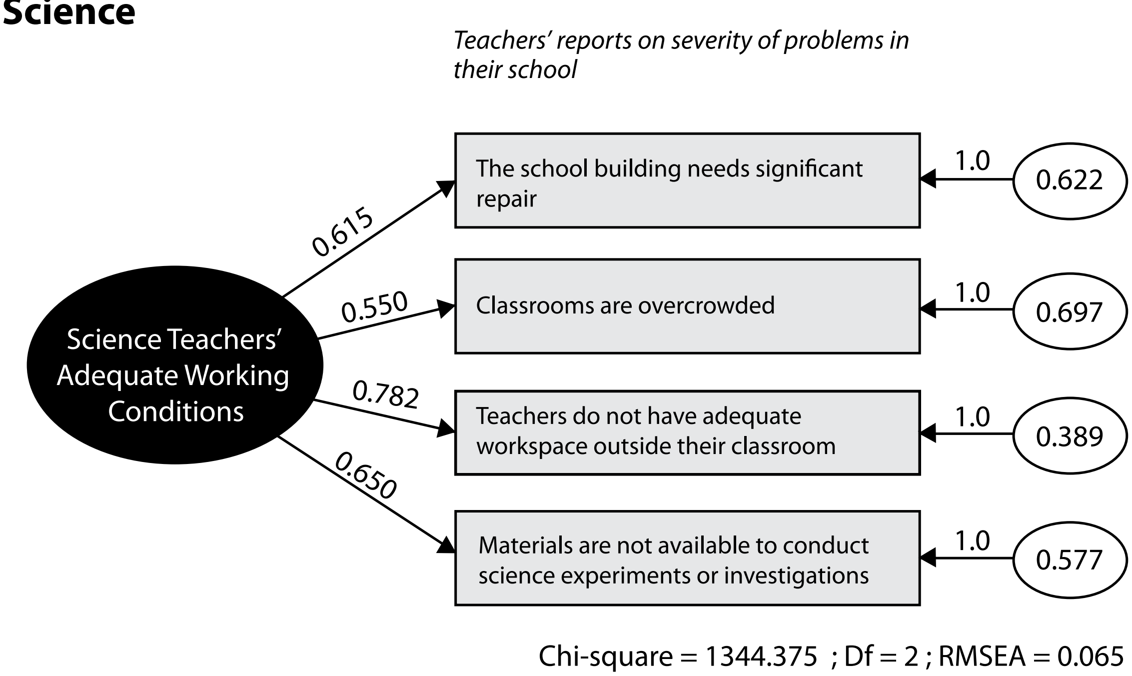
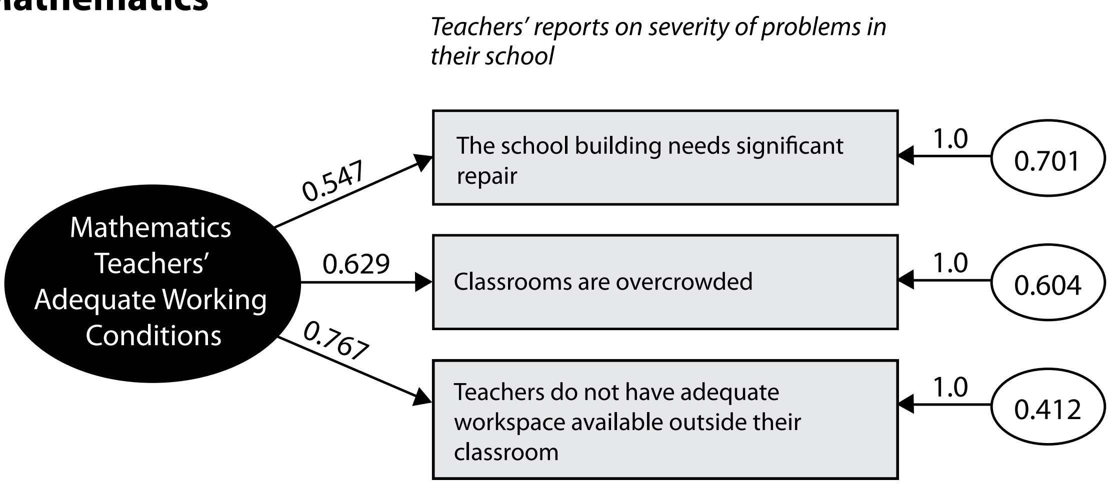
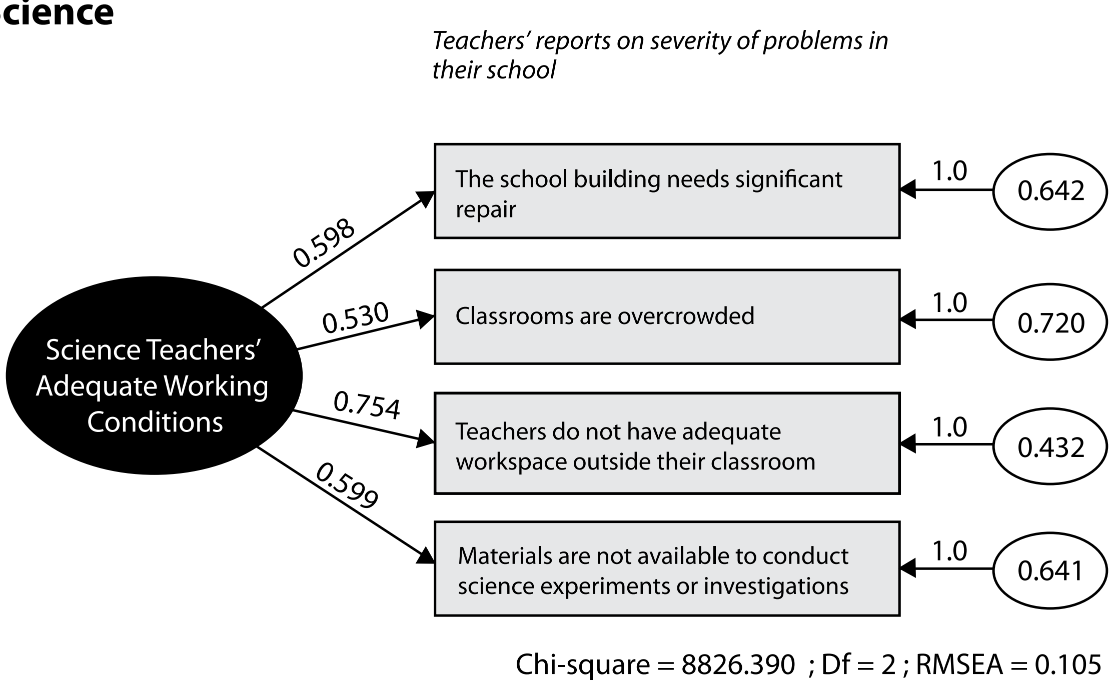


Exhibit 12.12 Latent Variable Model of Teachers' Adequate Working Conditions, Grade 8 (Continued)

Mathematics



Science



The Index of Mathematics Teachers' Perception of School Climate (TPSC) and the Index of Science Teachers Perception of School Climate (TPSC) summarize teachers' reports about their school and how supportive the climate is for learning. TIMSS asked teachers to rate their school on eight attributes:[2] 1) Teachers' job satisfaction; 2) Teachers' understanding of the school's curricular goals; 3) Teachers' degree of success in implementing the school's curriculum; 4) Teachers' expectations for student achievement; 5) Parental support for student achievement; 6) Parental involvement in school activities; 7) Students' regard for school property; and 8) Students' desire to do well in school. An average was computed across the eight items on a 5-point scale: *very high* = 1, *high* = 2, *medium* = 3, *low* = 4, and *very low* = 5. Students at the high level of the indices had teachers with an average score less than or equal to 2, meaning that they rated their school to be high or very high, on average, across the eight statements. Teacher ratings averaging greater than 2 but less than or equal to 3 corresponded to the medium level of the index, and teacher ratings greater than 3 corresponded to the low level. The index, developed in 2003, is presented in Exhibit 8.12 of the *TIMSS 2007 International Mathematics Report* and Exhibit 8.13 of the *TIMSS 2007 International Science Report,* including trends from 2003.

As shown in Exhibit 12.13, the eight components form reliable scales, with median reliability coefficients ranging from 0.81 to 0.83 for mathematics and science at fourth and eighth grades. Also, median multiple correlations between the eight statements and student achievement ranged from 0.21 to 0.23, corresponding to R-squares of 0.04 to 0.05, across the subjects and grades.

Exhibit 12.14 presents the latent factor models for the indices for mathematics and science at fourth and eighth grades. In each case, all component variables loaded relatively highly on the teachers' perception of school climate factor. Highest loadings (above 0.7) were associated with "parental support for student achievement", "parental involvement in school activities", and "teachers' degree of success in implementing the school curriculum."

2 TIMSS also asked school principals to rate their schools on these eight attributes. Indices based on principals' ratings are presented in Exhibits 12.21 and 12.22.

Exhibit 12.13 Index of Mathematics / Science Teachers' Perception of School Climate (TPSC)—Reliability and Validity Indicators

Country	Grade 4						Grade 8					
	Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables		Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science
Algeria	0.79	0.79	0.14	0.15	0.02	0.02	0.75	0.79	0.08	0.05	0.01	0.00
Armenia	0.67	0.67	0.19	0.22	0.04	0.05	0.64	0.67	0.10	0.07	0.01	0.01
Australia	0.85	0.86	0.30	0.29	0.09	0.08	0.89	0.88	0.40	0.35	0.16	0.12
Austria	0.80	0.80	0.18	0.19	0.03	0.04	◊	◊	◊	◊	◊	◊
Bahrain	◊	◊	◊	◊	◊	◊	0.84	0.77	0.16	0.15	0.03	0.02
Bosnia and Herzegovina	◊	◊	◊	◊	◊	◊	0.83	0.81	0.08	0.09	0.01	0.01
Botswana	◊	◊	◊	◊	◊	◊	0.81	0.85	0.27	0.27	0.07	0.08
Bulgaria	◊	◊	◊	◊	◊	◊	0.88	0.83	0.34	0.20	0.12	0.04
Chinese Taipei	0.82	0.86	0.16	0.10	0.02	0.01	0.85	0.86	0.22	0.24	0.05	0.06
Colombia	0.86	0.85	0.28	0.27	0.08	0.07	0.90	0.87	0.25	0.24	0.06	0.06
Cyprus	◊	◊	◊	◊	◊	◊	0.82	0.85	0.06	0.06	0.00	0.00
Czech Republic	0.75	0.77	0.13	0.14	0.02	0.02	0.80	0.77	0.25	0.21	0.06	0.04
Denmark	0.81	0.82	0.17	0.21	0.03	0.04	◊	◊	◊	◊	◊	◊
Egypt	◊	◊	◊	◊	◊	◊	0.85	0.86	0.20	0.24	0.04	0.06
El Salvador	0.85	0.87	0.16	0.17	0.02	0.03	0.85	0.81	0.16	0.23	0.03	0.05
England	0.83	0.82	0.24	0.24	0.06	0.06	0.88	0.88	0.43	0.39	0.19	0.15
Georgia	0.85	0.83	0.21	0.23	0.05	0.05	0.75	0.80	0.15	0.12	0.02	0.01
Germany	0.81	0.81	0.28	0.30	0.08	0.09	◊	◊	◊	◊	◊	◊
Ghana	◊	◊	◊	◊	◊	◊	0.68	0.77	0.27	0.26	0.08	0.07
Hong Kong SAR	0.88	0.86	0.30	0.18	0.09	0.03	0.85	0.85	0.44	0.36	0.19	0.13
Hungary	0.85	0.85	0.35	0.33	0.12	0.11	0.79	0.81	0.29	0.26	0.08	0.07
Indonesia	◊	◊	◊	◊	◊	◊	0.86	0.87	0.14	0.15	0.02	0.02
Iran, Islamic Rep. of	0.81	0.81	0.31	0.32	0.09	0.10	0.89	0.86	0.42	0.31	0.18	0.09
Israel	◊	◊	◊	◊	◊	◊	0.87	0.86	0.44	0.30	0.19	0.09
Italy	0.81	0.81	0.15	0.15	0.02	0.02	0.80	0.80	0.15	0.16	0.02	0.02
Japan	0.79	0.78	0.12	0.10	0.01	0.01	0.82	0.87	0.22	0.23	0.05	0.05
Jordan	◊	◊	◊	◊	◊	◊	0.80	0.84	0.31	0.24	0.09	0.06
Kazakhstan	0.82	0.82	0.21	0.15	0.04	0.02	◊	◊	◊	◊	◊	◊
Korea, Rep. of	◊	◊	◊	◊	◊	◊	0.81	0.81	0.18	0.13	0.03	0.02
Kuwait	0.75	0.86	0.19	0.20	0.03	0.04	0.75	0.85	0.07	0.16	0.01	0.03
Latvia	0.78	0.80	0.11	0.14	0.01	0.02	◊	◊	◊	◊	◊	◊
Lebanon	◊	◊	◊	◊	◊	◊	0.84	0.86	0.34	0.29	0.12	0.09
Lithuania	0.80	0.80	0.18	0.15	0.03	0.02	0.74	0.78	0.22	0.13	0.05	0.02
Malaysia	◊	◊	◊	◊	◊	◊	0.86	0.87	0.40	0.32	0.16	0.11
Malta	◊	◊	◊	◊	◊	◊	0.90	0.89	0.53	0.57	0.29	0.32
Morocco	0.85	0.82	0.29	0.34	0.08	0.12	0.82	0.85	0.18	0.23	0.03	0.05
Netherlands	0.76	0.76	0.26	0.25	0.07	0.06	◊	◊	◊	◊	◊	◊
New Zealand	0.83	0.83	0.29	0.30	0.09	0.09	◊	◊	◊	◊	◊	◊
Norway	0.77	0.77	0.14	0.13	0.02	0.02	0.75	0.76	0.16	0.14	0.03	0.02
Oman	◊	◊	◊	◊	◊	◊	0.79	0.75	0.22	0.25	0.05	0.06
Palestinian Nat'l Auth.	◊	◊	◊	◊	◊	◊	0.80	0.77	0.15	0.19	0.02	0.04
Qatar	0.82	0.84	0.15	0.31	0.02	0.10	0.77	0.85	0.20	0.37	0.04	0.14
Romania	◊	◊	◊	◊	◊	◊	0.89	0.83	0.26	0.19	0.07	0.03
Russian Federation	0.83	0.83	0.26	0.25	0.07	0.06	0.79	0.81	0.19	0.21	0.04	0.04
Saudi Arabia	◊	◊	◊	◊	◊	◊	0.83	0.82	0.12	0.18	0.01	0.03
Scotland	0.83	0.82	0.22	0.22	0.05	0.05	0.87	0.89	0.30	0.29	0.09	0.08
Serbia	◊	◊	◊	◊	◊	◊	0.78	0.78	0.16	0.13	0.03	0.02
Singapore	0.81	0.83	0.29	0.31	0.08	0.10	0.88	0.87	0.49	0.46	0.24	0.21
Slovak Republic	0.75	0.78	0.24	0.29	0.06	0.09	◊	◊	◊	◊	◊	◊
Slovenia	0.70	0.70	0.08	0.08	0.01	0.01	0.77	0.78	0.17	0.10	0.03	0.01
Sweden	0.77	0.76	0.21	0.25	0.04	0.06	0.79	0.79	0.16	0.16	0.03	0.03
Syrian Arab Republic	◊	◊	◊	◊	◊	◊	0.76	0.79	0.11	0.16	0.01	0.02
Thailand	◊	◊	◊	◊	◊	◊	0.83	0.87	0.25	0.33	0.06	0.11
Tunisia	0.64	0.62	0.20	0.19	0.04	0.04	0.81	0.77	0.20	0.13	0.04	0.02
Turkey	◊	◊	◊	◊	◊	◊	0.83	0.88	0.38	0.40	0.15	0.16
Ukraine	0.80	0.80	0.09	0.09	0.01	0.01	0.78	0.76	0.24	0.15	0.06	0.02
United States	0.88	0.87	0.37	0.37	0.14	0.14	0.88	0.87	0.37	0.33	0.14	0.11
Yemen	0.79	0.73	0.12	0.19	0.02	0.04	◊	◊	◊	◊	◊	◊
International Median	0.81	0.81	0.21	0.22	0.04	0.05	0.82	0.83	0.22	0.23	0.05	0.05
Benchmarking Participants												
Alberta, Canada	0.84	0.84	0.29	0.28	0.08	0.08	◊	◊	◊	◊	◊	◊
Basque Country, Spain	◊	◊	◊	◊	◊	◊	0.87	0.88	0.36	0.20	0.13	0.04
British Columbia, Canada	0.83	0.84	0.23	0.19	0.05	0.04	0.84	0.81	0.21	0.20	0.04	0.04
Dubai, UAE	0.86	0.83	0.37	0.41	0.14	0.17	0.82	0.79	0.44	0.35	0.19	0.12
Massachusetts, US	0.87	0.87	0.29	0.28	0.08	0.08	0.90	0.87	0.36	0.33	0.13	0.11
Minnesota, US	0.87	0.84	0.32	0.24	0.10	0.06	0.82	0.80	0.34	0.34	0.12	0.12
Ontario, Canada	0.81	0.84	0.28	0.25	0.08	0.06	0.86	0.85	0.32	0.30	0.10	0.09
Quebec, Canada	0.83	0.84	0.26	0.26	0.07	0.07	0.89	0.86	0.41	0.38	0.17	0.15

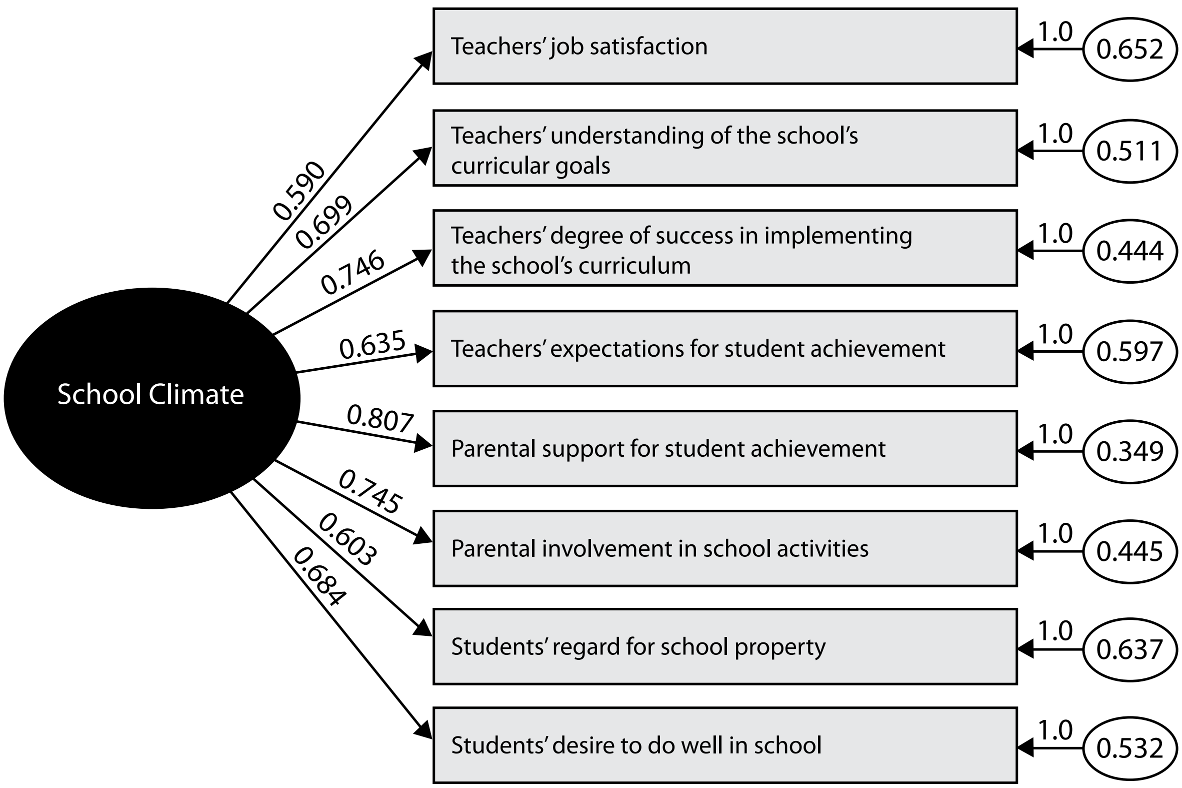
A diamond (◊) indicates the country did not participate in the assessment.

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

Exhibit 12.14 Latent Variable Model of Teachers' Perception of School Climate, Grade 4

Mathematics

How teachers characterize each of the following within their school



Chi-square = 91901.579 ; Df = 10 ; RMSEA = 0.241

Science

Factor: School Climate

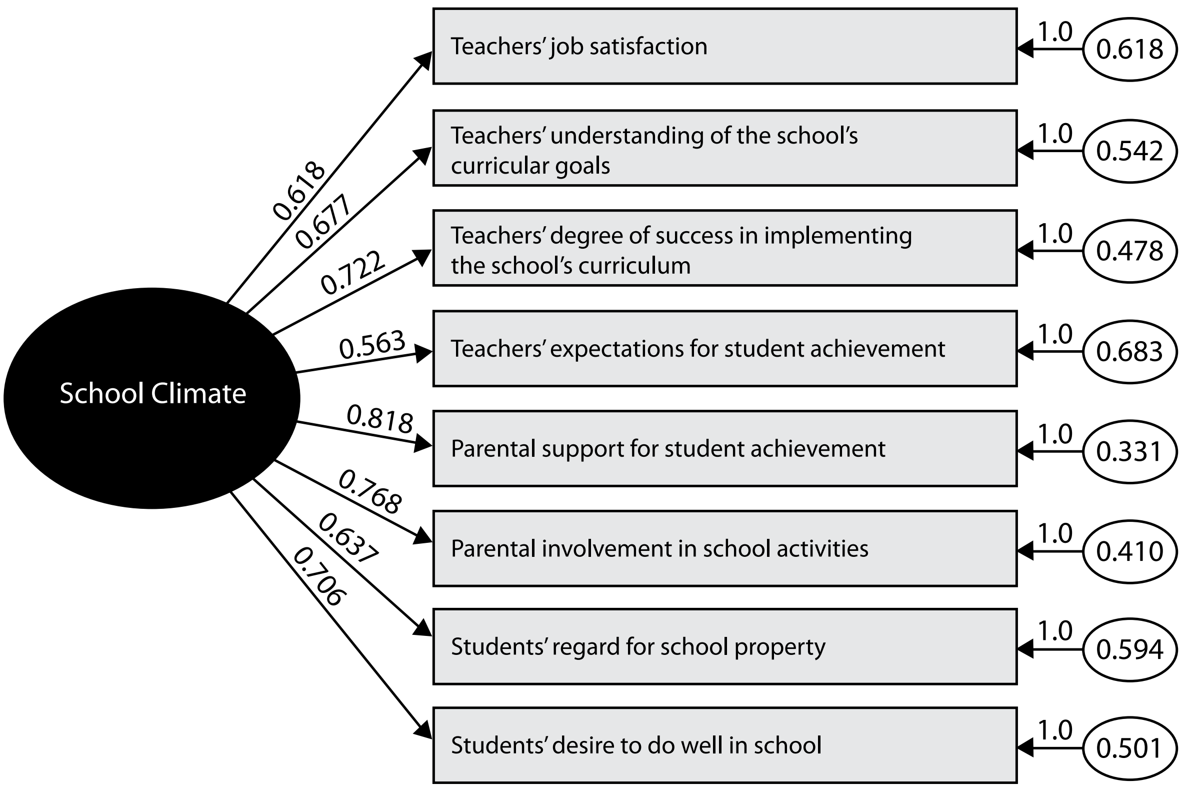
		School Climate
	Observed Variable	*Factor Loadings*
How teachers characterize each of the following within their school	Teachers' job satisfaction	0.593
	Teachers' understanding of the school's curricular goals	0.686
	Teachers' degree of success in implementing the school's curriculum	0.731
	Teachers' expectations for student achievement	0.638
	Parental support for student achievement	0.814
	Parental involvement in school activities	0.752
	Students' regard for school property	0.608
	Students' desire to do well in school	0.694

Chi-square=98349.091; Df=11; RMSEA=0.239

Exhibit 12.14 Latent Variable Model of Teachers' Perception of School Climate, Grade 8 (Continued)

Mathematics

How teachers characterize each of the following within their school



Chi-square = 160637.399 ; Df = 11 ; RMSEA = 0.259

Science

Factor: School Climate

		School Climate
	Observed Variable	*Factor Loadings*
How teachers characterize each of the following within their school	Teachers' job satisfaction	0.633
	Teachers' understanding of the school's curricular goals	0.664
	Teachers' degree of success in implementing the school's curriculum	0.713
	Teachers' expectations for student achievement	0.594
	Parental support for student achievement	0.787
	Parental involvement in school activities	0.726
	Students' regard for school property	0.627
	Students' desire to do well in school	0.699

Chi-square=258053.244; Df=11; RMSEA=0.242

The Index of Mathematics Teachers' Perception of Safety in School (TPSS) and the Index of Science Teachers' Perception of Safety in School (TPSS) summarize teachers' reports of how safe and secure they feel in their schools. The indices group students according to their teachers' responses to three statements about their school: 1) This school is located in a safe neighborhood; 2) I feel safe at this school; and 3) This school's security policies and practices are sufficient. Teachers responded on a 4-point scale: *agree a lot* = 1, *agree* = 2, *disagree* = 3, and *disagree a lot* = 4. Students were assigned to the high level of the indices if their teacher agreed with all three statements, on average (i.e., an average score of 2 or less), and to the low level if their teacher disagreed, on average, with the three statements (i.e., an average score of 3 or more). The medium level included all other response combinations. The indices, developed in 2003, are presented in Exhibit 8.13 of the *TIMSS 2007 International Mathematics Report* and Exhibit 8.14 of the *TIMSS 2007 International Science Report*.

For both subjects and at both grades, as shown in Exhibit 12.15, the three components form a reliable scale, with median reliability coefficients of 0.79 and 0.80 for mathematics and science, respectively, at the fourth grade, and 0.83 for both subjects at the eighth grade. The median multiple correlation between the three components and student achievement was 0.12 for both subjects at the fourth grade (R-square of 0.01) and 0.10 and 0.11, respectively, for mathematics and science at the eighth grade (again, R-square of 0.01, after rounding).

As shown in Exhibit 12.16, the three component variables loaded highly on the teachers' perception of safety factor at both grades and for both subjects, with all loadings above 0.8. No fit statistics could be computed because the model was just-identified, yielding trivially perfect fit. Essentially, when teachers report that they "feel safe at school", this summarizes effectively their overall perceptions of safety very well.

Exhibit 12.15 Index of Mathematics / Science Teachers' Perception of Safety in School (TPSS)—Reliability and Validity Indicators

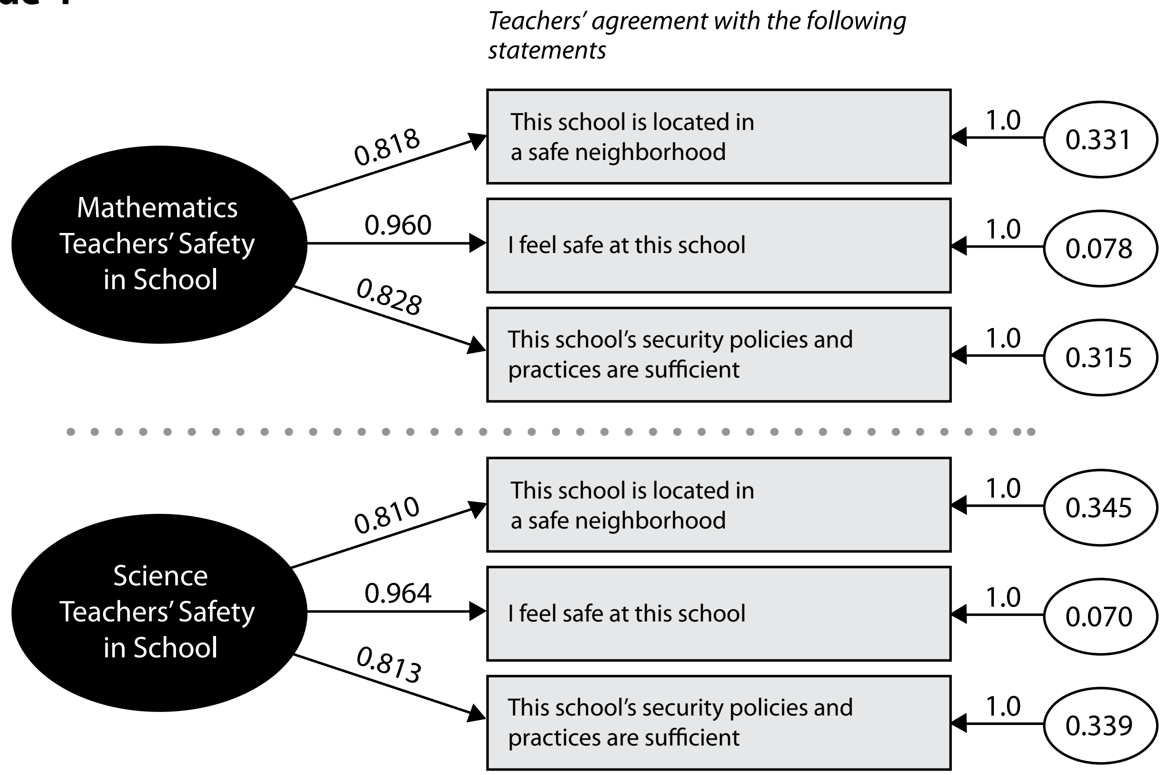
Country	Grade 4						Grade 8					
	Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables		Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science
Algeria	0.83	0.83	0.19	0.17	0.04	0.03	0.84	0.84	0.05	0.04	0.00	0.00
Armenia	0.89	0.89	0.03	0.03	0.00	0.00	0.84	0.89	0.07	0.04	0.00	0.00
Australia	0.79	0.79	0.24	0.24	0.06	0.06	0.86	0.82	0.24	0.21	0.06	0.04
Austria	0.69	0.71	0.06	0.09	0.00	0.01	◊	◊	◊	◊	◊	◊
Bahrain	◊	◊	◊	◊	◊	◊	0.90	0.80	0.07	0.05	0.00	0.00
Bosnia and Herzegovina	◊	◊	◊	◊	◊	◊	0.83	0.84	0.04	0.04	0.00	0.00
Botswana	◊	◊	◊	◊	◊	◊	0.78	0.72	0.09	0.14	0.01	0.02
Bulgaria	◊	◊	◊	◊	◊	◊	0.77	0.84	0.16	0.06	0.03	0.00
Chinese Taipei	0.83	0.84	0.09	0.03	0.01	0.00	0.80	0.79	0.09	0.07	0.01	0.01
Colombia	0.85	0.84	0.20	0.14	0.04	0.02	0.84	0.87	0.19	0.23	0.04	0.05
Cyprus	◊	◊	◊	◊	◊	◊	0.83	0.82	0.04	0.06	0.00	0.00
Czech Republic	0.74	0.71	0.03	0.01	0.00	0.00	0.75	0.79	0.09	0.05	0.01	0.00
Denmark	0.72	0.76	0.16	0.14	0.02	0.02	◊	◊	◊	◊	◊	◊
Egypt	◊	◊	◊	◊	◊	◊	0.79	0.79	0.09	0.09	0.01	0.01
El Salvador	0.88	0.87	0.11	0.13	0.01	0.02	0.86	0.80	0.06	0.14	0.00	0.02
England	0.81	0.82	0.23	0.26	0.05	0.07	0.81	0.79	0.25	0.22	0.06	0.05
Georgia	0.84	0.87	0.13	0.14	0.02	0.02	0.82	0.88	0.10	0.04	0.01	0.00
Germany	0.78	0.75	0.18	0.21	0.03	0.04	◊	◊	◊	◊	◊	◊
Ghana	◊	◊	◊	◊	◊	◊	0.79	0.83	0.20	0.23	0.04	0.05
Hong Kong SAR	0.84	0.82	0.12	0.10	0.02	0.01	0.83	0.91	0.26	0.11	0.07	0.01
Hungary	0.77	0.77	0.16	0.16	0.03	0.03	0.80	0.79	0.09	0.04	0.01	0.00
Indonesia	◊	◊	◊	◊	◊	◊	0.74	0.80	0.16	0.12	0.03	0.01
Iran, Islamic Rep. of	0.85	0.85	0.09	0.09	0.01	0.01	0.88	0.84	0.14	0.21	0.02	0.04
Israel	◊	◊	◊	◊	◊	◊	0.88	0.85	0.24	0.20	0.06	0.04
Italy	0.77	0.77	0.13	0.13	0.02	0.02	0.84	0.84	0.12	0.12	0.02	0.02
Japan	0.66	0.67	0.01	0.03	0.00	0.00	0.78	0.84	0.10	0.11	0.01	0.01
Jordan	◊	◊	◊	◊	◊	◊	0.86	0.87	0.10	0.11	0.01	0.01
Kazakhstan	0.69	0.69	0.20	0.17	0.04	0.03	◊	◊	◊	◊	◊	◊
Korea, Rep. of	◊	◊	◊	◊	◊	◊	0.80	0.75	0.10	0.05	0.01	0.00
Kuwait	0.74	0.83	0.06	0.08	0.00	0.01	0.87	0.82	0.09	0.12	0.01	0.01
Latvia	0.67	0.59	0.05	0.05	0.00	0.00	◊	◊	◊	◊	◊	◊
Lebanon	◊	◊	◊	◊	◊	◊	0.82	0.86	0.28	0.22	0.08	0.05
Lithuania	0.84	0.84	0.04	0.03	0.00	0.00	0.89	0.83	0.08	0.05	0.01	0.00
Malaysia	◊	◊	◊	◊	◊	◊	0.85	0.84	0.15	0.11	0.02	0.01
Malta	◊	◊	◊	◊	◊	◊	0.86	0.83	0.21	0.25	0.05	0.06
Morocco	0.87	0.87	0.14	0.13	0.02	0.02	0.83	0.87	0.12	0.14	0.01	0.02
Netherlands	0.88	0.88	0.22	0.23	0.05	0.05	◊	◊	◊	◊	◊	◊
New Zealand	0.78	0.78	0.30	0.32	0.09	0.10	◊	◊	◊	◊	◊	◊
Norway	0.81	0.81	0.08	0.08	0.01	0.01	0.76	0.78	0.04	0.02	0.00	0.00
Oman	◊	◊	◊	◊	◊	◊	0.76	0.80	0.07	0.16	0.00	0.02
Palestinian Nat'l Auth.	◊	◊	◊	◊	◊	◊	0.86	0.89	0.09	0.14	0.01	0.02
Qatar	0.80	0.67	0.02	0.10	0.00	0.01	0.87	0.77	0.10	0.10	0.01	0.01
Romania	◊	◊	◊	◊	◊	◊	0.81	0.79	0.14	0.09	0.02	0.01
Russian Federation	0.76	0.76	0.07	0.07	0.00	0.01	0.70	0.73	0.05	0.06	0.00	0.00
Saudi Arabia	◊	◊	◊	◊	◊	◊	0.82	0.83	0.09	0.12	0.01	0.02
Scotland	0.78	0.78	0.19	0.20	0.04	0.04	0.75	0.80	0.19	0.20	0.04	0.04
Serbia	◊	◊	◊	◊	◊	◊	0.78	0.84	0.04	0.04	0.00	0.00
Singapore	0.89	0.87	0.16	0.09	0.03	0.01	0.89	0.89	0.30	0.19	0.09	0.04
Slovak Republic	0.64	0.59	0.10	0.04	0.01	0.00	◊	◊	◊	◊	◊	◊
Slovenia	0.81	0.81	0.02	0.03	0.00	0.00	0.76	0.82	0.06	0.02	0.00	0.00
Sweden	0.77	0.77	0.16	0.19	0.02	0.04	0.76	0.72	0.11	0.10	0.01	0.01
Syrian Arab Republic	◊	◊	◊	◊	◊	◊	0.83	0.78	0.11	0.10	0.01	0.01
Thailand	◊	◊	◊	◊	◊	◊	0.85	0.84	0.16	0.15	0.03	0.02
Tunisia	0.92	0.90	0.07	0.06	0.00	0.00	0.88	0.84	0.04	0.12	0.00	0.02
Turkey	◊	◊	◊	◊	◊	◊	0.83	0.88	0.16	0.19	0.03	0.04
Ukraine	0.73	0.73	0.05	0.05	0.00	0.00	0.72	0.76	0.10	0.07	0.01	0.01
United States	0.80	0.80	0.33	0.36	0.11	0.13	0.82	0.84	0.27	0.29	0.07	0.08
Yemen	0.78	0.56	0.09	0.16	0.01	0.03	◊	◊	◊	◊	◊	◊
International Median	0.79	0.80	0.12	0.12	0.01	0.01	0.83	0.83	0.10	0.11	0.01	0.01
Benchmarking Participants												
Alberta, Canada	0.84	0.85	0.19	0.20	0.04	0.04	◊	◊	◊	◊	◊	◊
Basque Country, Spain	◊	◊	◊	◊	◊	◊	0.76	0.77	0.16	0.11	0.02	0.01
British Columbia, Canada	0.75	0.74	0.14	0.16	0.02	0.02	0.75	0.78	0.08	0.10	0.01	0.01
Dubai, UAE	0.75	0.69	0.14	0.09	0.02	0.01	0.81	0.86	0.13	0.10	0.02	0.01
Massachusetts, US	0.81	0.83	0.24	0.31	0.06	0.10	0.78	0.74	0.28	0.31	0.08	0.09
Minnesota, US	0.75	0.83	0.27	0.29	0.07	0.08	0.73	0.75	0.23	0.16	0.05	0.03
Ontario, Canada	0.81	0.82	0.13	0.13	0.02	0.02	0.72	0.80	0.17	0.20	0.03	0.04
Quebec, Canada	0.85	0.86	0.20	0.18	0.04	0.03	0.80	0.79	0.27	0.16	0.08	0.02

A diamond (◊) indicates the country did not participate in the assessment.

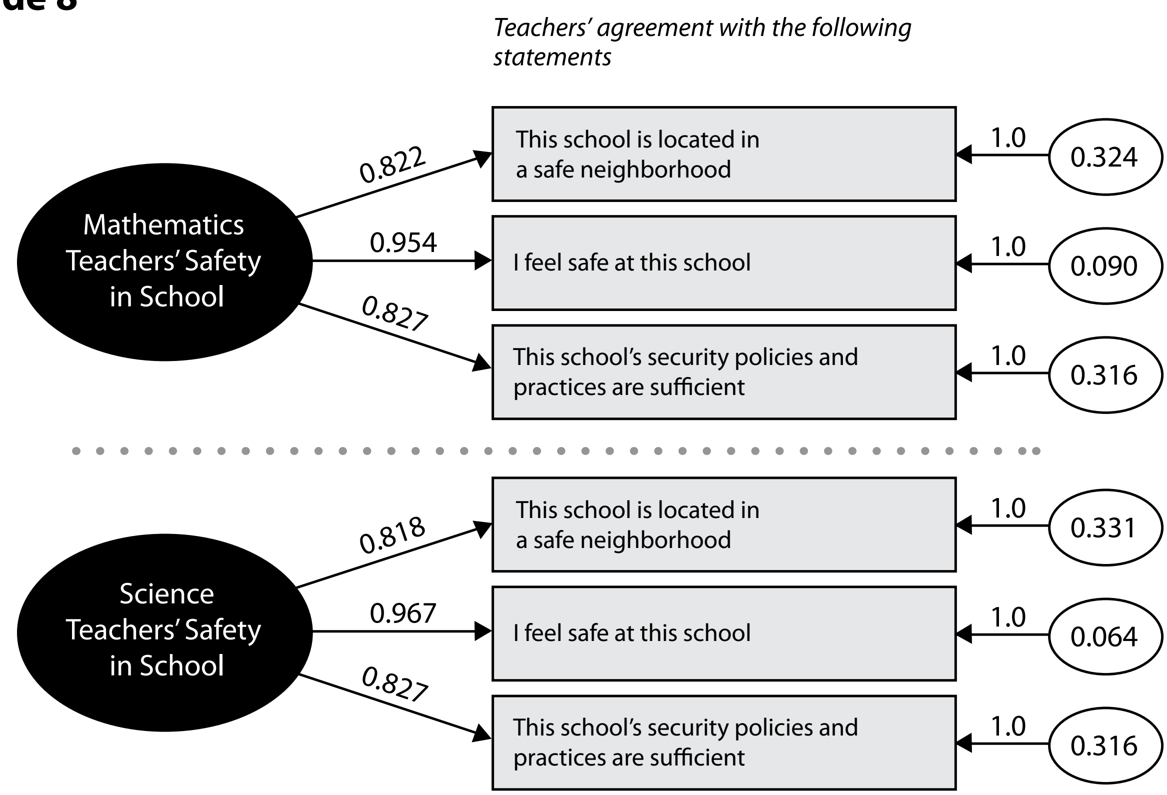
SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

Exhibit 12.16 Latent Variable Model of Teachers' Perception of Safety in School

Grade 4



Grade 8



12.4.3 School-level Indices

In the *TIMSS 2007 School Questionnaire*, school principals were asked to provide information about the school context and the resources available for mathematics and science instruction. Three indices presented in the TIMSS 2007 international reports were based on questions in the school questionnaires.

The Index of Good Attendance at School (GAS) categorizes students according to their school principals' reports on the frequency of students' absenteeism and its severity as a disruptive influence on continuity in the classroom and time for learning. The index was based on principals' reports on the frequency of occurrence (rated on a 5-point scale: 1 = *never*, 2 = *rarely*, 3 = *monthly*, 4 = *weekly*, and 5 = *daily*) and severity (rated on a 3-point scale: 1 = *not a problem*, 2 = *minor problem*, and 3 = *serious problem*) of three aspects of attendance problems: 1) Arriving late at school; 2) Absenteeism (i.e., unjustified absences); and 3) Skipping class. Students were assigned to the high level of the index if their school principal reported that all three behaviors either never occur or that they are not a serious problem. Students were assigned to the low level if their principal indicated that two or more of the behaviors were a serious problem, or two behaviors were minor problems and a third behavior a serious problem. The medium level of the indices included all other response combinations. The percentage of students at each level of the index together with achievement is presented in Exhibit 8.3 of the TIMSS 2007 international reports. Exhibit 8.4 reports the percentage of students at the high level of the index with trends from 2003 and 1999 (for eighth grade). The index, developed in 1999, was originally named Index of Good School and Class Attendance.

As shown in Exhibit 12.17, the six component variables (three addressing frequency and three addressing severity) form a fairly reliable scale, with an international median reliability coefficient of 0.76 at the fourth grade and 0.81 at the eighth grade. The median multiple correlation between the component variables and student achievement was 0.15 at the fourth grade and 0.17 at the eighth grade for both mathematics and science, corresponding to R-squares of 0.02 and 0.03.

The latent factor models presented in Exhibit 12.18 show that the index of good attendance at school may be considered as two correlated factors, one consisting of the three frequency variables and the other of the three severity variables. The correlation is higher at fourth grade than at eighth grade (0.920 compared to 0.791). In general, the component variables loaded relatively highly on the two underlying factors of frequency and severity of

Exhibit 12.17 Index of Good Attendance at School (GAS)—Reliability and Validity Indicators

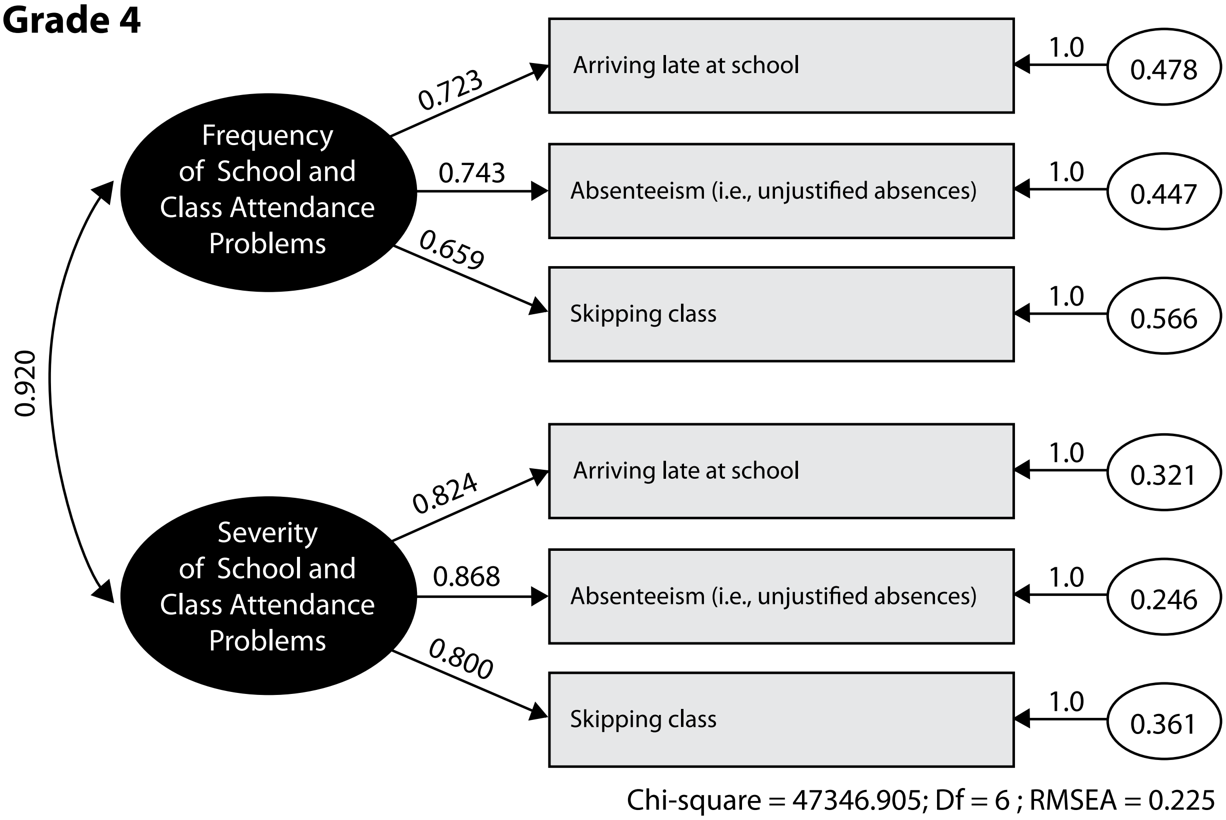
Country	Grade 4					Grade 8				
	Cronbach's Alpha Between the Component Variables	Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables		Cronbach's Alpha Between the Component Variables	Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
		Mathematics	Science	Mathematics	Science		Mathematics	Science	Mathematics	Science
Algeria	0.71	0.30	0.28	0.09	0.08	0.80	0.04	0.05	0.00	0.00
Armenia	0.75	0.13	0.14	0.02	0.02	0.73	0.15	0.18	0.02	0.03
Australia	0.69	0.29	0.27	0.09	0.08	0.86	0.44	0.39	0.19	0.15
Austria	0.69	0.11	0.12	0.01	0.01	◊	◊	◊	◊	◊
Bahrain	◊	◊	◊	◊	◊	0.77	0.18	0.28	0.03	0.08
Bosnia and Herzegovina	◊	◊	◊	◊	◊	0.86	0.11	0.11	0.01	0.01
Botswana	◊	◊	◊	◊	◊	0.83	0.24	0.23	0.06	0.05
Bulgaria	◊	◊	◊	◊	◊	0.88	0.20	0.15	0.04	0.02
Chinese Taipei	0.63	0.09	0.06	0.01	0.00	0.85	0.15	0.14	0.02	0.02
Colombia	0.80	0.25	0.23	0.06	0.05	0.86	0.24	0.23	0.06	0.05
Cyprus	◊	◊	◊	◊	◊	0.73	0.05	0.09	0.00	0.01
Czech Republic	0.77	0.13	0.13	0.02	0.02	0.78	0.19	0.18	0.04	0.03
Denmark	0.84	0.15	0.15	0.02	0.02	◊	◊	◊	◊	◊
Egypt	◊	◊	◊	◊	◊	0.70	0.13	0.12	0.02	0.02
El Salvador	0.83	0.16	0.19	0.03	0.04	0.83	0.15	0.16	0.02	0.02
England	0.79	0.24	0.25	0.06	0.06	0.87	0.35	0.36	0.13	0.13
Georgia	0.79	0.10	0.09	0.01	0.01	0.80	0.13	0.11	0.02	0.01
Germany	0.79	0.28	0.30	0.08	0.09	◊	◊	◊	◊	◊
Ghana	◊	◊	◊	◊	◊	0.74	0.23	0.24	0.05	0.06
Hong Kong SAR	0.67	0.16	0.14	0.03	0.02	0.79	0.43	0.39	0.18	0.15
Hungary	0.89	0.26	0.25	0.07	0.06	0.86	0.16	0.12	0.03	0.02
Indonesia	◊	◊	◊	◊	◊	0.72	0.25	0.24	0.06	0.06
Iran, Islamic Rep. of	0.69	0.15	0.16	0.02	0.02	0.69	0.21	0.22	0.04	0.05
Israel	◊	◊	◊	◊	◊	0.85	0.17	0.16	0.03	0.03
Italy	0.80	0.11	0.11	0.01	0.01	0.82	0.13	0.14	0.02	0.02
Japan	0.74	0.04	0.05	0.00	0.00	0.82	0.20	0.19	0.04	0.04
Jordan	◊	◊	◊	◊	◊	0.76	0.21	0.23	0.04	0.05
Kazakhstan	0.66	0.16	0.12	0.03	0.01	◊	◊	◊	◊	◊
Korea, Rep. of	◊	◊	◊	◊	◊	0.81	0.10	0.10	0.01	0.01
Kuwait	0.76	0.17	0.18	0.03	0.03	0.76	0.08	0.12	0.01	0.01
Latvia	0.69	0.14	0.13	0.02	0.02	◊	◊	◊	◊	◊
Lebanon	◊	◊	◊	◊	◊	0.70	0.28	0.27	0.08	0.07
Lithuania	0.80	0.15	0.12	0.02	0.02	0.89	0.16	0.14	0.03	0.02
Malaysia	◊	◊	◊	◊	◊	0.81	0.25	0.25	0.06	0.06
Malta	◊	◊	◊	◊	◊	0.78	0.55	0.52	0.30	0.27
Morocco	0.74	0.11	0.11	0.01	0.01	0.78	0.18	0.18	0.03	0.03
Netherlands	0.72	0.21	0.21	0.05	0.04	◊	◊	◊	◊	◊
New Zealand	0.80	0.31	0.32	0.10	0.10	◊	◊	◊	◊	◊
Norway	0.73	0.09	0.08	0.01	0.01	0.86	0.13	0.14	0.02	0.02
Oman	◊	◊	◊	◊	◊	0.76	0.09	0.10	0.01	0.01
Palestinian Nat'l Auth.	◊	◊	◊	◊	◊	0.81	0.21	0.21	0.04	0.04
Qatar	0.57	0.12	0.17	0.01	0.03	0.71	0.23	0.28	0.05	0.08
Romania	◊	◊	◊	◊	◊	0.87	0.19	0.17	0.03	0.03
Russian Federation	0.67	0.14	0.14	0.02	0.02	0.84	0.17	0.17	0.03	0.03
Saudi Arabia	◊	◊	◊	◊	◊	0.79	0.15	0.17	0.02	0.03
Scotland	0.71	0.23	0.23	0.05	0.05	0.83	0.29	0.28	0.08	0.08
Serbia	◊	◊	◊	◊	◊	0.89	0.11	0.12	0.01	0.01
Singapore	0.72	0.13	0.14	0.02	0.02	0.80	0.35	0.37	0.12	0.13
Slovak Republic	0.84	0.15	0.17	0.02	0.03	◊	◊	◊	◊	◊
Slovenia	0.71	0.08	0.08	0.01	0.01	0.89	0.09	0.09	0.01	0.01
Sweden	0.76	0.15	0.16	0.02	0.03	0.86	0.11	0.11	0.01	0.01
Syrian Arab Republic	◊	◊	◊	◊	◊	0.79	0.15	0.14	0.02	0.02
Thailand	◊	◊	◊	◊	◊	0.78	0.17	0.15	0.03	0.02
Tunisia	0.79	0.24	0.24	0.06	0.06	0.75	0.08	0.07	0.01	0.01
Turkey	◊	◊	◊	◊	◊	0.81	0.17	0.17	0.03	0.03
Ukraine	0.76	0.17	0.15	0.03	0.02	0.86	0.18	0.18	0.03	0.03
United States	0.76	0.21	0.24	0.05	0.06	0.85	0.23	0.25	0.05	0.06
Yemen	0.76	0.15	0.15	0.02	0.02	◊	◊	◊	◊	◊
International Median	0.76	0.15	0.15	0.02	0.02	0.81	0.17	0.17	0.03	0.03
Benchmarking Participants										
Alberta, Canada	0.79	0.18	0.17	0.03	0.03	◊	◊	◊	◊	◊
Basque Country, Spain	◊	◊	◊	◊	◊	0.76	0.23	0.21	0.05	0.05
British Columbia, Canada	0.75	0.20	0.18	0.04	0.03	0.87	0.22	0.20	0.05	0.04
Dubai, UAE	0.81	0.19	0.13	0.04	0.02	0.76	0.27	0.23	0.07	0.05
Massachusetts, US	0.64	0.14	0.15	0.02	0.02	0.78	0.33	0.29	0.11	0.09
Minnesota, US	0.65	0.21	0.23	0.05	0.05	0.81	0.23	0.20	0.05	0.04
Ontario, Canada	0.79	0.22	0.20	0.05	0.04	0.82	0.12	0.11	0.01	0.01
Quebec, Canada	0.76	0.16	0.15	0.03	0.02	0.88	0.36	0.32	0.13	0.10

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

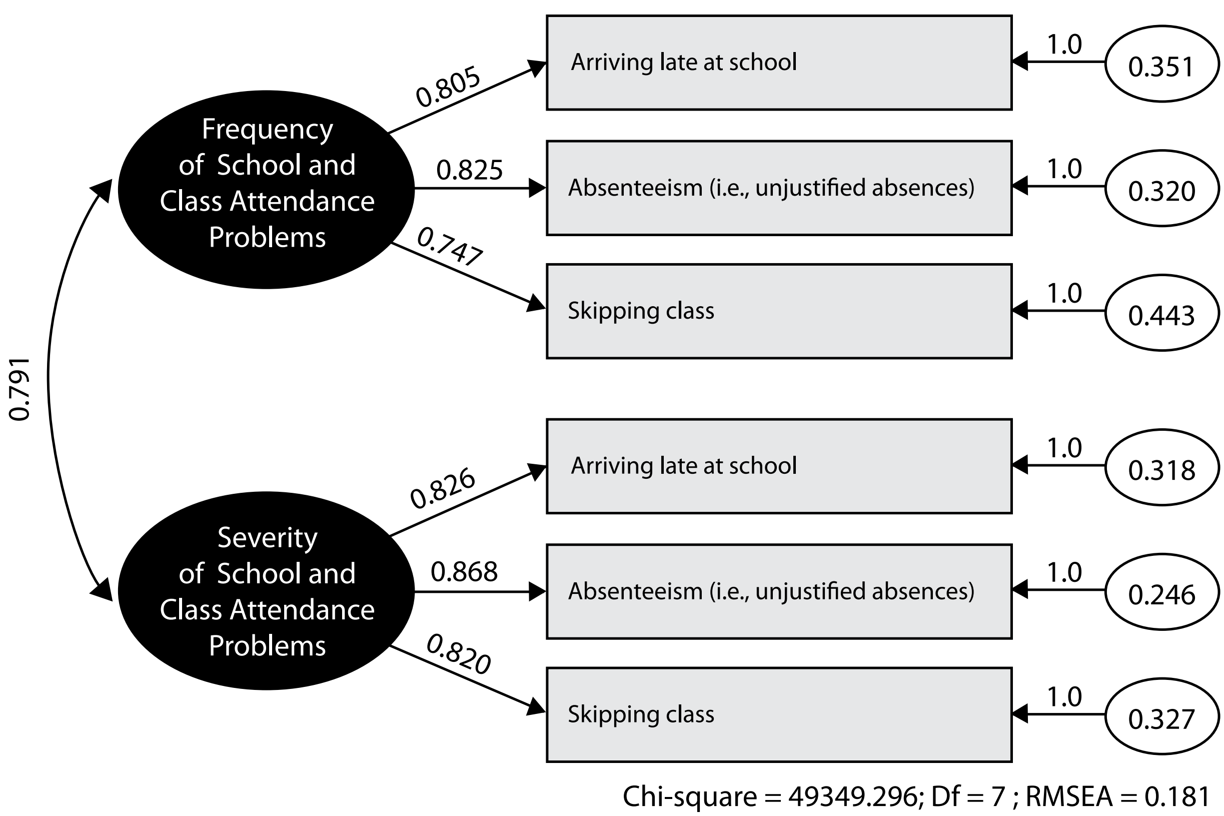
A diamond (◊) indicates the country did not participate in the assessment.

Exhibit 12.18 Latent Variable Model of Good Attendance at School

Grade 4



Grade 8



class attendance problems, with the somewhat higher loadings associated with the severity factor.

The Index of Availability of School Resources for Mathematics Instruction (ASRMI) and the Index of Availability of School Resources for Science Instruction (ASRSI) categorize students according to their principals' reports of the extent to which their schools' capacity to provide instruction is impacted by a lack of important resources. The index is based on principals' responses to a series of questions about shortages affecting schools' general capacity to provide instruction, and to provide mathematics and science instruction in particular.

Five areas where shortages or inadequacies could affect the school's general capacity to provide instruction were included in the index computation for both subjects: 1) Instructional materials (e.g., textbook); 2) Budget for supplies (e.g., paper, pencils); 3) School buildings and grounds; 4) Heating/cooling and lighting systems; and 5) Instructional space (e.g., classrooms). To make the index for mathematics, these were combined with five areas where shortages or inadequacies could affect the school's capacity to provide instruction in mathematics specifically: 1) Computers for mathematics instruction; 2) Computer software for mathematics instruction; 3) Calculators for mathematics instruction; 4) Library materials relevant to mathematics instruction; and 5) Audio-visual resources for mathematics instruction. Similarly, to make the indices for science, the five general areas were combined with six areas where shortages or inadequacies could affect the school's capacity to provide instruction in science: 1) Science laboratory equipment and materials; 2) Computers for science instruction; 3) Computer software for science instruction; 4) Calculators for science instruction; 5) Library materials relevant to science instruction; and 6) Audio-visual resources for science instruction. School principals rated each area on a 4-point scale: *none* = 1, *a little* = 2, *some* = 3, and *a lot* = 4. Students were assigned to the high level of the indices if their school principals reported that their school's capacity to provide instruction was not affected or affected only a little, on average, by shortages in both general and subject-specific areas (i.e., an average rating of less than 2 on both sets). Students at the low level had principals with average ratings greater than or equal to 3. The medium level included all other combinations of ratings.

The percentage of students at each level of the index together with achievement is presented in Exhibit 8.7 of the TIMSS 2007 international reports. Exhibit 8.8 reports the percentage of students at the high level of the index with trends from 2003, 1999 (for eighth grade) and 1995.

As shown in Exhibit 12.19, the components form reliable scales, with median reliability coefficients of 0.85 and 0.86 for mathematics and science at the fourth grade, and 0.84 and 0.85, respectively, at the eighth grade. The median multiple correlation between the statements and student achievement ranged from 0.16 to 0.18 across the subjects and grades, corresponding to an R-square value of about 0.03.

The factor loadings presented in Exhibit 12.20 all are strongly positive (0.6 or greater). Loadings for the mathematics- and science-specific areas were somewhat higher than for the general areas. For example, for fourth grade mathematics, loadings for the mathematics-specific areas ranged from 0.790 to 0.906, compared to a range of from 0.618 to 0.726 for the general areas. With a RMSEA value above 0.2 indicating not good fit for the single factor model, it may be useful to explore a two-factor model in the future, incorporating a general resource factor and a subject specific resource factor.

The Index of Principals' Perception of School Climate (PPSC) summarizes school principals' perceptions of their school's climate. This index is based on the same eight statements rated by teachers and reported in Exhibits 12.13 and 12.14. These were: 1) Teachers' job satisfaction; 2) Teachers' understanding of the school's curricular goals; 3) Teachers' degree of success in implementing the school's curriculum; 4) Teachers' expectations for student achievement; 5) Parental support for student achievement; 6) Parental involvement in school activities; and 7) Students' regard for school property; and 8) Students' desire to do well in school. Principals rated each attribute of their school on a 5-point scale: *very high* = 1, *high* = 2, *medium* = 3; *low* = 4; and *very low* = 5. Students were assigned to the high level of the index if their school principal rated each attribute as at least high, on average (i.e., an average rating of less than or equal to 2). The medium level of the index corresponds to an average rating greater than 2 but less than or equal to 3. The low level corresponds to an average rating of greater than 3 (i.e., ratings of low or very low, on average). The index, developed in 2003, is presented in Exhibit 8.11 of the *TIMSS 2007 International Mathematics Report* and Exhibit 8.12 of the *TIMSS 2007 International Science Report*, including trends from 2003.

Exhibit 12.19 Index of Availability of School Resources for Mathematics (ASRMI) / Science (ASRSI) Instruction—Reliability and Validity Indicators

Country	Grade 4						Grade 8					
	Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables		Cronbach's Alpha Between the Component Variables		Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science	Mathematics	Science
Algeria	0.80	0.84	0.22	0.18	0.05	0.03	0.74	0.79	0.07	0.11	0.01	0.01
Armenia	0.79	0.79	0.12	0.20	0.02	0.04	0.77	0.78	0.12	0.14	0.02	0.02
Australia	0.84	0.86	0.17	0.17	0.03	0.03	0.87	0.90	0.34	0.34	0.12	0.12
Austria	0.84	0.84	0.13	0.15	0.02	0.02	◊	◊	◊	◊	◊	◊
Bahrain	◊	◊	◊	◊	◊	◊	0.82	0.83	0.16	0.26	0.02	0.07
Bosnia and Herzegovina	◊	◊	◊	◊	◊	◊	0.83	0.85	0.13	0.17	0.02	0.03
Botswana	◊	◊	◊	◊	◊	◊	0.81	0.85	0.18	0.16	0.03	0.03
Bulgaria	◊	◊	◊	◊	◊	◊	0.83	0.85	0.34	0.23	0.12	0.05
Chinese Taipei	0.89	0.91	0.10	0.11	0.01	0.01	0.90	0.91	0.09	0.09	0.01	0.01
Colombia	0.88	0.91	0.38	0.35	0.14	0.12	0.92	0.92	0.27	0.24	0.07	0.06
Cyprus	◊	◊	◊	◊	◊	◊	0.80	0.87	0.07	0.09	0.01	0.01
Czech Republic	0.70	0.77	0.13	0.11	0.02	0.01	0.79	0.84	0.18	0.16	0.03	0.03
Denmark	0.85	0.86	0.14	0.16	0.02	0.03	◊	◊	◊	◊	◊	◊
Egypt	◊	◊	◊	◊	◊	◊	0.84	0.85	0.19	0.15	0.04	0.02
El Salvador	0.88	0.90	0.23	0.27	0.06	0.07	0.88	0.89	0.27	0.30	0.08	0.09
England	0.85	0.88	0.12	0.16	0.01	0.02	0.87	0.90	0.14	0.19	0.02	0.04
Georgia	0.82	0.83	0.20	0.20	0.04	0.04	0.81	0.80	0.16	0.17	0.03	0.03
Germany	0.86	0.87	0.14	0.15	0.02	0.02	◊	◊	◊	◊	◊	◊
Ghana	◊	◊	◊	◊	◊	◊	0.79	0.84	0.18	0.24	0.03	0.06
Hong Kong SAR	0.87	0.89	0.16	0.17	0.02	0.03	0.85	0.87	0.27	0.22	0.07	0.05
Hungary	0.86	0.88	0.18	0.24	0.03	0.06	0.85	0.88	0.20	0.13	0.04	0.02
Indonesia	◊	◊	◊	◊	◊	◊	0.88	0.89	0.27	0.31	0.07	0.09
Iran, Islamic Rep. of	0.80	0.81	0.25	0.24	0.06	0.06	0.77	0.81	0.28	0.23	0.08	0.05
Israel	◊	◊	◊	◊	◊	◊	0.87	0.88	0.31	0.30	0.10	0.09
Italy	0.85	0.85	0.16	0.15	0.02	0.02	0.81	0.84	0.17	0.17	0.03	0.03
Japan	0.90	0.91	0.07	0.09	0.01	0.01	0.87	0.89	0.17	0.15	0.03	0.02
Jordan	◊	◊	◊	◊	◊	◊	0.84	0.84	0.16	0.24	0.03	0.06
Kazakhstan	0.88	0.89	0.23	0.26	0.05	0.07	◊	◊	◊	◊	◊	◊
Korea, Rep. of	◊	◊	◊	◊	◊	◊	0.87	0.89	0.10	0.10	0.01	0.01
Kuwait	0.78	0.79	0.18	0.19	0.03	0.04	0.82	0.80	0.12	0.13	0.02	0.02
Latvia	0.77	0.80	0.12	0.15	0.01	0.02	◊	◊	◊	◊	◊	◊
Lebanon	◊	◊	◊	◊	◊	◊	0.78	0.86	0.36	0.39	0.13	0.15
Lithuania	0.85	0.85	0.16	0.12	0.02	0.01	0.80	0.83	0.11	0.14	0.01	0.02
Malaysia	◊	◊	◊	◊	◊	◊	0.93	0.94	0.18	0.16	0.03	0.02
Malta	◊	◊	◊	◊	◊	◊	0.84	0.85	0.31	0.34	0.10	0.11
Morocco	0.89	0.91	0.28	0.33	0.08	0.11	0.79	0.82	0.20	0.21	0.04	0.04
Netherlands	0.82	0.80	0.14	0.12	0.02	0.01	◊	◊	◊	◊	◊	◊
New Zealand	0.88	0.88	0.16	0.17	0.03	0.03	◊	◊	◊	◊	◊	◊
Norway	0.83	0.86	0.13	0.14	0.02	0.02	0.82	0.85	0.12	0.12	0.01	0.01
Oman	◊	◊	◊	◊	◊	◊	0.87	0.88	0.19	0.22	0.03	0.05
Palestinian Nat'l Auth.	◊	◊	◊	◊	◊	◊	0.84	0.86	0.23	0.17	0.05	0.03
Qatar	0.78	0.78	0.20	0.32	0.04	0.10	0.88	0.90	0.27	0.44	0.08	0.19
Romania	◊	◊	◊	◊	◊	◊	0.82	0.86	0.15	0.17	0.02	0.03
Russian Federation	0.88	0.89	0.20	0.22	0.04	0.05	0.84	0.85	0.15	0.16	0.02	0.03
Saudi Arabia	◊	◊	◊	◊	◊	◊	0.83	0.84	0.14	0.14	0.02	0.02
Scotland	0.84	0.84	0.17	0.18	0.03	0.03	0.82	0.85	0.20	0.22	0.04	0.05
Serbia	◊	◊	◊	◊	◊	◊	0.85	0.88	0.19	0.17	0.04	0.03
Singapore	0.90	0.90	0.12	0.13	0.01	0.02	0.69	0.73	0.19	0.11	0.03	0.01
Slovak Republic	0.85	0.84	0.23	0.22	0.05	0.05	◊	◊	◊	◊	◊	◊
Slovenia	0.84	0.86	0.09	0.10	0.01	0.01	0.86	0.88	0.14	0.14	0.02	0.02
Sweden	0.84	0.87	0.16	0.15	0.03	0.02	0.85	0.85	0.07	0.10	0.01	0.01
Syrian Arab Republic	◊	◊	◊	◊	◊	◊	0.74	0.76	0.18	0.21	0.03	0.04
Thailand	◊	◊	◊	◊	◊	◊	0.90	0.92	0.24	0.29	0.06	0.08
Tunisia	0.81	0.84	0.25	0.26	0.06	0.07	0.84	0.84	0.20	0.18	0.04	0.03
Turkey	◊	◊	◊	◊	◊	◊	0.82	0.86	0.27	0.25	0.08	0.06
Ukraine	0.85	0.84	0.17	0.18	0.03	0.03	0.84	0.84	0.20	0.20	0.04	0.04
United States	0.87	0.88	0.17	0.21	0.03	0.04	0.89	0.90	0.22	0.19	0.05	0.04
Yemen	0.91	0.92	0.17	0.25	0.03	0.06	◊	◊	◊	◊	◊	◊
International Median	0.85	0.86	0.16	0.17	0.03	0.03	0.84	0.85	0.18	0.17	0.03	0.03
Benchmarking Participants												
Alberta, Canada	0.89	0.91	0.15	0.14	0.02	0.02	◊	◊	◊	◊	◊	◊
Basque Country, Spain	◊	◊	◊	◊	◊	◊	0.89	0.91	0.15	0.13	0.02	0.02
British Columbia, Canada	0.81	0.83	0.14	0.17	0.02	0.03	0.89	0.89	0.11	0.12	0.01	0.02
Dubai, UAE	0.86	0.87	0.28	0.33	0.08	0.11	0.89	0.92	0.44	0.39	0.20	0.15
Massachusetts, US	0.90	0.92	0.18	0.24	0.03	0.06	0.87	0.89	0.39	0.34	0.15	0.12
Minnesota, US	0.83	0.85	0.18	0.17	0.03	0.03	0.88	0.89	0.25	0.22	0.06	0.05
Ontario, Canada	0.86	0.87	0.20	0.22	0.04	0.05	0.84	0.86	0.23	0.22	0.05	0.05
Quebec, Canada	0.78	0.82	0.09	0.06	0.01	0.00	0.82	0.87	0.29	0.27	0.08	0.07

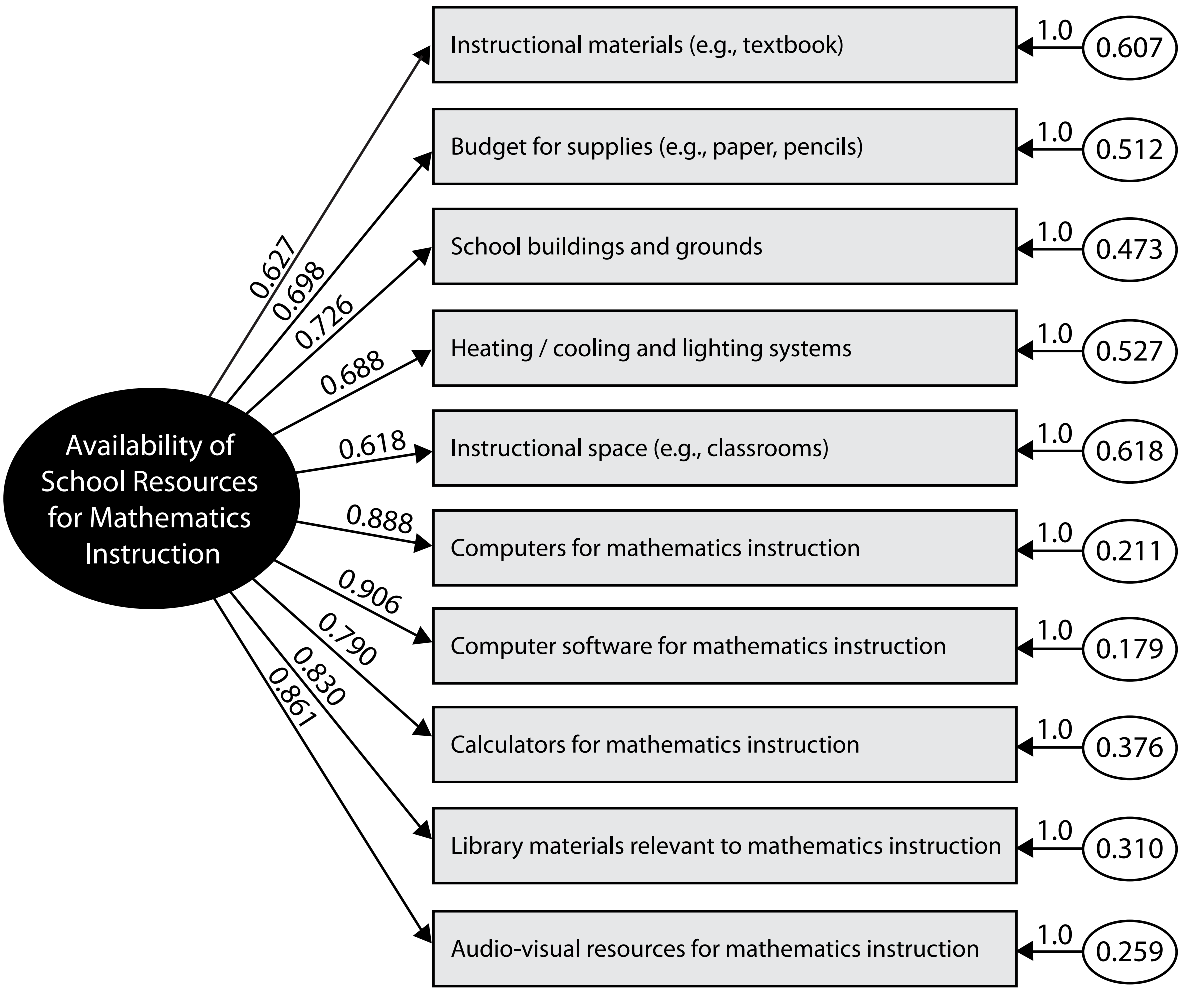
A diamond (◊) indicates the country did not participate in the assessment.

SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

Exhibit 12.20 Latent Variable Model of Availability of School Resources for Instruction, Grade 4

Mathematics

Principals' reports on shortage or inadequacy of…



Chi-square = 105955.037; Df = 15 ; RMSEA = 0.212

Science

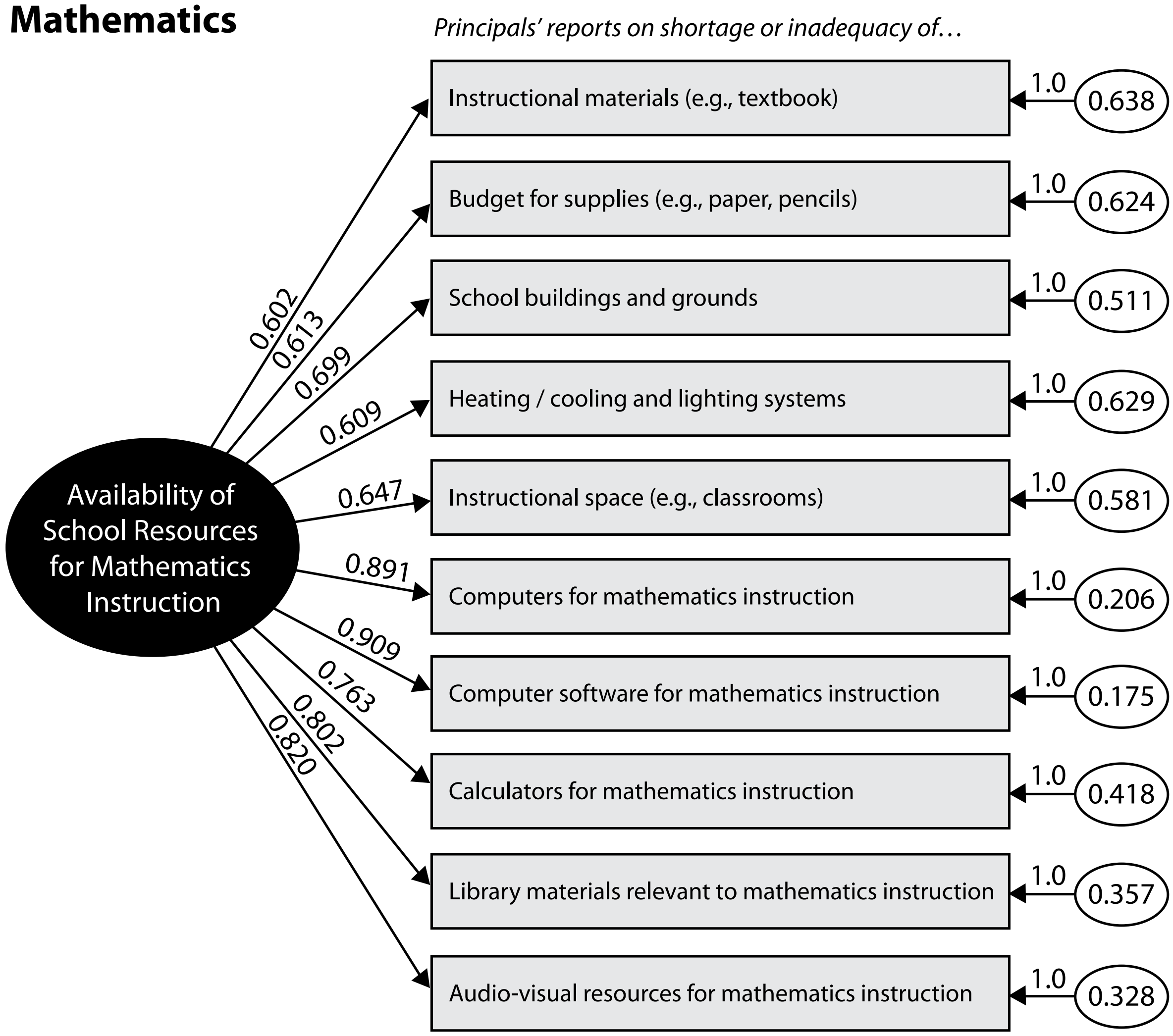
Factor: Availability of School Resources for Science Instruction

		Availability of School Resources for Science Instruction
	Observed Variable	*Factor Loadings*
Principals' reports on shortage or inadequacy of…	Instructional materials (e.g., textbook)	0.625
	Budget for supplies (e.g., paper, pencils)	0.683
	School buildings and grounds	0.711
	Heating / cooling and lighting systems	0.668
	Instructional space (e.g., classrooms)	0.603
	Science laboratory equipment and materials	0.734
	Computers for science instruction	0.902
	Computer software for science instruction	0.907
	Calculators for science instruction	0.764
	Library materials relevant to science instruction	0.845
	Audio-visual resources for science instruction	0.878

Chi-square= 114192.107; Df= 18; RMSEA= 0.201

Exhibit 12.20 Latent Variable Model of Availability of School Resources for Instruction, Grade 8 (Continued)

Mathematics



Chi-square = 189839.382; Df = 15 ; RMSEA = 0.242

Science

Factor: Availability of School Resources for Science Instruction

		Availability of School Resources for Science Instruction
Principals' reports on shortage or inadequacy of…	*Observed Variable*	*Factor Loadings*
	Instructional materials (e.g., textbook)	0.632
	Budget for supplies (e.g., paper, pencils)	0.635
	School buildings and grounds	0.688
	Heating / cooling and lighting systems	0.604
	Instructional space (e.g., classrooms)	0.638
	Science laboratory equipment and materials	0.696
	Computers for science instruction	0.911
	Computer software for science instruction	0.917
	Calculators for science instruction	0.740
	Library materials relevant to science instruction	0.816
	Audio-visual resources for science instruction	0.829

Chi-square= 199749.218; Df= 19; RMSEA= 0.221

As shown in Exhibit 12.21, the eight components form a reliable scale, with median reliability coefficients across countries of 0.79 for fourth grade and 0.81 for eighth grade. The median multiple correlation between the attributes and student achievement was 0.20 and 0.21 for mathematics and science, respectively, at the fourth grade, and 0.23 and 0.22, respectively, at the eighth grade, corresponding to R-square values of between 0.04 and 0.05.

As shown in Exhibit 12.22 all component variables loaded relatively highly on the school climate factors. Similar to the teacher perception factors (Exhibit 12.14) the highest loadings were associated with "parental support for student achievement". "Teachers' degree of success in implementing the school's curriculum" and "students' desire to do well in school" also loaded relatively highly on the underlying factors.

Exhibit 12.21 Index of Principals' Perception of School Climate (PPSC)—Reliability and Validity Indicators

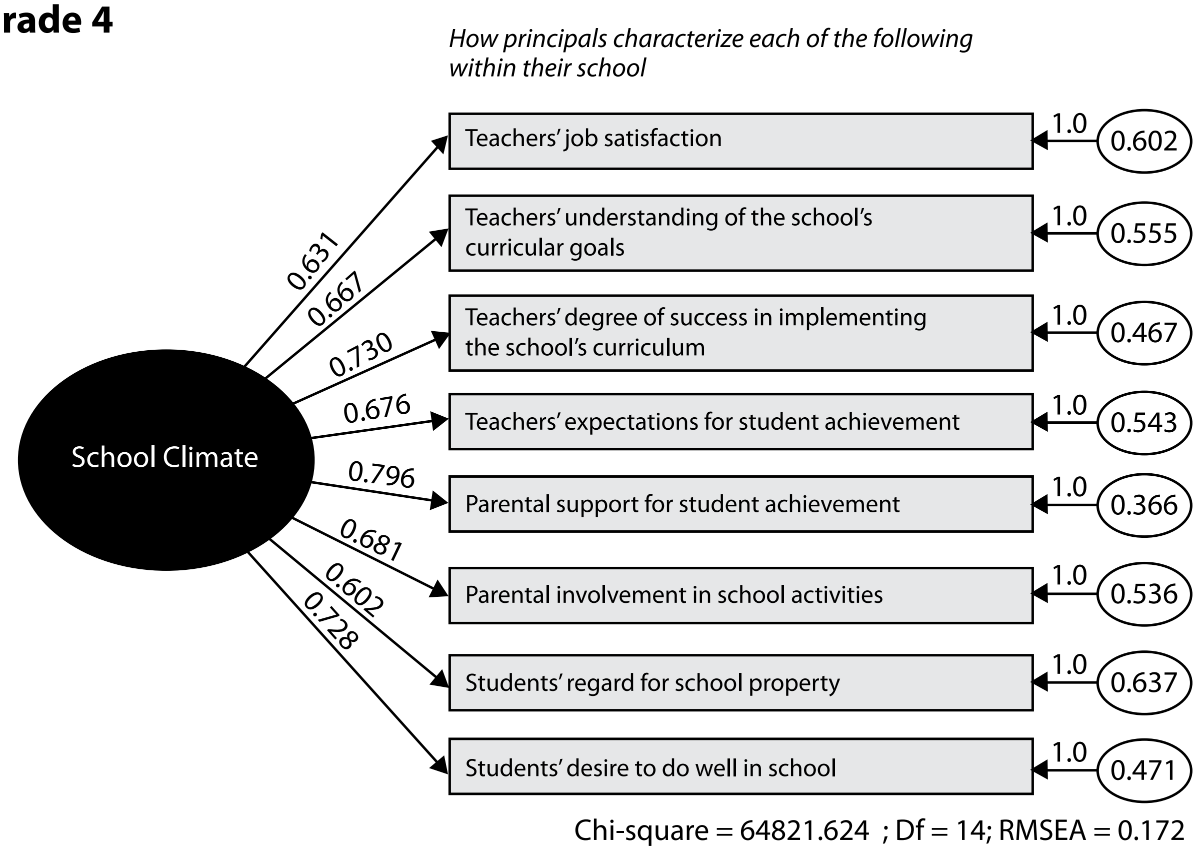
Country	Grade 4					Grade 8				
	Cronbach's Alpha Between the Component Variables	Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables		Cronbach's Alpha Between the Component Variables	Multiple R Between Student Achievement and Component Variables		Percent of Variance in Student Achievement Accounted for by the Component Variables	
		Mathematics	Science	Mathematics	Science		Mathematics	Science	Mathematics	Science
Algeria	0.79	0.15	0.16	0.02	0.03	0.82	0.08	0.08	0.01	0.01
Armenia	0.71	0.17	0.19	0.03	0.04	0.70	0.10	0.13	0.01	0.02
Australia	0.83	0.30	0.27	0.09	0.07	0.88	0.45	0.42	0.20	0.18
Austria	0.73	0.15	0.18	0.02	0.03	◊	◊	◊	◊	◊
Bahrain	◊	◊	◊	◊	◊	0.76	0.23	0.27	0.05	0.08
Bosnia and Herzegovina	◊	◊	◊	◊	◊	0.76	0.13	0.12	0.02	0.01
Botswana	◊	◊	◊	◊	◊	0.75	0.21	0.21	0.04	0.04
Bulgaria	◊	◊	◊	◊	◊	0.83	0.37	0.24	0.13	0.06
Chinese Taipei	0.84	0.13	0.12	0.02	0.01	0.85	0.19	0.18	0.04	0.03
Colombia	0.85	0.23	0.21	0.05	0.04	0.86	0.25	0.24	0.06	0.06
Cyprus	◊	◊	◊	◊	◊	0.77	0.06	0.09	0.00	0.01
Czech Republic	0.59	0.19	0.16	0.04	0.03	0.72	0.26	0.22	0.07	0.05
Denmark	0.82	0.17	0.16	0.03	0.03	◊	◊	◊	◊	◊
Egypt	◊	◊	◊	◊	◊	0.83	0.21	0.19	0.04	0.04
El Salvador	0.83	0.19	0.22	0.04	0.05	0.84	0.19	0.20	0.04	0.04
England	0.87	0.22	0.24	0.05	0.06	0.86	0.29	0.29	0.09	0.08
Georgia	0.77	0.17	0.16	0.03	0.03	0.78	0.20	0.16	0.04	0.03
Germany	0.75	0.24	0.24	0.06	0.06	◊	◊	◊	◊	◊
Ghana	◊	◊	◊	◊	◊	0.79	0.29	0.30	0.09	0.09
Hong Kong SAR	0.81	0.17	0.13	0.03	0.02	0.84	0.42	0.36	0.18	0.13
Hungary	0.82	0.33	0.30	0.11	0.09	0.80	0.32	0.30	0.10	0.09
Indonesia	◊	◊	◊	◊	◊	0.86	0.21	0.20	0.04	0.04
Iran, Islamic Rep. of	0.78	0.27	0.28	0.07	0.08	0.81	0.39	0.40	0.16	0.16
Israel	◊	◊	◊	◊	◊	0.79	0.30	0.28	0.09	0.08
Italy	0.76	0.13	0.11	0.02	0.01	0.73	0.17	0.17	0.03	0.03
Japan	0.77	0.12	0.09	0.01	0.01	0.81	0.27	0.25	0.07	0.06
Jordan	◊	◊	◊	◊	◊	0.81	0.25	0.26	0.06	0.07
Kazakhstan	0.86	0.24	0.22	0.06	0.05	◊	◊	◊	◊	◊
Korea, Rep. of	◊	◊	◊	◊	◊	0.82	0.11	0.10	0.01	0.01
Kuwait	0.80	0.19	0.20	0.04	0.04	0.83	0.17	0.19	0.03	0.03
Latvia	0.68	0.11	0.11	0.01	0.01	◊	◊	◊	◊	◊
Lebanon	◊	◊	◊	◊	◊	0.85	0.43	0.48	0.19	0.23
Lithuania	0.75	0.20	0.17	0.04	0.03	0.68	0.19	0.17	0.03	0.03
Malaysia	◊	◊	◊	◊	◊	0.83	0.35	0.33	0.13	0.11
Malta	◊	◊	◊	◊	◊	0.85	0.60	0.57	0.36	0.33
Morocco	0.87	0.26	0.27	0.07	0.07	0.80	0.17	0.16	0.03	0.03
Netherlands	0.68	0.27	0.27	0.07	0.07	◊	◊	◊	◊	◊
New Zealand	0.85	0.29	0.30	0.09	0.09	◊	◊	◊	◊	◊
Norway	0.73	0.14	0.14	0.02	0.02	0.65	0.12	0.13	0.01	0.02
Oman	◊	◊	◊	◊	◊	0.81	0.21	0.21	0.05	0.05
Palestinian Nat'l Auth.	◊	◊	◊	◊	◊	0.77	0.20	0.17	0.04	0.03
Qatar	0.78	0.22	0.25	0.05	0.06	0.79	0.16	0.33	0.02	0.11
Romania	◊	◊	◊	◊	◊	0.85	0.32	0.28	0.11	0.08
Russian Federation	0.79	0.21	0.20	0.04	0.04	0.78	0.26	0.24	0.07	0.06
Saudi Arabia	◊	◊	◊	◊	◊	0.83	0.13	0.14	0.02	0.02
Scotland	0.81	0.22	0.22	0.05	0.05	0.88	0.27	0.24	0.07	0.06
Serbia	◊	◊	◊	◊	◊	0.70	0.16	0.15	0.03	0.02
Singapore	0.83	0.23	0.23	0.05	0.05	0.88	0.44	0.45	0.19	0.21
Slovak Republic	0.75	0.23	0.24	0.05	0.06	◊	◊	◊	◊	◊
Slovenia	0.74	0.09	0.10	0.01	0.01	0.72	0.13	0.12	0.02	0.02
Sweden	0.81	0.20	0.21	0.04	0.04	0.75	0.15	0.15	0.02	0.02
Syrian Arab Republic	◊	◊	◊	◊	◊	0.79	0.19	0.14	0.04	0.02
Thailand	◊	◊	◊	◊	◊	0.83	0.31	0.30	0.10	0.09
Tunisia	0.77	0.29	0.29	0.08	0.09	0.73	0.27	0.22	0.07	0.05
Turkey	◊	◊	◊	◊	◊	0.84	0.42	0.39	0.18	0.15
Ukraine	0.70	0.18	0.15	0.03	0.02	0.77	0.27	0.23	0.07	0.05
United States	0.88	0.31	0.30	0.09	0.09	0.88	0.33	0.32	0.11	0.10
Yemen	0.72	0.12	0.13	0.01	0.02	◊	◊	◊	◊	◊
International Median	0.79	0.20	0.21	0.04	0.04	0.81	0.23	0.22	0.05	0.05
Benchmarking Participants										
Alberta, Canada	0.83	0.25	0.26	0.06	0.07	◊	◊	◊	◊	◊
Basque Country, Spain	◊	◊	◊	◊	◊	0.87	0.27	0.21	0.07	0.05
British Columbia, Canada	0.86	0.22	0.20	0.05	0.04	0.82	0.23	0.21	0.05	0.04
Dubai, UAE	0.79	0.22	0.27	0.05	0.08	0.77	0.42	0.35	0.18	0.12
Massachusetts, US	0.82	0.26	0.30	0.07	0.09	0.89	0.40	0.40	0.16	0.16
Minnesota, US	0.89	0.36	0.36	0.13	0.13	0.87	0.31	0.29	0.09	0.08
Ontario, Canada	0.84	0.29	0.27	0.08	0.07	0.84	0.26	0.22	0.07	0.05
Quebec, Canada	0.67	0.22	0.21	0.05	0.04	0.81	0.39	0.37	0.15	0.13

A diamond (◊) indicates the country did not participate in the assessment.

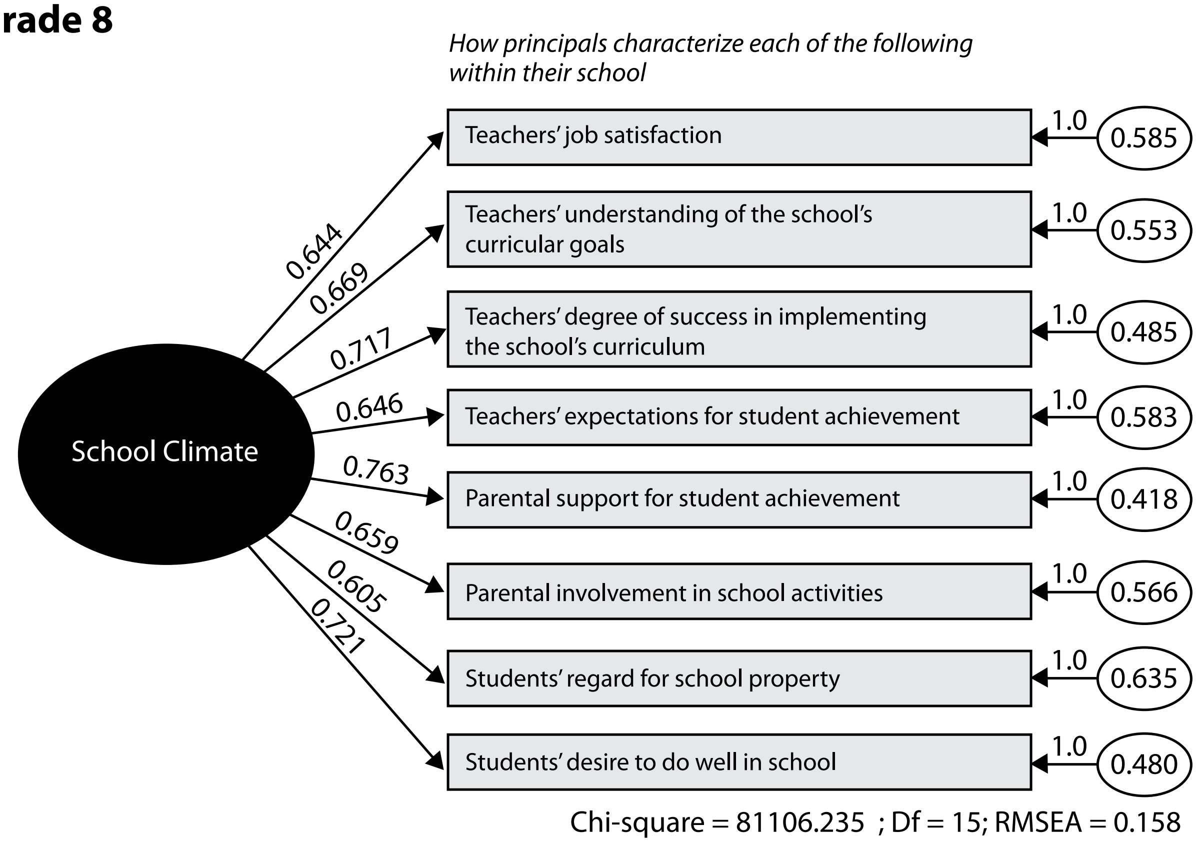
SOURCE: IEA's Trends in International Mathematics and Science Study (TIMSS) 2007

Exhibit 12.22 Latent Variable Model of Principals' Perception of School Climate

Grade 4



Grade 8



References

Byrne, B.M. (2001). *Structural equation modeling with AMOS—basic concepts, applications, and programming*. London: LEA Publishers.

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

Martin, M.O., Mullis, I.V.S., & Foy, P. (with Olson, J.F., Erberber, E., Preuschoff, C., & Galia, J.). (2008). *TIMSS 2007 international science report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

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

Martin, M.O., Mullis, I.V.S., Gonzales, E.J., Gregory, K.D., Smith, T.A., Chrostowski, S.J., Garden, R.A., & O'Connor, K.M. (2000). *TIMSS 1999 international science report: findings from IEA's repeat of the Third International Mathematics and Science Study at the eighth grade*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., & Foy, P. (with Olson, J.F., Preuschoff, C., Erberber, E., Arora, A., & Galia, J.). (2008). *TIMSS 2007 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

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

Mullis, I.V.S., Martin, M.O., Gonzales, E.J., Gregory, K.D., Garden, R.A., O'Connor, K.M., Chrostowski, S.J., & Smith, T.A. (2000). *TIMSS 1999 international mathematics report: Findings from IEA's Repeat of the Third International Mathematics and Science Study at the Eighth Grade*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Muthén, L.K. & Muthén, L.K. (2007). *Mplus—statistical analysis with latent variables user's guide*. Los Angeles, CA: Muthén & Muthén.

Schumacker R.E. & Lonax, R.G. (2004). *A beginner's guide to structural equation modeling*. Mahaw, NJ: LEA.

Chapter 13

The TIMSS 2007 International Benchmarks of Student Achievement in Mathematics and Science

Ina V.S. Mullis, Ebru Erberber, and Corinna Preuschoff

13.1 Overview

It is important for users of the TIMSS achievement results to understand what the scores on the TIMSS mathematics and science achievement scales mean. That is, what does it mean to have a scale score of 513 or 426? To describe student performance at various points along the TIMSS mathematics and science achievement scales, TIMSS 2007 used scale anchoring to summarize and describe student achievement at four points on the mathematics and science scales—Advanced International Benchmark (625), High International Benchmark (550), Intermediate International Benchmark (475), and Low International Benchmark (400). For the description of performance at the international benchmarks please see *TIMSS 2007 International Mathematics Report* (Mullis, Martin, & Foy, 2008) and *TIMSS 2007 International Science Report* (Martin, Mullis, & Foy, 2008).

This chapter describes the scale anchoring procedures that were applied to describe student performance at these benchmarks. Information about the TIMSS 2007 achievement scales and details about the methods used for scaling were presented in Chapter 11. In brief, scale anchoring involves selecting benchmarks (scale points) on the TIMSS achievement scales to be described in terms of student performance and then identifying items that students scoring at the anchor points (the international benchmarks) can answer correctly. The items, so identified, are grouped by content domain within benchmarks for review by mathematics and science experts. For TIMSS 2007, the Science and Mathematics Item Review Committee conducted the review. The committee members examined the content of each item and determined the kind of mathematics or science knowledge and/or skill demonstrated by students answering the item correctly. They then summarized the detailed list of item competencies in a brief description of achievement at each international benchmark. This procedure resulted in a content-referenced interpretation of the achievement results that can be considered in light of the TIMSS 2007 mathematics and science frameworks. The item-by-item descriptions developed as part of the scale anchoring procedures are provided in Appendix F.

13.2 History of Identifying the International Benchmarks[1]

Identifying the scale points to serve as international benchmarks initially was a challenge for TIMSS in the context of measuring trends. For the TIMSS 1995 and 1999 assessments, the scales were anchored using percentiles. That is, the scale anchoring analysis was conducted using the Top 10 percent (90^{th} percentile), the Top Quarter (75^{th} percentile), the Top Half (50^{th} percentile), and the Bottom Quarter (25^{th} percentile). However, with different participating countries in each TIMSS cycle and different achievement for countries participating in previous cycles, the percentile points had changed between 1995 and 1999.

In planning for reporting the results of TIMSS 2003, it was clear that TIMSS needed a set of points to serve as benchmarks, that would not change in the future, that would look sensible, and that were similar to the points used in 1999. After much consideration, a set of four points with equal intervals on the mathematics and science achievement scales was identified to be used as the international benchmarks, namely 400, 475, 550, and 625. These points were selected to be as close as possible to the percentile points anchored in 1999 at the eighth grade (i.e., Top 10 percent was 616 for mathematics and science, Top Quarter was 555 for mathematics and 558 for science, Top Half was 479 for mathematics and 488 for science, and Bottom Quarter was 396 for mathematics and 410 for science). The newly defined benchmark scale points were used as the basis for the scale anchoring descriptions in TIMSS 2003 and again in TIMSS 2007.

1 The description of the scale anchoring procedure was adapted from Kelly (1999), and Gregory and Mullis (2000).

13.3 Identifying the Students Achieving at the International Benchmarks

The first step in the scale-anchoring procedure was to identify those students scoring at the international benchmarks. Following the procedure used in previous IEA studies, students scoring within plus and minus 5 scale score points of each benchmark were identified for the benchmark analysis. The score ranges around each international benchmark and the number of students scoring in each range at the fourth and eighth grades for mathematics are shown in Exhibit 13.1 and for science in Exhibit 13.2. The range of plus and minus 5 points around a benchmark is intended to provide an adequate sample in each group, yet be small enough so that performance at each benchmark anchor point is still distinguishable from the next. The data analysis for the scale anchoring was based on these students scoring at each benchmark range.

Exhibit 13.1 Range Around Each International Benchmark and Number of Students Within Each Range – Mathematics

	Low Benchmark	Intermediate Benchmark	High Benchmark	Advanced Benchmark
Range of Scale Scores	**395–405**	**470–480**	**545–555**	**620–630**
Fourth Grade	3151	5243	5732	2755
Eighth Grade	6969	7649	5639	2335

Exhibit 13.2 Range Around Each International Benchmark and Number of Students Within Each Range – Science

	Low Benchmark	Intermediate Benchmark	High Benchmark	Advanced Benchmark
Range of Scale Scores	**395–405**	**470–480**	**545–555**	**620–630**
Fourth Grade	2950	5091	6321	2981
Eighth Grade	6393	8366	6749	2767

13.4 The Scale Anchoring Criteria

Having identified the number of students scoring at each benchmark anchor point, the next step was determining which particular items anchored at each of the anchor points. An important feature of the scale anchoring method is that it yields descriptions of the performance demonstrated by students reaching each of the benchmarks on the TIMSS mathematics and science achievement scales, and that these descriptions reflect demonstrably different accomplishments by students reaching each successively higher benchmark. The process entails the delineation of sets of items that students at each international benchmark are very likely to answer correctly and that discriminate between one benchmark and the next. Criteria were applied to identify the items that were answered correctly by most of the students at a particular benchmark, but by fewer students at the next lower benchmark.

In scale anchoring, the anchor items for each point are intended to be those that differentiate between adjacent anchor points (e.g., between the Advanced and the High International Benchmarks). To meet this goal, the criteria for identifying the items must take into consideration performance at more than one benchmark. Therefore, in addition to a criterion for the percentage of students at a particular benchmark correctly answering an item, it also was necessary to use a criterion for the percentage of students scoring at the next lower benchmark who correctly answer an item. For multiple-choice items, the criterion of 65 percent was used for the benchmark, since students would be likely (about two thirds of the time) to answer the item correctly. The criterion of less than 50 percent was used for the next lower benchmark, because with this response probability, students were more likely to have answered the item incorrectly than correctly. A somewhat less strict criterion was used for constructed-response items, because students have much less possibility of guessing. For constructed-response items, the criterion of 50 percent was used for the benchmark without any discrimination criterion for the next lower benchmark.

The criteria used to identify multiple-choice items that "anchored" are outlined below:

For the Low International Benchmark (400), a multiple-choice item anchored if

- At least 65 percent of students scoring in the range answered the item correctly (because this was the lowest benchmark described, there were no further criteria).

For the Intermediate International Benchmark (475), a multiple-choice item anchored if

- At least 65 percent of students scoring in the range answered the item correctly and

- Less than 50 percent of students at the Low International Benchmark answered the item correctly.

For the High International Benchmark (550), a multiple-choice item anchored if

- At least 65 percent of students scoring in the range answered the item correctly and
- Less than 50 percent of students at the Intermediate International Benchmark answered the item correctly.

For the Advanced International Benchmark (625), a multiple-choice item anchored if

- At least 65 percent of students scoring in the range answered the item correctly and
- Less than 50 percent of students at the High International Benchmark answered the item correctly.

To include all of the items in the anchoring process and provide information about content domains and cognitive processes that might not have had many items anchor exactly, items that met a slightly less stringent set of criteria were also identified. The criteria to identify multiple-choice items that "almost anchored" were the following:

For the Low International Benchmark (400), a multiple-choice item almost anchored if

- At least 60 percent of students scoring in the range answered the item correctly (because this was the lowest benchmark no further criteria were used).

For the Intermediate International Benchmark (475), a multiple-choice item almost anchored if

- At least 60 percent of students scoring in the range answered the item correctly and
- Less than 50 percent of students at the Low International Benchmark answered the item correctly.

For the High International Benchmark (550), a multiple-choice item almost anchored if

- At least 60 percent of students scoring in the range answered the item correctly and

- Less than 50 percent of students at the Intermediate International Benchmark answered the item correctly.

For the Advanced International Benchmark (625), a multiple-choice item almost anchored if

- At least 60 percent of students scoring in the range answered the item correctly and
- Less than 50 percent of students at the High International Benchmark answered the item correctly.

To be completely inclusive for all items, items that met only the criterion that at least 60 percent of the students answered correctly (regardless of the performance of students at the next lower point) were also identified. The three categories of items were mutually exclusive, and ensured that all of the items were available to inform the descriptions of student achievement at the anchor levels. A multiple-choice item was considered to be "too difficult" to anchor if less than 60 percent of students at the advanced benchmark answered the item correctly.

Different criteria were used to identify constructed-response items that "anchored." A constructed-response item anchored at one of the international benchmarks if at least 50 percent of students at that benchmark answer the item correctly. A constructed-response item was considered to be "too difficult" to anchor if less than 50 percent of students at the advanced benchmark answered the item correctly.

13.5 Identifying the Anchor Items at Each International Benchmark

For the students scoring in the range around each international benchmark, the percentage of those students that answered each item correctly was computed. To compute these percentages, students in each country were weighted to contribute proportional to the size of the student population in a country. Most of the TIMSS 2007 items were scored 1-point for a correct answer and 0 points for other answers. For these items, the percentage of students at each benchmark who answered each item correctly was computed. For relatively few constructed-response items scored for partial or full credit, percentages were computed for the students receiving full credit.

The criteria described above were applied to identify the items that anchored, almost anchored, and met only the 60 to 65 percent criteria. For mathematics at the fourth grade 118 items anchored, 19 almost anchored, and 40 met the 60 to 65 percent criteria. At the eighth grade, 151 mathematics items anchored, 27 almost anchored, and 36 met the 60 to 65 percent criteria. For science 111 items anchored, 16 almost anchored, and 43 met the 60 to 65 percent criteria at the fourth grade. At the eighth grade 152 science items anchored, 16 almost anchored, and 42 met the 60 to 65 percent criteria, respectively.

Broadening the anchor criteria on each benchmark to include items meeting the less stringent criteria, enabled the Science and Mathematics Item Review Committee to use all of the items included in the TIMSS 2007 assessment to characterize performance at each benchmark. Even though these items did not meet the 65 percent anchoring criteria, they were still items that students scoring at the benchmarks had a high degree of probability of answering correctly.

Exhibit 13.3 presents the number of mathematics items by content domain that anchored at each international benchmark at the fourth grade. Exhibit 13.4 presents the corresponding information for the eighth grade. Exhibit 13.5 and Exhibit 13.6 present the number of science items by content domain at each international benchmark at fourth and the eighth grades, respectively.

Exhibit 13.3 Number of Items Anchoring at Each International Benchmark by Content Domain – Fourth Grade Mathematics*

	Low (400)	Intermediate (475)	High (550)	Advanced (625)	Too Difficult to Anchor	Total
Number	6	15	36	30	4	91
Geometric Shapes and Measures	5	13	20	18	4	60
Data Display	3	11	9	3	–	26
Total	**14**	**39**	**65**	**51**	**8**	**177**

* Following the item review, 2 items were deleted out of 179 items in the mathematics fourth grade test, resulting in 177 items (see Chapter 10 for more details on the item review process).

Exhibit 13.4 Number of Items Anchoring at Each International Benchmark by Content Domain – Eighth Grade Mathematics*

	Low (400)	Intermediate (475)	High (550)	Advanced (625)	Too Difficult to Anchor	Total
Number	5	17	24	14	3	63
Algebra	1	7	29	26	1	64
Geometry	–	7	22	17	1	47
Data and Chance	3	9	19	8	1	40
Total	**9**	**40**	**94**	**65**	**6**	**214**

* Following the item review, 1 item was deleted out of 215 items in the mathematics eighth grade test, resulting in 214 items (see Chapter 10 for more details on the item review process).

Exhibit 13.5 Number of Items Anchoring at Each International Benchmark by Content Domain – Fourth Grade Science*

	Low (400)	Intermediate (475)	High (550)	Advanced (625)	Too Difficult to Anchor	Total
Life Science	7	17	15	20	12	71
Physical Science	7	9	28	15	5	64
Earth Science	1	6	11	13	4	35
Total	**15**	**32**	**54**	**48**	**21**	**170**

* Following the item review, 3 items were deleted out of 174 items in the science fourth grade test, resulting in 171 items. Also, 1 two-part item was combined to form a single item, further reducing the number of items to 170 (see Chapter 10 for more details on the item review process).

Exhibit 13.6 Number of Items Anchoring at Each International Benchmark by Content Domain – Eighth Grade Science*

	Low (400)	Intermediate (475)	High (550)	Advanced (625)	Too Difficult to Anchor	Total
Biology	2	11	26	25	11	75
Chemistry	3	4	11	16	7	41
Physics	2	2	14	24	12	54
Earth Science	-	6	17	12	5	40
Total	**7**	**23**	**68**	**77**	**35**	**210**

* Following the item review, 4 items were deleted out of 214 items in the science eighth grade test, resulting in 210 items (see Chapter 10 for more details on the item review process).

13.6 Experts Review Anchor Items by International Benchmark and Content Domains to Develop the Descriptions of Achievement

Having identified the items that anchored at each of the international benchmarks, the next step was to have the items reviewed by the TIMSS 2007 Science and Mathematics Item Review Committee to develop descriptions of student performance. In preparation for the review by the members of the TIMSS 2007 Science and Mathematics Item Review Committee, the mathematics and science items, respectively, were organized in binders grouped by international benchmark and within benchmark, the items were sorted by content area and then by the anchoring criteria they met - items that anchored, followed by items that almost anchored, followed by items that met only the 60 to 65 percent criteria. The following information was included for each item: content area, topic area, cognitive domain, maximum points, answer key, release status, percent correct at each benchmark, and overall international percent correct. For constructed-response items, the scoring guides were included.

The TIMSS & PIRLS International Study Center staff convened the TIMSS 2007 Science and Mathematics Item Review Committee for a four-day meeting in Kaohsiung, Taiwan. The work involved in completing the scale anchoring for the international benchmarks consisted of three tasks: (1) work through each item in each binder and arrive at a short description of the knowledge, understanding, and/or skills demonstrated by students answering the item correctly; (2) based on the items that anchored, almost anchored, and met only the 60 to 65 percent criterion, develop a description (in detailed and summary form) of the level of mathematics or science proficiency demonstrated by students at each of the four international benchmarks to publish in the TIMSS 2007 international reports; and (3) select example items that supported and illustrated the benchmark descriptions to publish together with the descriptions.

References

Gregory, K. D. & Mullis, I. V. S. (2000). Describing international benchmarks of student achievement. In M. O. Martin, K. D. Gregory, & S. E. Stemler (Eds.), *TIMSS 1999 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Kelly, D. L. (1999). *Interpreting the Third International Mathematics and Science Study (TIMSS) achievement scales using scale anchoring*. Unpublished doctoral dissertation, Boston College, Chestnut Hill, MA.

Martin, M.O., Mullis, I.V.S., & Foy, P. (with Olson, J.F., Erberber, E., Preuschoff, C., & Galia, J.). (2008). *TIMSS 2007 international science report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., & Foy, P. (with Olson, J.F., Preuschoff, C., Erberber, E., Arora, A., & Galia, J.). (2008). *TIMSS 2007 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Appendix A

Organizations and Individuals Responsible for TIMSS 2007

Introduction

TIMSS 2007 was a collaborative effort involving hundreds of individuals around the world. This appendix recognizes the individuals and organizations for their contributions. Given the work on TIMSS 2007 has spanned approximately five years and has involved so many people and organizations, this list may not include all who contributed. Any omission is inadvertent.

Of the first importance, TIMSS 2007 is deeply indebted to the students, teachers, and school principals who contributed their time and effort to the study.

Management and Coordination

TIMSS is a major undertaking of IEA, and together with PIRLS, comprises the core of IEA's regular cycle of studies. PIRLS, which regularly assesses reading at the fourth grade, complements the TIMSS assessments.

The TIMSS & PIRLS International Study Center at Boston College has responsibility for the overall direction and management of the TIMSS and PIRLS projects. Headed by Drs. Michael O. Martin and Ina V.S. Mullis, the study center is located in the Lynch School of Education. In carrying out the project, the TIMSS & PIRLS International Study Center worked closely with the IEA Secretariat in Amsterdam, which provided guidance overall and was responsible for verification of all translations produced by the participating countries. The IEA Data Processing and Research Center in Hamburg was responsible for processing and verifying the internal consistency and accuracy of the data submitted by the participants. Statistics Canada in Ottawa was responsible for school and student sampling activities. Educational Testing Service (ETS) in Princeton, New Jersey provided psychometric methodology recommendations addressing calibration, scaling, and survey design changes implemented in TIMSS 2007, and assisted in executing the item calibration analyses and made available software for scaling the achievement data.

The Project Management Team, comprised of the Directors and Senior Management from the TIMSS & PIRLS International Study Center, the IEA Secretariat, the IEA Data Processing and Research Center, Statistics Canada, and ETS met twice a year throughout the study to discuss the study's progress, procedures, and schedule. In addition, the Directors of the TIMSS & PIRLS International Study Center met with members of IEA's Technical Executive Group twice yearly to review technical issues.

Dr. Graham Ruddock from the National Foundation for Educational Research in England (NFER) was the TIMSS 2007 Mathematics Coordinator and Dr. Christine O'Sullivan from K–12 Consulting was the TIMSS 2007 Science Coordinator. Together with the Science and Mathematics Item Review Committee, a panel of internationally recognized experts in mathematics and science research, curriculum, instructions, and assessments, they provided excellent guidance throughout TIMSS 2007.

To work with the international team and coordinate within-country activities, each participating country designated one or two individuals to be the TIMSS National Research Coordinator or Co-Coordinators, known as the NRCs. The NRCs had the complicated and challenging task of implementing the TIMSS 2007 study in their countries in accordance with TIMSS guidelines and procedures. The quality of the TIMSS 2007 assessment and data depends on the work of the NRCs and their colleagues in carrying out the very complex sampling, data collection, and scoring tasks involved. In addition, the Questionnaire Development Group, comprised of NRCs, provided advice on questionnaire development.

Continuing the tradition of truly exemplary work established in previous TIMSS assessments, the TIMSS 2007 NRCs (often the same NRCs as in previous assessments), performed their many tasks with dedication, competence, energy, and goodwill, and have been commended by the IEA Secretariat, the TIMSS & PIRLS International Study Center, the IEA Data Processing and Research Center, and Statistics Canada for their commitment to the project and the high quality of their work.

Funding

A project of this magnitude requires considerable financial support. IEA's major funding partners for TIMSS 2007 included the World Bank, the U.S. Department of Education through the National Center for Education Statistics, the United Nations Development Programme (UNDP) and those countries that contributed by way of fees. The financial support provided by Boston College and NFER also is gratefully acknowledged.

IEA Secretariat

Hans Wagemaker, Executive Director
Barbara Malak, Manager, Membership Relations
Juriaan Hartenburg, Financial Manager
David Ebbs, Manager Assistant

TIMSS & PIRLS International Study Center at Boston College

Ina V.S. Mullis, Co-Director
Michael O. Martin, Co-Director
Pierre Foy, Director of Sampling and Data Analysis
John F. Olson, TIMSS Coordinator
Ann M. Kennedy, Coordinator of Project Development and Operations, PIRLS Coordinator
Alka Arora, TIMSS Advanced 2008 Project Coordinator
Debra Berger, Production Editor
Marcie Bligh, Manager of Office Administration
Ebru Erberber, TIMSS Research Associate
Susan Farrell, Co-Manager of Publications
Joseph Galia, Senior Statistician/Programmer
Christine Hoage, Manager of Finance
Ieva Johansone, Survey Operations Coordinator
Betty Kioroglou, Administrative Coordinator
Isaac Li, Statistician/Programmer
Dana Milne, TIMSS Graduate Assistant
Jennifer Moher, Data Graphics Specialist
Mario Pita, Co-Manager of Publications
Corinna Preuschoff, TIMSS Research Associate
Ruthanne Ryan, Data Graphics Specialist
Gabrielle Stanco, TIMSS Graduate Assistant

Feng Tian, TIMSS Graduate Assistant
Kathleen L. Trong, PIRLS Research Associate

IEA Data Processing and Research Center

Dirk Hastedt, Co-Director
Juliane Barth, Co-Manager, TIMSS and PIRLS Data Processing
Oliver Neuschmidt, Co-Manager, TIMSS and PIRLS Data Processing
Yasin Afana, Researcher
Alena Becker, Researcher
Christine Busch, Researcher
Ralph Carstens, Researcher
Tim Daniel, Researcher
Keith Hanmer, Researcher
Hauke Heyen, Programmer
Michael Jung, Researcher
Marta Kostek-Drosihn, Researcher
Sabine Meinck, Researcher
Dirk Oehler, Researcher
Stephan Petzchen, Senior Programmer
Anke Sielemann, Researcher
Harpreet Singh Choudry, Programmer
Milena Taneva, Researcher
Sabine Tieck, Researcher
Simone Uecker, Researcher
Bettina Wietzorek, Researcher
Olaf Zuehlke, Researcher

Statistics Canada

Marc Joncas, Senior Methodologist

Educational Testing Service

Matthias Von Davier, Principal Research Scientist
Scott Davis, Data Analysis and Computational Research Specialist
Edward Kulick, Director, Data Analysis and Computational Research

Science and Mathematics Item Review Committee

Mathematics
Graham Ruddock, TIMSS 2007 Mathematics Coordinator
Kiril Bankov, Bulgaria
Hanako Senuma, Japan
Khattab Mohammad Ahmad AbuLibdeh, Jordan
Robert Garden, New Zealand
Liv Sissel Gronmo, Norway
Mary Lindquist, United States

Science
Christine O'Sullivan, TIMSS 2007 Science Coordinator
Ahmed Muhammed Rafea, Bahrain
Chen-Yung Lin, Chinese Taipei
Jophus Anamuah-Mensah, Ghana
Gabriela Noveanu, Romania
Galina Kovaleva, Russian Federation
Jackie Heaton, Scotland
Audrey Champagne, United States

Questionnaire Item Review Committee

Sue Thomson, Australia
Serge Baillargeon, Quebec Province, Canada
Solaiman El-Khodary El-Sheikh, Egypt
Martina Meelissen, The Netherlands
Boey Kok Leong, Singapore
Barbara Japelj, Slovenia
Patrick Gonzales, United States

TIMSS 2007 National Research Coordinators (NRCs)

Algeria

Samia Mezaib
Sous-directrice de l'Evaluation
Ministere de l'Education Nationale

Armenia

Arsen Baghdasaryan
Yerevan State University

Australia

Sue Thomson
Australian Council for Educational Research

Austria

Gunter Haider
Birgit Suchan
Bundesinstitut fuer Bildungsforschung, Innovation und Entwicklung des Oesterreichischen Schulwesens (BIFIE)

Bahrain

Huda Al-Awadi
Ministry of Education

Bosnia and Herzegovina

Zaneta Dzumhur
Standard and Assessment Agency

Botswana

Sheila Barongwi
Ministry of Education

Bulgaria

Kiril Bankov
Georgi Simidchiev
Faculty of Mathematics and Informatics
University of Sofia

Chinese Taipei

Chun-Yen Chang
National Taiwan Normal University

Colombia

Francisco Ernesto Reyes Jimenez
Instituto Colombiano para el Fomento de la Educacion Superior

Cyprus

Constantinos Papanastasiou
Department of Education
University of Cyprus

Czech Republic

Vladislav Tomasek
Institute for Information on Education (UIV)

Denmark

Peter Allerup
The Danish University of Education

Egypt

Naguib Khouzam
National Center of Examinations and Educational Evaluation

El Salvador

Renán Rápalo
Ministry of Education

England

Linda Sturman

National Foundation for Educational Research

Georgia

Dito Pataraia

Mamuka Jibladze

National Assessment and Examinations Center

Germany

Martin Bonsen

Wilfried Bos

Centre for School Development Research

University of Dortmund

Ghana

Clara Rosaline Anumel

Inspectorate Division

Ghana Education Service

Hong Kong SAR

Frederick Leung

Faculty of Education

University of Hong Kong

Hungary

Ildiko Szepesi

Educational Authority

Department of Assessment and Evaluation

Indonesia

Burhanuddin Tola

Ministry of National Education

Center for National Assessment

Iran, Islamic Republic of

Abdol'azim Karimi

Ministry of Education

Institute for Educational Research

Israel

Rafi Nachmias

Ruth Zuzovsky

Center for Science & Technology Education

Tel Aviv University

Italy

Anna Maria Caputo

Instituto Nazionale per la Valuatazione del Sistema Educativo di Istruzione e di Formazione

Japan

Hanako Senuma

Yuji Saruta

National Institute for Educational Policy Research

Jordan

Khattab Mohammad Ahmad AbuLibdeh

National Center for Human Resources Development

Kazakhstan

Bazar Kabdoshevich Damitov

The National Centre for Assessment of the Quality of Education

Korea, Republic of

Kyunghee Kim

Korean Institute of Curriculum & Evaluation

Kuwait

Ebraheem Al-Qattan
Ministry of Education

Latvia

Andrejs Geske
Institute for Educational Research
University of Latvia

Lebanon

Leila Maliha Fayad
Educational Center for Research and Development
Ministry of Education

Lithuania

Aiste Elijio
Jolita Dudaite
National Examinations Centre

Malaysia

Amir bin Salleh
Ministry of Education
Educational Planning & Research Division

Malta

Raymond Camilleri
Department of Planning and Development
Education Division

Mongolia

Tseenoidov Oyunsaikhan
Mongolian Education Evaluation Center

Morocco

Mohammed Sassi
Department de l'Education Nationale
Centre National de l'Evaluation et des Examens

Netherlands

Martina Meelissen
Centre for Applied Research in Education
University of Twente

New Zealand

Robyn Caygill
New Zealand Ministry of Education
Comparative Education Research Unit

Norway

Liv Sissel Grønmo
University of Oslo, ILS

Oman

Zuwaina Saleh Al-Maskari
Ministry of Education

Palestinian National Authority

Mohammed O. Matar Mustafa
Ministry of Education and Higher Education, Assessment and Evaluation Center

Qatar

Abdulsattar Mohammed Nagi
Student Assessment Office

Romania

Gabriela Noveanu
Institute of Educational Sciences
Curriculum Department

Russian Federation

Galina Kovaleva
Center for Evaluating the Quality of Secondary General Education
Institute of Content and Methods of Learning
Russian Academy of Education

Saudi Arabia

Fahad A. Al-Muhaiza
Ministry of Education
Department of Measurement and Evaluation

Scotland

Linda Sturman
National Foundation for Educational Research

Serbia

Slobodanka Gasic-Pavisic
Radovan Antonijevic
Institute for Educational Research

Singapore

Gary Quek
Ministry of Education
Planning Division

Slovak Republic

Patricia Jelemenska
Jozef Kuraj (through 2006)
National Institute for Education (SPU)

Slovenia

Barbara Japelj Pavesic
Educational Research Institute

Sweden

Camilla Thinsz Fjellstrom
Director of Education
National Agency for Education
Per-Olof Bentley (Mathematics)
University of Gothenburg
Christina Kärrqvist (Science)
University of Gothenburg

Syrian Arab Republic

Omar Abou Awn
Ministry of Education

Thailand

Precharn Dechsri
The Institute for the Promotion of Teaching Science and Technology

Tunisia

Nejib Ayed
Centre National d'Innovation Pedagogique et de Recherche en Education

Turkey

Ibrahim Demirer
Educational Research and Development Directorate
Ministry of National Education

Ukraine

Nataliia Prokopenko
Ministry of Education and Science of Ukraine

United States

Patrick Gonzales
National Center for Education Statistics

Yemen

Tawfiq Ahmad Al-Mekhlafy
Ministry of Education
Educational Research & Development Center

Benchmarking Participants

Alberta, Canada
Ping Yang
Learner Assessment Branch
Alberta Education

Basque Country, Spain
Josu Sierra Orrantia
ISEI-IVEI Basque Institute for Research and Evaluation in Education

British Columbia, Canada
Diane Lalancette
Provincial, National and International Indicators Branch
Ministry of Education

Dubai, United Arab Emirates
Zulaikha Mohamed
Knowledge & Human Development Authority
Government of Dubai

Massachusetts, United States
Robert Lee
Massachusetts Department of Education

Minnesota, United States
Dirk Mattson
Minnesota Department of Education

Ontario, Canada
Michael Kozlow
Education Quality and Accountability Office

Quebec, Canada
Robert Marcotte
MELS Direction de la Sanction des Etudes

Appendix B

Characteristics of National Samples

B.1 Algeria

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<15), private schools, and schools with multi-level classes
- No within-school exclusions

Sample Design

- No explicit stratification
- Implicit stratification by region (48 regions), for a total of 48 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.1.1 Allocation of School Sample in Algeria – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Algeria	150	0	149	0	0	1
Total	**150**	**0**	**149**	**0**	**0**	**1**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of private schools
- No within-school exclusions

Sample Design

- No explicit stratification
- Implicit stratification by region (48 regions), for a total of 48 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.1.2 Allocation of School Sample in Algeria – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Algeria	150	0	149	0	0	1
Total	**150**	**0**	**149**	**0**	**0**	**1**

B.2 Armenia

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (Number of students in the school <50) and special boarding schools for disabled students
- Within-sample exclusions consisted of students with special needs

Sample Design

- No explicit stratification
- Implicit stratification by region (11 regions), for a total of 11 implicit strata
- Sampled two classrooms per school whenever possible
- Fourth and eighth grade students found in same schools: maximum school sample overlap

B.2.1 Allocation of School Sample in Armenia – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Armenia	150	2	143	5	0	0
Total	**150**	**2**	**143**	**5**	**0**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (Number of students in the school <50) and special boarding schools for disabled students
- Within-school exclusions consisted of students with special needs

Sample Design

- No explicit stratification
- Implicit stratification by region (11 regions), for a total of 11 implicit strata
- Sampled two classrooms per school whenever possible
- Small schools sampled with equal probabilities
- Fourth and eighth grade students found in same schools: maximum school sample overlap

B.2.2 Allocation of School Sample in Armenia – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Armenia	150	2	143	5	0	0
Total	**150**	**2**	**143**	**5**	**0**	**0**

B.3 Australia

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<5), remote indigenous schools, special education schools, hospital schools, language centers and other schools with 'radically different curricula'
- Within-school exclusions consisted of disabled students, and English as second language (ESL) students

Sample Design

- Explicit stratification by states and territories, for a total of eight explicit strata
- School sample allocation is not proportional at the explicit stratum level
- Implicit stratification by school type (government, Catholic, independent; in small explicit strata, 'Catholic' and 'independent' schools were grouped into one 'non-government' stratum) and geographic location (metro, non-metro) in large 'school type' strata, for a total of 31 implicit strata
- Sampled two classrooms per school having at least 86 students (MOS≥86), and one classroom otherwise
- Small schools sampled with equal probabilities

B.3.1 Allocation of School Sample in Australia – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Australian Capital Territory	15	0	15	0	0	0
New South Wales	40	0	40	0	0	0
Northern Territory	15	0	14	0	1	0
Queensland	35	0	35	0	0	0
South Australia	30	0	30	0	0	0
Tasmania	30	0	30	0	0	0
Victoria	35	0	34	1	0	0
Western Australia	30	1	28	1	0	0
Total	**230**	**1**	**226**	**2**	**1**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<5), remote indigenous schools, special education schools, hospital schools, language centers and other schools with 'radically different curricula'
- Within schools exclusions consisted of disabled students, and English as second language (ESL) students

Sample Design

- Explicit stratification by states and territories, for a total of eight explicit strata
- School sample allocation is not proportional at the explicit stratum level
- Implicit stratification by school type (government, Catholic, independent; in small explicit strata, 'Catholic' and 'independent' schools were grouped into one 'non-government' stratum) and geographic location (metro, non-metro) in large 'school type' strata, for a total of 31 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities
- Within the states of Queensland and Victoria, student sampling weights were post-stratified by gender

B.3.2 Allocation of School Sample in Australia – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Australian Capital Territory	15	0	15	0	0	0
New South Wales	40	1	39	0	0	0
Northern Territory	15	1	14	0	0	0
Queensland	35	0	35	0	0	0
South Australia	30	0	30	0	0	0
Tasmania	30	0	30	0	0	0
Victoria	35	0	35	0	0	0
Western Australia	30	0	30	0	0	0
Total	**230**	**2**	**228**	**0**	**0**	**0**

B.4 Austria

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<3) and special education schools
- Within-school exclusions consisted of students with special needs in normal schools (special education classrooms, intellectually or functionally disabled students, and non-native language speakers)

Sample Design

- Explicit stratification by region (nine regions), for a total of nine explicit strata
- Implicit stratification by district (4-25 districts, depending on the region), for a total of 116 implicit strata
- Sampled two classrooms per school whenever possible
- Small schools sampled with equal probabilities

B.4.1 Allocation of School Sample in Austria – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Burgenland	6	0	6	0	0	0
Kärnten	14	1	13	0	0	0
Niederösterreich	38	0	37	0	0	1
Oberösterreich	37	0	37	0	0	0
Salzburg	14	0	14	0	0	0
Steiermark	28	0	28	0	0	0
Tirol	18	0	18	0	0	0
Vorarlberg	10	0	10	0	0	0
Wien	34	1	31	2	0	0
Total	**199**	**2**	**194**	**2**	**0**	**1**

B.5 Bahrain

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<8) and one school under the umbrella of the Secondary Education
- Within schools exclusions consisted of functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by school type (public, private) and governorate (capital, central, Muharraq, Northern, Southern) in the 'public' strata, for a total of six explicit strata
- Implicit stratification by gender (boys, girls) in the 'public' strata, for a total of 11 implicit strata

- Sampled three classrooms whenever possible
- Sampled all schools

B.5.1 Allocation of School Sample in Bahrain – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Public – Capital	10	0	10	0	0	0
Public – Central	12	0	12	0	0	0
Public – Muharraq	11	0	11	0	0	0
Public – Northern	15	0	15	0	0	0
Public – Southern	9	0	9	0	0	0
Private	17	0	17	0	0	0
Total	**74**	**0**	**74**	**0**	**0**	**0**

B.6 Bosnia and Herzegovina

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<6), special needs schools, music schools, and international schools
- Within schools exclusions consisted of functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by cantons and regions, for a total of 16 explicit strata
- No implicit stratification
- Sampled two classrooms per school having at least 155 students (MOS≥155), and one classroom otherwise
- Small schools sampled with equal probabilities

B.6.1 Allocation of School Sample in Bosnia and Herzegovina – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Kanton Sarajevo	15	0	15	0	0	0
Tuzlanski Kanton	23	0	23	0	0	0
Zeničko-Dobojski Kanton	18	0	18	0	0	0
Srednjebosanski Kanton	12	0	12	0	0	0
Hercegovačko-Neretvanski Kanton	9	0	9	0	0	0
Zapadno-Hercegovacki Kanton	4	0	4	0	0	0
Herceg-Bosanski Kanton	3	0	3	0	0	0
Unsko-Sanski Kanton	12	0	12	0	0	0
Posavski Kanton	2	0	2	0	0	0
Bosansko-Podrinjski Kanton	2	0	2	0	0	0
Prijedorska Regija	5	0	5	0	0	0
Banjalucka Regija	18	0	18	0	0	0
Dobojska Regija	6	0	6	0	0	0
Focanska Regija	8	0	8	0	0	0
Bijeljinska Regija	10	0	10	0	0	0
Brcko Distrikt	3	0	3	0	0	0
Total	**150**	**0**	**150**	**0**	**0**	**0**

B.7 Botswana

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<10)
- Within-school exclusions consisted of students in special education classes and special education students in regular classes

Sample Design

- Explicit stratification by region (central, north/northwest, south central, south, west), for a total of five explicit strata
- Implicit stratification by school type (public, private) in the 'south central' stratum, and urbanization (urban, semi-urban, rural) in the 'public' strata, for a total of 15 implicit strata
- Sampled one classroom per school
- Schools sampled with equal probabilities

B.7.1 Allocation of School Sample in Botswana – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Central	40	0	40	0	0	0
North / Northwest	26	0	26	0	0	0
South Central	43	0	43	0	0	0
South	25	0	25	0	0	0
West	16	0	16	0	0	0
Total	**150**	**0**	**150**	**0**	**0**	**0**

B.8 Bulgaria

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<5) and special needs schools
- Within schools exclusions consisted of special education classes, functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by school type (general, vocational, profiled), for a total of three explicit strata
- Implicit stratification by location (Sofia, large city, other location), for a total of nine implicit strata
- Sampled two classrooms per school having at least 55 students (MOS≥55), and one classroom otherwise
- Small schools sampled with equal probabilities

B.8.1 Allocation of School Sample in Bulgaria – Eighth Grade Mathematics

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
General	118	4	108	4	0	2
Vocational	25	0	23	0	1	1
Profiled	27	0	27	0	0	0
Total	**170**	**4**	**158**	**4**	**1**	**3**

B.8.2 Allocation of School Sample in Bulgaria – Eighth Grade Science

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
General	118	9	103	4	0	2
Vocational	25	0	23	0	1	1
Profiled	27	19	8	0	0	0
Total	**170**	**28**	**134**	**4**	**1**	**3**

B.9 Chinese Taipei

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<8) and special needs schools
- Within-school exclusions consisted of special education classes, functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by geographic location (north, middle, south, east, isolated islands), for a total of five explicit strata
- No implicit stratification
- Sampled two classrooms per school having at least 450 students (MOS≥450), and one classroom otherwise
- Small schools sampled with equal probabilities

B.9.1 Allocation of School Sample in Chinese Taipei – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
North	66	0	66	0	0	0
Middle	38	0	38	0	0	0
South	40	0	40	0	0	0
East	4	0	4	0	0	0
Isolated Islands	2	0	2	0	0	0
Total	**150**	**0**	**150**	**0**	**0**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<8) and special needs schools
- Within-school exclusions consisted of special education classes, functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by geographic location (north, middle, south, east, isolated island), for a total of five explicit strata
- No implicit stratification
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.9.2 Allocation of School Sample in Chinese Taipei – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
North	66	0	66	0	0	0
Middle	38	0	38	0	0	0
South	40	0	40	0	0	0
East	4	0	4	0	0	0
Isolated Islands	2	0	2	0	0	0
Total	**150**	**0**	**150**	**0**	**0**	**0**

B.10 Colombia

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<3), evening schools and weekend schools
- Within-school exclusions consisted of disabled students

Sample Design

- Explicit stratification by grade (grade 4 only, grade 4 & grade 8), for a total of two explicit strata
- Implicit stratification by school type (public, private), urbanization (urban, rural) and school calendar (A, B), for a total of 16 implicit strata
- Sampled two classrooms per school having at least 120 students (MOS≥120), and one classroom otherwise
- Small schools sampled with equal probabilities

B.10.1 Allocation of School Sample in Colombia – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 4 Only	110	6	94	8	1	1
Grade 4 & 8	40	1	38	1	0	0
Total	**150**	**7**	**132**	**9**	**1**	**1**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<3), evening schools and weekend school
- Within-school exclusions consisted of disabled students

Sample Design

- Explicit stratification by grade (grade 8, grade 4 & grade 8), for a total of two explicit strata
- Implicit stratification by school type (public, private), urbanization (urban, rural) and school calendar (A, B), for a total of 16 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.10.2 Allocation of School Sample in Colombia – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 8 Only	86	1	80	5	0	0
Grade 4 & 8	64	1	62	1	0	0
Total	**150**	**2**	**142**	**6**	**0**	**0**

B.11 Cyprus

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- No school-level exclusions
- Within-school exclusions consisted of intellectually disabled students, functionally disabled students and non-native language speakers (not able to read or write in the Greek language)

Sample Design

- Explicit stratification by district (Nicosia, Limassol, Larnaca-Famagusta, Paphos), for a total of four explicit strata
- Implicit stratification by urbanization (urban, rural), for a total of eight implicit strata
- Sampled four classrooms per school having at least 100 students (MOS≥100), three classrooms whenever possible otherwise
- Sampled all schools

B.11.1 Allocation of School Sample in Cyprus – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Nicosia	23	0	23	0	0	0
Limassol	19	0	19	0	0	0
Larnaca-Famagusta	15	0	15	0	0	0
Paphos	10	0	10	0	0	0
Total	**67**	**0**	**67**	**0**	**0**	**0**

B.12 The Czech Republic

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<5), special education schools, and Polish language schools
- Within-school exclusions consisted of disabled students

Sample Design

- Explicit stratification by grade (grade 4 only, grade 4 & grade 8) and school type (only primary level, complete basic school) in the 'grade 4 only' stratum, for a total of three explicit strata
- Implicit stratification by region (14 regions), for a total of 41 implicit strata
- Sampled one or two classrooms per school
- Small schools sampled with equal probabilities

B.12.1 Allocation of School Sample in the Czech Republic – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Gr4 only – Only primary level	16	2	11	1	0	2
Gr4 only – Complete Basic school	2	0	2	0	0	0
Gr4 & Gr8 – Basic school	132	1	119	11	0	1
Total	**150**	**3**	**132**	**12**	**0**	**3**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<6), special education schools, and Polish language schools
- Within-school exclusions consisted of disabled students

Sample Design

- Explicit stratification by grade (grade 8 only, grade 4 & grade 8) and school type (basic school; 6, 8-year gymnasium), for a total of three explicit strata
- Implicit stratification by region (14 regions), for a total of 39 implicit strata
- Sampled one or two classrooms per school
- Small schools sampled with equal probabilities

B.12.2 Allocation of School Sample in the Czech Republic – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Gr8 only – Basic school	2	0	2	0	0	0
Gr8 only – 6, 8-year gymnasium	14	0	12	2	0	0
Gr4 & Gr8 – Basic school	134	3	121	10	0	0
Total	**150**	**3**	**135**	**12**	**0**	**0**

B.13 Denmark

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of special needs schools, and schools with less than 70% of students organized specifically in grade levels
- Within-school exclusions consisted of functionally or intellectually disabled students and non-native language speakers

Sample Design

- No explicit stratification
- No implicit stratification
- Sampled two classrooms, if possible, in schools having at least 55 students (MOS≥55), and one classroom otherwise
- Small schools sampled with equal probabilities

B.13.1 Allocation of School Sample in Denmark – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Denmark	150	0	105	27	5	13
Total	**150**	**0**	**105**	**27**	**5**	**13**

B.14 Egypt

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<12)
- Within-sample exclusions consisted of special education classes and French teaching language schools

Sample Design

- Explicit stratification by school type (public, experimental language, free private, private, private language) and region (Cairo, Alexandria, other regions) in the 'public' stratum, for a total of seven strata
- School sample allocation is not proportional at the stratum level
- Implicit stratification in the 'public' strata by urbanization (urban, rural), shift (morning, noon, afternoon 2nd, full day), and gender (boys, girls, mixed), for a total of 53 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.14.1 Allocation of School Sample in Egypt – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Public – Cairo	18	0	18	0	0	0
Public – Alexandria	22	0	22	0	0	0
Public – All other regions	120	0	119	1	0	0
Experimental Language	25	0	25	0	0	0
Free Private	2	0	2	0	0	0
Private	25	0	24	1	0	0
Private Language	25	4	21	0	0	0
Total	**237**	**4**	**231**	**2**	**0**	**0**

B.15 El Salvador

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<7)
- Within-sample exclusions consisted of jail schools

Sample Design

- Explicit stratification by grade (grade 4 only, grade 4 & grade 8) and urbanization (urban, rural), for a total of four explicit strata
- Implicit stratification by school type (public, private) in the 'urban' strata, school type (community-based – ACE, council-based – CDE, other) in the 'rural' strata, and region (14 regions) in the 'grade 4 & grade 8' strata, for a total of 36 implicit strata. ACE schools are run by parents, CDE schools are run by parents and teacher councils
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.15.1 Allocation of School Sample in El Salvador – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Gr4 Only – Urban	4	0	4	0	0	0
Gr4 Only – Rural	28	0	28	0	0	0
Gr4 & Gr8 – Urban	62	1	60	1	0	0
Gr4 & Gr8 – Rural	56	1	54	1	0	0
Total	**150**	**2**	**146**	**2**	**0**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<7)
- Within-sample exclusions consisted of jail schools

Sample Design

- Explicit stratification by grade (grade 8 only, grade 4 & grade 8) and urbanization (urban, rural), for a total of three explicit strata
- Implicit stratification by school type (public, private) in the 'urban' stratum, school type (community-based – ACE, council-based – CDE, other) in the 'rural' stratum, and region (14 regions) in the 'grade 4 & grade 8' strata, for a total of 33 implicit strata. ACE schools are run by parents, CDE schools are run by parents and teacher councils
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.15.2 Allocation of School Sample in El Salvador – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Gr8 only – Urban	4	2	2	0	0	0
Gr4 & Gr8 – Urban	90	1	88	1	0	0
Gr4 & Gr8 – Rural	56	2	53	1	0	0
Total	**150**	**5**	**143**	**2**	**0**	**0**

B.16 England

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- The population covered is actually the fifth grade, due to early start of schooling
- School-level exclusions consisted of very small schools (MOS<8) and special schools
- Within-school exclusions consisted of students with special needs and non-native language speakers

Sample Design

- Explicit stratification by grade (grade 4 only and grade 4 & grade 8), for a total of two explicit strata
- Implicit stratification by school performance based on national test (5 levels and a category for 'missing'), and school type (primary/combined, junior, middle, independent), for a total of 25 implicit strata

- Selected up to four classrooms per school
- Small schools sampled with equal probabilities

B.16.1 Allocation of School Sample in England – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 4 Only	154	0	127	12	0	15
Grade 4 & 8	6	1	4	0	0	1
Total	**160**	**1**	**131**	**12**	**0**	**16**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- The population covered is actually the ninth grade, due to early start of schooling
- School-level exclusions consisted of very small schools (MOS<7) and special schools
- Within-school exclusions consisted of students with special needs and non-native language speakers

Sample Design

- Explicit stratification by grade (grade 8 only, grade 4 & grade 8), for a total of two explicit strata
- Implicit stratification by school performance based on national test (5 levels and a category for 'missing'), and school type (comprehensive to 16, comprehensive to 18, independent, grammar, other), for a total of 27 implicit strata
- Selected up to four classrooms per school
- Small schools sampled with equal probabilities

B.16.2 Allocation of School Sample in England – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 8 Only	154	0	122	10	0	22
Grade 4 & 8	6	0	4	1	0	1
Total	**160**	**0**	**126**	**11**	**0**	**23**

B.17 Georgia

FOURTH GRADE

Coverage and Exclusions

- Coverage in Georgia was restricted to students whose language of instruction is Georgian (85% of the International Desired Target Population)
- School-level exclusions consisted of very small schools (MOS<4), and specialized schools for disabled students
- Within-school exclusions consisted of special education classes, functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by regions, for a total of 12 explicit strata
- Implicit stratification by urbanization (urban, rural), for a total of 23 implicit strata
- Sampled two classrooms per school having at least 60 students (MOS≥60), and one classroom otherwise
- Small schools sampled with equal probabilities
- The same sample of schools was used for the eighth graders

B.17.1 Allocation of School Sample in Georgia – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Kvemo Kartli	10	1	8	0	1	0
Adjara	14	1	13	0	0	0
Apxazeti	2	0	2	0	0	0
Guria	6	0	5	0	1	0
Imereti	27	1	22	3	1	0
Kaxeti	14	1	12	0	1	0
Mckheta-Mtianeti	5	0	4	0	1	0
Racha-Lechkhumi	2	0	1	1	0	0
Samckhe-Javakheti	5	1	4	0	0	0
Shida Kartli	12	0	10	1	1	0
Tbilisi	39	2	36	1	0	0
Samegrelo	16	1	14	0	1	0
Total	**152**	**8**	**131**	**6**	**7**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage in Georgia was restricted to students whose language of instruction is Georgian (85% of the International Desired Target Population)
- School-level exclusions consisted of very small schools (MOS<4) and specialized schools for disabled students
- Within-school exclusions consisted of special education classes, functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by region, for a total of 12 explicit strata
- Implicit stratification by urbanization (urban, rural), for a total of 23 implicit strata
- Sampled two classrooms per school having at least 60 students (MOS≥60), and one classroom otherwise
- Small schools sampled with equal probabilities
- The same sample of schools was used for the fourth graders

B.17.2 Allocation of School Sample in Georgia – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Kvemo Kartli	10	2	8	0	0	0
Adjara	14	1	13	0	0	0
Apxazeti	2	0	2	0	0	0
Guria	6	1	5	0	0	0
Imereti	27	4	22	1	0	0
Kaxeti	14	2	12	0	0	0
Mckheta-Mtianeti	5	0	4	0	1	0
Racha-Lechkhumi	2	1	1	0	0	0
Samckhe-Javakheti	5	1	4	0	0	0
Shida Kartli	12	2	10	0	0	0
Tbilisi	39	2	36	1	0	0
Samegrelo	16	1	14	0	1	0
Total	**152**	**17**	**131**	**2**	**2**	**0**

B.18 Germany

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<6), and schools for children with disabilities
- Within-school exclusions consisted of special needs students and non-native language speakers

Sample Design

- Explicit stratification by state, for a total of 16 explicit strata
- Implicit stratification by school type (primary, remedial education) and region in the 'primary' strata of the larger states, for a total of 45 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.18.1 Allocation of School Sample in Germany – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Baden-Württemberg	36	0	34	1	1	0
Bayern	41	0	41	0	0	0
Berlin	8	0	7	0	0	1
Brandenburg	5	1	4	0	0	0
Bremen	2	0	2	0	0	0
Hamburg	5	1	4	0	0	0
Hessen	19	0	19	0	0	0
Mecklenburg-Vorpommern	3	0	3	0	0	0
Niedersachsen	28	0	28	0	0	0
Nordrhein-Westfalen	60	1	58	1	0	0
Rheinland-Pfalz	13	0	11	2	0	0
Saarland	3	0	3	0	0	0
Sachsen	9	0	8	1	0	0
Sachsen-Anhalt	5	0	5	0	0	0
Schleswig-Holstein	9	0	8	1	0	0
Thüringen	4	0	4	0	0	0
Total	**250**	**3**	**239**	**6**	**1**	**1**

B.19 Ghana

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<11)
- No within-school exclusions

Sample Design

- Explicit stratification by school type (public, private), for a total of two explicit strata
- Implicit stratification by region (10 regions), for a total of 20 implicit strata
- Sampled two classrooms in 11 large schools, and one classroom in the remaining schools
- Small schools sampled with equal probabilities

B.19.1 Allocation of School Sample in Ghana – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Private	30	0	30	0	0	0
Public	133	0	133	0	0	0
Total	**163**	**0**	**163**	**0**	**0**	**0**

B.20 Hong Kong SAR

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<8), special education schools, international schools, and private independent schools (non-local curriculum)
- Within-school exclusions consisted of functionally or intellectually disabled students

Sample Design

- Explicit stratification by form of financing (government, aided, direct subsidy scheme, private), for a total of four explicit strata
- Implicit stratification by gender (co-educational, boys, girls) and shift (AM, PM, whole day), for a total of 21 implicit strata
- Sampled two classrooms per school having at least 185 students (MOS≥185), and one classroom otherwise
- Small schools sampled with equal probabilities

B.20.1 Allocation of School Sample in Hong Kong SAR – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Aided	126	0	102	2	1	21
Direct Subsidy Scheme	4	0	4	0	0	0
Government	10	0	8	0	0	2
Private	10	0	8	1	0	1
Total	**150**	**0**	**122**	**3**	**1**	**24**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of special education schools, international schools, private independent schools (non-local curriculum)
- No within-school exclusions

Sample Design

- Explicit stratification by form of financing (government, aided, direct subsidy scheme, private) for at total of four explicit strata
- Implicit stratification by gender (co-educational, boys, girls) and language (Chinese, English), for a total of 18 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.20.2 Allocation of School Sample in Hong Kong SAR – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Aided	126	0	94	2	3	27
Direct Subsidy Scheme	12	0	8	2	0	2
Government	12	0	8	1	0	3
Private	2	0	2	0	0	0
Total	**152**	**0**	**112**	**5**	**3**	**32**

B.21 Hungary

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<6) and special needs schools
- Within-school exclusions consisted of students with special needs, and non-native language speakers

Sample Design

- Explicit stratification by grade (grade 4 only, grade 4 & grade 8), school status (already existed in 2004, established after 2004) in the 'grade 4 & grade 8' stratum, and location (capital, county town, town, rural) in the 'already existed in 2004' strata, for a total of six explicit strata
- Implicit stratification by region (7 regions), for a total of 35 implicit strata
- Sampled two classrooms per school having at least 40 students (MOS≥40), and one classroom otherwise
- Small schools sampled with equal probabilities

B.21.1 Allocation of School Sample in Hungary – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 4 Only	3	1	2	0	0	0
Gr4 & Gr8 – Already existed in 2004 – Capital	20	0	18	1	0	1
Gr4 & Gr8 – Already existed in 2004 – County Town	26	0	22	4	0	0
Gr4 & Gr8 – Already existed in 2004 – Town	49	1	48	0	0	0
Gr4 & Gr8 – Already existed in 2004 – Rural	50	1	45	4	0	0
Gr4 & Gr8 – after 2004	2	2	0	0	0	0
Total	**150**	**5**	**135**	**9**	**0**	**1**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<6) and special needs schools
- Within-school exclusions consisted of special education classes, students with special needs, and non-native language speakers

Sample Design

- Explicit stratification by grade (grade 8 only, grade 4 & grade 8), school status (already existed in 2004, established after 2004) in the 'grade 4 & grade 8' stratum, and location (capital, county town, town, rural) in the 'already existed in 2004' strata, for a total of six explicit strata
- Implicit stratification by region (7 regions), for a total of 35 implicit strata
- Sampled two classrooms per school having at least 40 students (MOS≥40), and one classroom otherwise
- Small schools sampled with equal probabilities

B.21.2 Allocation of School Sample in Hungary – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 8 Only	11	1	8	1	1	0
Gr4 & Gr8 – Already existed in 2004 – Capital	19	0	17	1	0	1
Gr4 & Gr8 – prior to 2005 – County Town	26	0	22	4	0	0
Gr4 & Gr8 – Already existed in 2004 – Town	47	1	46	0	0	0
Gr4 & Gr8 – Already existed in 2004 – Rural	45	1	40	4	0	0
Gr4 & Gr8 – after 2004	2	2	0	0	0	0
Total	**150**	**5**	**133**	**10**	**1**	**1**

B.22 Indonesia

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<10), schools in Papua, and schools in Nanggroe Aceh Darussalam
- No within-school exclusions

Sample Design

- No explicit stratification
- Implicit stratification by school type (general, Islamic), school type (public, private), and national examination score (high performance, middle performance, low performance), for a total of 12 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.22.1 Allocation of School Sample in Indonesia – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Indonesia	150	1	149	0	0	0
Total	**150**	**1**	**149**	**0**	**0**	**0**

B.23 Islamic Republic of Iran

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<5) and schools in the Bam area due to recent severe earthquakes
- No within-school exclusions

Sample Design

- Explicit stratification by school type (public, private) and gender (boys, girls, mixed), for a total of five explicit strata
- School sample allocation is not proportional at the explicit stratum level (Private schools are over-represented)
- Implicit stratification by province (29 provinces), for a total of 145 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.23.1 Allocation of School Sample in Islamic Republic of Iran – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Public – Boys	70	3	67	0	0	0
Public – Girls	70	4	66	0	0	0
Public – Mixed	50	4	46	0	0	0
Private – Boys	30	3	27	0	0	0
Private – Girls	20	2	18	0	0	0
Total	**240**	**16**	**224**	**0**	**0**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<5), schools in the Bam area due to recent severe earthquakes, and schools for distance learning students
- No within-school exclusions

Sample Design

- Explicit stratification by school type (public, private) and gender (boys, girls, mixed), for a total of five explicit strata
- School sample allocation is not proportional at the explicit stratum level (Private schools are over-represented)
- Implicit stratification by province (29 provinces), for a total of 144 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.23.2 Allocation of School Sample in Islamic Republic of Iran – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Public – Boys	80	5	75	0	0	0
Public – Girls	80	3	77	0	0	0
Public – Mixed	10	0	10	0	0	0
Private – Boys	35	3	32	0	0	0
Private – Girls	15	1	14	0	0	0
Total	**220**	**12**	**208**	**0**	**0**	**0**

B.24 Israel

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<9), special education schools, and Ultra Orthodox schools
- Within-school exclusions consisted of special education classes, functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by school type (Hebrew secular, Hebrew religious, Arab), for a total of three explicit strata
- Implicit stratification by socio-economic status (high SES, medium SES, low or unknown SES), for a total of nine implicit strata

- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.24.1 Allocation of School Sample in Israel – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Hebrew Secular	70	0	64	3	1	2
Hebrew Religious	40	0	37	1	1	1
Arab Secular	40	0	39	0	0	1
Total	**150**	**0**	**140**	**4**	**2**	**4**

B.25 Italy

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<8)
- Within-school exclusions consisted of functionally or intellectually disabled students and non-native language speakers

Sample Design

- No explicit stratification
- Implicit stratification by regions (20 regions) and type of municipality (province capital town, other), for a total of 40 implicit strata
- Sampled one or two classrooms per school
- Sampled two classrooms per school in most schools
- All schools sampled with probabilities proportional to size

B.25.1 Allocation of School Sample in Italy – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Italy	170	0	155	14	1	0
Total	**170**	**0**	**155**	**14**	**1**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<8)
- Within-school exclusions consisted of intellectually or functionally disabled students and non-native language speakers

Sample Design

- No explicit stratification
- Implicit stratification by regions (20 regions) and type of municipality (province capital town, other), for a total of 40 implicit strata
- Sampled one or two classrooms per school
- All schools sampled with probabilities proportional to size

B.25.2 Allocation of School Sample in Italy – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Italy	170	0	159	10	1	0
Total	**170**	**0**	**159**	**10**	**1**	**0**

B.26 Japan

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of special needs schools
- Within-school exclusions consisted of classes within general schools for multi-grade setting, and classes within general schools for disabled children

Sample Design

- Explicit stratification by location (very large city, large city, small city, non-city area), for a total of four explicit strata
- No implicit stratification

- Sampled two classrooms per school having at least 110 students (MOS≥110), and one classroom otherwise
- Small schools sampled with equal probabilities

B.26.1 Allocation of School Sample in Japan – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Very Large City	30	0	29	0	0	1
Large City	26	0	26	0	0	0
Small City	68	0	66	1	0	1
Non-City Area	26	0	24	2	0	0
Total	**150**	**0**	**145**	**3**	**0**	**2**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of special needs schools
- Within-school exclusions consisted of classes within general schools for multi-grade setting, and classes within general schools for disabled children

Sample Design

- Explicit stratification by school type (public, private or national) and location (very large city, large city, small city, non-city area) in the 'public' stratum, for a total of five explicit strata
- No implicit stratification
- Sampled two classrooms per school having at least 230 students (MOS≥230), and one classroom otherwise
- Small schools sampled with equal probabilities

B.26.2 Allocation of School Sample in Japan – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Public – Very large City	25	0	23	0	0	2
Public – Large City	24	0	24	0	0	0
Public – Small City	64	0	62	1	0	1
Public – Non-city Area	26	0	26	0	0	0
Private or National	11	0	9	1	0	1
Total	**150**	**0**	**144**	**2**	**0**	**4**

B.27 Jordan

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<9)
- Within-school exclusions consisted of disabled students and non-native language speakers

Sample Design

- Explicit stratification by school type (Discovery, public, UNRWA, private), for a total of four explicit strata
- All Discovery schools were selected
- Implicit stratification by urbanization (urban, rural), gender (boys, girls, mixed), and school form (basic, secondary) in the ‘public’, ‘UNRWA’ and ‘private’ strata, for a total of 27 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.27.1 Allocation of School Sample in Jordan – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Discovery	61	0	61	0	0	0
Public	110	0	110	0	0	0
UNRWA	17	0	17	0	0	0
Private	12	0	12	0	0	0
Total	**200**	**0**	**200**	**0**	**0**	**0**

B.28 Kazakhstan

FOURTH GRADE

Coverage and Exclusions

- Coverage in Kazakhstan was restricted to students whose language of instruction is Kazakh or Russian (94% of the International Desired Target Population)
- School-level exclusions consisted of very small schools (MOS<5) and special education schools

- Within-sample exclusions consisted of students in very remote schools, disabled students and non-native language speakers

Sample Design

- Explicit stratification by urbanization (urban, rural) and size (large schools, small school) in the 'rural' stratum, for a total of three explicit strata
- Implicit stratification by region (16 regions) and language (Kazakh, Russian, Kazakh & Russian), for a total of 126 implicit strata
- Sampled two classrooms per school having at least 135 students (MOS≥135), and one classroom otherwise
- Small schools sampled with equal probabilities

B.28.1 Allocation of School Sample in Kazakhstan – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Urban Schools	92	3	89	0	0	0
Rural – Large Schools	52	6	45	0	1	0
Rural – Small Schools	6	0	6	0	0	0
Total	**150**	**9**	**140**	**0**	**1**	**0**

B.29 Republic of Korea

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<10), geographically inaccessible schools, and physical eduation middle schools
- Within-school exclusions consisted of functionally or intellectually disabled students

Sample Design

- Explicit stratification by province, for a total of 16 explicit strata
- Implicit stratification by urbanization (urban, suburban, rural), and gender (boys, girls, co-educational), for a total of 81 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.29.1 Allocation of School Sample in Republic of Korea – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Seoul	28	0	28	0	0	0
Pusan	11	0	11	0	0	0
Taegu	8	0	8	0	0	0
Inchon	9	0	9	0	0	0
Kwangju	5	0	5	0	0	0
Taejon	5	0	5	0	0	0
Ulsan	4	0	4	0	0	0
Kyunggi-do	34	0	34	0	0	0
Kangwon-do	4	0	4	0	0	0
Chungchongbuk-do	5	0	5	0	0	0
Chungchongnam-do	6	0	6	0	0	0
Chollabuk-do	6	0	6	0	0	0
Chollanam-do	6	0	6	0	0	0
Kyongsangbuk-do	7	0	7	0	0	0
Kyongsangnam-do	10	0	10	0	0	0
Cheju-do	2	0	2	0	0	0
Total	**150**	**0**	**150**	**0**	**0**	**0**

B.30 Kuwait

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- The population covered is actually the fifth grade, due to late data collection
- School-level exclusions consisted of very small schools (MOS<10)
- No within-school exclusions

Sample Design

- No explicit stratification
- Implicit stratification by region (six regions) and gender (boys, girls), for a total of 12 implicit strata
- Sampled two classrooms per school having at least 174 students (MOS≥174), and one classroom otherwise

B.30.1 Allocation of School Sample in Kuwait – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Kuwait	150	0	150	0	0	0
Total	**150**	**0**	**150**	**0**	**0**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- No school-level exclusions
- Within-school exclusions consisted of intellectually disabled students

Sample Design

- No explicit stratification
- Implicit stratification by region (six regions) and gender (boys, girls), for a total of 12 implicit strata
- Sampled one classroom per school
- Sampled all schools

B.30.2 Allocation of School Sample in Kuwait – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Kuwait	163	0	158	0	0	5
Total	**163**	**0**	**158**	**0**	**0**	**5**

B.31 Latvia

FOURTH GRADE

Coverage and Exclusions

- Coverage in Latvia was restricted to students whose language of instruction is Latvian (72% of the International Desired Target Population)
- School-level exclusions consisted of very small schools (MOS<4) and schools for functionally or intellectually disabled students
- Within-school exclusions consisted of functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by location (Riga, rural, small town, town), for a total of four explicit strata
- Implicit stratification by urbanization code, for a total of 12 implicit strata
- Sampled three classrooms per school whenever possible
- Small schools sampled with equal probabilities

B.31.1 Allocation of School Sample in Latvia – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Riga	28	0	25	2	0	1
Rural	58	0	55	1	1	1
Small Town	45	0	42	2	0	1
Town	19	0	18	0	0	1
Total	**150**	**0**	**140**	**5**	**1**	**4**

B.32 Lebanon

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<9)
- No within-school exclusions

Sample Design

- Explicit stratification by school type (public, private), for a total of two explicit strata
- No implicit stratification
- Sampled two classrooms per school having at least 60 students (MOS≥60), and one classroom otherwise
- Small schools sampled with equal probabilities

B.32.1 Allocation of School Sample in Lebanon – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Public	64	0	59	1	0	4
Private	86	2	61	12	3	8
Total	**150**	**2**	**120**	**13**	**3**	**12**

B.33 Lithuania

FOURTH GRADE

Coverage and Exclusions

- Coverage in Lithuania was restricted to students whose language of instruction is Lithuanian (93% of the International Desired Target Population)
- School-level exclusions consisted of very small schools (MOS<5) and special education schools
- Within-school exclusions consisted of functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by grade (grade 4 only, grade 4 & grade 8), for a total of two explicit strata
- Implicit stratification by county (10 counties) and location (Vilnius-capital, other major cities, cities, small cities, others), for a total of 61 implicit strata
- Sampled two classrooms whenever possible
- Small schools sampled with equal probabilities

B.33.1 Allocation of School Sample in Lithuania – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 4 Only	37	3	33	1	0	0
Grade 4 & 8	126	4	121	1	0	0
Total	**163**	**7**	**154**	**2**	**0**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage in Lithuania was restricted to students whose language of instruction is Lithuanian (92% of the International Desired Target Population)
- School-level exclusions consisted of very small schools (MOS<6) and special education schools

- Within-school exclusions consisted of functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by grade (grade 8 only, grade 4 & grade 8), for a total of two explicit strata
- Implicit stratification by county (10 counties) and location (Vilnius-capital, other major cities, cities, small cities, other), for a total of 62 implicit strata
- Sampled two classrooms whenever possible
- Small schools sampled with equal probabilities

B.33.2 Allocation of School Sample in Lithuania – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 8 Only	24	3	19	0	0	2
Grade 4 & 8	126	3	122	1	0	0
Total	**150**	**6**	**141**	**1**	**0**	**2**

B.34 Malaysia

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of religious schools, special education schools, private schools, and private Chinese schools
- No within-school exclusions

Sample Design

- Explicit stratification by state, for a total of 14 explicit strata
- Implicit stratification by urbanization (urban, rural), for a total of 27 implicit strata
- Sampled two classrooms per school having at least 450 students (MOS≥450), and one classroom otherwise
- Small schools sampled with equal probabilities

B.34.1 Allocation of School Sample in Malaysia – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Johor	18	0	18	0	0	0
Kedah	12	0	12	0	0	0
Kelantan	11	0	11	0	0	0
Melaka	5	0	5	0	0	0
Negeri Sembilan	6	0	6	0	0	0
Pahang	9	0	9	0	0	0
Perak	15	0	15	0	0	0
Perlis	2	0	2	0	0	0
Pulau Pinang	7	0	7	0	0	0
Sabah	13	0	13	0	0	0
Sarawak	14	0	14	0	0	0
Selangor	22	0	22	0	0	0
Terengganu	8	0	8	0	0	0
WP Kuala Lumpur	8	0	8	0	0	0
Total	**150**	**0**	**150**	**0**	**0**	**0**

B.35 Malta

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- The population covered is actually the ninth grade, due to early start of schooling
- School-level exclusions consisted of schools not following the mainstream curriculum
- Within-school exclusions consisted of special education students and non-native language speakers

Sample Design

- Explicit stratification by island (Malta, Gozo), for a total of two explicit strata
- Implicit stratification by sector (state, church, independent), school type (junior lyceums, secondary schools) in the ‘state’ sector,and gender (boys, girls, mixed), for a total of 15 implicit strata
- Sampled all schools and all classrooms

B.35.1 Allocation of School Sample in Malta – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Malta	54	1	53	0	0	0
Gozo	6	0	6	0	0	0
Total	**60**	**1**	**59**	**0**	**0**	**0**

B.36 Morocco

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (Number of students in the school <30)
- No within-school exclusions

Sample Design

- Explicit stratification by school type (public, private) and groups of regions (eight groups) in the ‘public’ strata, for a total of nine explicit strata
- Implicit stratification by regions (16 regions) and urbanization (urban, rural) in the ‘public’ stratum, for a total of 33 implicit strata
- Sampled one classroom per school

B.36.1 Allocation of School Sample in Morocco – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Public – Region Stratum 1	22	0	22	0	0	0
Public – Region Stratum 2	24	0	18	0	0	6
Public – Region Stratum 3	30	0	29	0	0	1
Public – Region Stratum 4	22	0	17	0	0	5
Public – Region Stratum 5	36	2	27	0	0	7
Public – Region Stratum 6	30	0	25	0	0	5
Public – Region Stratum 7	22	0	12	0	0	10
Public – Region Stratum 8	28	0	26	0	0	2
Private	12	0	8	0	0	4
Total	**226**	**2**	**184**	**0**	**0**	**40**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (Number of students in the school <33)
- Within-school exclusions consisted of special education classes

Sample Design

- Explicit stratification by school type (public, private) and groups of regions (eight groups) in the 'public' stratum, for a total of nine explicit strata
- Implicit stratification by regions (16 regions) and urbanization (urban, rural) in the 'public' strata, for a total of 31 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.36.2 Allocation of School Sample in Morocco – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Public – Region Stratum 1	9	0	7	0	0	2
Public – Region Stratum 2	20	0	11	0	0	9
Public – Region Stratum 3	42	0	38	0	0	4
Public – Region Stratum 4	17	0	11	0	0	6
Public – Region Stratum 5	30	0	25	0	0	5
Public – Region Stratum 6	34	0	24	0	0	10
Public – Region Stratum 7	20	0	3	0	0	17
Public – Region Stratum 8	28	0	8	0	0	20
Private	5	0	4	0	0	1
Total	**205**	**0**	**131**	**0**	**0**	**74**

B.37 The Netherlands

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<6) and special schools for primary education

- Within-school exclusions consisted of special education classes, disabled students and non-native language speakers (less than one year of instruction in Dutch)

Sample Design

- Explicit stratification by schools average socio-economic status (low mean SES, medium mean SES, high mean SES), for a total of three explicit strata
- No implicit stratification
- Sampled all classrooms
- Small schools sampled with equal probabilities

B.37.1 Allocation of School Sample in the Netherlands – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Low Mean SES	38	2	12	12	10	2
Medium Mean SES	51	0	27	16	5	3
High Mean SES	61	0	33	17	9	2
Total	**150**	**2**	**72**	**45**	**24**	**7**

B.38 New Zealand

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- The population covered is actually the fifth grade, due to early start of schooling
- School-level exclusions consisted of very small schools (MOS<4), special education schools, Rudolf Steiner schools, the Correspondence School, and schools that provide more than 80% of their instruction in the Māori language
- Within-school exclusions consisted of special education classes, special needs students, foreign fee paying students, students with insufficient instruction in test language, and units within schools that provide more than 80% of their instruction in the Māori language

Sample Design

- No explicit stratification

- Implicit stratification by school decile indicator (which gives the extent to which a school draws students from low socio-economic communities – high, medium, low, not assigned) and level of urbanization (urban, rural), for a total of eight implicit strata
- Sampled at least two classrooms whenever possible

B.38.1 Allocation of School Sample in New Zealand – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
New Zealand	220	0	213	6	1	0
Total	**220**	**0**	**213**	**6**	**1**	**0**

B.39 Norway

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<5), special needs schools, Sami schools, and a very remote school in Longyearbyen
- Within-school exclusions consisted of functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by grade (grade 4 only, grade 4 & grade 8), for a total of two explicit strata
- No implicit stratification
- Sampled all classrooms
- Small schools sampled with equal probabilities

B.39.1 Allocation of School Sample in Norway – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 4 Only	119	0	105	9	2	3
Grade 4 & 8	31	0	26	2	1	2
Total	**150**	**0**	**131**	**11**	**3**	**5**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<5), special needs schools, Sami schools, and a very remote school in Longyearbyen
- Within-school exclusions consisted of functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by grade (grade 8 only, grade 4 & grade 8), for a total of two explicit strata
- No implicit stratification
- Sampled up to four classrooms per school
- Small schools sampled with equal probabilities

B.39.2 Allocation of School Sample in Norway – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 8 Only	119	0	108	4	0	7
Grade 4 & 8	31	0	25	2	0	4
Total	**150**	**0**	**133**	**6**	**0**	**11**

B.40 Oman

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<8) and special needs schools
- Within-school exclusions consisted of functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by region (11 regions), for a total of 11 explicit strata

- Implicit stratification by gender (boys, girls, mixed) in the eight largest explicit strata, for a total of 23 implicit strata
- All classrooms were selected in five of the schools, and one classroom otherwise
- Small schools sampled with equal probabilities

B.40.1 Allocation of School Sample in Oman – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Muscat	26	0	26	0	0	0
Al-Batinah North	33	0	33	0	0	0
Al-Batinah South	20	0	20	0	0	0
Al-Dakhiliya	20	2	18	0	0	0
Al-Sharqiya South	12	1	11	0	0	0
Al-Sharqiya North	11	0	11	0	0	0
Al-Dhahara South	9	0	9	0	0	0
Al-Dhahara North	3	0	3	0	0	0
Dhofar	12	1	11	0	0	0
Musandam	2	0	2	0	0	0
Wosta	2	0	2	0	0	0
Total	**150**	**4**	**146**	**0**	**0**	**0**

B.41 Palestinian National Authority

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<10)
- Within-school exclusions consisted of special education classes, functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by school type (private, public, UNWRA (United Nations Relief and Works Agency) – regular program, UNRWA – special program), for a total of three explicit strata
- All schools in the ‘UNRWA – special program’ stratum (Schools of Excellence) were selected

- Implicit stratification by gender (boys, girls, mixed) in the 'public' and 'UNRWA' strata and region (West Bank, Gaza) in the 'public', 'UNRWA-boys' and 'UNRWA-girls' strata, for a total of 13 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.41.1 Allocation of School Sample in Palestinian National Authority – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Private	6	0	5	1	0	0
Public	100	7	93	0	0	0
UNRWA – Regular Program	44	0	44	0	0	0
UNRWA – Special Program	5	0	5	0	0	0
Total	**155**	**7**	**147**	**1**	**0**	**0**

B.42 Qatar

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<10)
- Within-school exclusions consisted of functionally or intellectually disabled students

Sample Design

- Explicit stratification by grade (grade 4 only, grade 4 & grade 8), for a total of two explicit strata
- Implicit stratification by school type (independent, Ministry of Education, private Arabic) and gender (boys, girls), for a total of nine implicit strata
- Sampled all schools and all classrooms

B.42.1 Allocation of School Sample in Qatar – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 4 Only	102	0	102	0	0	0
Grade 4 & 8	12	0	12	0	0	0
Total	**114**	**0**	**114**	**0**	**0**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<10) and international schools
- Within-school exclusions consisted of functionally or intellectually disabled students

Sample Design

- Explicit stratification by grade (grade 8 only, grade 4 & grade 8), for a total of two explicit strata
- Implicit stratification by school type (independent, Ministry of Education, private Arabic) and gender (boys, girls), for a total of nine implicit strata
- Sampled all schools and all classrooms

B.42.2 Allocation of School Sample in Qatar – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 8 Only	55	0	54	0	0	1
Grade 4 & 8	12	0	12	0	0	0
Total	**67**	**0**	**66**	**0**	**0**	**1**

B.43 Romania

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<7) and special education schools
- Within-school exclusions consisted of very few intellectually disabled students

Sample Design

- No explicit stratification
- Implicit stratification by region (7 regions) and urbanization (rural, urban), for a total of 14 implicit strata

- Sampled two or three classrooms per school having at least 45 students (MOS≥45), and one classroom otherwise
- Small schools sampled with equal probabilities

B.43.1 Allocation of School Sample in Romania – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Romania	150	0	149	0	0	1
Total	**150**	**0**	**149**	**0**	**0**	**1**

B.44 Russian Federation

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (definition depending on the region) and special needs schools
- Within-school exclusions consisted of special education classes, functionally or intellectually disabled students and non-native language speakers

Sample Design

- A sample of 45 regions out of 86 (Regions 81, 85 and 88 were collapsed prior to region sampling) was first sampled with PPS. The largest 17 regions were sampled with certainty. A sample of schools was then drawn within each region
- Explicit stratification by region type (certainty versus sampled) for a total of 18 explicit strata. However, the table below gives the school sample allocation for the 45 regions in the sample
- Implicit stratification by urbanization (seven different types) in all but the ‘Moscow’ and ‘St. Petersburg’ strata, for a total of 232 implicit strata
- Sampled two classrooms per school out of the two largest schools in any given certainty region and sampled two classrooms per school out of the largest school in any sampled region having at least a certain number of students depending on the region, and one classroom otherwise
- Small schools sampled with equal probabilities

B.44.1 Allocation of School Sample in Russian Federation – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Adygea	4	0	4	0	0	0
Alania	4	0	4	0	0	0
Marii Al	4	0	4	0	0	0
Amur oblast	4	0	4	0	0	0
Kabardino oblast	4	0	4	0	0	0
Kaliningrad oblast	4	0	4	0	0	0
Moscow	10	0	10	0	0	0
St. Petersburg	4	0	4	0	0	0
Chelyabinsk oblast	4	0	4	0	0	0
Bashkortostan	10	0	10	0	0	0
Irkutsk oblast	4	0	4	0	0	0
Krasnodar kr	8	0	8	0	0	0
Krasnoyarsk kr	4	0	4	0	0	0
N Novgorod oblast	4	0	4	0	0	0
Orenburg oblast	4	0	4	0	0	0
Belgorod oblast	4	0	4	0	0	0
Altai kr	4	0	4	0	0	0
Kirov oblast	4	0	4	0	0	0
Kurgan oblast	4	0	4	0	0	0
Kurst oblast	4	0	4	0	0	0
Novosibirsk oblast	4	0	4	0	0	0
Omsk oblast	4	0	4	0	0	0
Chita oblast	4	0	4	0	0	0
Hakasia	4	0	4	0	0	0
Hanty-Mansii ok	4	0	4	0	0	0
Pskov oblast	4	0	4	0	0	0
Razan oblast	4	0	4	0	0	0
Saratov oblast	4	0	4	0	0	0
Tambov oblast	4	0	4	0	0	0
Perm oblast	4	0	4	0	0	0
Rostov oblast	8	0	8	0	0	0
Stavropol kr	4	0	4	0	0	0
Sverdlovsk oblast	6	0	6	0	0	0
Tatarstan	6	0	6	0	0	0
Arhangelsk oblast	4	0	4	0	0	0
Chuvashia	4	0	4	0	0	0
Kemerovo oblast	4	0	4	0	0	0
Lipstek oblast	4	0	4	0	0	0
Sakha	4	0	4	0	0	0
Tula oblast	4	0	4	0	0	0
Udmurtia	4	0	4	0	0	0
Volvograd oblast	4	0	4	0	0	0
Orel oblast	4	0	4	0	0	0
Moskva oblast	8	0	8	0	0	0
Dagestan	2	0	2	0	0	0
Total	**206**	**0**	**206**	**0**	**0**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (definition depending on the region) and special needs schools
- Within-school exclusions consisted of special education schools, functionally or intellectually disabled students and non-native language speakers

Sample Design

- A sample of 45 regions out of 86 (Regions 81, 85 and 88 were collapsed prior to region sampling) was first sampled with PPS. The largest 16 regions were sampled with certainty. A sample of schools was then drawn within each region
- Explicit stratification by region type (certainty versus sampled) for a total of 17 explicit strata. However, the table below gives the school sample allocation for the 45 regions in the sample
- Implicit stratification by urbanization (seven different types) in all but the 'Moscow' and 'St. Petersburg' strata, for a total of 231 implicit strata
- Sampled two classrooms per school out of the two largest schools in any given certainty region and sampled two classrooms per school out of the largest school in any sampled region and one classroom otherwise
- Small schools sampled with equal probabilities

B.44.2 Allocation of School Sample in Russian Federation – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Primorskii kray	4	0	4	0	0	0
Rostov oblast	6	0	6	0	0	0
Samara oblast	4	0	4	0	0	0
Saratov oblast	4	0	4	0	0	0
Perm oblast	4	0	4	0	0	0
Hanty-Mansii ok.	4	0	4	0	0	0
Udmurtia	4	0	4	0	0	0
Chelyabinsk oblast	6	0	6	0	0	0
Chita oblast	4	0	4	0	0	0
Smolensk oblast	4	0	4	0	0	0
Sverdlovsk oblast	6	0	6	0	0	0
St. Petersburg	4	0	4	0	0	0
Stavropol Kray	4	0	4	0	0	0
Tatarstan	8	0	8	0	0	0
Pensa oblast	4	0	4	0	0	0
Arhangelsk oblast	4	0	4	0	0	0
Evrey-Auto oblast	4	0	4	0	0	0
Novosibirsk oblast	4	0	4	0	0	0
Omsk oblast	4	0	4	0	0	0
Orenburg oblast	4	0	4	0	0	0
Sakha	4	0	4	0	0	0
Altay kray	4	0	4	0	0	0
Astrahan oblast	4	0	4	0	0	0
Bashkortostan	8	0	8	0	0	0
Belgorod oblast	4	0	4	0	0	0
Bransk oblast	4	0	4	0	0	0
Vladimir oblast	4	0	4	0	0	0
Volgograd oblast	4	0	4	0	0	0
Irkutsk oblast	4	0	4	0	0	0
Karelia	4	0	4	0	0	0
Kemerovo oblast	4	0	4	0	0	0
Kirov oblast	4	0	4	0	0	0
Komi	4	0	4	0	0	0
Krasnodar kray	8	0	8	0	0	0
Krasnoyarsk kray	6	0	6	0	0	0
Kurgan oblast	4	0	4	0	0	0
Kursk oblast	4	0	4	0	0	0
Karachaevo-Cherkessia	4	0	4	0	0	0
Lipetsk oblast	4	0	4	0	0	0
Moscow	10	0	10	0	0	0
N. Novgorod oblast	6	0	6	0	0	0
Kostroma oblast	4	0	4	0	0	0
Moskva olb.	8	0	8	0	0	0
Dagestan	2	0	2	0	0	0
Voronezh obl.	4	0	4	0	0	0
Total	**210**	**0**	**210**	**0**	**0**	**0**

B.45 Saudi Arabia

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<7)
- Within-school exclusions consisted of functionally or intellectually disabled students

Sample Design

- Explicit stratification by gender (boys, girls) and school type (government, private), for a total of four explicit strata
- Implicit stratification by urbanization (rural, suburban, urban) and school type (general, Quranic), for a total of 21 implicit strata
- Sampled two classrooms per school having at least 140 students (MOS≥140), and one classroom otherwise
- Small schools sampled with equal probabilities

B.45.1 Allocation of School Sample in Saudi Arabia – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Boys – Government	76	0	75	0	0	1
Boys – Private	10	1	9	0	0	0
Girls – Government	75	0	75	0	0	0
Girls – Private	6	0	6	0	0	0
Total	**167**	**1**	**165**	**0**	**0**	**1**

B.46 Scotland

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- The population covered is actually the fifth grade, due to early start of schooling
- School-level exclusions consisted of very small schools (MOS<6), special schools, and Gaelic schools
- Within-school exclusions consisted of students with special needs and non-native language speakers

Sample Design

- No explicit stratification
- Implicit stratification by grade (grade 4 only, grade 4 & grade 8), urbanization (large urban area, other urban area, accessible small town, remote small town, accessible rural area, remote rural area), and school deprivation index (low, middle, high, unknown), for a total of 33 implicit strata
- Sampled up to four classrooms per school
- Small schools sampled with equal probabilities

B.46.1 Allocation of School Sample in Scotland – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Scotland	150	2	114	18	7	9
Total	**150**	**2**	**114**	**18**	**7**	**9**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- The population covered is actually the ninth grade, due to early start of schooling
- School-level exclusions consisted of very small schools (MOS<7), and special schools
- Within-school exclusions consisted of students with special needs and non-native language speakers

Sample Design

- No explicit stratification
- Implicit stratification by grade (grade 8 only, grade 4 & grade 8), urbanization (large urban area, other urban area, accessible small town, remote small town, accessible rural area, remote rural area), and school deprivation index (low, middle, high, unknown), for a total of 29 implicit strata
- Sampled up to three classrooms per school
- Small schools sampled with equal probabilities

B.46.2 Allocation of School Sample in Scotland – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Scotland	150	0	109	14	6	21
Total	**150**	**0**	**109**	**14**	**6**	**21**

B.47 Serbia

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<10), special education schools, schools for talented students, Albanian schools and Romanian schools
- Within-school exclusions consisted of special education classes, functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by region (Belgrade, Central Serbia, Vojvodina), for a total of three explicit strata
- Implicit stratification by urbanization (rural, urban), for a total of six implicit strata
- Sampled two classrooms per school having at least 95 students (MOS≥95), and one classroom otherwise
- Small schools sampled with equal probabilities

B.47.1 Allocation of School Sample in Serbia – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Belgrade	30	0	30	0	0	0
Central Serbia	79	2	77	0	0	0
Vojvodina	41	1	40	0	0	0
Total	**150**	**3**	**147**	**0**	**0**	**0**

B.48 Singapore

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of special education schools, private schools, and religious schools
- No within-school exclusions

Sample Design

- No explicit stratification
- Implicit stratification by school rank (22 different ranks based on students performance), for a total of 23 implicit strata
- Sampled two classrooms per school. Classrooms were sampled with PPS. A sample of 19 students was drawn in each class
- Sampled all schools

B.48.1 Allocation of School Sample in Singapore – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Singapore	177	0	177	0	0	0
Total	**177**	**0**	**177**	**0**	**0**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of special education schools, private schools, and religious schools
- No within-school exclusions

Sample Design

- No explicit stratification
- Implicit stratification by school rank (30 different ranks based on students performance), for a total of 30 implicit strata

- Sampled two classrooms per school. Classrooms were sampled with PPS. A sample of 19 students was drawn in each class
- Sampled all schools

B.48.2 Allocation of School Sample in Singapore – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Singapore	164	0	164	0	0	0
Total	**164**	**0**	**164**	**0**	**0**	**0**

B.49 The Slovak Republic

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<5), schools where students taught in another language than Slovak or Hungarian, and private Slovak schools
- Within-school exclusions consisted of functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by region (eight regions) and urbanization (urban, rural), for a total of 16 explicit strata
- School sample allocation is not proportional at the explicit stratum level (minimum school sample size of 10 per stratum)
- Implicit stratification by language (Slovak only, Hungarian only, Slovak and Hungarian), for a total of 30 implicit strata
- Selected one or two classrooms per school, depending on the explicit stratum
- Small schools sampled with equal probabilities

B.49.1 Allocation of School Sample in the Slovak Republic – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Bratislavsky – Rural	10	0	10	0	0	0
Bratislavsky – Urban	16	0	16	0	0	0
Trnavsky – Rural	10	0	10	0	0	0
Trnavsky – Urban	10	0	10	0	0	0
Trenciansky – Rural	10	0	10	0	0	0
Trenciansky – Urban	12	0	12	0	0	0
Nitriansky – Rural	10	0	10	0	0	0
Nitriansky – Urban	10	0	10	0	0	0
Zilinsky – Rural	12	0	12	0	0	0
Zilinsky – Urban	12	0	12	0	0	0
Banskobystricky – Rural	10	0	9	0	1	0
Banskobystricky – Urban	12	0	12	0	0	0
Presovsky – Rural	12	0	11	1	0	0
Presovsky – Urban	14	0	14	0	0	0
Kosicky – Rural	10	0	9	1	0	0
Kosicky – Urban	14	0	14	0	0	0
Total	**184**	**0**	**181**	**2**	**1**	**0**

B.50 Slovenia

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<8), special needs schools, Italian language schools, and Waldorf schools
- Within-school exclusions consisted of students with functional disabilities and students not able to talk or understand Slovene language

Sample Design

- No explicit stratification
- Implicit stratification by region (eight regions), for a total of eight implicit strata
- Sampled three classrooms per school having at least 75 students (MOS≥75), two classrooms whenever possible otherwise
- Small schools sampled with equal probabilities

B.50.1 Allocation of School Sample in Slovenia – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Slovenia	150	0	138	9	1	2
Total	**150**	**0**	**138**	**9**	**1**	**2**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- The population covered is eighth grade. Some students, who where put into the new school system after they started school in the old system, were promoted from their grade 5 of the old system to grade 7 of the new system. Those students were in their 7th year of schooling in grade 8 of the new system
- School-level exclusions consisted of very small schools (MOS<8), special needs schools, and Waldorf schools
- Within-school exclusions consisted of students with functional disabilities and students not able to talk or understand Slovene language

Sample Design

- No explicit stratification
- Implicit stratification by region (eight regions), for a total of eight implicit strata
- Sampled two classrooms per school having at least 40 students (MOS≥40), one classroom otherwise
- Small schools sampled with equal probabilities

B.50.2 Allocation of School Sample in Slovenia – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Slovenia	150	0	138	9	1	2
Total	**150**	**0**	**138**	**9**	**1**	**2**

B.51 Sweden

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<5), special schools for disabled students, and non-Swedish language schools (international schools)
- Within-school exclusions consisted special education classes, disabled students and non-native language speakers (less than one year instruction in Swedish)

Sample Design

- Explicit stratification by principal organiser (public, independent), for a total of two explicit strata
- No implicit stratification
- Sampled two classrooms per school whenever possible
- Small schools sampled with equal probabilities

B.51.1 Allocation of School Sample in Sweden – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Public	130	5	122	3	0	0
Independent	30	0	29	1	0	0
Total	**160**	**5**	**151**	**4**	**0**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<5), special schools for disabled students, and non-Swedish language schools (international schools)
- Within-school exclusions consisted special education classes, disabled students and non-native language speakers (less than one year instruction in Swedish)

Sample Design

- Explicit stratification by principal organizer (public, independent), for a total of two explicit strata
- No implicit stratification
- Sampled two classrooms per school whenever possible
- Small schools sampled with equal probabilities

B.51.2 Allocation of School Sample in Sweden – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Public	130	1	129	0	0	0
Independent	30	0	29	1	0	0
Total	**160**	**1**	**158**	**1**	**0**	**0**

B.52 Syrian Arab Republic

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<15)
- No within-school exclusions

Sample Design

- Explicit stratification by urbanization (urban, rural) and gender (boys, girls, mixed), for a total of six explicit strata
- Implicit stratification by governorate (14 governorates) and school type (public, private), for a total of 91 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.52.1 Allocation of School Sample in Syrian Arab Republic – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Urban – Girls	20	0	20	0	0	0
Urban – Boys	20	0	20	0	0	0
Urban – Mixed	13	0	13	0	0	0
Rural – Girls	14	0	14	0	0	0
Rural – Boys	13	0	13	0	0	0
Rural – Mixed	70	0	70	0	0	0
Total	**150**	**0**	**150**	**0**	**0**	**0**

B.53 Thailand

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<15), special education schools, special curriculum schools, and religious schools
- No within-school exclusions

Sample Design

- Explicit stratification by geographical location (Bangkok, northern, northeastern, central, southern), for a total of five explicit strata
- Implicit stratification by school type (Office of the Basic Education Commission, Office of the Private Education Commission, Office of the Higher Education Commission, Department of Local Administration, Department of Education – Bangkok Metropolitan Administration) and urbanization (urban, rural), for a total of 29 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities
- Post-stratification adjustments to sampling weights were done by gender in Bangkok, Northern and Central strata

B.53.1 Allocation of School Sample in Thailand – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Bangkok	14	0	11	3	0	0
Northern	28	0	26	2	0	0
North-Eastern	36	0	32	4	0	0
Central	54	0	49	4	1	0
Southern	18	0	16	2	0	0
Total	**150**	**0**	**134**	**15**	**1**	**0**

B.54 Tunisia

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<8) and schools with multilevel classes
- Within-school exclusions consisted of functionally or intellectually disabled student

Sample Design

- No explicit stratification
- Implicit stratification by school type (public, private), school priority education programme (PEP, non-PEP), and geographic location (northeast, northwest, southeast, southwest), for a total of ten implicit strata
- Sampled two classrooms per school having at least 70 students (MOS≥70), and one classroom otherwise
- Small schools sampled with equal probabilities

B.54.1 Allocation of School Sample in Tunisia – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Tunisia	150	0	150	0	0	0
Total	**150**	**0**	**150**	**0**	**0**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- No school-level exclusions
- No within-school exclusions

Sample Design

- No explicit stratification
- Implicit stratification by school priority education zone programme (PEP, non-PEP) and geographic location (northeast, northwest, southeast, southwest), for a total of eight implicit strata
- Sampled two classrooms per school having at least 375 students (MOS≥375), and one classroom otherwise
- Small schools sampled with equal probabilities

B.54.2 Allocation of School Sample in Tunisia – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Tunisia	150	0	150	0	0	0
Total	**150**	**0**	**150**	**0**	**0**	**0**

B.55 Turkey

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<10), special education schools, and schools that were difficult to reach (travelling difficulties)
- Within-sample exclusions consisted of functionally or intellectually disabled students

Sample Design

- Explicit stratification by region, for a total of seven explicit strata
- Implicit stratification by school type (public, private), for a total of 14 implicit strata
- Sampled one classroom per school
- Small schools sampled with equal probabilities

B.55.1 Allocation of School Sample in Turkey – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Marmara Bölgesi	40	1	39	0	0	0
Iç Anadolu	26	0	26	0	0	0
Ege Bölgesi	18	0	18	0	0	0
Akdeniz Bölgesi	18	0	18	0	0	0
Karadeniz Bölgesi	16	0	16	0	0	0
Dogu Anadolu	14	2	12	0	0	0
Güney Dogu Anadolu	18	1	17	0	0	0
Total	**150**	**4**	**146**	**0**	**0**	**0**

B.56 Ukraine

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<6)
- No within-school exclusions

Sample Design

- Explicit stratification by grade (grade 4 only, grade 4 & grade 8) and urbanization (towns, villages, big villages) in the 'grade 4 & grade 8' stratum, for a total of four explicit strata
- Implicit stratification by region (27 regions) in the 'grade 4 & grade 8 – towns' and 'grade 4 & grade 8 – villages' strata , for a total of 55 implicit strata
- Sampled two classrooms per school having at least 65 students (MOS≥65), and one classroom otherwise
- Small schools sampled with equal probabilities

B.56.1 Allocation of School Sample in Ukraine – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 4 Only	4	0	4	0	0	0
Grade 4 & 8 – Towns	90	0	87	0	0	3
Grade 4 & 8 – Villages	54	0	52	0	0	2
Grade 4 & 8 – Big Villages	2	0	1	0	0	1
Total	**150**	**0**	**144**	**0**	**0**	**6**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<6)
- No within-school exclusions

Sample Design

- Explicit stratification by grade (grade 8 only, grade 4 & grade 8) and urbanization (towns, villages, big villages) in the 'grade 4 & grade 8' stratum, for a total of four explicit strata
- Implicit stratification by region (27 regions) in the 'grade 4 & grade 8 – towns' and 'grade 4 & grade 8 – villages' strata, for a total of 56 implicit strata
- Sampled one or two classrooms per school
- Small schools sampled with equal probabilities

B.56.2 Allocation of School Sample in Ukraine – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 8 Only	10	0	10	0	0	0
Grade 4 & 8 – Towns	86	0	85	0	0	1
Grade 4 & 8 – Villages	52	0	50	0	0	2
Grade 4 & 8 – Big Villages	2	0	1	0	0	1
Total	**150**	**0**	**146**	**0**	**0**	**4**

B.57 United States

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- No school-level exclusions
- Within-school exclusions consisted of special education classes, disabled students within regular classes, and students unable to be tested in English

Sample Design

- No explicit stratification

- Implicit stratification by school type (public, private), geographic location (northeast, southeast, mid-west, west), location indicator relative to populous areas (8 categories), and minority status (above/below 15% minority), for a total of 128 implicit strata
- Sampled two classrooms per school
- Small schools sampled with equal probabilities

B.57.1 Allocation of School Sample in United States – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
United States	300	10	202	46	9	33
Total	**300**	**10**	**202**	**46**	**9**	**33**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- No school-level exclusions
- Within-school exclusions consisted of special education classes, disabled students within regular classes, and students unable to be tested in English

Sample Design

- No explicit stratification
- Implicit stratification by school type (public, private), geographic location (northeast, southeast, mid-west, west), location indicator relative to populous areas (8 categories), and minority status (above/below 15% minority), for a total of 128 implicit strata
- Sampled two classrooms per school
- Small schools sampled with equal probabilities

B.57.2 Allocation of School Sample in United States – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
United States	300	13	197	33	9	48
Total	**300**	**13**	**197**	**33**	**9**	**48**

B.58 Yemen

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<8) and private schools (English sections only)
- Within-school exclusions consisted of functionally or intellectually disabled students

Sample Design

- Explicit stratification by urbanization (urban, rural), for a total of two explicit strata
- Implicit stratification by school type (pubic, private), gender (boys, girls, mixed) in the 'public' strata, and region (22 regions) in the 'rural – public – mixed' stratum, for a total of 29 implicit strata
- Sampled one classroom per school (in two of the sampled schools, two classrooms were selected)
- Small schools sampled with equal probabilities

B.58.1 Allocation of School Sample in Yemen – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Rural	113	6	106	1	0	0
Urban	37	0	37	0	0	0
Total	**150**	**6**	**143**	**1**	**0**	**0**

Benchmarking Participants

B.59 Alberta, Canada

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<6) and schools for online/correspondence students
- Within-school exclusions consisted of physically or cognitvely disabled students and non-native language speakers

Sample Design

- No explicit stratification
- Implicit stratification by school type (charter, francophone, public, private, separate), for a total of five implicit strata
- Sampled two classrooms per school having at least 60 students (MOS≥60), and one classroom otherwise
- Small schools sampled with equal probabilities

B.59.1 Allocation of School Sample in Alberta, Canada – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Canada, Alberta	150	2	146	0	0	2
Total	**150**	**2**	**146**	**0**	**0**	**2**

B.60 Basque Country, Spain

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<10), special needs schools and schools for students taught in another language than Basque or Castilian
- Within-school exclusions consisted of disabled students and non-official language speakers

Sample Design

- Prior to school selection, schools were split into up to three parts. These parts (based on languages) were then considered as primary sampling units (PSU). As a result, schools had chances to be selected up to three times in the sample of PSUs. In fact, 16 schools were sampled twice and one school was sampled three times
- Explicit stratification by school type (public, private) and language (type A: Castilian, type B: mixed, type D: Basque), for a total of six explicit strata
- School sample allocation is not proportional at the explicit stratum level
- Implicit stratification by province (Araba, Bizkaia, Gipuzkoa), for a total of 18 implicit strata
- Sampled two classrooms in schools with at least 60 students in the Private, mixed stratum, with at least 50 students in the Private, Basque stratum, and with at least 80 students in the Public, Basque stratum, and one classroom otherwise
- Small schools sampled with equal probabilities

B.60.1 Allocation of School Sample in Basque Country, Spain – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Private – Type A (Castilian)	20	0	20	0	0	0
Private – Type B (Mixed)	20	0	20	0	0	0
Private – Type D (Basque)	20	0	20	0	0	0
Public – Type A (Castilian)	20	0	20	0	0	0
Public – Type B (Mixed)	20	0	20	0	0	0
Public – Type D (Basque)	30	0	30	0	0	0
Total	**130**	**0**	**130**	**0**	**0**	**0**

B.61 British Columbia, Canada

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<6), alternate and distance education schools, district distance education schools, and long term provincial resource program (PRP) schools
- Within-school exclusions consisted of functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by grade (grade 4 only, grade 4 & grade 8), for a total of two explicit strata
- Implicit stratification by school type (public, independent), for a total of four implicit strata
- Sampled two classrooms per school having at least 52 students (MOS≥52), and one classroom otherwise
- Small schools sampled with equal probabilities

B.61.1 Allocation of School Sample in British Columbia, Canada – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 4 Only	140	0	137	3	0	0
Grade 4 & 8	10	0	10	0	0	0
Total	**150**	**0**	**147**	**3**	**0**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<6), alternate and distance education schools, district distance education schools, and long term provincial resource program (PRP) schools
- Within-school exclusions consisted of special education classes, functionally or intellectually disabled students, and non-native language speakers. Classes from semester schools were also considered excluded

since they were not considered for sampling as they were not attending a mathematics class during the semester

Sample Design

- Explicit stratification by grade (grade 8 only, grade 4 & grade 8), for a total of two explicit strata
- Implicit stratification by school type (public, independent) and facility type (continuing education, standard) in the 'grade 8 only' stratum, for a total of five implicit strata
- Sampled two classrooms per school having at least 270 students (MOS≥270), and one classroom otherwise
- Small schools sampled with equal probabilities

B.61.2 Allocation of School Sample in British Columbia, Canada – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 8 Only	140	0	137	3	0	0
Grade 4 & 8	10	0	10	0	0	0
Total	**150**	**0**	**147**	**3**	**0**	**0**

B.62 Dubai, United Arab Emirates

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of schools teaching neither in Arabic nor in English language
- Within-school exclusions consisted of disabled students

Sample Design

- Explicit stratification by grade (grade 4/5 only, grade 4/5 & grade 8/9) and by schedule (Indian, Non-Indian)
- No implicit stratification
- All schools sampled
- Sampled one or two classrooms per school
- In schools that followed the Indian time schedule, students were tested at the end of their fourth school year. In schools that followed the non-Indian time schedule, students were tested at the beginning of their fifth school year

B.62.1 Allocation of School Sample in Dubai, United Arab Emirates – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 4 Only - Indian	4	1	3	0	0	0
Grade 5 Only - Non- Indian	49	2	38	0	0	9
Grade 4 & 8 - Indian	15	3	10	0	0	2
Grade 5 & 9 - Non- Indian	75	5	46	0	0	24
Total	**143**	**11**	**97**	**0**	**0**	**35**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of schools teaching neither in Arabic nor in English language
- Within-school exclusions consisted of disabled students

Sample Design

- Explicit stratification by grade (grade 8/9 only, grade 4/5 & grade 8/9) and by schedule (Indian, Non-Indian)
- No implicit stratification
- All schools sampled
- Sampled one or two classrooms per school
- In schools that followed the Indian time schedule, students were tested at the end of their eight school year. In schools that followed the non-Indian time schedule, students were tested at the beginning of their ninth school year

B.62.2 Allocation of School Sample in Dubai, United Arab Emirates – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 8 Only - Indian	2	0	2	0	0	0
Grade 9 Only - Non- Indian	31	2	28	0	0	1
Grade 4 & 8 - Indian	15	2	11	0	0	2
Grade 5 & 9 - Non- Indian	74	3	47	0	0	24
Total	**122**	**7**	**88**	**0**	**0**	**27**

B.63 Massachusetts, United States

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100% of public schools
- No school-level exclusions
- Within-school exclusions consisted of special education classes, disabled students within regular classes, and students unable to be tested in English

Sample Design

- Sample design based on the 2007 NAEP sample design. To reduce duplication, the sample of schools was drawn at the midpoints between the 2007 NAEP sample schools. A sub-sample of schools was then drawn to bring the final school sample to 50
- Sampled two classrooms per school whenever possible
- Sampling variance was computed by pairing these 50 schools following the order of selection. Sampling weights were derived from the NAEP design
- Sampling weights were derived from the NAEP design including adjustments for sub-sampling

B.63.1 Allocation of School Sample in Massachusetts, United States – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
USA (Massachusetts)	50	1	45	2	0	2
Total	**50**	**1**	**45**	**2**	**0**	**2**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100% of public schools
- No school-level exclusions
- Within-school exclusions consisted of special education classes, disabled students within regular classes, and students unable to be tested in English

Sample Design

- Sample design based on the 2007 NAEP sample design. To reduce duplication, the sample of schools was drawn at the midpoints between the 2007 NAEP sample schools. A sub-sample of schools was then drawn to bring the final school sample to 50
- Sampled two classrooms per school whenever possible
- Sampling variance was computed by pairing these 50 schools following the order of selection. Sampling weights were derived from the NAEP design
- Sampling weights were derived from the NAEP design including adjustments for sub-sampling

B.63.2 Allocation of School Sample in Massachusetts, United States – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
USA (Massachusetts)	50	1	45	3	0	1
Total	**50**	**1**	**45**	**3**	**0**	**1**

B.64 Minnesota , United States

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100% of public schools
- No school-level exclusions
- Within-school exclusions consisted of special education classes, special education students needing accommodations within regular classes, and students unable to be tested in English

Sample Design

- Sample design based on the 2007 NAEP sample design. To reduce duplication, the sample of schools was drawn at the midpoints between the 2007 NAEP sample schools. A sub-sample of schools was then drawn to bring the final school sample to 50
- Sampled two classrooms per school whenever possible

- Sampling variance was computed by pairing these 50 schools following the order of selection. Sampling weights were derived from the NAEP design
- Sampling weights were derived from the NAEP design including adjustments for sub-sampling

B.64.1 Allocation of School Sample in Minnesota, United States – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
USA (Minnesota)	50	0	30	15	5	0
Total	**50**	**0**	**30**	**15**	**5**	**0**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100% of public schools
- No school-level exclusions
- Within-school exclusions consisted of special education classes, special education students needing accommodations within regular classes, and students unable to be tested in English

Sample Design

- Sample design based on the 2007 NAEP sample design. To reduce duplication, the sample of schools was drawn at the midpoints between the 2007 NAEP sample schools. A sub-sample of schools was then drawn to bring the final school sample to 50
- Sampled two classrooms per school whenever possible
- Sampling variance was computed by pairing these 50 schools following the order of selection Sampling weights were derived from the NAEP design
- Sampling weights were derived from the NAEP design including adjustments for sub-sampling

B.64.2 Allocation of School Sample in Minnesota, United States – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
USA (Minnesota)	50	0	32	8	9	1
Total	**50**	**0**	**32**	**8**	**9**	**1**

B.65 Ontario, Canada

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<10)
- Within-school exclusions consisted of special education classes, functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by grade (grade 4 only, grade 4 & grade 8) and language (French, English), for a total of four explicit strata
- School sample allocation is not proportional at the explicit strata level
- Implicit stratification by school type (public, private, separate), for a total of 11 implicit strata
- Sampled two classrooms per school having at least 75 students (MOS≥75), and one classroom otherwise
- Small schools sampled with equal probabilities

B.65.1 Allocation of School Sample in Ontario, Canada – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 4 Only – English	23	2	21	0	0	0
Grade 4 Only – French	32	0	28	4	0	0
Grade 4 & 8 – English	97	1	89	2	0	5
Grade 4 & 8 – French	48	0	41	3	0	4
Total	**200**	**3**	**179**	**9**	**0**	**9**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<10)
- Within-school exclusions consisted of special education classes, functionally or intellectually disabled students and non-native language speakers

Sample Design

- Explicit stratification by grade (grade 8 only, grade 4 & grade 8) language (French, English), for a total of four explicit strata
- School sample allocation is not proportional at the explicit strata level
- Implicit stratification by school type (public, private, separate), for a total of 11 implicit strata
- Sampled two classrooms per school having at least 140 students (MOS≥140), and one classroom otherwise
- Small schools sampled with equal probabilities

B.65.2 Allocation of School Sample in Ontario, Canada – Eighth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 8 Only – English	23	0	20	2	0	1
Grade 8 Only – French	32	1	26	1	0	4
Grade 4 & 8 – English	97	2	87	2	0	6
Grade 4 & 8 – French	48	6	35	3	0	4
Total	**200**	**9**	**168**	**8**	**0**	**15**

B.66 Québec, Canada

FOURTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<11), special needs schools, native schools, and non-ministry schools
- Within-school exclusions consisted of special needs students. Some special need classes as well as some classes with international curricula were also excluded.

Sample Design

- Explicit stratification by grade (grade 4 only, grade 4 & grade 8) and language (French, English), for a total of four explicit strata
- School sample allocation is not proportional at the explicit stratum level
- Implicit stratification by school type (public, private), for a total of eight implicit strata
- Sampled two classrooms per school having at least 80 students (MOS≥80), and one classroom otherwise
- Small schools sampled with equal probabilities

B.66.1 Allocation of School Sample in Québec, Canada – Fourth Grade

Explicit Stratum	Total sampled schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 4 Only – English	74	7	63	0	0	4
Grade 4 Only – French	114	1	111	0	0	2
Grade 4 & 8 – English	6	0	6	0	0	0
Grade 4 & 8 – French	6	0	5	1	0	0
Total	**200**	**8**	**185**	**1**	**0**	**6**

EIGHTH GRADE

Coverage and Exclusions

- Coverage is 100%
- School-level exclusions consisted of very small schools (MOS<11), special needs schools, native schools, and non-ministry schools
- Within-school exclusions consisted of special needs students. Some special need classes as well as some classes with international curricula were also excluded.

Sample Design

- Explicit stratification by grade (grade 8 only, grade 4 & grade 8) and language (French, English), for a total of four explicit strata
- School sample allocation is not proportional at the explicit stratum level
- Implicit stratification by school type (public, private), for a total of eight implicit strata
- Sampled two classrooms per school having at least 280 students (MOS≥280), and one classroom otherwise
- Small schools sampled with equal probabilities

B.66.2 Allocation of School Sample in Québec, Canada – Eighth Grade

Explicit Stratum	Total Sampled Schools	Ineligible Schools	Participating Schools			Non-Participating Schools
			Sampled	1st Replacement	2nd Replacement	
Grade 8 Only – English	65	5	56	0	0	4
Grade 8 Only – French	114	3	103	0	0	8
Grade 4 & 8 – English	6	0	6	0	0	0
Grade 4 & 8 – French	6	0	5	0	0	1
Total	**191**	**8**	**170**	**0**	**0**	**13**

Appendix C

Country Adaptations to Items and Item Scoring

Fourth Grade

Items deleted

ALL COUNTRIES

M09_04, M13_04 Mathematics (faulty distracters)
S08_07, S09_06, S12_10 Science (faulty distracters)

ALGERIA

M08_08, M10_09, M12_13B Mathematics (scorer reliability less than 70%)
S06_05 Science (poor discrimination)
S08_05 Science (scorer reliability less than 70%)

AUSTRIA

S07_02 Science (translation error)

COLOMBIA

S11_04 Science (too few valid responses)

DENMARK

M11_03 Mathematics (printing error)

DUBAI, UAE

M09_05 Mathematics (negative discrimination)

EL SALVADOR

M05_03 Mathematics (poor discrimination)

GEORGIA

M03_03, M05_06 Mathematics (negative discrimination)
S04_04 Science (scorer reliability less than 70%)

HONG KONG SAR

S10_05 Science (negative discrimination)

IRAN, ISLAMIC REP. OF

S10_05 Science (negative discrimination)

KUWAIT

S07_10 Science (negative discrimination)
S11_04 Science (erroneous application of scoring guide)

LATVIA

M04_07 Mathematics (scorer reliability less than 70%)
S01_01A, S02_13B, S03_07, S04_04, S10_03, S10_06A, S10_06B Science (scorer reliability less than 70%)

LITHUANIA

M09_08, M09_09 Mathematics (printing error in Booklet 8 only)
S06_12 Science (translation error)

MONGOLIA

M02_08A, M02_08B, M02_08C, M02_08D, M03_08A, M03_08B, M03_08C Mathematics (administered without manipulatives)
M08_04B Mathematics (erroneous application of scoring guide)
S04_11, S06_13A, S06_13B, S08_09A, S11_06, S14_10B, S14_12 Science (erroneous application of scoring guide)
S05_02A, S05_02B Science (scorer reliability less than 70%)
S10_04 Science (negative discrimination)

MOROCCO

M10_09, M12_13B Mathematics (scorer reliability less than 70%)

QATAR

S07_10 Science (negative discrimination)

SWEDEN

S04_04, S08_09B Science (scorer reliability less than 70%)

TUNISIA

M08_03 Mathematics (poor discrimination)
S04_14 Science (negative discrimination)
S07_10 Science (poor discrimination)
S08_08 Science (too few valid responses)

Constructed-response items needing category recoding

ALL COUNTRIES

M02_11, M03_10 Mathematics (recode 11 to 70)
M04_09A Mathematics (recode 11 to 71)
M08_08 Mathematics (recode 20 to 10, 10 to 70)
M09_12 Mathematics (recode 20 to 10, 10 to 71)
M13_01C Mathematics (recode 21 to 12)
S06_03 Science (recode 20 to 10, 10 to 79)
S06_04 Science (recode 11 to 70)
S06_06 Science (recode 19 to 70)
S06_12, S06_13A, S06_13B Science (recode 19 to 10)
S08_09A Science (recode 11 to 10)
S09_08 Science (recode 20 to 10 and 10, 11, 12 to 79)
S09_10 Science (recode 19 to 10)
S10_10 Science (recode 13 to 11)
S11_04 Science (recode 19 to 79)
S11_06 Science (recode 70 to 79)
S12_09 Science (recode 70 to 11)
S13_04 (recode 19 to 10)
S13_10 Science (recode 71 to 70)
S14_06 Science (recode 21 to 10 and 10, 11, 70 to 79)
S14_10A Science (recode 19 to 10, 11 to 79)
S14_10B Science (recode 19 to 79)
S14_10A and S14_10B combined to form S14_10D,
a new 2-point item, as follows:
If S14_10A=10 and S14_10B=10 or 11, S14_10D → 20.
If S14_10A=10 or S14_10B=10 or 11, S14_10D → 10.
Otherwise, S14_10D → 70.

Eighth Grade

Items deleted

ALL COUNTRIES

M02_15 Mathematics (poor discrimination)
S03_08 Science (poor discrimination)
S06_12C Science (faulty scoring guide)
S10_16, S14_13 Science (faulty distracters)

ALGERIA

M02_04 Mathematics (negative discrimination)
M04_05B Mathematics (scorer reliability less than 70%)
S14_07 Science (negative discrimination)

BAHRAIN

S08_09 Science (negative discrimination)

BASQUE COUNTRY, SPAIN

M02_04 Mathematics (negative discrimination)

BOSNIA AND HERZEGOVINA

M10_12B Mathematics (negative discrimination)

BOTSWANA

M07_03, M09_02, M09_10 Mathematics (negative discrimination)
S01_01, S08_10 Science (poor discrimination)
S11_09 Science (negative discrimination)

BULGARIA

S13_12, S14_02 Science (scorer reliability less than 70%)

CHINESE TAIPEI

S04_09 Science (scorer reliability less than 70%)
S09_07 Science (negative discrimination)

COLOMBIA

M11_13A Mathematics (too few valid responses)

CYPRUS

M09_13 Mathematics (erroneous application of scoring guide)

EGYPT

S08_09 Science (negative discrimination)

GEORGIA

M07_06, M12_13 Mathematics (negative discrimination)
S09_02 Science (poor discrimination)
S10_17, S14_02 Science (scorer reliability less than 70%)

GHANA

S01_01 Science (poor discrimination)
S06_02, S10_03, S11_09 Science (negative discrimination)

INDONESIA

S11_09 Science (poor discrimination)

ITALY

M03_10 Mathematics (poor discrimination)
S14_04A, S14_04C Science (scorer reliability less than 70%)

JAPAN

S10_17, S13_03, S13_12 Science (scorer reliability less than 70%)

JORDAN

M13_08 Mathematics (negative discrimination)

KOREA, REP. OF

S08_04 Science (negative discrimination)

KUWAIT

M04_07 Mathematics (negative discrimination)
S08_09 Science (negative discrimination)

LEBANON

S04_11C Science (negative discrimination)

MINNESOTA, US

S09_07 Science (negative discrimination)

MONGOLIA

M03_10, M06_13, M11_08 Mathematics (poor discrimination)
M05_07C Mathematics (erroneous application of scoring guide)
M08_12 Mathematics (negative discrimination)
S04_11B, S05_08, S06_03, S06_07 Science (erroneous application of scoring guide)
S06_08 Science (poor discrimination)
S07_08, S10_07 Science (negative discrimination)
S10_06, S10_09, S10_10, S10_12, S10_17 Science (scorer reliability less than 70%)

MOROCCO

M02_04, M04_13 Mathematics (negative discrimination)
S01_06, S03_12, S04_04, S10_09, S10_12, S11_04 Science (scorer reliability less than 70%)
S08_09, S09_07 Science (negative discrimination)

NORWAY

M02_02 Mathematics (printing error in Booklet 2 only)
S11_10 Science (too few valid responses in Booklets 1-14 only)

OMAN

S07_08 Science (poor discrimination)

PALESTINIAN NAT'L AUTH.

M05_06 Mathematics (printing error)
M07_03, M13_08 Mathematics (negative discrimination)
S06_02 Science (poor discrimination)
S07_08, S08_09 Science (negative discrimination)

QATAR

S06_02, S07_10 Science (poor discrimination)
S10_06 Science (translation error)

ROMANIA

M12_14 Mathematics (negative discrimination)

SAUDI ARABIA

S08_09 Science (negative discrimination)

SERBIA

M10_12B Mathematics (negative discrimination)
S08_04 Science (poor discrimination)

SWEDEN

M10_08 Mathematics (negative discrimination)
S11_10 Science (too few valid responses in Booklets 1-14 only)

SYRIAN ARAB REPUBLIC

M07_03 Mathematics (negative discrimination)
S14_07 Science (negative discrimination)

THAILAND

S08_04 Science (negative discrimination)

TUNISIA

M11_02 Mathematics (translation error in Booklets 1-14 only)
S04_04, S10_17 Science (scorer reliability less than 70%)

UKRAINE

S04_04 Science (scorer reliability less than 70%)

Constructed-response items needing category recoding

ALL COUNTRIES

M04_12A Mathematics (recode 11 to 70, 70 to 71)
M05_05 Mathematics (recode 20 to 10, 10 to 70)
M08_05 Mathematics (recode 11 to 70)
S01_08B Science (recode 20 to 11, 29 to 19)
S04_06 Science (recode 12 to 71)
S05_14B Science (recode 71 to 12, 72 to 71)
S06_07 Science (recode 10 to 20, 11 to 10, 70 to 11)
S07_11 Science (recode 20 to 10, 29 to 19, 10 to 12)
S08_12 Science (recode 20 to 10, 21 to 11, 29 to 19, 10 to 12)
S08_13 Science (recode 20 to 10, 10 to 70, 11 to 71)
S09_10B Science (recode 20 to 10, 29 to 19, 10 to 12)
S10_10 Science (recode 11 to 70, 70 to 71)
S12_12 Science (recode 11 to 70)
S14_02 Science (recode 21 to 20, 11 to 10)
S14_05 Science (recode 20 to 10, 21 to 11, 10 to 12)

Country Adaptations to Bridge Items and Bridge Item Scoring

Fourth Grade

Items deleted

ALL COUNTRIES

S02_08 Science (deleted in TIMSS 2003)

ARMENIA

S01_03, S01_04 Science (deleted in TIMSS 2003)
S11_04 Science (trend difference greater than two logits)

CHINESE TAIPEI

M05_10 Mathematics (scorer reliability less than 70%)

HUNGARY

M09_06A, M09_06B Mathematics (too few valid responses)
M14_06 Mathematics (deleted in TIMSS 2003)
S10_02 Science (deleted in TIMSS 2003)

IRAN, ISLAMIC REP. OF

S05_02 Science (trend difference greater than two logits)
S05_06 Science (deleted in TIMSS 2003)

LITHUANIA

M12_06C Mathematics (deleted in TIMSS 2003)

MOROCCO

M11_10 Mathematics (deleted in TIMSS 2003)

NETHERLANDS

S13_03 Science (deleted in TIMSS 2003)

TUNISIA

S05_06 Science (deleted in TIMSS 2003)

Eighth Grade

Items deleted

ALL COUNTRIES

M06_07, M06_11, M06_12, M06_13 Mathematics (calculator item in TIMSS 2003)

ARMENIA

S13_10 Science (erroneous application of scoring guide)

BAHRAIN

M10_10 Mathematics (poor discrimination)

BOTSWANA

M02_14 Mathematics (negative discrimination)
S02_01, S05_01 Science (deleted in TIMSS 2003)

CYPRUS

S02_01 Science (negative discrimination)

EGYPT

M01_12 Mathematics (deleted in TIMSS 2003)
S09_13, S14_01 Science (deleted in TIMSS 2003)

GHANA

M01_01 Mathematics (deleted in TIMSS 2003 in Booklet 12 (B4 in TIMSS 2007 only)
M01_05, M01_12, M01_14, M02_06, M02_07, M02_12, M02_14, M02_15, M05_09, M09_05, M09_10, M10_02 , M10_03, M10_04, M10_05, M11_02, M11_05, M11_06, M11_10, M11_12, M12_04, M12_05, M12_09, M12_10, M13_02, M13_04, M13_05, M13_06, M13_09, M13_10, M13_11, M14_01, M14_05, M14_09 Mathematics (deleted in TIMSS 2003)
M02_04 Mathematics (negative discrimination)
M10_10 Mathematics (poor discrimination)
M12_03 Mathematics (not administered in TIMSS 2003)
S01_03, S09_03 Science (deleted in TIMSS 2003)
S09_06, S09_13 Science (negative discrimination)

HONG KONG SAR

S01_01, S11_05 Science (negative discrimination)

HUNGARY

S01_13 Science (negative discrimination)

INDONESIA

M02_04 Mathematics (deleted in TIMSS 2003)
M10_10 Mathematics (poor discrimination)
S01_08, S05_01 Science (negative discrimination)

IRAN, ISLAMIC REP. OF

S10_09A Science (negative discrimination)

ISRAEL

S09_06, S11_07 Science (negative discrimination)

ITALY

S10_10B Science (poor discrimination)

JAPAN

S01_09 Science (scorer reliability less than 70%)
S10_06 Science (poor discrimination)

JORDAN

M09_01, M14_02 Mathematics (deleted in TIMSS 2003)
S09_05, S09_06 Science (deleted in TIMSS 2003)

KOREA, REP. OF

S02_11 Science (deleted in TIMSS 2003)
S11_07 Science (negative discrimination)

LEBANON

M01_06 Mathematics (deleted in TIMSS 2003)

LITHUANIA

M02_14 Mathematics (translation error)
M13_08 Mathematics (negative discrimination)
S14_09 Science (deleted in TIMSS 2003)

MALAYSIA

M10_02 Mathematics (trend difference greater than two logits)

MOROCCO

S01_08, S01_12 Science (deleted in TIMSS 2003)
S01_11, S13_06, S14_03 Science (negative discrimination)

NORWAY

M13_08 Mathematics (negative discrimination)
M13_10 Mathematics (poor discrimination)

PALESTINIAN NAT'L AUTH.

S01_06, S01_16, S09_05, S09_06 Science (deleted in TIMSS 2003)

QUEBEC, CANADA

M05_08 Mathematics (deleted in TIMSS 2003)

ROMANIA

M01_05, M01_07, M01_14 Mathematics (printing error in B4 only)
M01_09, M10_07, M10_08, M10_09, M10_11, M12_01, M12_10, M12_12, M12_13A, M12_13B, M12_13C, M14_03, M14_04A, M14_04B, M14_04C, M14_05, M14_06 Mathematics (printing error)
M02_03, M02_04, M02_05, M02_08, M02_09, M02_11, M12_04, M12_05, M12_08 Mathematics (printing error in Hungarian version only)
M12_11 Mathematics (deleted in TIMSS 2003)
S01_07 Science (trend difference greater than two logits)
S01_08 Science (printing error in B4 only)
S01_10 Science (printing error in Hungarian B4 only)
S05_11, S05_13, S06_05, S06_06 Science (printing error in Hungarian version only)
S06_08 Science (printing error in Hungarian B3 only)
S10_04, S10_09A, S10_09B, S10_10A, S10_10B, S12_11, S14_08A, S14_08B Science (printing error)

SINGAPORE

M11_09 Mathematics (trend difference greater than two logits)

SLOVENIA

M10_09 Mathematics (too few valid responses)
S14_08B, S14_09 Science (deleted in TIMSS 2003)

TUNISIA

M11_06, M12_13A Mathematics (trend difference greater than two logits)
S01_06, S05_01, S09_11, S10_01 Science (negative discrimination)
S09_13 Science (poor discrimination)
S13_01 Science (deleted in TIMSS 2003)

Constructed-response items needing category recoding

ALL COUNTRIES

M13_14 Mathematics (recode 20 to 10, 10 to 70)
S09_03 Science (recode 20 to 10, 29 to 19, 10 to 12)
S10_06 Science (recode 20 to 10, 21 to 11, 22 to 12, 29 to 19, 10 to 71, 11 to 72, 19 to 79)
S13_10 Science (recode 20 to 10, 29 to 19, 10 to 70, 19 to 79, 70 to 71, 71 to 72)

Appendix D

Item Parameters for IRT Analyses of TIMSS 2007 Data

Exhibit D.1 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Mathematics

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M011009	M01_01	1.200 (0.060)	-1.118 (0.059)	0.148 (0.031)		
M011010	M01_02	1.261 (0.059)	-0.201 (0.038)	0.136 (0.019)		
M012044	M01_03	0.955 (0.046)	-0.056 (0.044)	0.073 (0.019)		
M011011	M01_04	1.159 (0.059)	-0.769 (0.057)	0.170 (0.029)		
M011017	M01_05	0.710 (0.039)	-0.403 (0.075)	0.113 (0.028)		
M011018	M01_06	0.764 (0.039)	-1.066 (0.083)	0.117 (0.033)		
M011019	M01_07	0.718 (0.044)	-0.412 (0.090)	0.120 (0.035)		
M011020	M01_08	0.948 (0.074)	0.903 (0.047)	0.189 (0.018)		
M012065	M01_09	0.952 (0.057)	0.418 (0.046)	0.152 (0.019)		
M011023	M01_10	0.588 (0.039)	-0.999 (0.142)	0.171 (0.047)		
M011024	M01_11	0.824 (0.043)	-1.814 (0.097)	0.122 (0.039)		
M012048	M01_12	0.887 (0.052)	-0.064 (0.061)	0.142 (0.026)		
M011012	M02_01	0.888 (0.062)	-1.054 (0.087)	0.114 (0.035)		
M011013	M02_02	0.680 (0.079)	0.135 (0.129)	0.210 (0.043)		
M011014	M02_03	0.817 (0.062)	-1.645 (0.122)	0.136 (0.044)		
M011015	M02_04	0.662 (0.060)	-0.323 (0.117)	0.113 (0.041)		
M011016	M02_05	0.901 (0.087)	0.161 (0.081)	0.171 (0.033)		
M012078	M02_06	0.918 (0.065)	-0.888 (0.085)	0.122 (0.035)		
M012119	M02_07	0.625 (0.064)	-0.174 (0.136)	0.162 (0.045)		
M011021	M02_08	0.751 (0.065)	-0.634 (0.120)	0.136 (0.045)		
M012023	M02_09	0.800 (0.075)	-0.478 (0.122)	0.208 (0.046)		
M011022	M02_10	0.576 (0.047)	-1.006 (0.129)	0.119 (0.039)		
M011003	M02_11	0.682 (0.060)	-0.407 (0.115)	0.114 (0.041)		
M011004	M02_12	0.859 (0.065)	-0.913 (0.099)	0.138 (0.040)		
M011005	M02_13	0.513 (0.047)	-1.735 (0.197)	0.150 (0.050)		
M012126	M03_01	0.873 (0.065)	-0.604 (0.084)	0.127 (0.034)		
M011006	M03_02	0.465 (0.045)	-0.164 (0.140)	0.116 (0.037)		
M012117	M03_03	0.935 (0.101)	0.464 (0.077)	0.209 (0.029)		

( ) Standard errors appear in parentheses

Exhibit D.1 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M011007	M03_04	0.973 (0.074)	-1.650 (0.111)	0.146 (0.045)		
M011008	M03_05	1.048 (0.083)	-0.436 (0.077)	0.166 (0.035)		
M011001	M03_06	0.664 (0.055)	-1.361 (0.147)	0.152 (0.048)		
M011002	M03_07	0.801 (0.082)	0.285 (0.089)	0.173 (0.033)		
M012069	M03_08	0.445 (0.071)	0.946 (0.182)	0.166 (0.046)		
M011025	M03_09	0.673 (0.071)	0.266 (0.106)	0.148 (0.036)		
M011026	M03_10	0.642 (0.059)	-0.528 (0.136)	0.161 (0.046)		
M011027	M03_11	0.897 (0.068)	-0.596 (0.086)	0.133 (0.036)		
M011028	M03_12	0.773 (0.063)	-0.535 (0.103)	0.140 (0.039)		
M031305	M04_01	0.665 (0.041)	-1.088 (0.068)			
M031310	M04_02	1.251 (0.093)	-0.720 (0.070)	0.163 (0.035)		
M031065	M04_03	1.003 (0.051)	0.001 (0.034)			
M031051	M04_04	0.939 (0.072)	-0.498 (0.081)	0.146 (0.034)		
M031220	M04_05	0.976 (0.067)	-0.837 (0.076)	0.114 (0.033)		
M031322	M04_06	0.588 (0.039)	-1.064 (0.075)			
M031298	M04_07	0.943 (0.054)	0.665 (0.043)			
M031327	M04_08	0.431 (0.034)	-0.017 (0.069)			
M031269	M04_09	0.403 (0.016)	-0.752 (0.051)		-1.551 (0.125)	1.551 (0.114)
M031264	M04_10	1.300 (0.069)	-0.923 (0.039)			
M031265	M04_11	0.636 (0.042)	0.296 (0.053)			
M031286	M05_01	0.815 (0.026)	0.183 (0.023)			
M031106	M05_02	0.810 (0.026)	0.079 (0.023)			
M031282	M05_03	0.711 (0.016)	0.781 (0.019)		-0.886 (0.042)	0.886 (0.046)
M031227	M05_04	1.103 (0.040)	1.198 (0.031)			
M031335	M05_05	1.178 (0.060)	0.018 (0.039)	0.218 (0.018)		
M031068	M05_06	1.174 (0.034)	0.244 (0.018)			
M031299	M05_07	1.151 (0.033)	-0.009 (0.018)			
M031301	M05_08	1.088 (0.033)	-0.660 (0.023)			
M031271	M05_09	0.737 (0.027)	-1.440 (0.047)			
M031134	M05_10	0.538 (0.024)	1.023 (0.050)			
M031045	M05_11	1.248 (0.052)	-0.463 (0.036)	0.117 (0.019)		
M031235	M06_01	0.752 (0.021)	0.434 (0.021)			
M031285	M06_02	0.712 (0.022)	0.807 (0.026)			
M031050	M06_03	1.197 (0.061)	0.598 (0.030)	0.227 (0.013)		
M031258	M06_04	0.975 (0.026)	0.758 (0.020)			
M031334	M06_05	1.126 (0.060)	0.716 (0.031)	0.220 (0.012)		
M031255	M06_06	0.860 (0.048)	0.144 (0.056)	0.232 (0.022)		
M031041	M06_07	0.803 (0.022)	0.108 (0.019)			
M031350A	M06_08A	1.099 (0.028)	0.470 (0.016)			
M031350B	M06_08B	1.114 (0.027)	0.067 (0.015)			
M031350C	M06_08C	0.797 (0.024)	0.825 (0.024)			
M031274	M06_09	0.837 (0.023)	-0.395 (0.022)			
M031240	M06_10	0.703 (0.020)	-0.361 (0.025)			
M031303	M07_01	1.064 (0.119)	-0.456 (0.109)	0.204 (0.046)		

( ) Standard errors appear in parentheses

Exhibit D.1 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M031309	M07_02	0.976 (0.071)	-0.372 (0.053)			
M031245	M07_03	1.302 (0.178)	1.047 (0.069)	0.105 (0.019)		
M031242A	M07_04A	1.056 (0.075)	-0.174 (0.048)			
M031242B	M07_04B	1.225 (0.086)	0.265 (0.042)			
M031242C	M07_04C	0.999 (0.124)	0.139 (0.097)	0.195 (0.038)		
M031247	M07_05	0.549 (0.041)	1.294 (0.088)		-0.282 (0.102)	0.282 (0.141)
M031219	M07_06	0.440 (0.072)	0.385 (0.218)	0.168 (0.051)		
M031173	M07_07	1.288 (0.124)	-0.273 (0.070)	0.130 (0.032)		
M031085	M07_08	0.566 (0.100)	0.638 (0.177)	0.189 (0.049)		
M031172	M07_09	1.227 (0.120)	-0.221 (0.073)	0.131 (0.033)		
M031029	M08_01	0.652 (0.081)	-0.205 (0.148)	0.159 (0.047)		
M031030	M08_02	0.646 (0.069)	1.468 (0.139)			
M031332	M08_03	0.914 (0.114)	0.065 (0.107)	0.190 (0.041)		
M031098	M08_04	1.394 (0.150)	0.209 (0.061)	0.151 (0.028)		
M031254	M08_05	1.155 (0.135)	0.180 (0.079)	0.176 (0.033)		
M031038	M08_06	0.673 (0.077)	-0.745 (0.162)	0.164 (0.050)		
M031276	M08_07	1.182 (0.139)	0.214 (0.077)	0.184 (0.032)		
M031064	M08_08	0.911 (0.123)	0.580 (0.089)	0.140 (0.032)		
M031006	M08_09	0.688 (0.074)	-0.893 (0.152)	0.150 (0.047)		
M031330	M08_10	0.563 (0.054)	-1.445 (0.136)			
M031351	M08_11	0.687 (0.082)	0.108 (0.117)	0.125 (0.038)		
M031135	M08_12	1.093 (0.102)	-0.612 (0.087)	0.124 (0.036)		
M031162	M09_01	0.641 (0.025)	-0.854 (0.041)			
M031341	M09_02	0.837 (0.042)	-0.815 (0.071)	0.129 (0.030)		
M031216	M09_03	0.913 (0.054)	-0.479 (0.075)	0.210 (0.032)		
M031249	M09_04	0.942 (0.040)	1.337 (0.042)			
M031347A	M09_05A	0.838 (0.030)	0.183 (0.025)			
M031347B	M09_05B	0.771 (0.029)	0.370 (0.028)			
M031347C	M09_05C	1.038 (0.036)	0.597 (0.023)			
M031348A	M09_06A	0.780 (0.031)	0.673 (0.032)			
M031348B	M09_06B	0.792 (0.028)	1.287 (0.030)		0.562 (0.029)	-0.562 (0.051)
M031190	M09_07	1.161 (0.065)	0.290 (0.038)	0.164 (0.017)		
M031306	M10_01	0.803 (0.028)	-0.420 (0.030)			
M031108	M10_02	1.312 (0.069)	0.362 (0.031)	0.149 (0.015)		
M031011	M10_03	0.854 (0.030)	0.053 (0.025)			
M031304	M10_04	0.977 (0.033)	-0.384 (0.025)			
M031023	M10_05	0.663 (0.052)	0.072 (0.102)	0.214 (0.034)		
M031008	M10_06	0.985 (0.092)	1.365 (0.052)	0.186 (0.014)		
M031338	M10_07	0.615 (0.039)	-0.197 (0.093)	0.128 (0.032)		
M031272A	M10_08A	1.032 (0.036)	-0.947 (0.032)			
M031272B	M10_08B	0.940 (0.037)	-1.436 (0.046)			
M031272C	M10_08C	1.092 (0.036)	0.146 (0.020)			
M031267	M10_09	0.628 (0.026)	0.369 (0.032)			
M031315	M10_10	1.048 (0.059)	0.398 (0.040)	0.147 (0.017)		

( ) Standard errors appear in parentheses

Exhibit D.1 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M031128	M11_01	0.372 (0.023)	-1.647 (0.110)			
M031016	M11_02	1.008 (0.041)	0.888 (0.031)			
M031183	M11_03	0.650 (0.023)	0.088 (0.023)		0.454 (0.040)	-0.454 (0.040)
M031187	M11_05	0.716 (0.055)	-0.511 (0.120)	0.216 (0.043)		
M031251	M11_06	1.654 (0.119)	0.689 (0.032)	0.253 (0.014)		
M031294	M11_07	1.162 (0.073)	0.040 (0.049)	0.198 (0.023)		
M031297	M11_08	0.691 (0.030)	0.410 (0.034)			
M031218	M11_09	1.189 (0.077)	0.149 (0.047)	0.212 (0.022)		
M031109	M11_10	0.651 (0.054)	-0.196 (0.123)	0.199 (0.042)		
M031159	M11_11	1.065 (0.065)	-0.237 (0.058)	0.178 (0.027)		
M031133	M11_12	0.843 (0.034)	-0.979 (0.041)			
M031210	M12_01	0.785 (0.076)	0.687 (0.072)	0.217 (0.026)		
M031009	M12_02	0.859 (0.035)	0.534 (0.029)			
M031252	M12_03	0.904 (0.054)	-0.224 (0.063)	0.141 (0.026)		
M031316	M12_04	0.558 (0.030)	-2.245 (0.106)			
M031317	M12_05	1.028 (0.071)	0.582 (0.042)	0.135 (0.018)		
M031079B	M12_06B	1.044 (0.038)	-0.711 (0.030)			
M031079C	M12_06C	0.718 (0.032)	0.559 (0.035)			
M031004	M12_07	1.164 (0.093)	1.068 (0.041)	0.152 (0.013)		
M031043	M12_08	1.256 (0.080)	0.358 (0.039)	0.198 (0.018)		
M031325	M12_09	0.826 (0.035)	0.680 (0.032)			
M031088	M12_10	0.838 (0.054)	-0.562 (0.087)	0.193 (0.035)		
M031093	M12_11	0.435 (0.041)	0.215 (0.149)	0.136 (0.040)		
M031155	M12_12	1.213 (0.075)	0.127 (0.044)	0.211 (0.020)		
M031344A	M13_01A	0.619 (0.055)	0.313 (0.075)			
M031344B	M13_01B	1.065 (0.078)	0.217 (0.047)			
M031344C	M13_01C	0.602 (0.029)	-0.023 (0.044)		-1.558 (0.139)	1.558 (0.138)
M031345A	M13_02A	0.776 (0.061)	-0.284 (0.062)			
M031345B	M13_02B	0.794 (0.063)	-0.169 (0.059)			
M031345C	M13_02C	0.515 (0.066)	1.924 (0.227)			
M031130	M13_03	0.913 (0.069)	-0.602 (0.061)			
M031097	M13_04	1.141 (0.144)	0.314 (0.081)	0.182 (0.033)		
M031178	M13_05	0.909 (0.118)	0.625 (0.087)	0.112 (0.029)		
M031333	M13_06	1.148 (0.151)	0.510 (0.076)	0.158 (0.030)		
M031346A	M14_01A	1.338 (0.092)	-0.423 (0.043)			
M031346B	M14_01B	1.318 (0.097)	0.480 (0.043)			
M031346C	M14_01C	0.976 (0.062)	0.251 (0.035)		0.373 (0.056)	-0.373 (0.060)
M031379	M14_02	0.996 (0.083)	0.846 (0.065)			
M031380	M14_03	0.967 (0.089)	1.179 (0.085)			
M031313	M14_05	0.664 (0.058)	-1.197 (0.104)			
M031083	M14_06	1.076 (0.113)	-0.361 (0.092)	0.148 (0.038)		
M031071	M14_07	1.043 (0.145)	0.708 (0.080)	0.139 (0.028)		
M031185	M14_08	1.500 (0.182)	0.255 (0.065)	0.212 (0.030)		
MF11009	M01F01	1.470 (0.076)	-1.044 (0.048)	0.127 (0.026)		

( ) Standard errors appear in parentheses

Exhibit D.1 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MF11010	M01F02	1.240 (0.062)	-0.236 (0.039)	0.107 (0.019)		
MF12044	M01F03	0.876 (0.043)	-0.110 (0.043)	0.049 (0.016)		
MF11011	M01F04	1.390 (0.065)	-0.785 (0.040)	0.090 (0.020)		
MF11017	M01F05	0.892 (0.040)	-0.469 (0.044)	0.055 (0.017)		
MF11018	M01F06	1.044 (0.047)	-0.861 (0.048)	0.069 (0.020)		
MF11019	M01F07	0.876 (0.043)	-0.403 (0.051)	0.059 (0.020)		
MF11020	M01F08	0.952 (0.070)	0.687 (0.045)	0.126 (0.018)		
MF12065	M01F09	0.729 (0.057)	0.664 (0.059)	0.107 (0.022)		
MF11023	M01F10	0.803 (0.038)	-0.840 (0.060)	0.073 (0.023)		
MF11024	M01F11	1.291 (0.064)	-1.141 (0.052)	0.105 (0.026)		
MF12048	M01F12	1.019 (0.046)	-0.195 (0.037)	0.051 (0.015)		
MF11012	M02F01	1.380 (0.061)	-0.729 (0.035)	0.064 (0.017)		
MF11013	M02F02	0.865 (0.049)	0.083 (0.050)	0.085 (0.020)		
MF11014	M02F03	1.238 (0.055)	-0.824 (0.040)	0.068 (0.019)		
MF11015	M02F04	0.797 (0.038)	-0.018 (0.041)	0.036 (0.014)		
MF11016	M02F05	1.041 (0.054)	0.281 (0.035)	0.063 (0.014)		
MF12078	M02F06	1.320 (0.056)	-0.337 (0.030)	0.053 (0.013)		
MF12119	M02F07	0.897 (0.042)	0.155 (0.036)	0.043 (0.013)		
MF11021	M02F08	1.212 (0.049)	-0.157 (0.028)	0.029 (0.010)		
MF12023	M02F09	1.199 (0.052)	-0.194 (0.031)	0.050 (0.013)		
MF11022	M02F10	1.083 (0.044)	-0.253 (0.030)	0.033 (0.010)		
MF11003	M02F11	1.052 (0.048)	0.120 (0.031)	0.037 (0.012)		
MF11004	M02F12	1.513 (0.066)	-0.144 (0.026)	0.056 (0.012)		
MF11005	M02F13	0.975 (0.044)	-0.419 (0.042)	0.062 (0.017)		
MF12126	M03F01	1.280 (0.119)	-0.085 (0.062)	0.109 (0.028)		
MF11006	M03F02	0.896 (0.096)	0.367 (0.075)	0.093 (0.027)		
MF12117	M03F03	1.192 (0.132)	0.564 (0.060)	0.101 (0.023)		
MF11007	M03F04	2.370 (0.251)	-0.335 (0.050)	0.206 (0.032)		
MF11008	M03F05	1.852 (0.167)	0.104 (0.042)	0.093 (0.021)		
MF11001	M03F06	1.546 (0.153)	-0.231 (0.064)	0.167 (0.032)		
MF11002	M03F07	1.328 (0.144)	0.518 (0.055)	0.110 (0.023)		
MF12069	M03F08	0.902 (0.115)	0.865 (0.084)	0.095 (0.024)		
MF11025	M03F09	1.237 (0.128)	0.518 (0.055)	0.087 (0.021)		
MF11026	M03F10	1.461 (0.144)	0.220 (0.052)	0.115 (0.025)		
MF11027	M03F11	1.815 (0.167)	0.134 (0.043)	0.097 (0.021)		
MF11028	M03F12	1.617 (0.147)	0.141 (0.045)	0.091 (0.021)		
MF31305	M04F01	0.811 (0.063)	-0.691 (0.072)			
MF31310	M04F02	1.440 (0.138)	-0.370 (0.068)	0.135 (0.033)		
MF31065	M04F03	1.179 (0.084)	0.198 (0.042)			
MF31051	M04F04	1.276 (0.121)	-0.109 (0.065)	0.118 (0.029)		
MF31220	M04F05	1.310 (0.116)	-0.323 (0.064)	0.098 (0.028)		
MF31322	M04F06	0.924 (0.070)	-0.332 (0.057)			
MF31298	M04F07	1.240 (0.096)	0.811 (0.050)			
MF31327	M04F08	0.726 (0.061)	0.413 (0.064)			

( ) Standard errors appear in parentheses

Exhibit D.1 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MF31269	M04F09	0.598 (0.030)	0.024 (0.044)		-1.178 (0.120)	1.178 (0.118)
MF31264	M04F10	1.456 (0.010)	-0.209 (0.039)			
MF31265	M04F11	0.920 (0.075)	0.620 (0.059)			
MF31286	M05F01	0.955 (0.071)	0.213 (0.050)			
MF31106	M05F02	0.960 (0.070)	0.054 (0.049)			
MF31282	M05F03	0.805 (0.043)	0.792 (0.042)		-0.779 (0.092)	0.779 (0.103)
MF31227	M05F04	1.120 (0.102)	1.252 (0.078)			
MF31335	M05F05	1.400 (0.142)	0.041 (0.062)	0.146 (0.028)		
MF31068	M05F06	1.538 (0.107)	0.317 (0.036)			
MF31299	M05F07	1.479 (0.100)	0.105 (0.036)			
MF31301	M05F08	1.435 (0.098)	-0.365 (0.041)			
MF31271	M05F09	1.001 (0.073)	-0.756 (0.061)			
MF31134	M05F10	0.686 (0.064)	0.963 (0.092)			
MF31045	M05F11	1.547 (0.136)	-0.134 (0.051)	0.089 (0.023)		
MF31235	M06F01	0.825 (0.065)	0.309 (0.057)			
MF31285	M06F02	0.817 (0.068)	0.723 (0.069)			
MF31050	M06F03	1.122 (0.153)	0.460 (0.084)	0.211 (0.033)		
MF31258	M06F04	0.930 (0.074)	0.683 (0.060)			
MF31334	M06F05	1.180 (0.160)	0.665 (0.075)	0.178 (0.028)		
MF31255	M06F06	0.851 (0.101)	-0.035 (0.111)	0.162 (0.041)		
MF31041	M06F07	0.816 (0.064)	0.165 (0.057)			
MF31350A	M06F08A	1.319 (0.094)	0.463 (0.041)			
MF31350B	M06F08B	1.426 (0.097)	0.098 (0.037)			
MF31350C	M06F08C	1.083 (0.084)	0.735 (0.055)			
MF31274	M06F09	1.084 (0.076)	-0.334 (0.049)			
MF31240	M06F10	0.831 (0.063)	-0.129 (0.056)			
MF31303	M07F01	1.500 (0.158)	-0.322 (0.073)	0.199 (0.036)		
MF31309	M07F02	1.427 (0.096)	-0.183 (0.039)			
MF31245	M07F03	1.566 (0.189)	0.878 (0.053)	0.081 (0.016)		
MF31242A	M07F04A	1.269 (0.088)	0.110 (0.040)			
MF31242B	M07F04B	1.291 (0.094)	0.467 (0.042)			
MF31242C	M07F04C	1.518 (0.187)	0.427 (0.060)	0.201 (0.026)		
MF31247	M07F05	0.732 (0.055)	1.341 (0.077)		-0.152 (0.083)	0.152 (0.123)
MF31219	M07F06	0.898 (0.129)	0.637 (0.096)	0.167 (0.033)		
MF31173	M07F07	1.767 (0.169)	0.130 (0.046)	0.109 (0.022)		
MF31085	M07F08	0.965 (0.151)	0.831 (0.093)	0.157 (0.030)		
MF31172	M07F09	1.515 (0.152)	0.239 (0.052)	0.107 (0.023)		
MF31029	M08F01	0.784 (0.096)	0.120 (0.111)	0.144 (0.039)		
MF31030	M08F02	0.714 (0.074)	1.445 (0.128)			
MF31332	M08F03	1.218 (0.136)	0.241 (0.070)	0.154 (0.030)		
MF31098	M08F04	1.722 (0.172)	0.228 (0.047)	0.117 (0.022)		
MF31254	M08F05	1.469 (0.161)	0.354 (0.056)	0.140 (0.025)		
MF31038	M08F06	1.311 (0.127)	-0.093 (0.064)	0.120 (0.029)		
MF31276	M08F07	1.376 (0.150)	0.341 (0.059)	0.143 (0.025)		

( ) Standard errors appear in parentheses

Exhibit D.1 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MF31064	M08F08	1.022 (0.126)	0.802 (0.075)	0.096 (0.022)		
MF31006	M08F09	1.147 (0.128)	-0.027 (0.082)	0.180 (0.035)		
MF31330	M08F10	0.859 (0.065)	-0.375 (0.058)			
MF31351	M08F11	1.073 (0.124)	0.556 (0.070)	0.106 (0.025)		
MF31135	M08F12	1.305 (0.125)	0.092 (0.058)	0.103 (0.025)		
MF31128	M11F01	0.633 (0.054)	-0.719 (0.083)			
MF31016	M11F02	1.267 (0.101)	0.870 (0.055)			
MF31183	M11F03	1.060 (0.066)	0.336 (0.034)		0.349 (0.051)	-0.349 (0.058)
MF31187	M11F05	1.113 (0.122)	-0.063 (0.083)	0.166 (0.035)		
MF31251	M11F06	1.483 (0.192)	0.672 (0.060)	0.159 (0.023)		
MF31294	M11F07	1.417 (0.138)	0.109 (0.055)	0.104 (0.024)		
MF31297	M11F08	1.186 (0.086)	0.412 (0.046)			
MF31218	M11F09	1.789 (0.165)	0.207 (0.041)	0.076 (0.018)		
MF31109	M11F10	1.087 (0.122)	0.253 (0.075)	0.130 (0.030)		
MF31159	M11F11	1.557 (0.151)	0.104 (0.051)	0.105 (0.023)		
MF31133	M11F12	1.082 (0.076)	-0.301 (0.047)			
MF31210	M12F01	1.214 (0.168)	0.611 (0.073)	0.173 (0.028)		
MF31009	M12F02	1.088 (0.082)	0.502 (0.050)			
MF31252	M12F03	0.982 (0.109)	-0.092 (0.094)	0.157 (0.037)		
MF31316	M12F04	0.843 (0.068)	-1.298 (0.089)			
MF31317	M12F05	1.074 (0.122)	0.429 (0.070)	0.105 (0.026)		
MF31079B	M12F06B	1.252 (0.086)	-0.470 (0.046)			
MF31079C	M12F06C	0.799 (0.069)	0.698 (0.073)			
MF31004	M12F07	1.036 (0.142)	0.910 (0.080)	0.108 (0.023)		
MF31043	M12F08	1.444 (0.163)	0.399 (0.057)	0.147 (0.024)		
MF31325	M12F09	1.047 (0.086)	0.855 (0.063)			
MF31088	M12F10	0.970 (0.109)	-0.070 (0.094)	0.162 (0.037)		
MF31093	M12F11	0.621 (0.097)	0.655 (0.140)	0.148 (0.040)		
MF31155	M12F12	1.310 (0.144)	0.215 (0.064)	0.151 (0.028)		
MF31344A	M13F01A	0.805 (0.034)	0.770 (0.035)			
MF31344B	M13F01B	1.415 (0.051)	0.566 (0.020)			
MF31344C	M13F01C	0.760 (0.018)	0.435 (0.019)		-1.313 (0.059)	1.313 (0.061)
MF31345A	M13F02A	0.995 (0.037)	0.230 (0.024)			
MF31345B	M13F02B	0.964 (0.036)	0.321 (0.025)			
MF31345C	M13F02C	0.663 (0.042)	2.083 (0.107)			
MF31130	M13F03	0.986 (0.036)	0.048 (0.024)			
MF31097	M13F04	1.685 (0.113)	0.823 (0.027)	0.151 (0.011)		
MF31178	M13F05	1.355 (0.093)	1.003 (0.032)	0.093 (0.010)		
MF31333	M13F06	1.542 (0.104)	0.913 (0.029)	0.119 (0.010)		
MF31346A	M14F01A	1.407 (0.049)	-0.196 (0.021)			
MF31346B	M14F01B	1.390 (0.051)	0.725 (0.021)			
MF31346C	M14F01C	1.088 (0.034)	0.495 (0.017)		0.407 (0.025)	-0.407 (0.028)
MF31379	M14F02	1.105 (0.046)	1.107 (0.032)			
MF31380	M14F03	1.090 (0.050)	1.375 (0.040)			

( ) Standard errors appear in parentheses

Exhibit D.1 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MF31313	M14F05	0.752 (0.030)	-0.554 (0.037)			
MF31083	M14F06	1.155 (0.064)	0.030 (0.042)	0.122 (0.020)		
MF31071	M14F07	0.985 (0.067)	0.845 (0.040)	0.099 (0.015)		
MF31185	M14F08	1.593 (0.087)	0.303 (0.029)	0.148 (0.015)		
MZ31286	M01_01	1.019 (0.057)	0.231 (0.037)			
MZ31106	M01_02	0.992 (0.056)	0.109 (0.037)			
MZ31282	M01_03	0.884 (0.036)	0.729 (0.029)		-0.629 (0.063)	0.629 (0.070)
MZ31227	M01_04	1.170 (0.075)	1.061 (0.048)			
MZ31335	M01_05	1.210 (0.107)	-0.035 (0.067)	0.208 (0.030)		
MZ31068	M01_06	1.325 (0.071)	0.239 (0.030)			
MZ31299	M01_07	1.282 (0.069)	-0.021 (0.032)			
MZ31301	M01_08	1.092 (0.061)	-0.626 (0.043)			
MZ31271	M01_09	0.660 (0.047)	-1.582 (0.102)			
MZ31134	M01_10	0.570 (0.044)	0.851 (0.076)			
MZ31045	M01_11	1.131 (0.084)	-0.622 (0.070)	0.115 (0.031)		
MZ41014	M02_01	0.795 (0.062)	-0.936 (0.103)	0.120 (0.037)		
MZ41039	M02_02	0.808 (0.077)	-0.225 (0.106)	0.177 (0.040)		
MZ41278	M02_03	0.497 (0.039)	0.239 (0.067)			
MZ41006	M02_04	1.107 (0.098)	0.351 (0.057)	0.127 (0.024)		
MZ41250	M02_05	0.920 (0.053)	-0.065 (0.040)			
MZ41094	M02_06	1.033 (0.127)	0.850 (0.072)	0.206 (0.025)		
MZ41330	M02_07	0.792 (0.075)	0.069 (0.091)	0.139 (0.034)		
MZ41300A	M02_08A	1.055 (0.060)	0.123 (0.036)			
MZ41300B	M02_08B	1.154 (0.065)	0.065 (0.034)			
MZ41300C	M02_08C	1.023 (0.061)	0.447 (0.039)			
MZ41300D	M02_08D	1.127 (0.068)	0.632 (0.039)			
MZ41173	M02_09	0.836 (0.107)	1.036 (0.085)	0.140 (0.026)		
MZ41274	M02_10	0.823 (0.052)	-0.441 (0.052)			
MZ41203	M02_11	0.865 (0.055)	0.041 (0.044)			
MZ31235	M03_01	0.786 (0.049)	0.496 (0.048)			
MZ31285	M03_02	0.791 (0.051)	0.752 (0.053)			
MZ31050	M03_03	1.171 (0.130)	0.606 (0.066)	0.235 (0.026)		
MZ31258	M03_04	1.060 (0.063)	0.656 (0.040)			
MZ31334	M03_05	1.261 (0.146)	0.782 (0.062)	0.237 (0.024)		
MZ31255	M03_06	0.945 (0.100)	0.113 (0.095)	0.225 (0.037)		
MZ31041	M03_07	0.950 (0.055)	0.099 (0.039)			
MZ31350A	M03_08A	1.061 (0.062)	0.486 (0.038)			
MZ31350B	M03_08B	1.056 (0.060)	0.051 (0.037)			
MZ31350C	M03_08C	0.811 (0.055)	0.857 (0.056)			
MZ31274	M03_09	0.798 (0.050)	-0.495 (0.054)			
MZ31240	M03_10	0.625 (0.043)	-0.489 (0.066)			
MZ41052	M04_01	0.780 (0.072)	-0.761 (0.136)	0.191 (0.049)		
MZ41056	M04_02	0.936 (0.054)	0.146 (0.039)			
MZ41069	M04_03	1.233 (0.121)	0.967 (0.051)	0.100 (0.017)		

( ) Standard errors appear in parentheses

Exhibit D.1 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MZ41076	M04_04	0.955 (0.054)	0.169 (0.039)			
MZ41281	M04_05	1.241 (0.095)	-0.249 (0.061)	0.133 (0.029)		
MZ41164	M04_06	0.684 (0.059)	-0.947 (0.136)	0.144 (0.045)		
MZ41146	M04_07	0.732 (0.047)	-0.443 (0.056)			
MZ41152	M04_08	1.090 (0.109)	0.511 (0.065)	0.177 (0.027)		
MZ41258A	M04_09A	0.932 (0.054)	-0.227 (0.043)			
MZ41258B	M04_09B	0.742 (0.047)	0.282 (0.048)			
MZ41131	M04_10	0.631 (0.109)	1.553 (0.142)	0.167 (0.031)		
MZ41275	M04_11	0.529 (0.021)	-0.634 (0.046)		-1.378 (0.118)	1.378 (0.107)
MZ41186	M04_12	0.973 (0.073)	-0.034 (0.062)	0.092 (0.025)		
MZ41336	M04_13	1.228 (0.143)	0.942 (0.060)	0.181 (0.022)		
MZ31303	M05_01	1.278 (0.111)	-0.312 (0.074)	0.221 (0.036)		
MZ31309	M05_02	1.175 (0.063)	-0.261 (0.036)			
MZ31245	M05_03	1.763 (0.163)	0.949 (0.037)	0.105 (0.013)		
MZ31242A	M05_04A	1.089 (0.060)	-0.226 (0.037)			
MZ31242B	M05_04B	1.212 (0.066)	0.256 (0.032)			
MZ31242C	M05_04C	1.028 (0.105)	0.156 (0.085)	0.239 (0.035)		
MZ31247	M05_05	0.552 (0.030)	1.228 (0.059)		-0.328 (0.077)	0.328 (0.100)
MZ31219	M05_06	0.507 (0.070)	0.597 (0.163)	0.167 (0.045)		
MZ31173	M05_07	1.559 (0.115)	-0.146 (0.047)	0.125 (0.024)		
MZ31085	M05_08	0.928 (0.116)	0.685 (0.087)	0.234 (0.032)		
MZ31172	M05_09	1.397 (0.103)	-0.243 (0.052)	0.113 (0.026)		
MZ41010	M06_01	1.027 (0.099)	-0.184 (0.094)	0.245 (0.039)		
MZ41098	M06_02	1.496 (0.144)	0.519 (0.050)	0.216 (0.023)		
MZ41064	M06_03	0.908 (0.053)	-0.524 (0.048)			
MZ41003	M06_04	0.783 (0.048)	-0.057 (0.046)			
MZ41104	M06_05	0.987 (0.056)	-0.121 (0.039)			
MZ41299	M06_06	1.193 (0.071)	0.819 (0.038)			
MZ41329	M06_07	0.711 (0.078)	-0.252 (0.148)	0.213 (0.050)		
MZ41143	M06_08	0.390 (0.016)	-0.352 (0.051)		-1.892 (0.141)	1.892 (0.134)
MZ41158	M06_09	0.799 (0.075)	-0.369 (0.115)	0.179 (0.043)		
MZ41328	M06_10	0.821 (0.049)	-0.339 (0.049)			
MZ41155	M06_11	0.890 (0.079)	0.210 (0.074)	0.129 (0.029)		
MZ41284	M06_12	0.749 (0.041)	0.700 (0.036)		0.499 (0.051)	-0.499 (0.063)
MZ41335	M06_13	0.829 (0.067)	-0.935 (0.111)	0.135 (0.041)		
MZ41184	M06_14	1.001 (0.082)	-0.530 (0.085)	0.144 (0.037)		
MZ31029	M07_01	0.847 (0.081)	-0.015 (0.094)	0.155 (0.037)		
MZ31030	M07_02	0.725 (0.057)	1.530 (0.096)			
MZ31332	M07_03	0.979 (0.097)	-0.017 (0.091)	0.229 (0.037)		
MZ31098	M07_04	1.335 (0.117)	0.302 (0.052)	0.174 (0.024)		
MZ31254	M07_05	1.213 (0.108)	0.186 (0.061)	0.181 (0.028)		
MZ31038	M07_06	0.844 (0.070)	-0.626 (0.103)	0.144 (0.039)		
MZ31276	M07_07	1.358 (0.125)	0.194 (0.059)	0.227 (0.028)		
MZ31064	M07_08	0.999 (0.104)	0.595 (0.068)	0.160 (0.027)		

( ) Standard errors appear in parentheses

Exhibit D.1 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MZ31006	M07_09	0.782 (0.069)	-0.767 (0.124)	0.164 (0.045)		
MZ31330	M07_10	0.695 (0.047)	-1.214 (0.084)			
MZ31351	M07_11	0.815 (0.076)	0.256 (0.081)	0.123 (0.031)		
MZ31135	M07_12	1.127 (0.086)	-0.587 (0.072)	0.128 (0.033)		
MZ41291	M08_01	0.557 (0.040)	-0.935 (0.084)			
MZ41289	M08_02	0.950 (0.091)	0.208 (0.079)	0.185 (0.032)		
MZ41068	M08_03	1.250 (0.106)	0.615 (0.046)	0.087 (0.019)		
MZ41065A	M08_04A	1.632 (0.142)	0.598 (0.039)	0.129 (0.018)		
MZ41065B	M08_04B	1.266 (0.076)	0.932 (0.040)			
MZ41096	M08_05	1.068 (0.097)	0.362 (0.062)	0.146 (0.026)		
MZ41125	M08_06	1.061 (0.103)	0.658 (0.058)	0.124 (0.023)		
MZ41135	M08_07	0.464 (0.048)	-1.103 (0.206)	0.155 (0.051)		
MZ41257	M08_08	0.729 (0.047)	0.500 (0.050)			
MZ41268	M08_09	1.726 (0.181)	0.922 (0.043)	0.180 (0.017)		
MZ41151	M08_10	0.608 (0.065)	-0.549 (0.172)	0.194 (0.053)		
MZ41264	M08_11	0.713 (0.078)	0.182 (0.114)	0.170 (0.040)		
MZ41182	M08_12	0.935 (0.060)	-1.369 (0.073)			
MZ41200	M08_13	0.584 (0.028)	-0.340 (0.041)		-0.280 (0.081)	0.280 (0.072)
MZ31128	M09_01	0.446 (0.037)	-1.531 (0.135)			
MZ31016	M09_02	1.182 (0.069)	0.734 (0.037)			
MZ31183	M09_03	0.795 (0.040)	0.170 (0.032)		0.632 (0.052)	-0.632 (0.052)
MZ31187	M09_05	0.731 (0.063)	-0.695 (0.121)	0.147 (0.042)		
MZ31251	M09_06	1.449 (0.136)	0.442 (0.050)	0.193 (0.023)		
MZ31294	M09_07	1.235 (0.091)	-0.148 (0.054)	0.112 (0.025)		
MZ31297	M09_08	0.875 (0.053)	0.379 (0.043)			
MZ31218	M09_09	1.387 (0.117)	-0.019 (0.056)	0.179 (0.027)		
MZ31109	M09_10	0.779 (0.076)	-0.229 (0.115)	0.188 (0.041)		
MZ31159	M09_11	1.070 (0.088)	-0.335 (0.076)	0.150 (0.033)		
MZ31133	M09_12	0.918 (0.057)	-1.096 (0.065)			
MZ41107	M10_01	0.947 (0.075)	-1.173 (0.107)	0.140 (0.041)		
MZ41011	M10_02	1.200 (0.099)	-0.243 (0.069)	0.181 (0.031)		
MZ41122	M10_03	0.473 (0.022)	0.598 (0.044)		-0.903 (0.095)	0.903 (0.103)
MZ41041	M10_04	0.956 (0.103)	0.126 (0.096)	0.260 (0.037)		
MZ41320	M10_05	1.605 (0.138)	0.336 (0.044)	0.166 (0.022)		
MZ41115A	M10_06A	0.861 (0.050)	-0.345 (0.047)			
MZ41115B	M10_06B	1.075 (0.060)	0.106 (0.035)			
MZ41160A	M10_07A	0.976 (0.059)	-1.231 (0.066)			
MZ41160B	M10_07B	1.167 (0.073)	-1.325 (0.062)			
MZ41327	M10_08	0.546 (0.039)	-0.261 (0.066)			
MZ41148	M10_09	0.401 (0.028)	-0.025 (0.055)		0.358 (0.103)	-0.358 (0.096)
MZ41265	M10_10	0.870 (0.088)	0.667 (0.070)	0.118 (0.025)		
MZ41175	M10_11	0.905 (0.073)	-1.103 (0.111)	0.139 (0.041)		
MZ41199	M10_12	1.343 (0.103)	-0.577 (0.065)	0.132 (0.031)		
MZ31210	M11_01	1.067 (0.107)	0.459 (0.066)	0.170 (0.027)		

( ) Standard errors appear in parentheses

Exhibit D.1 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MZ31009	M11_02	0.907 (0.054)	0.463 (0.041)			
MZ31252	M11_03	0.935 (0.077)	-0.329 (0.083)	0.138 (0.033)		
MZ31316	M11_04	0.646 (0.048)	-2.001 (0.128)			
MZ31317	M11_05	1.184 (0.110)	0.553 (0.053)	0.137 (0.023)		
MZ31079B	M11_06B	1.132 (0.063)	-0.726 (0.046)			
MZ31079C	M11_06C	0.743 (0.048)	0.465 (0.049)			
MZ31004	M11_07	0.988 (0.113)	1.085 (0.067)	0.120 (0.020)		
MZ31043	M11_08	1.242 (0.109)	0.232 (0.057)	0.176 (0.026)		
MZ31325	M11_09	0.940 (0.057)	0.595 (0.042)			
MZ31088	M11_10	0.732 (0.064)	-0.742 (0.126)	0.154 (0.043)		
MZ31093	M11_11	0.456 (0.060)	0.256 (0.189)	0.163 (0.049)		
MZ31155	M11_12	1.233 (0.118)	0.207 (0.066)	0.235 (0.029)		
MZ41298	M12_01	1.194 (0.092)	-0.634 (0.073)	0.141 (0.034)		
MZ41007	M12_02	0.773 (0.084)	0.459 (0.091)	0.151 (0.033)		
MZ41280	M12_03	0.765 (0.087)	0.518 (0.094)	0.164 (0.033)		
MZ41059	M12_04	1.051 (0.058)	-0.057 (0.036)			
MZ41046	M12_05	1.269 (0.097)	0.149 (0.048)	0.098 (0.022)		
MZ41048	M12_06	1.227 (0.125)	0.468 (0.062)	0.220 (0.026)		
MZ41169	M12_07	0.838 (0.075)	0.005 (0.085)	0.135 (0.033)		
MZ41333	M12_08	1.170 (0.115)	0.693 (0.055)	0.150 (0.022)		
MZ41262	M12_09	1.027 (0.114)	0.675 (0.070)	0.198 (0.027)		
MZ41267	M12_10	0.460 (0.039)	0.791 (0.086)			
MZ41177	M12_11	1.018 (0.080)	-0.617 (0.083)	0.136 (0.035)		
MZ41271	M12_12	1.000 (0.078)	-0.386 (0.075)	0.121 (0.031)		
MZ41276A	M12_13A	0.988 (0.057)	0.080 (0.038)			
MZ41276B	M12_13B	0.796 (0.054)	0.688 (0.051)			
MZ31346A	M13_01A	1.552 (0.083)	-0.356 (0.030)			
MZ31346B	M13_01B	1.759 (0.094)	0.474 (0.025)			
MZ31346C	M13_01C	1.345 (0.063)	0.280 (0.022)		0.438 (0.033)	-0.438 (0.034)
MZ31379	M13_02	1.189 (0.070)	0.834 (0.039)			
MZ31380	M13_03	1.048 (0.068)	1.134 (0.053)			
MZ31313	M13_05	0.617 (0.043)	-1.207 (0.090)			
MZ31083	M13_06	0.897 (0.083)	-0.424 (0.106)	0.195 (0.042)		
MZ31071	M13_07	0.905 (0.093)	0.641 (0.071)	0.136 (0.027)		
MZ31185	M13_08	1.314 (0.118)	0.170 (0.059)	0.192 (0.028)		
MZ41004	M14_01	0.963 (0.077)	-1.300 (0.110)	0.150 (0.043)		
MZ41023	M14_02	1.448 (0.118)	-0.886 (0.071)	0.178 (0.036)		
MZ41034	M14_03	0.787 (0.062)	-0.367 (0.087)	0.108 (0.032)		
MZ41087	M14_04	0.784 (0.047)	-0.007 (0.045)			
MZ41124	M14_05	0.907 (0.052)	-0.218 (0.042)			
MZ41302A	M14_06A	0.693 (0.059)	-1.169 (0.140)	0.148 (0.046)		
MZ41302B	M14_06B	0.568 (0.039)	-0.457 (0.066)			
MZ41302C	M14_06C	0.949 (0.054)	-0.466 (0.045)			
MZ41254	M14_07	0.764 (0.079)	0.215 (0.098)	0.161 (0.035)		

( ) Standard errors appear in parentheses

Exhibit D.1 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MZ41153	M14_08	0.928 (0.086)	0.214 (0.075)	0.152 (0.030)		
MZ41132	M14_09	0.506 (0.069)	0.763 (0.147)	0.149 (0.040)		
MZ41165	M14_10	0.295 (0.015)	0.365 (0.064)		-1.536 (0.150)	1.536 (0.155)
MZ41174	M14_11	1.022 (0.059)	-0.789 (0.050)			
MZ41191	M14_12	0.881 (0.073)	-1.384 (0.124)	0.149 (0.045)		

( ) Standard errors appear in parentheses

Exhibit D.2 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Science

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S011006	S01_01	0.571 (0.041)	-0.891 (0.146)	0.201 (0.044)		
S011007	S01_02	0.702 (0.041)	-0.817 (0.094)	0.162 (0.033)		
S011008	S01_03	0.806 (0.049)	-0.472 (0.080)	0.181 (0.031)		
S012033	S01_04	0.399 (0.038)	0.038 (0.167)	0.153 (0.041)		
S011001	S01_05	0.674 (0.039)	-1.602 (0.121)	0.155 (0.040)		
S011003	S01_06	0.783 (0.052)	-0.460 (0.092)	0.232 (0.034)		
S011004	S01_07	0.668 (0.040)	-0.656 (0.093)	0.149 (0.033)		
S011005	S01_08	0.858 (0.042)	-1.018 (0.070)	0.120 (0.028)		
S011021	S01_09	0.563 (0.034)	-1.556 (0.137)	0.143 (0.041)		
S011022	S01_10	0.526 (0.041)	-0.576 (0.149)	0.192 (0.043)		
S011023	S01_11	0.896 (0.047)	-0.454 (0.061)	0.134 (0.026)		
S011030	S02_01	0.520 (0.049)	-1.766 (0.214)	0.181 (0.053)		
S011031	S02_02	0.957 (0.074)	-1.649 (0.113)	0.148 (0.040)		
S011032	S02_03	0.656 (0.042)	0.209 (0.048)			
S011033	S02_04	0.300 (0.058)	1.547 (0.308)	0.179 (0.047)		
S011025	S02_05	0.578 (0.052)	-1.842 (0.198)	0.179 (0.052)		
S011026	S02_06	0.368 (0.040)	-2.873 (0.329)	0.163 (0.050)		
S011027	S02_07	0.781 (0.074)	-0.558 (0.125)	0.206 (0.045)		
S011029	S02_09	0.974 (0.079)	-1.762 (0.123)	0.165 (0.044)		
S011016	S02_10	0.716 (0.061)	-2.118 (0.173)	0.166 (0.049)		
S012007	S02_11	0.466 (0.046)	-1.436 (0.214)	0.171 (0.051)		
S011017	S03_01	0.586 (0.056)	-0.513 (0.140)	0.161 (0.042)		
S011018	S03_02	0.575 (0.053)	-2.814 (0.241)	0.167 (0.051)		
S012010	S03_03	0.733 (0.067)	-0.578 (0.124)	0.175 (0.043)		
S011019	S03_04	0.632 (0.040)	-1.194 (0.079)			
S011009	S03_05	0.477 (0.048)	-1.051 (0.194)	0.169 (0.049)		
S011010	S03_06	0.574 (0.054)	-2.913 (0.258)	0.176 (0.054)		
S011011	S03_07	0.878 (0.113)	0.733 (0.080)	0.222 (0.029)		
S011012	S03_08	0.649 (0.051)	-1.851 (0.149)	0.100 (0.039)		
S011013	S03_09	0.490 (0.066)	0.622 (0.138)	0.121 (0.039)		
S011014	S03_10	0.766 (0.069)	-0.253 (0.101)	0.168 (0.037)		
S011015	S03_11	0.506 (0.048)	-1.156 (0.185)	0.165 (0.048)		
S031017	S04_01	0.613 (0.056)	-0.871 (0.146)	0.163 (0.045)		
S031246	S04_02	0.723 (0.052)	1.044 (0.069)			
S031287	S04_03	0.662 (0.060)	-0.376 (0.116)	0.153 (0.039)		
S031251	S04_04	0.552 (0.044)	1.077 (0.089)			
S031053	S04_05	0.499 (0.024)	-0.089 (0.038)		-0.229 (0.078)	0.229 (0.075)
S031005	S04_06	0.642 (0.053)	1.447 (0.105)			
S031306	S04_07	0.920 (0.117)	0.863 (0.072)	0.181 (0.025)		
S031372A	S04_08A	0.887 (0.049)	-0.346 (0.041)			
S031372B	S04_08B	0.708 (0.033)	0.907 (0.039)		-0.341 (0.058)	0.341 (0.073)
S031082	S04_09	0.456 (0.048)	-0.896 (0.198)	0.167 (0.048)		
S031229	S05_01	1.266 (0.086)	0.654 (0.034)	0.259 (0.015)		
S031270	S05_02	0.501 (0.029)	1.827 (0.094)			

( ) Standard errors appear in parentheses

Exhibit D.2 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S031026	S05_03	0.510 (0.012)	-0.176 (0.022)		-0.793 (0.050)	0.793 (0.047)
S031319	S05_04	1.133 (0.085)	1.047 (0.037)	0.186 (0.013)		
S031414A	S05_05A	1.234 (0.036)	-0.315 (0.019)			
S031414B	S05_05B	1.036 (0.032)	-0.395 (0.022)			
S031078	S05_06	0.639 (0.052)	-0.065 (0.115)	0.247 (0.038)		
S031009	S05_07	0.694 (0.025)	-0.230 (0.028)			
S031401	S05_08	1.032 (0.072)	0.548 (0.043)	0.240 (0.018)		
S031384A	S05_09A	0.806 (0.029)	-1.336 (0.045)			
S031384B	S05_09B	0.928 (0.030)	-0.482 (0.026)			
S031255	S06_01	0.847 (0.045)	-0.191 (0.062)	0.254 (0.024)		
S031240D	S06_02D	0.640 (0.015)	-0.157 (0.017)		0.818 (0.030)	-0.818 (0.025)
S031239	S06_03	0.596 (0.050)	-0.083 (0.130)	0.372 (0.034)		
S031235A	S06_04A	0.943 (0.026)	0.388 (0.017)			
S031235B	S06_04B	1.000 (0.028)	0.641 (0.018)			
S031205	S06_05	0.710 (0.041)	0.144 (0.062)	0.218 (0.022)		
S031399A	S06_06A	0.924 (0.025)	0.233 (0.017)			
S031399B	S06_06B	0.936 (0.025)	-0.067 (0.017)			
S031393	S06_07	0.783 (0.023)	-1.347 (0.039)			
S031278	S06_08	0.615 (0.020)	-0.568 (0.030)			
S031317	S07_01	0.603 (0.083)	-0.530 (0.193)	0.204 (0.054)		
S031190	S07_02	0.817 (0.075)	0.752 (0.071)			
S031431	S07_03	0.691 (0.175)	1.789 (0.253)	0.161 (0.031)		
S031283	S07_04	0.510 (0.069)	-0.969 (0.231)	0.188 (0.055)		
S031426	S07_05	0.673 (0.090)	-0.389 (0.165)	0.201 (0.051)		
S031422	S07_06	0.787 (0.090)	-1.441 (0.179)	0.186 (0.052)		
S031427	S07_07	0.654 (0.084)	-0.265 (0.148)	0.170 (0.046)		
S031075	S07_08	0.327 (0.062)	0.139 (0.309)	0.194 (0.056)		
S031047	S07_09	0.614 (0.057)	0.025 (0.071)			
S031387	S07_10	0.948 (0.185)	1.208 (0.120)	0.170 (0.030)		
S031396D	S07_11D	0.490 (0.033)	-1.081 (0.084)		-0.401 (0.146)	0.401 (0.112)
S031340	S08_01	1.079 (0.208)	1.002 (0.102)	0.254 (0.031)		
S031236	S08_02	0.593 (0.070)	-1.390 (0.204)	0.163 (0.049)		
S031391D	S08_03D	0.485 (0.036)	0.390 (0.057)		-0.167 (0.106)	0.167 (0.114)
S031361	S08_04	0.564 (0.102)	0.560 (0.171)	0.202 (0.048)		
S031001	S08_05	0.780 (0.085)	-1.011 (0.148)	0.164 (0.046)		
S031410	S08_07	0.439 (0.065)	-0.502 (0.228)	0.172 (0.051)		
S031421	S08_08	0.386 (0.045)	-0.638 (0.128)			
S031298	S08_09	0.855 (0.204)	1.448 (0.173)	0.225 (0.031)		
S031076	S08_10	0.707 (0.068)	0.715 (0.078)			
S031275	S08_11	0.663 (0.159)	1.641 (0.225)	0.167 (0.033)		
S031349	S09_01	0.659 (0.036)	-1.419 (0.105)	0.127 (0.034)		
S031330	S09_02	0.888 (0.032)	-0.617 (0.031)			
S031212	S09_03	0.617 (0.041)	-0.688 (0.114)	0.167 (0.037)		
S031241D	S09_04D	0.658 (0.022)	0.611 (0.023)		0.682 (0.033)	-0.682 (0.041)

( ) Standard errors appear in parentheses

Exhibit D.2 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S031038	S09_05	0.581 (0.043)	-0.894 (0.147)	0.219 (0.044)		
S031252	S09_06	0.523 (0.019)	-1.053 (0.039)		0.366 (0.062)	-0.366 (0.041)
S031406A	S09_07A	1.015 (0.036)	-0.504 (0.027)			
S031406B	S09_07B	1.019 (0.043)	1.067 (0.033)			
S031383	S09_08	0.688 (0.047)	0.775 (0.050)	0.078 (0.017)		
S031379	S09_09	0.699 (0.051)	0.018 (0.081)	0.197 (0.029)		
S031060	S09_10	0.799 (0.093)	1.297 (0.072)	0.234 (0.020)		
S031269	S10_01	0.568 (0.054)	0.386 (0.105)	0.184 (0.034)		
S031284	S10_02	0.544 (0.079)	1.627 (0.119)	0.171 (0.026)		
S031338	S10_03	0.823 (0.045)	-0.618 (0.073)	0.157 (0.030)		
S031382	S10_04	0.750 (0.030)	0.147 (0.027)			
S031218	S10_05	0.748 (0.030)	-0.160 (0.029)			
S031326D	S10_06D	0.458 (0.017)	0.267 (0.026)		0.036 (0.050)	-0.036 (0.051)
S031003	S10_07	0.677 (0.040)	-0.526 (0.085)	0.128 (0.031)		
S031035	S10_08	0.750 (0.046)	-0.991 (0.106)	0.199 (0.039)		
S031420	S10_09	0.652 (0.061)	0.945 (0.067)	0.130 (0.024)		
S031370	S10_10	0.938 (0.035)	0.156 (0.022)			
S031313	S10_11	0.858 (0.084)	0.969 (0.057)	0.225 (0.021)		
S031254	S11_01	0.249 (0.030)	0.053 (0.257)	0.214 (0.040)		
S031266	S11_02	2.388 (0.162)	0.728 (0.026)	0.369 (0.013)		
S031233	S11_03	0.363 (0.023)	-0.519 (0.066)			
S031204	S11_04	0.310 (0.024)	1.035 (0.098)			
S031273	S11_05	2.536 (0.177)	0.698 (0.026)	0.433 (0.013)		
S031299	S11_06	0.425 (0.027)	0.898 (0.067)			
S031281	S11_07	0.480 (0.034)	-1.812 (0.188)	0.160 (0.048)		
S031077	S11_08	0.342 (0.032)	-1.391 (0.275)	0.181 (0.054)		
S031311	S11_09	2.837 (0.190)	0.654 (0.024)	0.457 (0.012)		
S031088D	S11_10D	0.294 (0.009)	0.491 (0.056)		2.280 (0.083)	-2.280 (0.096)
S031389	S11_11	1.521 (0.145)	1.124 (0.044)	0.282 (0.013)		
S031356	S12_01	0.457 (0.035)	-2.615 (0.254)	0.194 (0.059)		
S031291	S12_02	0.802 (0.049)	-1.174 (0.103)	0.161 (0.038)		
S031230	S12_03	0.598 (0.040)	-1.535 (0.149)	0.159 (0.044)		
S031325	S12_04	0.521 (0.029)	0.473 (0.044)			
S031068	S12_05	1.127 (0.096)	0.652 (0.044)	0.239 (0.019)		
S031418	S12_06	0.871 (0.089)	0.974 (0.057)	0.191 (0.020)		
S031197D	S12_07D	0.400 (0.013)	-0.838 (0.043)		-0.989 (0.085)	0.989 (0.072)
S031371	S12_08	1.106 (0.114)	0.959 (0.051)	0.278 (0.018)		
S031376	S12_09	1.076 (0.121)	1.170 (0.058)	0.253 (0.017)		
S031044	S12_10	0.618 (0.030)	0.120 (0.035)			
S031390D	S12_11D	0.516 (0.021)	0.353 (0.028)		0.293 (0.048)	-0.293 (0.052)
S031409	S13_01	0.817 (0.100)	-0.242 (0.124)	0.181 (0.044)		
S031398	S13_02	0.675 (0.092)	-0.140 (0.150)	0.190 (0.047)		
S031072	S13_03	0.690 (0.047)	-0.070 (0.049)		0.813 (0.082)	-0.813 (0.076)
S031061	S13_04	0.579 (0.078)	-0.491 (0.182)	0.186 (0.050)		

( ) Standard errors appear in parentheses

Exhibit D.2 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S031439A	S13_05A	0.828 (0.089)	1.349 (0.115)			
S031439B	S13_05B	0.668 (0.061)	0.006 (0.068)			
S031440	S13_06	0.872 (0.081)	0.831 (0.071)			
S031441A	S13_07A	1.204 (0.089)	-0.236 (0.047)			
S031441B	S13_07B	0.983 (0.067)	0.555 (0.039)		0.597 (0.054)	-0.597 (0.067)
S031442	S13_08	1.143 (0.091)	0.166 (0.044)			
S031443	S13_09	0.896 (0.091)	1.008 (0.081)			
S031446A	S14_01A	0.923 (0.078)	0.623 (0.059)			
S031446B	S14_01B	0.716 (0.070)	0.882 (0.087)			
S031446C	S14_01C	0.762 (0.063)	-0.152 (0.061)			
S031445A	S14_02A	1.318 (0.098)	0.353 (0.039)			
S031445B	S14_02B	1.139 (0.083)	-0.639 (0.056)			
S031447	S14_03	0.447 (0.042)	1.025 (0.093)		0.396 (0.105)	-0.396 (0.144)
S031193	S14_04	0.497 (0.070)	-0.425 (0.202)	0.174 (0.050)		
S031264	S14_05	0.807 (0.086)	-0.451 (0.113)	0.141 (0.039)		
S031347	S14_06	0.598 (0.073)	-0.942 (0.185)	0.170 (0.049)		
S031346	S14_07	0.722 (0.082)	1.420 (0.134)			
S031081	S14_08	0.755 (0.064)	-0.692 (0.080)			
SF11006	S01F01	1.148 (0.054)	-0.238 (0.036)	0.083 (0.016)		
SF11007	S01F02	1.174 (0.053)	-0.201 (0.033)	0.063 (0.014)		
SF11008	S01F03	1.375 (0.059)	-0.143 (0.027)	0.040 (0.011)		
SF12033	S01F04	0.851 (0.046)	0.220 (0.038)	0.055 (0.014)		
SF11001	S01F05	1.775 (0.077)	-0.441 (0.027)	0.076 (0.014)		
SF11003	S01F06	1.420 (0.059)	-0.206 (0.026)	0.047 (0.011)		
SF11004	S01F07	1.460 (0.062)	-0.059 (0.024)	0.048 (0.010)		
SF11005	S01F08	1.959 (0.081)	-0.237 (0.021)	0.041 (0.009)		
SF11021	S01F09	1.610 (0.069)	-0.301 (0.027)	0.061 (0.012)		
SF11022	S01F10	1.200 (0.057)	-0.077 (0.032)	0.072 (0.014)		
SF11023	S01F11	1.743 (0.074)	0.000 (0.021)	0.040 (0.009)		
SF11030	S02F01	1.526 (0.060)	-0.191 (0.023)	0.033 (0.008)		
SF11031	S02F02	2.477 (0.101)	-0.265 (0.018)	0.026 (0.007)		
SF11032	S02F03	1.398 (0.057)	0.636 (0.020)			
SF11033	S02F04	0.920 (0.045)	0.682 (0.031)	0.017 (0.005)		
SF11025	S02F05	2.299 (0.090)	-0.126 (0.017)	0.018 (0.005)		
SF11026	S02F06	2.027 (0.077)	-0.122 (0.018)	0.018 (0.005)		
SF11027	S02F07	2.155 (0.085)	0.119 (0.015)	0.013 (0.004)		
SF11029	S02F09	3.846 (0.179)	-0.081 (0.013)	0.021 (0.005)		
SF11016	S02F10	3.123 (0.135)	-0.094 (0.014)	0.021 (0.005)		
SF12007	S02F11	2.034 (0.081)	0.073 (0.016)	0.018 (0.005)		
SF11017	S03F01	1.281 (0.123)	0.098 (0.053)	0.086 (0.024)		
SF11018	S03F02	2.164 (0.199)	-0.461 (0.049)	0.130 (0.029)		
SF12010	S03F03	1.744 (0.152)	0.014 (0.040)	0.062 (0.019)		
SF11019	S03F04	1.649 (0.117)	-0.000 (0.032)			
SF11009	S03F05	1.868 (0.160)	-0.014 (0.037)	0.067 (0.018)		

( ) Standard errors appear in parentheses

Exhibit D.2 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SF11010	S03F06	2.703 (0.256)	-0.275 (0.038)	0.114 (0.024)		
SF11011	S03F07	1.467 (0.156)	0.590 (0.046)	0.069 (0.018)		
SF11012	S03F08	2.554 (0.217)	-0.139 (0.032)	0.053 (0.017)		
SF11013	S03F09	1.302 (0.130)	0.571 (0.048)	0.043 (0.015)		
SF11014	S03F10	2.268 (0.202)	0.265 (0.029)	0.054 (0.015)		
SF11015	S03F11	1.989 (0.176)	0.010 (0.035)	0.070 (0.018)		
SF31017	S04F01	1.069 (0.110)	-0.395 (0.091)	0.146 (0.039)		
SF31246	S04F02	1.104 (0.099)	0.948 (0.063)			
SF31287	S04F03	0.985 (0.103)	-0.037 (0.079)	0.121 (0.032)		
SF31251	S04F04	0.917 (0.086)	0.952 (0.074)			
SF31053	S04F05	0.836 (0.049)	0.224 (0.034)		-0.201 (0.069)	0.201 (0.070)
SF31005	S04F06	1.066 (0.110)	1.307 (0.093)			
SF31306	S04F07	1.298 (0.167)	0.805 (0.060)	0.108 (0.021)		
SF31372A	S04F08A	1.664 (0.118)	0.214 (0.032)			
SF31372B	S04F08B	1.258 (0.082)	1.021 (0.039)		-0.189 (0.055)	0.189 (0.073)
SF31082	S04F09	1.039 (0.119)	0.179 (0.077)	0.142 (0.032)		
SF31229	S05F01	0.963 (0.139)	0.547 (0.087)	0.172 (0.033)		
SF31270	S05F02	0.665 (0.080)	1.591 (0.162)			
SF31026	S05F03	0.637 (0.034)	0.095 (0.042)		-0.739 (0.099)	0.739 (0.098)
SF31319	S05F04	1.021 (0.148)	0.913 (0.083)	0.118 (0.024)		
SF31414A	S05F05A	2.472 (0.178)	0.027 (0.026)			
SF31414B	S05F05B	2.058 (0.144)	-0.005 (0.029)			
SF31078	S05F06	1.098 (0.127)	0.197 (0.074)	0.130 (0.032)		
SF31009	S05F07	1.117 (0.084)	0.162 (0.043)			
SF31401	S05F08	1.260 (0.151)	0.534 (0.059)	0.122 (0.025)		
SF31384A	S05F09A	1.489 (0.103)	-0.371 (0.041)			
SF31384B	S05F09B	1.670 (0.117)	0.067 (0.033)			
SF31255	S06F01	1.001 (0.111)	-0.208 (0.095)	0.164 (0.038)		
SF31240D	S06F02D	0.786 (0.053)	-0.057 (0.041)		0.582 (0.071)	-0.582 (0.064)
SF31239	S06F03	0.797 (0.098)	-0.209 (0.125)	0.184 (0.044)		
SF31235A	S06F04A	1.313 (0.099)	0.449 (0.040)			
SF31235B	S06F04B	1.522 (0.116)	0.577 (0.038)			
SF31205	S06F05	0.847 (0.101)	0.267 (0.088)	0.127 (0.032)		
SF31399A	S06F06A	1.661 (0.120)	0.354 (0.033)			
SF31399B	S06F06B	1.647 (0.117)	0.194 (0.032)			
SF31393	S06F07	1.289 (0.091)	-0.542 (0.049)			
SF31278	S06F08	0.933 (0.072)	-0.079 (0.051)			
SF31317	S07F01	0.882 (0.098)	-0.398 (0.113)	0.180 (0.042)		
SF31190	S07F02	1.094 (0.087)	0.605 (0.051)			
SF31431	S07F03	0.816 (0.171)	1.598 (0.181)	0.121 (0.024)		
SF31283	S07F04	0.788 (0.088)	-0.476 (0.124)	0.163 (0.043)		
SF31426	S07F05	1.054 (0.115)	-0.111 (0.085)	0.156 (0.035)		
SF31422	S07F06	1.444 (0.156)	-0.551 (0.086)	0.228 (0.042)		
SF31427	S07F07	0.941 (0.109)	0.101 (0.088)	0.146 (0.034)		

( ) Standard errors appear in parentheses

Exhibit D.2 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SF31075	S07F08	0.652 (0.092)	0.275 (0.131)	0.158 (0.040)		
SF31047	S07F09	0.879 (0.070)	0.234 (0.053)			
SF31387	S07F10	0.938 (0.162)	1.180 (0.112)	0.129 (0.026)		
SF31396D	S07F11D	0.531 (0.028)	-0.330 (0.051)		-1.026 (0.126)	1.026 (0.119)
SF31340	S08F01	0.848 (0.135)	0.779 (0.104)	0.179 (0.034)		
SF31236	S08F02	0.827 (0.087)	-1.049 (0.138)	0.163 (0.045)		
SF31391D	S08F03D	0.621 (0.043)	0.432 (0.048)		0.023 (0.083)	-0.023 (0.092)
SF31361	S08F04	0.703 (0.102)	0.373 (0.125)	0.180 (0.040)		
SF31001	S08F05	1.128 (0.109)	-0.637 (0.090)	0.155 (0.037)		
SF31410	S08F07	0.653 (0.087)	-0.121 (0.149)	0.182 (0.046)		
SF31421	S08F08	0.566 (0.052)	-0.219 (0.079)			
SF31298	S08F09	0.749 (0.162)	1.446 (0.170)	0.180 (0.031)		
SF31076	S08F10	0.832 (0.072)	0.702 (0.067)			
SF31275	S08F11	0.834 (0.167)	1.463 (0.153)	0.141 (0.026)		
SF31254	S11F01	1.039 (0.158)	0.491 (0.092)	0.245 (0.035)		
SF31266	S11F02	1.541 (0.161)	0.420 (0.047)	0.101 (0.020)		
SF31233	S11F03	1.067 (0.081)	0.271 (0.045)			
SF31204	S11F04	0.928 (0.082)	0.795 (0.066)			
SF31273	S11F05	2.176 (0.241)	0.524 (0.038)	0.133 (0.019)		
SF31299	S11F06	1.185 (0.100)	0.860 (0.057)			
SF31281	S11F07	2.420 (0.287)	0.232 (0.042)	0.227 (0.025)		
SF31077	S11F08	1.928 (0.231)	0.416 (0.047)	0.202 (0.024)		
SF31311	S11F09	1.966 (0.220)	0.521 (0.041)	0.132 (0.020)		
SF31088D	S11F10D	0.859 (0.061)	0.899 (0.049)		0.608 (0.056)	-0.608 (0.088)
SF31389	S11F11	1.445 (0.195)	1.031 (0.065)	0.088 (0.016)		
SF31356	S12F01	0.863 (0.097)	-1.179 (0.159)	0.209 (0.053)		
SF31291	S12F02	1.293 (0.124)	-0.613 (0.082)	0.154 (0.037)		
SF31230	S12F03	1.006 (0.099)	-0.729 (0.103)	0.154 (0.040)		
SF31325	S12F04	0.817 (0.070)	0.555 (0.063)			
SF31068	S12F05	1.155 (0.154)	0.700 (0.068)	0.128 (0.025)		
SF31418	S12F06	0.948 (0.121)	0.629 (0.077)	0.110 (0.026)		
SF31197D	S12F07D	0.700 (0.038)	-0.191 (0.042)		-0.422 (0.090)	0.422 (0.084)
SF31371	S12F08	1.173 (0.149)	0.704 (0.064)	0.107 (0.023)		
SF31376	S12F09	0.926 (0.161)	1.033 (0.102)	0.154 (0.028)		
SF31044	S12F10	0.833 (0.071)	0.565 (0.062)			
SF31390D	S12F11D	0.907 (0.059)	0.457 (0.036)		0.131 (0.059)	-0.131 (0.067)
SF31409	S13F01	1.403 (0.081)	-0.022 (0.038)	0.175 (0.020)		
SF31398	S13F02	0.992 (0.064)	0.065 (0.053)	0.183 (0.023)		
SF31072	S13F03	0.848 (0.029)	0.154 (0.020)		0.487 (0.033)	-0.487 (0.032)
SF31061	S13F04	0.799 (0.055)	-0.157 (0.074)	0.191 (0.028)		
SF31439A	S13F05A	1.061 (0.051)	1.205 (0.040)			
SF31439B	S13F05B	0.945 (0.039)	0.385 (0.025)			
SF31440	S13F06	1.139 (0.049)	0.893 (0.029)			
SF31441A	S13F07A	1.396 (0.050)	0.021 (0.019)			

( ) Standard errors appear in parentheses

Exhibit D.2 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SF31441B	S13F07B	1.118 (0.038)	0.648 (0.017)		0.421 (0.023)	-0.421 (0.030)
SF31442	S13F08	1.027 (0.041)	0.388 (0.023)			
SF31443	S13F09	0.929 (0.045)	1.148 (0.042)			
SF31446A	S14F01A	1.075 (0.046)	0.751 (0.027)			
SF31446B	S14F01B	0.984 (0.044)	0.895 (0.032)			
SF31446C	S14F01C	0.870 (0.036)	0.229 (0.026)			
SF31445A	S14F02A	1.489 (0.057)	0.576 (0.018)			
SF31445B	S14F02B	1.312 (0.047)	-0.222 (0.022)			
SF31447	S14F03	0.605 (0.026)	1.139 (0.037)		0.430 (0.039)	-0.430 (0.058)
SF31193	S14F04	0.838 (0.061)	0.247 (0.062)	0.180 (0.025)		
SF31264	S14F05	1.099 (0.062)	-0.044 (0.044)	0.127 (0.021)		
SF31347	S14F06	0.981 (0.069)	0.037 (0.061)	0.210 (0.027)		
SF31346	S14F07	1.030 (0.055)	1.443 (0.051)			
SF31081	S14F08	0.776 (0.033)	-0.073 (0.030)			
SZ31446A	S01_01A	1.141 (0.073)	0.568 (0.036)			
SZ31446B	S01_01B	0.984 (0.066)	0.748 (0.044)			
SZ31446C	S01_01C	0.706 (0.048)	-0.064 (0.050)			
SZ31445A	S01_02A	1.662 (0.093)	0.341 (0.024)			
SZ31445B	S01_02B	1.054 (0.063)	-0.757 (0.053)			
SZ31447	S01_03	0.486 (0.034)	0.962 (0.061)		0.459 (0.074)	-0.459 (0.099)
SZ31193	S01_04	0.503 (0.064)	-0.316 (0.191)	0.194 (0.050)		
SZ31264	S01_05	0.731 (0.067)	-0.646 (0.119)	0.154 (0.039)		
SZ31347	S01_06	0.470 (0.054)	-1.184 (0.234)	0.182 (0.053)		
SZ31346	S01_07	0.718 (0.061)	1.245 (0.086)			
SZ31081	S01_08	0.679 (0.048)	-0.782 (0.075)			
SZ41007	S02_01	0.664 (0.071)	-0.284 (0.133)	0.180 (0.043)		
SZ41164	S02_02	0.984 (0.115)	0.635 (0.069)	0.180 (0.028)		
SZ41018	S02_03	0.637 (0.036)	-0.188 (0.038)		0.504 (0.068)	-0.504 (0.058)
SZ41160	S02_04	0.592 (0.115)	1.523 (0.167)	0.174 (0.034)		
SZ41042	S02_05	0.625 (0.058)	-1.341 (0.165)	0.152 (0.044)		
SZ41079	S02_06	0.942 (0.078)	-0.930 (0.102)	0.159 (0.038)		
SZ41073	S02_07	0.592 (0.044)	-0.172 (0.060)			
SZ41217	S02_08	1.051 (0.107)	0.283 (0.068)	0.175 (0.030)		
SZ41196	S02_09	0.664 (0.070)	-0.411 (0.136)	0.181 (0.044)		
SZ41211	S02_10	0.907 (0.056)	-0.072 (0.041)			
SZ41051	S02_11	0.873 (0.135)	1.023 (0.092)	0.217 (0.030)		
SZ41089	S02_12	1.003 (0.101)	0.029 (0.082)	0.206 (0.035)		
SZ41156A	S02_13A	1.173 (0.102)	-0.203 (0.071)	0.173 (0.033)		
SZ41156B	S02_13B	1.126 (0.069)	0.145 (0.034)			
SZ31229	S03_01	1.187 (0.134)	0.557 (0.061)	0.212 (0.028)		
SZ31270	S03_02	0.589 (0.059)	1.686 (0.141)			
SZ31026	S03_03	0.534 (0.025)	-0.219 (0.041)		-0.601 (0.088)	0.601 (0.081)
SZ31319	S03_04	1.151 (0.160)	1.064 (0.071)	0.209 (0.023)		
SZ31414A	S03_05A	1.351 (0.075)	-0.285 (0.034)			

( ) Standard errors appear in parentheses

Exhibit D.2 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SZ31414B	S03_05B	1.190 (0.068)	-0.327 (0.038)			
SZ31078	S03_06	0.756 (0.082)	-0.026 (0.109)	0.189 (0.039)		
SZ31009	S03_07	0.713 (0.050)	-0.218 (0.054)			
SZ31401	S03_08	1.172 (0.138)	0.542 (0.066)	0.255 (0.029)		
SZ31384A	S03_09A	0.822 (0.056)	-1.292 (0.086)			
SZ31384B	S03_09B	0.924 (0.059)	-0.511 (0.053)			
SZ41165	S04_01	0.603 (0.086)	0.647 (0.125)	0.182 (0.039)		
SZ41023	S04_02	0.999 (0.061)	0.278 (0.036)			
SZ41047	S04_03	0.491 (0.052)	-0.907 (0.181)	0.150 (0.044)		
SZ41001	S04_04	0.365 (0.028)	0.446 (0.061)		0.338 (0.104)	-0.338 (0.113)
SZ41029	S04_05	0.755 (0.050)	-0.791 (0.067)			
SZ41054	S04_06	0.597 (0.062)	-1.402 (0.206)	0.192 (0.053)		
SZ41308	S04_07	0.977 (0.115)	0.526 (0.075)	0.207 (0.031)		
SZ41179	S04_08	0.804 (0.056)	0.569 (0.047)			
SZ41087	S04_09	1.193 (0.126)	0.504 (0.058)	0.189 (0.027)		
SZ41205	S04_10	0.866 (0.090)	0.152 (0.086)	0.170 (0.034)		
SZ41216	S04_11	0.726 (0.052)	0.309 (0.048)			
SZ41061	S04_12	0.866 (0.056)	-0.256 (0.047)			
SZ41202	S04_13	0.555 (0.025)	0.617 (0.038)		-1.099 (0.092)	1.099 (0.098)
SZ41215	S04_14	0.948 (0.116)	0.751 (0.074)	0.162 (0.029)		
SZ31255	S05_01	0.755 (0.078)	-0.293 (0.121)	0.194 (0.043)		
SZ31240D	S05_02D	0.634 (0.033)	-0.114 (0.040)		0.909 (0.070)	-0.909 (0.059)
SZ31239	S05_03	0.560 (0.069)	-0.518 (0.196)	0.220 (0.053)		
SZ31235A	S05_04A	1.184 (0.071)	0.382 (0.032)			
SZ31235B	S05_04B	1.241 (0.076)	0.573 (0.032)			
SZ31205	S05_05	0.550 (0.068)	0.040 (0.152)	0.182 (0.044)		
SZ31399A	S05_06A	1.210 (0.071)	0.222 (0.031)			
SZ31399B	S05_06B	1.101 (0.066)	-0.059 (0.036)			
SZ31393	S05_07	0.796 (0.057)	-1.275 (0.093)			
SZ31278	S05_08	0.534 (0.044)	-0.674 (0.089)			
SZ41117	S06_01	0.519 (0.053)	-2.498 (0.278)	0.180 (0.055)		
SZ41120	S06_02	0.938 (0.191)	1.271 (0.122)	0.372 (0.029)		
SZ41003	S06_03	0.669 (0.048)	0.196 (0.050)			
SZ41224	S06_04	1.039 (0.054)	0.495 (0.025)		0.374 (0.038)	-0.374 (0.044)
SZ41163	S06_05	0.527 (0.088)	1.055 (0.151)	0.181 (0.039)		
SZ41039	S06_06	0.750 (0.050)	-0.030 (0.047)			
SZ41014	S06_07	1.251 (0.217)	1.430 (0.092)	0.217 (0.020)		
SZ41181	S06_08	0.703 (0.048)	-0.323 (0.056)			
SZ41174	S06_09	0.872 (0.060)	0.675 (0.046)			
SZ41049	S06_10	1.073 (0.115)	0.433 (0.067)	0.206 (0.029)		
SZ41208	S06_11	0.362 (0.066)	0.785 (0.255)	0.204 (0.053)		
SZ41060	S06_12	1.016 (0.075)	1.078 (0.057)			
SZ41201A	S06_13A	1.188 (0.071)	0.244 (0.032)			
SZ41201B	S06_13B	1.198 (0.075)	0.392 (0.032)			

( ) Standard errors appear in parentheses

Exhibit D.2 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SZ31317	S07_01	0.768 (0.083)	-0.464 (0.140)	0.251 (0.047)		
SZ31190	S07_02	1.069 (0.067)	0.575 (0.037)			
SZ31431	S07_03	0.693 (0.125)	1.555 (0.147)	0.152 (0.027)		
SZ31283	S07_04	0.477 (0.054)	-0.957 (0.214)	0.177 (0.051)		
SZ31426	S07_05	0.756 (0.074)	-0.353 (0.115)	0.174 (0.040)		
SZ31422	S07_06	0.821 (0.078)	-1.255 (0.151)	0.209 (0.049)		
SZ31427	S07_07	0.745 (0.078)	-0.075 (0.109)	0.179 (0.038)		
SZ31075	S07_08	0.351 (0.057)	0.138 (0.274)	0.200 (0.055)		
SZ31047	S07_09	0.681 (0.048)	-0.036 (0.052)			
SZ31387	S07_10	0.866 (0.136)	1.279 (0.101)	0.160 (0.026)		
SZ31396D	S07_11D	0.510 (0.026)	-1.020 (0.066)		-0.462 (0.114)	0.462 (0.087)
SZ41009	S08_01	0.889 (0.086)	-0.566 (0.117)	0.227 (0.043)		
SZ41223	S08_02	0.927 (0.108)	0.183 (0.095)	0.261 (0.037)		
SZ41026	S08_03	0.587 (0.079)	0.544 (0.125)	0.174 (0.038)		
SZ41177	S08_04	0.451 (0.033)	1.024 (0.067)		0.377 (0.079)	-0.377 (0.105)
SZ41183	S08_05	0.692 (0.032)	0.193 (0.038)		1.032 (0.059)	-1.032 (0.059)
SZ41008	S08_06	1.048 (0.119)	0.590 (0.066)	0.192 (0.028)		
SZ41195	S08_08	0.748 (0.061)	1.133 (0.075)			
SZ41134A	S08_09A	0.888 (0.057)	0.324 (0.040)			
SZ41134B	S08_09B	0.750 (0.050)	-0.027 (0.047)			
SZ41134C	S08_09C	0.835 (0.090)	0.424 (0.078)	0.145 (0.030)		
SZ41191	S08_10	0.690 (0.010)	0.674 (0.113)	0.217 (0.038)		
SZ41107	S08_11	0.462 (0.021)	-0.576 (0.053)		-1.042 (0.114)	1.042 (0.100)
SZ41113	S08_12	0.886 (0.058)	0.274 (0.040)			
SZ31340	S09_01	0.851 (0.117)	0.789 (0.088)	0.222 (0.031)		
SZ31236	S09_02	0.624 (0.059)	-1.249 (0.164)	0.154 (0.043)		
SZ31391D	S09_03D	0.634 (0.036)	0.329 (0.034)		0.092 (0.063)	-0.092 (0.065)
SZ31361	S09_04	0.729 (0.094)	0.432 (0.108)	0.213 (0.038)		
SZ31001	S09_05	1.008 (0.082)	-0.829 (0.093)	0.153 (0.036)		
SZ31410	S09_07	0.544 (0.067)	-0.142 (0.161)	0.180 (0.046)		
SZ31421	S09_08	0.468 (0.039)	-0.598 (0.089)			
SZ31298	S09_09	1.162 (0.210)	1.408 (0.099)	0.240 (0.022)		
SZ31076	S09_10	0.934 (0.061)	0.393 (0.039)			
SZ31275	S09_11	0.815 (0.136)	1.441 (0.121)	0.160 (0.025)		
SZ41311	S10_01	0.687 (0.066)	-2.334 (0.226)	0.184 (0.055)		
SZ41178	S10_02	0.949 (0.110)	0.479 (0.076)	0.196 (0.031)		
SZ41182	S10_03	0.632 (0.048)	0.324 (0.055)			
SZ41180	S10_04	1.467 (0.137)	0.225 (0.051)	0.212 (0.026)		
SZ41187	S10_05	0.927 (0.176)	1.547 (0.131)	0.185 (0.024)		
SZ41013A	S10_06A	0.534 (0.045)	0.577 (0.069)			
SZ41013B	S10_06B	0.466 (0.049)	1.565 (0.150)			
SZ41067	S10_07	0.830 (0.052)	-0.407 (0.051)			
SZ41305	S10_08	1.141 (0.124)	0.570 (0.060)	0.187 (0.027)		
SZ41048	S10_09	0.848 (0.055)	0.274 (0.041)			

( ) Standard errors appear in parentheses

Exhibit D.2 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SZ41110	S10_10	0.739 (0.049)	-0.350 (0.055)			
SZ41069	S10_11	1.000 (0.119)	0.247 (0.090)	0.279 (0.036)		
SZ41100	S10_12	1.134 (0.109)	0.364 (0.058)	0.161 (0.026)		
SZ41092	S10_13	0.793 (0.092)	0.123 (0.109)	0.218 (0.040)		
SZ31254	S11_01	0.530 (0.077)	0.175 (0.180)	0.225 (0.050)		
SZ31266	S11_02	1.085 (0.100)	0.115 (0.065)	0.165 (0.029)		
SZ31233	S11_03	0.674 (0.046)	-0.335 (0.058)			
SZ31204	S11_04	0.646 (0.049)	0.497 (0.055)			
SZ31273	S11_05	1.345 (0.128)	0.200 (0.056)	0.207 (0.028)		
SZ31299	S11_06	0.672 (0.050)	0.487 (0.053)			
SZ31281	S11_07	0.880 (0.076)	-1.212 (0.124)	0.160 (0.041)		
SZ31077	S11_08	0.669 (0.070)	-0.667 (0.153)	0.194 (0.047)		
SZ31311	S11_09	1.012 (0.097)	0.043 (0.076)	0.176 (0.033)		
SZ31088D	S11_10D	0.616 (0.027)	0.196 (0.044)		1.319 (0.068)	-1.319 (0.069)
SZ31389	S11_11	0.888 (0.117)	0.820 (0.081)	0.184 (0.030)		
SZ41027	S12_01	0.754 (0.052)	-1.811 (0.111)			
SZ41043	S12_02	0.577 (0.042)	-0.458 (0.068)			
SZ41050	S12_03	0.492 (0.069)	0.609 (0.148)	0.163 (0.039)		
SZ41070	S12_04	0.818 (0.084)	0.213 (0.084)	0.145 (0.032)		
SZ41006	S12_05	0.519 (0.033)	0.505 (0.043)		0.213 (0.072)	-0.213 (0.080)
SZ41052	S12_06	0.732 (0.076)	-0.661 (0.147)	0.222 (0.047)		
SZ41301	S12_07	0.734 (0.054)	0.723 (0.055)			
SZ41080	S12_08	0.669 (0.110)	1.359 (0.127)	0.147 (0.029)		
SZ41033	S12_09	0.925 (0.066)	0.949 (0.054)			
SZ41077	S12_11	0.792 (0.053)	0.253 (0.043)			
SZ41209	S12_12	0.747 (0.103)	0.772 (0.098)	0.190 (0.034)		
SZ41081	S12_13	0.582 (0.030)	0.532 (0.037)		-0.363 (0.072)	0.363 (0.078)
SZ41102	S12_14	0.861 (0.094)	-0.041 (0.106)	0.236 (0.040)		
SZ31356	S13_01	0.792 (0.078)	-1.675 (0.186)	0.210 (0.054)		
SZ31291	S13_02	1.132 (0.089)	-1.012 (0.089)	0.150 (0.036)		
SZ31230	S13_03	0.665 (0.061)	-1.492 (0.171)	0.162 (0.046)		
SZ31325	S13_04	0.691 (0.049)	0.367 (0.049)			
SZ31068	S13_05	0.994 (0.105)	0.422 (0.069)	0.170 (0.029)		
SZ31418	S13_06	0.859 (0.094)	0.625 (0.072)	0.134 (0.027)		
SZ31197D	S13_07D	0.548 (0.027)	-0.714 (0.053)		-0.411 (0.097)	0.411 (0.077)
SZ31371	S13_08	0.951 (0.113)	0.566 (0.076)	0.198 (0.031)		
SZ31376	S13_09	0.952 (0.137)	1.027 (0.081)	0.204 (0.027)		
SZ31044	S13_10	0.738 (0.051)	0.176 (0.046)			
SZ31390D	S13_11D	0.857 (0.046)	0.230 (0.028)		0.326 (0.049)	-0.326 (0.048)
SZ41010	S14_01	1.137 (0.089)	-0.694 (0.081)	0.157 (0.035)		
SZ41034	S14_02	0.529 (0.058)	-0.636 (0.174)	0.166 (0.046)		
SZ41017	S14_03	0.699 (0.112)	0.990 (0.113)	0.213 (0.035)		
SZ41124	S14_04	0.876 (0.135)	1.060 (0.093)	0.232 (0.029)		
SZ41186	S14_05	0.623 (0.053)	1.193 (0.089)			

( ) Standard errors appear in parentheses

Exhibit D.2 IRT Parameters for TIMSS Joint 2003-2007 Fourth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SZ41037	S14_06	0.516 (0.028)	0.083 (0.040)		-0.224 (0.081)	0.224 (0.078)
SZ41119	S14_07	0.547 (0.071)	-0.580 (0.222)	0.251 (0.057)		
SZ41105	S14_08	0.923 (0.082)	-0.073 (0.077)	0.145 (0.031)		
SZ41193	S14_09	0.475 (0.049)	-0.815 (0.171)	0.137 (0.041)		
SZ41149D	S14_10D	0.535 (0.024)	0.867 (0.044)		-1.158 (0.094)	1.158 (0.104)
SZ41032	S14_11	0.956 (0.063)	-1.331 (0.080)			
SZ41068	S14_12	0.665 (0.049)	0.342 (0.051)			
SZ41303	S14_13	0.506 (0.077)	0.709 (0.154)	0.171 (0.043)		

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M012001	M01_01	1.721 (0.067)	0.247 (0.020)	0.149 (0.010)		
M012002	M01_02	0.738 (0.035)	-0.355 (0.070)	0.155 (0.028)		
M012003	M01_03	1.075 (0.038)	0.092 (0.026)	0.065 (0.012)		
M012004	M01_04	1.216 (0.066)	0.771 (0.032)	0.262 (0.012)		
M012005	M01_05	0.760 (0.045)	0.246 (0.066)	0.205 (0.025)		
M012006	M01_06	0.768 (0.044)	-0.326 (0.086)	0.232 (0.033)		
M012037	M01_07	0.546 (0.032)	0.305 (0.083)	0.110 (0.027)		
M012038	M01_08	0.987 (0.051)	-0.196 (0.058)	0.291 (0.024)		
M012039	M01_09	1.141 (0.054)	0.368 (0.034)	0.198 (0.015)		
M012040	M01_10	1.071 (0.049)	-0.231 (0.046)	0.217 (0.021)		
M012041	M01_11	1.234 (0.048)	0.103 (0.027)	0.115 (0.013)		
M012042	M01_12	1.380 (0.061)	0.378 (0.027)	0.173 (0.013)		
M032570	M01_13	1.307 (0.068)	0.398 (0.034)	0.302 (0.014)		
M032643	M01_14	1.247 (0.064)	0.777 (0.029)	0.190 (0.011)		
M032693	M01_15	0.713 (0.021)	0.672 (0.026)			
M012013	M02_01	1.249 (0.086)	0.296 (0.045)	0.185 (0.020)		
M012014	M02_02	0.938 (0.068)	-0.545 (0.092)	0.235 (0.039)		
M012015	M02_03	0.954 (0.059)	-0.251 (0.061)	0.123 (0.027)		
M012016	M02_04	1.280 (0.140)	0.954 (0.059)	0.372 (0.017)		
M012017	M02_05	0.803 (0.053)	0.203 (0.061)	0.097 (0.023)		
M022251	M02_06	1.052 (0.109)	1.302 (0.058)	0.166 (0.016)		
M022185	M02_07	0.847 (0.079)	0.295 (0.090)	0.249 (0.033)		
M022188	M02_08	0.812 (0.088)	0.973 (0.077)	0.238 (0.024)		
M022189	M02_09	0.913 (0.055)	-0.564 (0.071)	0.141 (0.032)		
M022191	M02_10	0.824 (0.066)	-0.033 (0.090)	0.208 (0.034)		
M022194	M02_11	0.863 (0.064)	0.285 (0.064)	0.140 (0.025)		
M022196	M02_12	1.302 (0.081)	-0.115 (0.045)	0.144 (0.023)		
M022198	M02_13	1.126 (0.100)	0.789 (0.053)	0.235 (0.019)		
M022199	M02_14	1.318 (0.109)	0.721 (0.045)	0.214 (0.018)		
M022202	M02_15	0.734 (0.036)	0.865 (0.046)			
M012025	M03_01	0.892 (0.058)	-0.342 (0.072)	0.137 (0.031)		
M012026	M03_02	1.098 (0.084)	0.476 (0.051)	0.192 (0.020)		
M012027	M03_03	1.227 (0.090)	0.256 (0.049)	0.213 (0.022)		
M012028	M03_04	1.182 (0.073)	-0.213 (0.051)	0.156 (0.025)		
M012029	M03_05	1.057 (0.072)	0.199 (0.051)	0.145 (0.022)		
M012030	M03_06	1.363 (0.094)	0.654 (0.036)	0.139 (0.014)		
M022135	M03_07	0.664 (0.048)	0.911 (0.065)	0.068 (0.016)		
M022139	M03_08	1.259 (0.103)	1.021 (0.043)	0.153 (0.013)		
M022142	M03_09	1.405 (0.103)	0.556 (0.039)	0.190 (0.016)		
M022144	M03_10	0.576 (0.072)	0.801 (0.126)	0.212 (0.038)		
M022146	M03_11	1.516 (0.091)	0.112 (0.034)	0.147 (0.017)		
M022148	M03_12	0.989 (0.040)	0.037 (0.028)			
M022253	M03_13	1.252 (0.050)	0.076 (0.024)			
M022154	M03_14	0.934 (0.072)	0.330 (0.062)	0.165 (0.025)		

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M022156	M03_15	1.258 (0.049)	0.261 (0.024)			
M022002	M04_01	1.700 (0.119)	1.222 (0.036)	0.128 (0.009)		
M022004	M04_02	1.421 (0.117)	0.583 (0.044)	0.276 (0.018)		
M022005	M04_03	1.065 (0.121)	1.211 (0.063)	0.260 (0.017)		
M022008	M04_04	0.624 (0.033)	0.795 (0.051)			
M022010	M04_05	0.811 (0.053)	-0.321 (0.078)	0.141 (0.032)		
M022012	M04_06	0.604 (0.030)	-0.416 (0.044)			
M022021	M04_07	1.755 (0.123)	0.611 (0.030)	0.167 (0.013)		
M022016	M04_08	0.872 (0.098)	1.158 (0.069)	0.195 (0.021)		
M022252	M04_09	1.104 (0.087)	0.056 (0.066)	0.272 (0.028)		
M022261A	M04_10A	1.219 (0.048)	0.355 (0.025)			
M022261B	M04_10B	1.285 (0.055)	0.852 (0.029)			
M022261C	M04_10C	0.749 (0.024)	1.142 (0.030)		-1.910 (0.107)	1.910 (0.113)
M022227A	M04_11A	1.058 (0.043)	-0.216 (0.027)			
M022227B	M04_11B	1.505 (0.061)	0.547 (0.023)			
M022227C	M04_11C	1.353 (0.059)	0.899 (0.029)			
M022127	M04_12	1.592 (0.121)	1.327 (0.043)	0.170 (0.001)		
M022043	M05_01	0.635 (0.024)	-0.480 (0.058)	0.075 (0.022)		
M022046	M05_02	0.844 (0.023)	-0.449 (0.021)			
M022049	M05_03	0.556 (0.040)	0.123 (0.116)	0.243 (0.034)		
M022050	M05_04	0.918 (0.046)	0.948 (0.031)	0.111 (0.011)		
M022055	M05_05	1.231 (0.028)	0.508 (0.015)			
M022057	M05_06	0.453 (0.025)	-0.352 (0.119)	0.153 (0.034)		
M022257	M05_07	1.519 (0.069)	0.551 (0.024)	0.252 (0.001)		
M022062	M05_08	0.925 (0.039)	0.625 (0.029)	0.095 (0.011)		
M022066	M05_09	1.347 (0.044)	0.097 (0.021)	0.081 (0.010)		
M022232	M05_10	0.529 (0.011)	1.592 (0.029)		-2.177 (0.069)	2.177 (0.078)
M022234A	M05_11A	0.804 (0.014)	0.769 (0.014)		-0.633 (0.029)	0.633 (0.032)
M022234B	M05_11B	0.897 (0.017)	1.082 (0.015)		-1.479 (0.051)	1.479 (0.053)
M022243	M05_12	1.157 (0.027)	0.510 (0.016)			
M022097	M06_01	1.007 (0.028)	-0.193 (0.026)	0.084 (0.012)		
M022101	M06_02	0.769 (0.025)	-0.497 (0.046)	0.106 (0.020)		
M022104	M06_03	0.887 (0.025)	-0.493 (0.032)	0.074 (0.015)		
M022105	M06_04	0.605 (0.027)	0.695 (0.044)	0.082 (0.015)		
M022106	M06_05	0.935 (0.019)	0.753 (0.017)			
M022108	M06_06	0.875 (0.036)	0.199 (0.040)	0.206 (0.016)		
M022110	M06_07	0.433 (0.027)	0.328 (0.057)			
M022181	M06_08	1.080 (0.037)	-0.683 (0.043)	0.222 (0.021)		
M032307	M06_09	1.394 (0.028)	0.951 (0.014)			
M032523	M06_10	1.789 (0.058)	1.132 (0.016)	0.165 (0.005)		
M032701	M06_11	0.986 (0.064)	-0.931 (0.088)	0.190 (0.043)		
M032704	M06_12	1.101 (0.075)	-0.005 (0.058)	0.181 (0.027)		
M032525	M06_13	1.009 (0.074)	0.314 (0.060)	0.173 (0.025)		
M032579	M06_14	1.113 (0.036)	-0.145 (0.029)	0.170 (0.014)		

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M032691	M06_15	0.850 (0.017)	0.417 (0.016)			
M032142	M07_01	1.479 (0.230)	1.049 (0.075)	0.385 (0.021)		
M032198	M07_02	0.734 (0.092)	0.637 (0.112)	0.182 (0.037)		
M032640	M07_03	0.613 (0.032)	1.557 (0.068)		-0.797 (0.095)	0.797 (0.125)
M032344	M07_04	1.157 (0.066)	0.551 (0.038)			
M032754	M07_05	0.855 (0.052)	-0.429 (0.048)			
M032755	M07_06	1.099 (0.054)	1.224 (0.036)		-0.287 (0.057)	0.287 (0.072)
M032753A	M07_07A	1.118 (0.048)	0.737 (0.027)		-0.286 (0.050)	0.286 (0.057)
M032753B	M07_07B	1.206 (0.058)	0.911 (0.029)		-0.029 (0.043)	0.029 (0.054)
M032753C	M07_07C	1.049 (0.061)	0.463 (0.040)			
M032756	M07_08	0.776 (0.049)	0.360 (0.049)			
M032205	M07_09	0.585 (0.067)	0.032 (0.158)	0.192 (0.048)		
M032163	M07_10	1.389 (0.152)	0.637 (0.060)	0.231 (0.024)		
M032381	M08_01	1.042 (0.059)	0.263 (0.038)			
M032416	M08_02	1.328 (0.138)	0.894 (0.053)	0.126 (0.019)		
M032160	M08_03	1.909 (0.158)	1.226 (0.046)	0.142 (0.012)		
M032273	M08_04	1.021 (0.098)	-0.228 (0.095)	0.221 (0.040)		
M032540	M08_05	0.761 (0.090)	0.048 (0.138)	0.240 (0.048)		
M032698	M08_06	1.010 (0.108)	0.526 (0.065)	0.140 (0.026)		
M032097	M08_07	1.391 (0.170)	1.298 (0.065)	0.190 (0.016)		
M032575	M08_08	2.090 (0.206)	0.513 (0.040)	0.216 (0.019)		
M032414	M08_09	1.078 (0.064)	0.597 (0.041)			
M032294	M08_10	0.848 (0.082)	-0.149 (0.107)	0.188 (0.041)		
M032688	M08_11	0.858 (0.054)	0.744 (0.052)			
M032529	M08_12	1.644 (0.169)	0.907 (0.045)	0.148 (0.016)		
M032637A	M08_13A	1.038 (0.059)	-0.291 (0.040)			
M032637B	M08_13B	1.354 (0.073)	-0.074 (0.032)			
M032637C	M08_13C	1.325 (0.073)	0.355 (0.033)			
M032079	M09_01	1.191 (0.074)	1.198 (0.033)	0.199 (0.009)		
M032652	M09_02	1.278 (0.035)	1.010 (0.020)			
M032228	M09_03	1.304 (0.060)	0.406 (0.028)	0.195 (0.012)		
M032044	M09_04	1.177 (0.061)	0.650 (0.032)	0.216 (0.012)		
M032046	M09_05	1.378 (0.074)	1.194 (0.026)	0.121 (0.007)		
M032545	M09_06	1.200 (0.032)	0.879 (0.019)			
M032649A	M09_07A	1.045 (0.028)	0.612 (0.019)			
M032649B	M09_07B	1.226 (0.036)	1.221 (0.024)			
M032533	M09_08	1.474 (0.067)	0.503 (0.025)	0.208 (0.011)		
M032678	M09_09	1.634 (0.056)	0.459 (0.016)	0.062 (0.007)		
M032403	M09_10	0.898 (0.024)	-0.028 (0.019)			
M032261	M09_11	0.947 (0.049)	0.633 (0.037)	0.150 (0.014)		
M032489	M09_12	0.873 (0.041)	-0.449 (0.062)	0.176 (0.027)		
M032588	M09_13	0.902 (0.047)	0.169 (0.051)	0.212 (0.020)		
M032271	M09_14	1.371 (0.072)	0.702 (0.028)	0.241 (0.011)		
M032671	M10_01	1.019 (0.026)	-0.319 (0.018)			

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M032612	M10_02	1.115 (0.063)	1.034 (0.032)	0.173 (0.011)		
M032557	M10_03	1.185 (0.033)	0.979 (0.020)			
M032208	M10_04	1.352 (0.065)	0.429 (0.030)	0.235 (0.013)		
M032210	M10_05	1.714 (0.083)	0.800 (0.022)	0.197 (0.009)		
M032699	M10_06	0.652 (0.034)	-0.755 (0.105)	0.184 (0.040)		
M032762	M10_07	0.411 (0.008)	1.204 (0.029)		-2.657 (0.081)	2.657 (0.087)
M032763	M10_08	0.837 (0.021)	1.732 (0.027)		-0.761 (0.043)	0.761 (0.055)
M032764	M10_09	0.853 (0.021)	1.634 (0.026)		-0.427 (0.035)	0.427 (0.047)
M032647	M10_10	0.872 (0.084)	1.449 (0.058)	0.326 (0.015)		
M032689	M10_11	0.795 (0.072)	1.310 (0.060)	0.306 (0.016)		
M032094	M11_01	1.442 (0.080)	0.288 (0.034)	0.302 (0.015)		
M032662	M11_02	1.657 (0.076)	1.398 (0.027)	0.096 (0.006)		
M032064	M11_03	1.302 (0.038)	0.674 (0.018)			
M032419	M11_04	1.324 (0.083)	0.855 (0.033)	0.252 (0.012)		
M032477	M11_05	1.500 (0.075)	0.524 (0.027)	0.191 (0.012)		
M032538	M11_06	1.231 (0.036)	0.289 (0.017)			
M032324	M11_07	1.372 (0.072)	0.800 (0.027)	0.162 (0.010)		
M032116	M11_08	0.973 (0.062)	0.720 (0.044)	0.225 (0.016)		
M032100	M11_09	0.982 (0.046)	0.264 (0.036)	0.106 (0.015)		
M032402	M11_10	0.706 (0.059)	0.796 (0.075)	0.233 (0.025)		
M032734	M11_11	0.861 (0.026)	-0.259 (0.023)			
M032397	M11_12	1.127 (0.070)	0.845 (0.037)	0.212 (0.013)		
M032695	M11_13	0.554 (0.011)	-0.202 (0.020)		-1.064 (0.050)	1.064 (0.048)
M032132	M11_14	0.700 (0.036)	0.353 (0.052)	0.090 (0.019)		
M032352	M12_01	1.449 (0.098)	0.633 (0.038)	0.393 (0.013)		
M032725	M12_02	1.072 (0.033)	0.914 (0.024)			
M032683	M12_03	0.492 (0.011)	0.868 (0.025)		-1.599 (0.061)	1.599 (0.067)
M032738	M12_04	1.329 (0.068)	-0.102 (0.040)	0.263 (0.020)		
M032295	M12_05	1.507 (0.082)	-0.268 (0.041)	0.336 (0.021)		
M032331	M12_06	2.035 (0.084)	1.298 (0.025)	0.191 (0.007)		
M032623	M12_07	1.860 (0.087)	0.643 (0.020)	0.136 (0.008)		
M032679	M12_08	1.130 (0.062)	0.397 (0.039)	0.230 (0.016)		
M032047	M12_09	1.905 (0.110)	1.169 (0.033)	0.438 (0.009)		
M032398	M12_10	1.783 (0.105)	0.964 (0.028)	0.302 (0.001)		
M032507	M12_11	1.788 (0.087)	1.123 (0.025)	0.180 (0.007)		
M032424	M12_12	1.172 (0.058)	0.425 (0.032)	0.153 (0.014)		
M032681A	M12_13A	0.567 (0.021)	-0.530 (0.036)			
M032681B	M12_13B	0.532 (0.022)	0.983 (0.045)			
M032681C	M12_13C	1.026 (0.030)	0.536 (0.021)			
M032609	M13_01	0.924 (0.077)	-0.288 (0.084)	0.145 (0.035)		
M032690	M13_02	1.024 (0.142)	0.999 (0.082)	0.217 (0.026)		
M032727	M13_03	1.461 (0.137)	0.458 (0.050)	0.172 (0.021)		
M032743	M13_04	0.616 (0.044)	-0.081 (0.058)			
M032744	M13_05	0.910 (0.057)	0.619 (0.048)			

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M032745	M13_06	0.505 (0.033)	2.289 (0.134)		-1.434 (0.154)	1.434 (0.213)
M032233	M13_07	1.060 (0.049)	1.083 (0.034)		-0.470 (0.062)	0.470 (0.074)
M032670	M13_08	0.919 (0.075)	-0.989 (0.104)	0.165 (0.044)		
M032447	M13_09	1.460 (0.141)	0.542 (0.050)	0.168 (0.022)		
M032036	M13_10	1.219 (0.133)	0.446 (0.069)	0.234 (0.029)		
M032728	M13_11	1.420 (0.177)	0.842 (0.062)	0.255 (0.023)		
M032732	M13_12	0.898 (0.111)	0.328 (0.107)	0.269 (0.038)		
M032166	M14_01	1.015 (0.102)	0.111 (0.084)	0.203 (0.035)		
M032721	M14_02	0.716 (0.134)	1.404 (0.141)	0.251 (0.032)		
M032757	M14_03	0.481 (0.019)	-0.187 (0.041)		-2.275 (0.147)	2.275 (0.145)
M032760A	M14_04A	0.813 (0.032)	0.671 (0.031)		-1.390 (0.102)	1.390 (0.107)
M032760B	M14_04B	1.452 (0.091)	1.000 (0.041)			
M032760C	M14_04C	1.601 (0.110)	1.233 (0.045)			
M032761	M14_05	1.047 (0.052)	1.253 (0.039)		-0.406 (0.063)	0.406 (0.080)
M032692	M14_06	0.690 (0.030)	0.984 (0.042)		-0.995 (0.089)	0.995 (0.102)
M032626	M14_07	0.877 (0.096)	0.454 (0.087)	0.171 (0.032)		
M032595	M14_08	1.421 (0.129)	0.323 (0.051)	0.153 (0.023)		
M032673	M14_09	1.527 (0.156)	0.559 (0.051)	0.197 (0.022)		
MF12001	M01F01	1.715 (0.074)	0.205 (0.022)	0.136 (0.012)		
MF12002	M01F02	0.798 (0.035)	-0.313 (0.053)	0.103 (0.022)		
MF12003	M01F03	1.105 (0.043)	0.153 (0.027)	0.066 (0.012)		
MF12004	M01F04	1.133 (0.068)	0.769 (0.036)	0.218 (0.013)		
MF12005	M01F05	0.877 (0.051)	0.331 (0.052)	0.172 (0.021)		
MF12006	M01F06	0.874 (0.044)	-0.223 (0.057)	0.137 (0.025)		
MF12037	M01F07	0.617 (0.035)	0.424 (0.062)	0.086 (0.021)		
MF12038	M01F08	1.091 (0.050)	-0.213 (0.042)	0.148 (0.020)		
MF12039	M01F09	1.159 (0.057)	0.350 (0.033)	0.148 (0.015)		
MF12040	M01F10	1.220 (0.055)	-0.128 (0.036)	0.163 (0.018)		
MF12041	M01F11	1.349 (0.051)	0.140 (0.023)	0.066 (0.010)		
MF12042	M01F12	1.371 (0.058)	0.402 (0.024)	0.078 (0.010)		
MF32570	M01F13	1.212 (0.058)	0.247 (0.033)	0.167 (0.015)		
MF32643	M01F14	1.174 (0.065)	0.770 (0.032)	0.152 (0.012)		
MF32693	M01F15	0.810 (0.025)	0.538 (0.025)			
MF12013	M02F01	1.209 (0.057)	0.353 (0.031)	0.141 (0.014)		
MF12014	M02F02	1.065 (0.051)	-0.303 (0.048)	0.188 (0.023)		
MF12015	M02F03	1.092 (0.043)	0.033 (0.029)	0.068 (0.013)		
MF12016	M02F04	1.213 (0.088)	0.898 (0.042)	0.321 (0.014)		
MF12017	M02F05	1.030 (0.049)	0.449 (0.033)	0.101 (0.013)		
MF22251	M02F06	1.089 (0.082)	1.431 (0.042)	0.157 (0.010)		
MF22185	M02F07	0.892 (0.049)	0.294 (0.049)	0.153 (0.020)		
MF22188	M02F08	0.903 (0.062)	0.983 (0.045)	0.192 (0.015)		
MF22189	M02F09	1.105 (0.048)	-0.051 (0.035)	0.121 (0.016)		
MF22191	M02F10	0.945 (0.044)	0.171 (0.039)	0.104 (0.016)		
MF22194	M02F11	0.965 (0.050)	0.527 (0.037)	0.122 (0.015)		

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MF22196	M02F12	1.605 (0.066)	0.219 (0.021)	0.098 (0.011)		
MF22198	M02F13	1.203 (0.067)	0.871 (0.031)	0.152 (0.011)		
MF22199	M02F14	1.459 (0.079)	0.829 (0.026)	0.147 (0.010)		
MF22202	M02F15	0.877 (0.029)	0.989 (0.029)			
MF12025	M03F01	0.960 (0.078)	-0.118 (0.075)	0.118 (0.031)		
MF12026	M03F02	1.088 (0.110)	0.633 (0.065)	0.147 (0.025)		
MF12027	M03F03	1.180 (0.106)	0.315 (0.061)	0.141 (0.027)		
MF12028	M03F04	1.269 (0.105)	0.113 (0.056)	0.131 (0.026)		
MF12029	M03F05	1.331 (0.110)	0.378 (0.049)	0.104 (0.021)		
MF12030	M03F06	1.584 (0.152)	0.801 (0.044)	0.132 (0.017)		
MF22135	M03F07	0.951 (0.089)	1.002 (0.064)	0.073 (0.017)		
MF22139	M03F08	1.434 (0.164)	1.117 (0.054)	0.148 (0.017)		
MF22142	M03F09	1.358 (0.125)	0.669 (0.049)	0.122 (0.019)		
MF22144	M03F10	0.740 (0.091)	0.859 (0.098)	0.141 (0.032)		
MF22146	M03F11	1.278 (0.099)	0.497 (0.045)	0.076 (0.017)		
MF22148	M03F12	1.213 (0.068)	0.431 (0.035)			
MF22253	M03F13	1.426 (0.079)	0.422 (0.031)			
MF22154	M03F14	1.220 (0.115)	0.618 (0.056)	0.136 (0.022)		
MF22156	M03F15	1.644 (0.091)	0.551 (0.029)			
MF22002	M04F01	1.634 (0.148)	1.359 (0.053)	0.111 (0.012)		
MF22004	M04F02	1.104 (0.131)	0.707 (0.075)	0.235 (0.029)		
MF22005	M04F03	0.930 (0.145)	1.271 (0.096)	0.241 (0.027)		
MF22008	M04F04	0.770 (0.054)	1.115 (0.068)			
MF22010	M04F05	0.953 (0.085)	0.216 (0.074)	0.139 (0.030)		
MF22012	M04F06	0.796 (0.050)	-0.003 (0.046)			
MF22021	M04F07	1.598 (0.158)	0.775 (0.045)	0.157 (0.019)		
MF22016	M04F08	0.808 (0.111)	1.250 (0.093)	0.132 (0.026)		
MF22252	M04F09	1.120 (0.109)	0.330 (0.070)	0.189 (0.030)		
MF22261A	M04F10A	1.519 (0.086)	0.711 (0.032)			
MF22261B	M04F10B	1.847 (0.117)	1.076 (0.034)			
MF22261C	M04F10C	1.094 (0.053)	1.310 (0.035)		-0.914 (0.088)	0.914 (0.098)
MF22227A	M04F11A	1.326 (0.073)	0.329 (0.032)			
MF22227B	M04F11B	1.832 (0.107)	0.808 (0.029)			
MF22227C	M04F11C	1.739 (0.109)	1.054 (0.035)			
MF22127	M04F12	1.626 (0.148)	1.449 (0.057)	0.117 (0.012)		
MF22043	M05F01	0.754 (0.062)	-0.339 (0.096)	0.115 (0.035)		
MF22046	M05F02	0.815 (0.049)	-0.384 (0.049)			
MF22049	M05F03	0.637 (0.073)	0.044 (0.149)	0.200 (0.047)		
MF22050	M05F04	0.921 (0.104)	0.976 (0.072)	0.111 (0.022)		
MF22055	M05F05	1.231 (0.070)	0.575 (0.036)			
MF22057	M05F06	0.583 (0.060)	-0.190 (0.148)	0.152 (0.046)		
MF22257	M05F07	1.536 (0.167)	0.644 (0.054)	0.239 (0.022)		
MF22062	M05F08	1.128 (0.109)	0.757 (0.057)	0.114 (0.020)		
MF22066	M05F09	1.452 (0.121)	0.317 (0.047)	0.120 (0.021)		

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MF22232	M05F10	0.585 (0.030)	1.652 (0.068)		-2.112 (0.171)	2.112 (0.190)
MF22234A	M05F11A	0.932 (0.040)	0.865 (0.032)		-0.467 (0.061)	0.467 (0.070)
MF22234B	M05F11B	1.053 (0.049)	1.134 (0.032)		-1.130 (0.101)	1.130 (0.107)
MF22243	M05F12	1.243 (0.071)	0.570 (0.036)			
MF22097	M06F01	1.243 (0.103)	-0.174 (0.064)	0.158 (0.031)		
MF22101	M06F02	0.887 (0.076)	-0.368 (0.094)	0.159 (0.038)		
MF22104	M06F03	0.911 (0.075)	-0.339 (0.086)	0.145 (0.036)		
MF22105	M06F04	0.718 (0.081)	0.808 (0.093)	0.113 (0.029)		
MF22106	M06F05	0.967 (0.059)	0.756 (0.048)			
MF22108	M06F06	0.836 (0.084)	0.130 (0.097)	0.157 (0.037)		
MF22110	M06F07	0.475 (0.039)	0.320 (0.074)			
MF22181	M06F08	1.323 (0.108)	-0.433 (0.066)	0.168 (0.034)		
MF32307	M06F09	1.403 (0.086)	0.987 (0.041)			
MF32523	M06F10	1.634 (0.171)	1.163 (0.052)	0.163 (0.015)		
MF32701	M06F11	1.165 (0.094)	-0.544 (0.076)	0.160 (0.036)		
MF32704	M06F12	1.137 (0.104)	0.182 (0.067)	0.161 (0.029)		
MF32525	M06F13	0.980 (0.090)	0.362 (0.070)	0.130 (0.028)		
MF32579	M06F14	1.114 (0.099)	0.136 (0.067)	0.155 (0.029)		
MF32691	M06F15	0.993 (0.058)	0.511 (0.042)			
MF32142	M07F01	1.971 (0.221)	1.028 (0.056)	0.370 (0.017)		
MF32198	M07F02	1.000 (0.108)	0.571 (0.076)	0.166 (0.029)		
MF32640	M07F03	0.655 (0.033)	1.374 (0.058)		-0.651 (0.085)	0.651 (0.110)
MF32344	M07F04	1.336 (0.076)	0.611 (0.035)			
MF32754	M07F05	0.818 (0.049)	-0.170 (0.046)			
MF32755	M07F06	1.036 (0.054)	1.397 (0.044)		-0.295 (0.062)	0.295 (0.083)
MF32753A	M07F07A	0.796 (0.037)	1.147 (0.043)		-0.560 (0.073)	0.560 (0.089)
MF32753B	M07F07B	0.870 (0.044)	1.304 (0.046)		-0.302 (0.064)	0.302 (0.086)
MF32753C	M07F07C	0.763 (0.052)	1.003 (0.067)			
MF32756	M07F08	0.608 (0.045)	0.914 (0.077)			
MF32205	M07F09	0.686 (0.106)	0.877 (0.133)	0.234 (0.040)		
MF32163	M07F10	1.469 (0.183)	1.014 (0.059)	0.215 (0.019)		
MF32381	M08F01	1.178 (0.065)	0.348 (0.036)			
MF32416	M08F02	1.209 (0.122)	0.830 (0.057)	0.112 (0.019)		
MF32160	M08F03	2.064 (0.169)	1.135 (0.042)	0.127 (0.011)		
MF32273	M08F04	1.161 (0.112)	-0.107 (0.079)	0.235 (0.034)		
MF32540	M08F05	0.821 (0.010)	0.231 (0.121)	0.246 (0.042)		
MF32698	M08F06	1.266 (0.122)	0.620 (0.055)	0.133 (0.021)		
MF32097	M08F07	1.312 (0.175)	1.413 (0.075)	0.183 (0.016)		
MF32575	M08F08	1.966 (0.195)	0.545 (0.042)	0.199 (0.018)		
MF32414	M08F09	1.237 (0.072)	0.676 (0.039)			
MF32294	M08F10	1.078 (0.010)	0.182 (0.070)	0.161 (0.028)		
MF32688	M08F11	0.995 (0.062)	0.920 (0.052)			
MF32529	M08F12	1.709 (0.183)	1.059 (0.050)	0.180 (0.015)		
MF32637A	M08F13A	0.971 (0.055)	0.216 (0.041)			

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MF32637B	M08F13B	1.107 (0.061)	0.145 (0.037)			
MF32637C	M08F13C	1.132 (0.065)	0.566 (0.040)			
MF32094	M11F01	1.379 (0.139)	0.262 (0.061)	0.238 (0.027)		
MF32662	M11F02	1.808 (0.158)	1.361 (0.051)	0.090 (0.010)		
MF32064	M11F03	1.623 (0.091)	0.509 (0.030)			
MF32419	M11F04	1.569 (0.188)	0.915 (0.054)	0.223 (0.019)		
MF32477	M11F05	1.675 (0.172)	0.719 (0.046)	0.171 (0.019)		
MF32538	M11F06	1.320 (0.075)	0.419 (0.033)			
MF32324	M11F07	1.361 (0.146)	1.011 (0.054)	0.124 (0.016)		
MF32116	M11F08	1.247 (0.143)	0.743 (0.064)	0.221 (0.023)		
MF32100	M11F09	1.126 (0.108)	0.582 (0.059)	0.136 (0.022)		
MF32402	M11F10	0.962 (0.127)	0.890 (0.084)	0.207 (0.028)		
MF32734	M11F11	0.974 (0.056)	0.166 (0.040)			
MF32397	M11F12	1.446 (0.174)	0.964 (0.057)	0.199 (0.019)		
MF32695	M11F13	0.557 (0.022)	0.354 (0.038)		-1.354 (0.103)	1.354 (0.107)
MF32132	M11F14	0.913 (0.100)	0.818 (0.074)	0.125 (0.024)		
MF32352	M12F01	1.432 (0.181)	0.572 (0.069)	0.355 (0.025)		
MF32725	M12F02	1.322 (0.078)	0.780 (0.039)			
MF32683	M12F03	0.694 (0.029)	0.732 (0.037)		-1.002 (0.087)	1.002 (0.095)
MF32738	M12F04	1.137 (0.099)	-0.168 (0.073)	0.166 (0.034)		
MF32295	M12F05	1.171 (0.103)	-0.339 (0.079)	0.185 (0.038)		
MF32331	M12F06	1.984 (0.180)	1.407 (0.057)	0.199 (0.013)		
MF32623	M12F07	1.757 (0.163)	0.619 (0.040)	0.136 (0.017)		
MF32679	M12F08	1.096 (0.119)	0.461 (0.072)	0.200 (0.029)		
MF32047	M12F09	1.797 (0.213)	1.243 (0.068)	0.378 (0.017)		
MF32398	M12F10	1.773 (0.203)	0.978 (0.052)	0.253 (0.018)		
MF32507	M12F11	1.998 (0.179)	1.151 (0.046)	0.167 (0.013)		
MF32424	M12F12	1.099 (0.126)	0.796 (0.067)	0.172 (0.024)		
MF32681A	M12F13A	0.555 (0.041)	-0.078 (0.062)			
MF32681B	M12F13B	0.627 (0.047)	0.988 (0.080)			
MF32681C	M12F13C	0.992 (0.060)	0.629 (0.045)			
MF32609	M13F01	1.012 (0.043)	-0.296 (0.041)	0.103 (0.019)		
MF32690	M13F02	1.157 (0.080)	1.131 (0.037)	0.195 (0.011)		
MF32727	M13F03	1.696 (0.082)	0.649 (0.022)	0.152 (0.009)		
MF32743	M13F04	0.607 (0.022)	0.283 (0.030)			
MF32744	M13F05	0.874 (0.029)	0.955 (0.029)			
MF32745	M13F06	0.564 (0.019)	2.402 (0.066)		-1.468 (0.083)	1.468 (0.113)
MF32233	M13F07	1.056 (0.027)	1.383 (0.021)		-0.519 (0.036)	0.519 (0.044)
MF32670	M13F08	0.728 (0.035)	-0.592 (0.073)	0.127 (0.029)		
MF32447	M13F09	1.422 (0.076)	0.901 (0.027)	0.139 (0.009)		
MF32036	M13F10	1.236 (0.076)	0.797 (0.035)	0.215 (0.013)		
MF32728	M13F11	1.616 (0.089)	1.248 (0.030)	0.219 (0.008)		
MF32732	M13F12	0.814 (0.063)	0.770 (0.060)	0.262 (0.020)		
MF32166	M14F01	1.024 (0.052)	0.201 (0.043)	0.177 (0.019)		

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MF32721	M14F02	0.709 (0.068)	1.483 (0.068)	0.206 (0.017)		
MF32757	M14F03	0.457 (0.009)	0.175 (0.022)		-2.463 (0.079)	2.463 (0.080)
MF32760A	M14F04A	0.864 (0.018)	0.973 (0.017)		-1.289 (0.052)	1.289 (0.056)
MF32760B	M14F04B	1.481 (0.051)	1.352 (0.025)			
MF32760C	M14F04C	1.742 (0.064)	1.514 (0.024)			
MF32761	M14F05	1.285 (0.036)	1.472 (0.019)		-0.214 (0.029)	0.214 (0.037)
MF32692	M14F06	0.646 (0.014)	1.239 (0.025)		-1.247 (0.054)	1.247 (0.061)
MF32626	M14F07	1.023 (0.059)	0.670 (0.040)	0.186 (0.015)		
MF32595	M14F08	1.327 (0.059)	0.474 (0.026)	0.108 (0.011)		
MF32673	M14F09	1.461 (0.075)	0.606 (0.027)	0.185 (0.012)		
MC22046	M05_02	0.818 (0.051)	-0.612 (0.052)			
MC22110	M06_07	0.528 (0.041)	-1.031 (0.089)			
MC32701	M06_11	1.219 (0.101)	-1.118 (0.087)	0.177 (0.044)		
MC32704	M06_12	1.184 (0.108)	-0.215 (0.075)	0.196 (0.034)		
MC32525	M06_13	1.008 (0.098)	0.094 (0.082)	0.187 (0.034)		
MZ22043	M01_01	0.652 (0.046)	-0.523 (0.105)	0.125 (0.037)		
MZ22046	M01_02	0.803 (0.037)	-0.555 (0.040)			
MZ22049	M01_03	0.619 (0.069)	0.174 (0.150)	0.252 (0.045)		
MZ22050	M01_04	0.961 (0.091)	1.015 (0.058)	0.143 (0.019)		
MZ22055	M01_05	1.236 (0.054)	0.538 (0.028)			
MZ22057	M01_06	0.452 (0.044)	-0.321 (0.191)	0.165 (0.051)		
MZ22257	M01_07	1.426 (0.117)	0.501 (0.046)	0.241 (0.019)		
MZ22062	M01_08	1.011 (0.081)	0.710 (0.052)	0.130 (0.019)		
MZ22066	M01_09	1.441 (0.093)	0.200 (0.037)	0.121 (0.017)		
MZ22232	M01_10	0.546 (0.021)	1.625 (0.053)		-2.323 (0.141)	2.323 (0.154)
MZ22234A	M01_11A	0.837 (0.029)	0.802 (0.027)		-0.429 (0.050)	0.429 (0.057)
MZ22234B	M01_11B	0.931 (0.034)	1.177 (0.029)		-1.381 (0.094)	1.381 (0.100)
MZ22243	M01_12	1.010 (0.050)	0.427 (0.031)			
MZ42003	M02_01	0.681 (0.051)	-0.352 (0.108)	0.161 (0.038)		
MZ42079	M02_02	0.908 (0.076)	-0.449 (0.106)	0.265 (0.042)		
MZ42018	M02_03	0.987 (0.044)	0.376 (0.032)			
MZ42055	M02_04	1.228 (0.121)	0.803 (0.058)	0.293 (0.020)		
MZ42039	M02_05	0.693 (0.049)	0.360 (0.070)	0.097 (0.023)		
MZ42199	M02_06	1.184 (0.087)	-0.181 (0.063)	0.224 (0.029)		
MZ42301A	M02_07A	0.522 (0.030)	-0.087 (0.051)			
MZ42301B	M02_07B	0.897 (0.042)	0.655 (0.037)			
MZ42301C	M02_07C	1.560 (0.073)	1.020 (0.030)			
MZ42263	M02_08	1.313 (0.061)	0.973 (0.033)			
MZ42265	M02_09	0.823 (0.067)	0.371 (0.074)	0.149 (0.027)		
MZ42137	M02_10	1.151 (0.102)	0.682 (0.056)	0.226 (0.020)		
MZ42148	M02_11	0.866 (0.075)	-0.195 (0.102)	0.255 (0.039)		
MZ42254	M02_12	0.659 (0.047)	-1.362 (0.133)	0.171 (0.047)		
MZ42250	M02_13	1.113 (0.050)	-0.726 (0.034)			
MZ42220	M02_14	0.691 (0.021)	0.441 (0.026)		-1.521 (0.084)	1.521 (0.086)

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MZ22097	M03_01	1.106 (0.068)	-0.142 (0.050)	0.113 (0.023)		
MZ22101	M03_02	0.881 (0.060)	-0.379 (0.077)	0.148 (0.032)		
MZ22104	M03_03	0.953 (0.062)	-0.348 (0.067)	0.138 (0.029)		
MZ22105	M03_04	0.630 (0.059)	0.754 (0.090)	0.117 (0.028)		
MZ22106	M03_05	0.979 (0.045)	0.657 (0.035)			
MZ22108	M03_06	0.909 (0.080)	0.185 (0.082)	0.224 (0.032)		
MZ22110	M03_07	0.593 (0.033)	-0.695 (0.054)			
MZ22181	M03_08	1.069 (0.077)	-0.762 (0.085)	0.213 (0.040)		
MZ32307	M03_09	1.378 (0.063)	0.900 (0.031)			
MZ32523	M03_10	1.922 (0.139)	1.076 (0.036)	0.177 (0.011)		
MZ32701	M03_11	1.211 (0.082)	-1.098 (0.077)	0.185 (0.041)		
MZ32704	M03_12	1.190 (0.084)	-0.235 (0.059)	0.187 (0.029)		
MZ32525	M03_13	0.993 (0.073)	0.080 (0.064)	0.166 (0.027)		
MZ32579	M03_14	1.140 (0.083)	-0.205 (0.064)	0.199 (0.030)		
MZ32691	M03_15	0.873 (0.041)	0.327 (0.034)			
MZ42001	M04_01	0.741 (0.059)	-0.500 (0.112)	0.194 (0.042)		
MZ42022	M04_02	0.915 (0.097)	0.773 (0.075)	0.261 (0.024)		
MZ42082	M04_03	1.290 (0.108)	0.788 (0.045)	0.171 (0.016)		
MZ42088	M04_04	1.115 (0.090)	0.274 (0.057)	0.206 (0.024)		
MZ42304A	M04_05A	1.403 (0.059)	-0.473 (0.025)			
MZ42304B	M04_05B	0.966 (0.038)	0.670 (0.025)		0.412 (0.033)	-0.412 (0.044)
MZ42304C	M04_05C	1.414 (0.064)	0.723 (0.028)			
MZ42304D	M04_05D	0.580 (0.017)	0.233 (0.028)		-1.411 (0.080)	1.411 (0.082)
MZ42267	M04_06	1.310 (0.099)	0.664 (0.041)	0.142 (0.015)		
MZ42239	M04_07	1.243 (0.010)	0.967 (0.044)	0.109 (0.013)		
MZ42238	M04_08	1.171 (0.120)	1.155 (0.057)	0.188 (0.016)		
MZ42279	M04_09	1.065 (0.069)	-0.017 (0.051)	0.118 (0.022)		
MZ42036	M04_10	1.470 (0.115)	0.741 (0.038)	0.141 (0.014)		
MZ42130	M04_11	0.528 (0.031)	-0.128 (0.050)			
MZ42303A	M04_12A	1.158 (0.050)	0.193 (0.027)			
MZ42303B	M04_12B	0.373 (0.019)	0.787 (0.055)		-0.050 (0.081)	0.050 (0.097)
MZ42222	M04_13	0.998 (0.074)	0.387 (0.053)	0.131 (0.021)		
MZ32142	M05_01	1.617 (0.178)	1.068 (0.055)	0.409 (0.015)		
MZ32198	M05_02	0.770 (0.073)	0.461 (0.086)	0.183 (0.031)		
MZ32640	M05_03	0.607 (0.023)	1.394 (0.048)		-0.752 (0.070)	0.752 (0.090)
MZ32344	M05_04	1.184 (0.052)	0.518 (0.029)			
MZ32754	M05_05	0.828 (0.039)	-0.440 (0.037)			
MZ32755	M05_06	1.081 (0.043)	1.214 (0.030)		-0.185 (0.042)	0.185 (0.056)
MZ32753A	M05_07A	1.033 (0.036)	0.779 (0.023)		-0.246 (0.040)	0.246 (0.047)
MZ32753B	M05_07B	1.157 (0.044)	0.922 (0.024)		-0.032 (0.035)	0.032 (0.044)
MZ32753C	M05_07C	1.024 (0.047)	0.518 (0.033)			
MZ32756	M05_08	0.700 (0.036)	0.441 (0.043)			
MZ32205	M05_09	0.570 (0.061)	0.181 (0.150)	0.207 (0.045)		
MZ32163	M05_10	1.518 (0.135)	0.692 (0.044)	0.238 (0.017)		

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MZ42032	M06_01	0.762 (0.060)	-0.560 (0.119)	0.205 (0.045)		
MZ42031	M06_02	1.406 (0.113)	0.582 (0.044)	0.215 (0.018)		
MZ42186	M06_03	1.037 (0.045)	0.336 (0.030)			
MZ42059	M06_04	0.772 (0.026)	0.081 (0.024)		-0.179 (0.047)	0.179 (0.047)
MZ42236	M06_05	1.240 (0.098)	0.270 (0.055)	0.239 (0.024)		
MZ42226	M06_06	1.211 (0.051)	0.260 (0.026)			
MZ42103	M06_07	0.753 (0.042)	1.376 (0.063)			
MZ42086	M06_08	1.134 (0.050)	0.631 (0.030)			
MZ42228	M06_09	0.770 (0.038)	0.664 (0.042)			
MZ42245	M06_10	1.580 (0.124)	1.212 (0.040)	0.138 (0.011)		
MZ42270	M06_11	0.909 (0.041)	-0.045 (0.032)			
MZ42201	M06_12	1.211 (0.051)	0.098 (0.026)			
MZ42152	M06_13	0.752 (0.079)	0.840 (0.085)	0.187 (0.028)		
MZ42269	M06_14	0.800 (0.079)	0.218 (0.106)	0.272 (0.037)		
MZ42179	M06_15	0.925 (0.079)	0.396 (0.071)	0.197 (0.027)		
MZ42177	M06_16	1.160 (0.087)	0.212 (0.056)	0.203 (0.024)		
MZ42207	M06_17	0.436 (0.013)	-0.024 (0.034)		-2.991 (0.145)	2.991 (0.144)
MZ32381	M07_01	1.098 (0.047)	0.260 (0.028)			
MZ32416	M07_02	1.370 (0.102)	0.855 (0.038)	0.111 (0.013)		
MZ32160	M07_03	1.884 (0.119)	1.258 (0.035)	0.126 (0.001)		
MZ32273	M07_04	0.955 (0.076)	-0.251 (0.088)	0.239 (0.036)		
MZ32540	M07_05	1.058 (0.101)	0.311 (0.077)	0.330 (0.028)		
MZ32698	M07_06	1.025 (0.079)	0.566 (0.053)	0.144 (0.020)		
MZ32097	M07_07	1.273 (0.133)	1.361 (0.055)	0.189 (0.014)		
MZ32575	M07_08	1.890 (0.145)	0.574 (0.033)	0.208 (0.015)		
MZ32414	M07_09	1.083 (0.048)	0.602 (0.031)			
MZ32294	M07_10	0.909 (0.076)	0.008 (0.085)	0.218 (0.034)		
MZ32688	M07_11	0.815 (0.040)	0.764 (0.042)			
MZ32529	M07_12	1.529 (0.131)	1.010 (0.040)	0.166 (0.013)		
MZ32637A	M07_13A	0.827 (0.039)	-0.408 (0.038)			
MZ32637B	M07_13B	1.070 (0.050)	-0.432 (0.034)			
MZ32637C	M07_13C	1.039 (0.048)	0.177 (0.031)			
MZ42183	M08_01	0.674 (0.051)	-0.233 (0.103)	0.155 (0.036)		
MZ42060	M08_02	1.218 (0.087)	0.188 (0.051)	0.192 (0.022)		
MZ42019	M08_03	0.734 (0.036)	0.451 (0.040)			
MZ42023	M08_04	1.179 (0.051)	0.453 (0.028)			
MZ42197	M08_05	1.132 (0.052)	0.818 (0.033)			
MZ42234	M08_06	1.419 (0.100)	0.319 (0.041)	0.184 (0.019)		
MZ42066	M08_07	0.678 (0.034)	0.274 (0.041)			
MZ42243	M08_08	1.828 (0.110)	0.394 (0.027)	0.091 (0.012)		
MZ42248	M08_09	1.377 (0.060)	0.686 (0.027)			
MZ42229A	M08_10A	1.666 (0.073)	0.700 (0.024)			
MZ42229B	M08_10B	1.862 (0.083)	0.731 (0.022)			
MZ42080A	M08_11A	0.819 (0.039)	0.467 (0.037)			

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MZ42080B	M08_11B	1.175 (0.061)	1.330 (0.044)			
MZ42120	M08_12	0.947 (0.078)	-0.109 (0.087)	0.243 (0.035)		
MZ42203	M08_13	1.417 (0.094)	0.156 (0.041)	0.160 (0.019)		
MZ42264	M08_14	0.737 (0.042)	1.384 (0.065)			
MZ42255	M08_15	0.639 (0.051)	-0.182 (0.112)	0.156 (0.038)		
MZ42224	M08_16	0.884 (0.040)	-0.080 (0.034)			
MZ32094	M09_01	1.174 (0.089)	0.021 (0.061)	0.226 (0.027)		
MZ32662	M09_02	1.836 (0.115)	1.319 (0.037)	0.122 (0.009)		
MZ32064	M09_03	1.302 (0.057)	0.704 (0.028)			
MZ32419	M09_04	1.363 (0.122)	0.805 (0.047)	0.239 (0.017)		
MZ32477	M09_05	1.713 (0.137)	0.628 (0.037)	0.230 (0.016)		
MZ32538	M09_06	1.231 (0.052)	0.244 (0.026)			
MZ32324	M09_07	1.179 (0.095)	0.772 (0.047)	0.170 (0.017)		
MZ32116	M09_08	1.079 (0.010)	0.771 (0.059)	0.238 (0.021)		
MZ32100	M09_09	0.955 (0.070)	0.404 (0.057)	0.132 (0.022)		
MZ32402	M09_10	0.784 (0.091)	0.777 (0.095)	0.266 (0.030)		
MZ32734	M09_11	0.773 (0.036)	-0.339 (0.039)			
MZ32397	M09_12	1.143 (0.106)	0.858 (0.055)	0.223 (0.019)		
MZ32695	M09_13	0.540 (0.017)	-0.149 (0.031)		-1.068 (0.078)	1.068 (0.076)
MZ32132	M09_14	0.675 (0.057)	0.430 (0.086)	0.130 (0.029)		
MZ42041	M10_01	1.160 (0.094)	-0.342 (0.080)	0.302 (0.035)		
MZ42024	M10_02	1.483 (0.097)	0.118 (0.039)	0.159 (0.019)		
MZ42016	M10_03	0.847 (0.074)	0.506 (0.073)	0.175 (0.027)		
MZ42002	M10_04	0.716 (0.038)	0.930 (0.051)			
MZ42198A	M10_05A	1.027 (0.047)	-0.863 (0.039)			
MZ42198B	M10_05B	0.983 (0.044)	0.338 (0.031)			
MZ42198C	M10_05C	1.517 (0.074)	1.096 (0.031)			
MZ42077	M10_06	1.308 (0.105)	0.361 (0.051)	0.238 (0.022)		
MZ42235	M10_07	1.559 (0.095)	0.174 (0.033)	0.112 (0.016)		
MZ42067	M10_08	1.484 (0.149)	1.112 (0.049)	0.270 (0.015)		
MZ42150	M10_09	0.818 (0.084)	0.930 (0.075)	0.190 (0.024)		
MZ42300A	M10_10A	1.216 (0.051)	0.084 (0.026)			
MZ42300B	M10_10B	1.266 (0.054)	0.316 (0.026)			
MZ42260	M10_11	0.847 (0.078)	0.046 (0.101)	0.269 (0.036)		
MZ42169A	M10_12A	1.033 (0.045)	0.277 (0.030)			
MZ42169B	M10_12B	0.352 (0.031)	1.897 (0.163)			
MZ42169C	M10_12C	0.759 (0.052)	1.929 (0.096)			
MZ32352	M11_01	1.376 (0.129)	0.384 (0.059)	0.379 (0.022)		
MZ32725	M11_02	1.226 (0.057)	0.830 (0.032)			
MZ32683	M11_03	0.533 (0.017)	0.790 (0.035)		-1.370 (0.083)	1.370 (0.090)
MZ32738	M11_04	1.210 (0.085)	-0.259 (0.060)	0.192 (0.029)		
MZ32295	M11_05	1.189 (0.086)	-0.631 (0.075)	0.222 (0.037)		
MZ32331	M11_06	1.957 (0.132)	1.209 (0.037)	0.188 (0.010)		
MZ32623	M11_07	1.676 (0.118)	0.627 (0.032)	0.132 (0.013)		

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MZ32679	M11_08	1.065 (0.084)	0.277 (0.060)	0.198 (0.025)		
MZ32047	M11_09	1.495 (0.180)	1.058 (0.061)	0.432 (0.016)		
MZ32398	M11_10	1.360 (0.132)	0.815 (0.051)	0.285 (0.018)		
MZ32507	M11_11	1.647 (0.137)	1.120 (0.041)	0.187 (0.012)		
MZ32424	M11_12	1.149 (0.082)	0.363 (0.047)	0.144 (0.020)		
MZ32681A	M11_13A	0.558 (0.032)	-0.595 (0.056)			
MZ32681B	M11_13B	0.589 (0.034)	0.794 (0.057)			
MZ32681C	M11_13C	1.101 (0.050)	0.481 (0.030)			
MZ42015	M12_01	0.923 (0.069)	-0.437 (0.088)	0.195 (0.038)		
MZ42196	M12_02	1.035 (0.067)	0.132 (0.049)	0.109 (0.021)		
MZ42194	M12_03	1.150 (0.050)	-0.446 (0.030)			
MZ42114A	M12_04A	1.388 (0.058)	-0.153 (0.024)			
MZ42114B	M12_04B	1.403 (0.059)	0.135 (0.023)			
MZ42112	M12_05	0.430 (0.064)	0.678 (0.230)	0.248 (0.054)		
MZ42109	M12_06	1.318 (0.130)	1.039 (0.051)	0.228 (0.016)		
MZ42050	M12_07	1.024 (0.047)	0.637 (0.033)			
MZ42074A	M12_08A	1.153 (0.051)	0.571 (0.030)			
MZ42074B	M12_08B	1.098 (0.050)	0.741 (0.033)			
MZ42074C	M12_08C	1.744 (0.084)	0.991 (0.027)			
MZ42151	M12_09	0.895 (0.041)	-0.077 (0.033)			
MZ42132	M12_10	1.392 (0.137)	1.195 (0.049)	0.192 (0.014)		
MZ42257	M12_11	0.712 (0.068)	0.782 (0.081)	0.145 (0.026)		
MZ42158	M12_12	0.737 (0.084)	0.332 (0.123)	0.307 (0.039)		
MZ42252	M12_13	1.120 (0.090)	0.746 (0.048)	0.149 (0.018)		
MZ42261	M12_14	0.694 (0.058)	0.092 (0.098)	0.159 (0.034)		
MZ32166	M13_01	0.996 (0.072)	-0.048 (0.065)	0.172 (0.028)		
MZ32721	M13_02	0.807 (0.102)	1.209 (0.086)	0.239 (0.023)		
MZ32757	M13_03	0.442 (0.014)	-0.435 (0.036)		-2.315 (0.119)	2.315 (0.115)
MZ32760A	M13_04A	0.898 (0.027)	0.592 (0.022)		-1.184 (0.069)	1.184 (0.072)
MZ32760B	M13_04B	1.564 (0.074)	0.935 (0.029)			
MZ32760C	M13_04C	1.752 (0.090)	1.162 (0.031)			
MZ32761	M13_05	1.203 (0.046)	1.178 (0.026)		-0.173 (0.039)	0.173 (0.050)
MZ32692	M13_06	0.601 (0.020)	0.974 (0.035)		-1.359 (0.082)	1.359 (0.091)
MZ32626	M13_07	0.826 (0.079)	0.393 (0.085)	0.209 (0.031)		
MZ32595	M13_08	1.398 (0.092)	0.181 (0.039)	0.124 (0.018)		
MZ32673	M13_09	1.427 (0.106)	0.354 (0.042)	0.197 (0.019)		
MZ42182	M14_01	1.214 (0.104)	0.098 (0.065)	0.309 (0.027)		
MZ42081	M14_02	1.017 (0.047)	0.705 (0.035)			
MZ42049	M14_03	0.868 (0.074)	-0.216 (0.010)	0.242 (0.038)		
MZ42052	M14_04	1.594 (0.105)	-0.046 (0.039)	0.166 (0.020)		
MZ42076	M14_05	1.071 (0.086)	0.368 (0.057)	0.186 (0.023)		
MZ42302A	M14_06A	0.847 (0.028)	0.359 (0.023)		-0.302 (0.044)	0.302 (0.047)
MZ42302B	M14_06B	0.840 (0.026)	0.480 (0.023)		-0.728 (0.053)	0.728 (0.057)
MZ42302C	M14_06C	0.529 (0.021)	1.426 (0.054)		-0.810 (0.076)	0.810 (0.097)

( ) Standard errors appear in parentheses

Exhibit D.3 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Mathematics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
MZ42100	M14_07	1.253 (0.102)	0.191 (0.056)	0.265 (0.024)		
MZ42202	M14_08	1.403 (0.117)	0.485 (0.046)	0.259 (0.019)		
MZ42240	M14_09	1.322 (0.088)	0.150 (0.042)	0.147 (0.019)		
MZ42093	M14_10	1.463 (0.073)	1.129 (0.034)			
MZ42271	M14_11	1.092 (0.077)	0.272 (0.049)	0.135 (0.021)		
MZ42268	M14_12	1.364 (0.119)	1.056 (0.043)	0.140 (0.013)		
MZ42159	M14_13	0.559 (0.032)	-0.808 (0.060)			
MZ42164	M14_14	1.213 (0.054)	0.563 (0.029)			
MZ42167	M14_15	1.210 (0.058)	0.771 (0.032)			

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S012025	S01_01	0.474 (0.077)	1.698 (0.131)	0.333 (0.028)		
S012026	S01_02	0.681 (0.050)	-0.438 (0.129)	0.400 (0.036)		
S012027	S01_03	0.759 (0.035)	-0.793 (0.076)	0.148 (0.031)		
S012028	S01_04	0.766 (0.039)	0.394 (0.044)	0.121 (0.017)		
S012029	S01_05	0.553 (0.061)	0.774 (0.116)	0.366 (0.029)		
S012030	S01_06	0.572 (0.063)	1.254 (0.085)	0.262 (0.024)		
S022035	S01_07	0.372 (0.017)	0.296 (0.042)			
S022225	S01_08	1.112 (0.090)	1.800 (0.059)	0.130 (0.007)		
S022117	S01_09	0.682 (0.052)	0.869 (0.059)	0.214 (0.019)		
S022235	S01_10	0.687 (0.057)	0.463 (0.086)	0.324 (0.026)		
S022188	S01_11	0.987 (0.100)	1.207 (0.054)	0.413 (0.014)		
S022074	S01_12	1.117 (0.065)	0.513 (0.038)	0.279 (0.015)		
S022240	S01_13	1.163 (0.099)	1.544 (0.051)	0.279 (0.009)		
S022206	S01_14	0.728 (0.064)	1.071 (0.057)	0.235 (0.019)		
S022160	S01_15	0.585 (0.021)	0.795 (0.034)			
S022058	S01_16	0.895 (0.064)	0.358 (0.064)	0.380 (0.021)		
S012013	S02_01	0.674 (0.104)	1.785 (0.133)	0.164 (0.021)		
S012014	S02_02	0.834 (0.059)	-0.704 (0.096)	0.182 (0.038)		
S012015	S02_03	0.718 (0.066)	-0.119 (0.115)	0.262 (0.038)		
S012016	S02_04	0.498 (0.051)	-0.717 (0.221)	0.263 (0.056)		
S012017	S02_05	1.336 (0.103)	0.548 (0.041)	0.200 (0.018)		
S012018	S02_06	0.358 (0.045)	-0.058 (0.258)	0.248 (0.052)		
S012001	S02_07	0.540 (0.046)	-0.104 (0.121)	0.139 (0.036)		
S012002	S02_08	0.580 (0.058)	-0.123 (0.148)	0.242 (0.043)		
S012003	S02_09	0.993 (0.070)	-0.429 (0.076)	0.236 (0.032)		
S012004	S02_10	0.579 (0.056)	-0.303 (0.156)	0.240 (0.046)		
S012005	S02_11	0.660 (0.075)	0.339 (0.121)	0.282 (0.036)		
S012006	S02_12	0.817 (0.074)	0.415 (0.075)	0.214 (0.027)		
S032131	S02_13	0.863 (0.038)	-0.112 (0.032)			
S032202	S02_14	0.598 (0.024)	-0.041 (0.028)		0.216 (0.052)	-0.216 (0.050)
S012037	S03_01	0.548 (0.043)	-1.802 (0.198)	0.223 (0.056)		
S012038	S03_02	0.965 (0.097)	0.348 (0.082)	0.356 (0.028)		
S012039	S03_03	0.741 (0.069)	-0.450 (0.137)	0.322 (0.044)		
S012040	S03_04	1.472 (0.128)	0.599 (0.043)	0.300 (0.018)		
S012041	S03_05	0.642 (0.075)	0.416 (0.124)	0.282 (0.037)		
S012042	S03_06	0.825 (0.101)	0.733 (0.088)	0.326 (0.028)		
S022086	S03_07	0.995 (0.043)	0.324 (0.028)			
S022198	S03_08	1.181 (0.150)	1.299 (0.062)	0.252 (0.016)		
S022275	S03_09	1.122 (0.123)	1.317 (0.058)	0.159 (0.014)		
S022041	S03_10	0.740 (0.055)	-0.620 (0.107)	0.197 (0.038)		
S022283	S03_11	0.890 (0.039)	-0.594 (0.037)			
S022202	S03_12	0.810 (0.116)	1.092 (0.090)	0.333 (0.025)		
S022152	S03_13	0.953 (0.042)	0.253 (0.029)			
S022154	S03_14	0.627 (0.032)	-0.134 (0.041)			

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S022187	S04_01	0.577 (0.073)	1.081 (0.104)	0.175 (0.030)		
S022161	S04_02	0.585 (0.032)	0.552 (0.047)			
S022222	S04_03	1.076 (0.089)	0.617 (0.049)	0.184 (0.020)		
S022191	S04_04	0.570 (0.019)	-0.283 (0.029)		-0.411 (0.061)	0.411 (0.056)
S022279	S04_05	0.628 (0.032)	0.294 (0.041)			
S022040	S04_06	0.674 (0.050)	-0.127 (0.088)	0.146 (0.030)		
S022088A	S04_07A	0.721 (0.034)	-0.454 (0.041)			
S022088B	S04_07B	0.532 (0.030)	0.342 (0.048)			
S022249D	S04_08D	0.836 (0.039)	0.356 (0.033)			
S022286	S04_09	0.786 (0.051)	1.703 (0.083)			
S032595	S04_10	1.157 (0.134)	1.420 (0.063)	0.153 (0.014)		
S032656	S04_11	1.069 (0.073)	0.335 (0.046)	0.132 (0.020)		
S032625A	S04_12A	0.822 (0.038)	0.249 (0.033)			
S032625B	S04_12B	1.005 (0.046)	0.615 (0.030)			
S022183	S05_01	1.222 (0.075)	0.967 (0.030)	0.263 (0.011)		
S022276	S05_02	0.876 (0.056)	0.582 (0.049)	0.305 (0.017)		
S022115	S05_03	0.887 (0.043)	0.074 (0.047)	0.218 (0.019)		
S022022	S05_04	0.612 (0.018)	0.136 (0.024)			
S022019	S05_05	0.826 (0.042)	-0.326 (0.068)	0.283 (0.025)		
S022002	S05_06	0.968 (0.049)	0.455 (0.037)	0.220 (0.015)		
S022294	S05_07	0.786 (0.048)	0.055 (0.072)	0.345 (0.023)		
S022106	S05_08	0.624 (0.078)	1.858 (0.114)	0.118 (0.017)		
S022244	S05_09	1.087 (0.030)	1.002 (0.021)			
S022150	S05_10	0.842 (0.052)	0.666 (0.045)	0.245 (0.016)		
S022042	S05_11	1.108 (0.047)	0.361 (0.029)	0.169 (0.012)		
S022289	S05_12	0.737 (0.017)	0.910 (0.018)		0.807 (0.021)	-0.807 (0.032)
S022069	S05_13	0.851 (0.023)	0.496 (0.019)			
S022268	S05_14	0.614 (0.019)	0.730 (0.029)			
S022290	S06_01	1.004 (0.040)	0.327 (0.031)	0.225 (0.013)		
S022292	S06_02	0.599 (0.015)	0.556 (0.022)			
S022054	S06_03	0.985 (0.046)	0.627 (0.031)	0.255 (0.012)		
S022181	S06_04	1.066 (0.054)	0.896 (0.028)	0.264 (0.001)		
S022208	S06_05	1.119 (0.061)	1.072 (0.027)	0.265 (0.009)		
S022078	S06_06	1.086 (0.021)	0.111 (0.012)			
S022126	S06_07	0.528 (0.030)	0.344 (0.076)	0.166 (0.024)		
S022281	S06_08	0.554 (0.016)	1.165 (0.033)			
S032385	S06_09	0.783 (0.035)	0.030 (0.052)	0.255 (0.019)		
S032035	S06_10	1.181 (0.040)	0.377 (0.022)	0.167 (0.001)		
S032519	S06_11	0.695 (0.016)	0.458 (0.018)			
S032683	S06_12	0.996 (0.044)	0.767 (0.027)	0.198 (0.010)		
S032258	S06_13	0.831 (0.033)	-0.063 (0.044)	0.201 (0.017)		
S032120A	S06_14A	0.745 (0.021)	1.433 (0.032)			
S032120B	S06_14B	0.912 (0.022)	1.165 (0.022)			
S032606	S07_01	0.777 (0.077)	-0.873 (0.146)	0.227 (0.050)		

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S032015	S07_02	0.732 (0.052)	0.735 (0.059)			
S032310D	S07_03D	0.468 (0.026)	-0.064 (0.047)		-0.266 (0.095)	0.266 (0.091)
S032680	S07_04	0.598 (0.031)	-0.375 (0.042)		-0.066 (0.081)	0.066 (0.070)
S032672	S07_05	0.392 (0.062)	0.064 (0.282)	0.250 (0.059)		
S032392	S07_06	0.441 (0.051)	-1.584 (0.274)	0.227 (0.059)		
S032425	S07_07	0.828 (0.123)	0.816 (0.103)	0.256 (0.033)		
S032257	S07_08	1.174 (0.177)	1.079 (0.075)	0.239 (0.024)		
S032663	S07_09	0.471 (0.106)	1.703 (0.239)	0.228 (0.042)		
S032660	S07_10	0.822 (0.170)	1.641 (0.150)	0.216 (0.026)		
S032555	S07_11	1.049 (0.068)	0.848 (0.047)			
S032122	S07_12	0.591 (0.048)	1.004 (0.084)			
S032542	S08_01	1.271 (0.176)	0.859 (0.070)	0.293 (0.025)		
S032645	S08_02	0.959 (0.141)	0.884 (0.089)	0.265 (0.029)		
S032530D	S08_03D	0.478 (0.029)	0.525 (0.054)		0.895 (0.080)	-0.895 (0.095)
S032007	S08_04	0.728 (0.050)	0.398 (0.052)			
S032502	S08_05	1.028 (0.118)	0.754 (0.068)	0.164 (0.025)		
S032679	S08_06	0.809 (0.068)	1.524 (0.097)			
S032184	S08_07	0.463 (0.097)	1.394 (0.222)	0.248 (0.045)		
S032394	S08_08	0.893 (0.119)	0.591 (0.097)	0.262 (0.034)		
S032151	S08_09	1.197 (0.140)	0.817 (0.061)	0.186 (0.023)		
S032651A	S08_10A	1.243 (0.073)	0.380 (0.034)			
S032651B	S08_10B	0.967 (0.071)	1.149 (0.062)			
S032665A	S08_11A	0.954 (0.061)	0.661 (0.046)			
S032665B	S08_11B	0.958 (0.071)	1.235 (0.067)			
S032665C	S08_11C	0.825 (0.064)	1.294 (0.079)			
S032607	S09_01	0.856 (0.048)	0.055 (0.059)	0.249 (0.022)		
S032063	S09_02	0.677 (0.018)	1.319 (0.027)		-0.239 (0.032)	0.239 (0.044)
S032206	S09_03	1.128 (0.036)	1.175 (0.025)			
S032008	S09_04	0.946 (0.052)	0.160 (0.050)	0.246 (0.020)		
S032083	S09_05	0.789 (0.055)	1.242 (0.045)	0.115 (0.013)		
S032564	S09_06	1.668 (0.080)	1.238 (0.026)	0.186 (0.007)		
S032057	S09_07	1.211 (0.035)	0.941 (0.020)			
S032055	S09_08	1.021 (0.050)	-0.911 (0.069)	0.277 (0.031)		
S032626	S09_09	1.096 (0.029)	0.260 (0.016)			
S032281	S09_10	1.595 (0.070)	0.180 (0.025)	0.232 (0.013)		
S032150	S09_11	0.637 (0.039)	0.043 (0.081)	0.191 (0.027)		
S032301	S09_12	1.574 (0.087)	1.033 (0.025)	0.237 (0.009)		
S032446	S09_13	0.882 (0.059)	0.389 (0.060)	0.301 (0.022)		
S032637	S10_01	0.833 (0.059)	1.051 (0.044)	0.190 (0.015)		
S032386	S10_02	1.097 (0.085)	1.383 (0.040)	0.183 (0.010)		
S032682	S10_03	1.370 (0.083)	1.014 (0.027)	0.213 (0.001)		
S032652	S10_04	0.906 (0.055)	0.718 (0.040)	0.187 (0.015)		
S032437	S10_05	0.806 (0.074)	1.067 (0.058)	0.325 (0.017)		
S032242	S10_06	0.640 (0.024)	1.117 (0.039)			

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S032709	S10_07	1.232 (0.037)	1.003 (0.020)			
S032711	S10_08	0.797 (0.017)	1.126 (0.019)		-0.554 (0.032)	0.554 (0.039)
S032712A	S10_09A	0.820 (0.026)	0.752 (0.024)			
S032712B	S10_09B	1.185 (0.042)	1.413 (0.030)			
S032713A	S10_10A	1.063 (0.037)	1.391 (0.032)			
S032713B	S10_10B	0.896 (0.043)	2.091 (0.069)			
S032465	S11_01	0.864 (0.064)	0.220 (0.072)	0.342 (0.024)		
S032315	S11_02	1.156 (0.083)	0.791 (0.040)	0.297 (0.015)		
S032306	S11_03	0.423 (0.009)	0.648 (0.026)		-1.812 (0.067)	1.812 (0.071)
S032640	S11_04	0.522 (0.021)	-0.039 (0.034)			
S032579	S11_05	0.818 (0.101)	1.552 (0.077)	0.297 (0.016)		
S032570	S11_06	0.703 (0.026)	0.925 (0.035)			
S032024	S11_07	1.227 (0.113)	1.453 (0.052)	0.305 (0.011)		
S032272	S11_08	0.930 (0.039)	1.617 (0.048)			
S032141	S11_09	1.665 (0.092)	1.133 (0.027)	0.201 (0.009)		
S032060	S11_10	0.737 (0.024)	-0.452 (0.028)			
S032463	S11_11	1.628 (0.101)	0.587 (0.029)	0.325 (0.013)		
S032650D	S11_12D	0.517 (0.014)	0.357 (0.022)		-0.179 (0.041)	0.179 (0.044)
S032514	S11_13	0.850 (0.079)	1.074 (0.057)	0.288 (0.018)		
S032611	S12_01	1.118 (0.098)	1.315 (0.045)	0.233 (0.012)		
S032614	S12_02	0.701 (0.024)	0.230 (0.026)			
S032451	S12_03	0.574 (0.012)	0.154 (0.019)		-1.345 (0.053)	1.345 (0.053)
S032156	S12_04	1.206 (0.073)	0.743 (0.032)	0.193 (0.013)		
S032056	S12_05	0.785 (0.027)	0.606 (0.026)			
S032087	S12_06	0.654 (0.060)	1.018 (0.068)	0.198 (0.022)		
S032279	S12_07	0.771 (0.076)	1.429 (0.063)	0.185 (0.016)		
S032238	S12_08	1.121 (0.066)	0.641 (0.034)	0.183 (0.014)		
S032369	S12_09	0.531 (0.015)	0.818 (0.026)		-0.259 (0.040)	0.259 (0.048)
S032160	S12_10	0.831 (0.065)	0.423 (0.070)	0.335 (0.023)		
S032654	S12_11	0.936 (0.064)	0.739 (0.044)	0.213 (0.017)		
S032126	S12_12	0.643 (0.024)	0.316 (0.029)			
S032510	S12_13	0.806 (0.051)	-0.282 (0.083)	0.270 (0.030)		
S032158	S12_14	0.837 (0.065)	0.470 (0.066)	0.315 (0.022)		
S032574	S13_01	0.900 (0.153)	0.945 (0.105)	0.309 (0.033)		
S032532	S13_02	0.747 (0.049)	-0.189 (0.052)			
S032562	S13_03	0.695 (0.030)	0.264 (0.033)		-0.549 (0.072)	0.549 (0.074)
S032422	S13_04	1.061 (0.107)	0.198 (0.076)	0.202 (0.031)		
S032375	S13_05	0.562 (0.025)	0.716 (0.043)		-1.025 (0.095)	1.025 (0.104)
S032714	S13_06	0.997 (0.110)	-0.294 (0.117)	0.322 (0.043)		
S032704	S13_07	0.847 (0.060)	0.777 (0.054)			
S032705A	S13_08A	0.879 (0.059)	0.539 (0.046)			
S032705B	S13_08B	0.930 (0.059)	0.153 (0.042)			
S032706A	S13_09A	0.890 (0.060)	0.728 (0.050)			
S032706B	S13_09B	1.038 (0.069)	0.822 (0.047)			

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S032707	S13_10	1.363 (0.109)	1.417 (0.062)			
S032115	S14_01	1.088 (0.113)	0.548 (0.064)	0.148 (0.026)		
S032565	S14_02	0.783 (0.058)	0.973 (0.066)			
S032403	S14_03	1.102 (0.166)	1.024 (0.080)	0.264 (0.025)		
S032273	S14_04	0.644 (0.157)	1.852 (0.227)	0.261 (0.031)		
S032019A	S14_05A	0.969 (0.075)	1.339 (0.074)			
S032019B	S14_05B	1.147 (0.106)	1.748 (0.098)			
S032516	S14_06	0.709 (0.047)	-0.218 (0.054)			
S032620	S14_07	0.780 (0.161)	1.693 (0.160)	0.190 (0.025)		
S032693A	S14_08A	0.942 (0.058)	0.190 (0.041)			
S032693B	S14_08B	1.133 (0.067)	0.049 (0.037)			
S032695	S14_09	0.703 (0.037)	0.782 (0.042)		-0.211 (0.068)	0.211 (0.080)
S032697D	S14_10D	0.840 (0.043)	0.669 (0.033)		-0.018 (0.054)	0.018 (0.063)
SF12025	S01F01	0.454 (0.051)	1.028 (0.121)	0.190 (0.033)		
SF12026	S01F02	0.818 (0.048)	-0.534 (0.085)	0.259 (0.032)		
SF12027	S01F03	0.959 (0.041)	-0.378 (0.044)	0.094 (0.020)		
SF12028	S01F04	0.857 (0.042)	0.400 (0.037)	0.084 (0.014)		
SF12029	S01F05	0.789 (0.063)	0.789 (0.057)	0.233 (0.020)		
SF12030	S01F06	0.507 (0.043)	0.862 (0.088)	0.136 (0.026)		
SF22035	S01F07	0.516 (0.022)	0.588 (0.038)			
SF22225	S01F08	1.053 (0.088)	1.642 (0.055)	0.094 (0.009)		
SF22117	S01F09	0.710 (0.048)	0.782 (0.051)	0.124 (0.018)		
SF22235	S01F10	0.719 (0.052)	0.420 (0.068)	0.178 (0.025)		
SF22188	S01F11	0.703 (0.059)	0.738 (0.067)	0.214 (0.023)		
SF22074	S01F12	1.167 (0.064)	0.454 (0.034)	0.186 (0.015)		
SF22240	S01F13	0.966 (0.087)	1.327 (0.050)	0.201 (0.013)		
SF22206	S01F14	0.672 (0.051)	0.927 (0.056)	0.120 (0.019)		
SF22160	S01F15	0.754 (0.028)	0.883 (0.032)			
SF22058	S01F16	0.850 (0.051)	0.141 (0.057)	0.191 (0.023)		
SF12013	S02F01	0.550 (0.060)	1.596 (0.089)	0.137 (0.021)		
SF12014	S02F02	0.970 (0.044)	-0.337 (0.046)	0.121 (0.021)		
SF12015	S02F03	0.877 (0.045)	-0.030 (0.052)	0.159 (0.021)		
SF12016	S02F04	0.779 (0.037)	-0.457 (0.061)	0.129 (0.024)		
SF12017	S02F05	1.298 (0.059)	0.383 (0.026)	0.111 (0.012)		
SF12018	S02F06	0.544 (0.030)	-0.050 (0.075)	0.100 (0.024)		
SF12001	S02F07	0.774 (0.033)	0.031 (0.039)	0.056 (0.014)		
SF12002	S02F08	0.793 (0.038)	-0.205 (0.053)	0.117 (0.021)		
SF12003	S02F09	1.274 (0.049)	-0.187 (0.028)	0.095 (0.014)		
SF12004	S02F10	0.943 (0.043)	0.061 (0.038)	0.104 (0.016)		
SF12005	S02F11	0.859 (0.048)	0.410 (0.044)	0.132 (0.017)		
SF12006	S02F12	1.006 (0.054)	0.618 (0.033)	0.115 (0.013)		
SF32131	S02F13	1.174 (0.034)	0.308 (0.017)			
SF32202	S02F14	0.778 (0.019)	0.485 (0.016)		-0.015 (0.028)	0.015 (0.031)
SF12037	S03F01	0.978 (0.085)	-0.531 (0.092)	0.177 (0.038)		

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SF12038	S03F02	1.007 (0.098)	0.191 (0.075)	0.170 (0.032)		
SF12039	S03F03	1.064 (0.093)	-0.125 (0.074)	0.166 (0.032)		
SF12040	S03F04	1.248 (0.124)	0.506 (0.056)	0.162 (0.025)		
SF12041	S03F05	0.807 (0.084)	0.433 (0.083)	0.142 (0.030)		
SF12042	S03F06	0.922 (0.093)	0.443 (0.072)	0.140 (0.028)		
SF22086	S03F07	1.342 (0.079)	0.554 (0.032)			
SF22198	S03F08	0.969 (0.144)	1.243 (0.086)	0.158 (0.024)		
SF22275	S03F09	1.141 (0.139)	1.196 (0.066)	0.102 (0.017)		
SF22041	S03F10	1.204 (0.099)	0.164 (0.054)	0.119 (0.024)		
SF22283	S03F11	1.087 (0.065)	0.203 (0.036)			
SF22202	S03F12	0.796 (0.108)	0.883 (0.094)	0.177 (0.031)		
SF22152	S03F13	1.237 (0.075)	0.631 (0.036)			
SF22154	S03F14	0.956 (0.061)	0.465 (0.041)			
SF22187	S04F01	0.818 (0.113)	1.014 (0.092)	0.162 (0.028)		
SF22161	S04F02	0.766 (0.057)	0.861 (0.062)			
SF22222	S04F03	1.456 (0.151)	0.862 (0.046)	0.111 (0.017)		
SF22191	S04F04	0.816 (0.036)	0.273 (0.029)		-0.280 (0.059)	0.280 (0.060)
SF22279	S04F05	0.945 (0.064)	0.759 (0.048)			
SF22040	S04F06	1.253 (0.111)	0.499 (0.049)	0.117 (0.021)		
SF22088A	S04F07A	1.234 (0.072)	0.350 (0.033)			
SF22088B	S04F07B	0.931 (0.064)	0.852 (0.052)			
SF22249D	S04F08D	1.182 (0.077)	0.873 (0.043)			
SF22286	S04F09	1.182 (0.115)	1.819 (0.108)			
SF32595	S04F10	1.520 (0.176)	1.395 (0.068)	0.127 (0.014)		
SF32656	S04F11	1.629 (0.149)	0.711 (0.039)	0.099 (0.016)		
SF32625A	S04F12A	1.278 (0.080)	0.766 (0.038)			
SF32625B	S04F12B	1.707 (0.113)	0.991 (0.036)			
SF22183	S05F01	1.246 (0.159)	0.923 (0.063)	0.207 (0.023)		
SF22276	S05F02	0.955 (0.101)	0.477 (0.074)	0.159 (0.028)		
SF22115	S05F03	1.185 (0.111)	0.342 (0.059)	0.161 (0.025)		
SF22022	S05F04	0.885 (0.057)	0.530 (0.046)			
SF22019	S05F05	1.347 (0.124)	0.073 (0.061)	0.212 (0.029)		
SF22002	S05F06	1.247 (0.127)	0.595 (0.056)	0.167 (0.023)		
SF22294	S05F07	1.099 (0.117)	0.279 (0.077)	0.239 (0.031)		
SF22106	S05F08	0.895 (0.141)	1.577 (0.117)	0.105 (0.019)		
SF22244	S05F09	1.503 (0.102)	1.101 (0.042)			
SF22150	S05F10	1.080 (0.117)	0.752 (0.062)	0.147 (0.023)		
SF22042	S05F11	1.453 (0.139)	0.620 (0.047)	0.145 (0.019)		
SF22289	S05F12	1.004 (0.056)	1.070 (0.037)		0.593 (0.041)	-0.593 (0.068)
SF22069	S05F13	1.225 (0.076)	0.802 (0.040)			
SF22268	S05F14	0.887 (0.064)	1.116 (0.065)			
SF22290	S06F01	1.187 (0.120)	0.392 (0.064)	0.200 (0.027)		
SF22292	S06F02	0.744 (0.052)	0.616 (0.055)			
SF22054	S06F03	1.260 (0.130)	0.566 (0.057)	0.182 (0.024)		

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SF22181	S06F04	1.114 (0.134)	0.790 (0.067)	0.195 (0.025)		
SF22208	S06F05	1.032 (0.139)	0.975 (0.075)	0.195 (0.025)		
SF22078	S06F06	1.393 (0.078)	0.352 (0.031)			
SF22126	S06F07	0.769 (0.086)	0.538 (0.091)	0.163 (0.031)		
SF22281	S06F08	0.765 (0.059)	1.212 (0.079)			
SF32385	S06F09	1.067 (0.107)	0.227 (0.075)	0.209 (0.031)		
SF32035	S06F10	1.368 (0.134)	0.581 (0.051)	0.161 (0.021)		
SF32519	S06F11	0.961 (0.062)	0.702 (0.046)			
SF32683	S06F12	1.172 (0.125)	0.841 (0.057)	0.135 (0.020)		
SF32258	S06F13	1.191 (0.112)	0.322 (0.061)	0.173 (0.026)		
SF32120A	S06F14A	0.998 (0.078)	1.414 (0.076)			
SF32120B	S06F14B	1.177 (0.086)	1.271 (0.059)			
SF32606	S07F01	0.919 (0.084)	-0.669 (0.107)	0.199 (0.042)		
SF32015	S07F02	0.783 (0.055)	0.855 (0.063)			
SF32310D	S07F03D	0.496 (0.026)	-0.077 (0.044)		-0.325 (0.091)	0.325 (0.089)
SF32680	S07F04	0.698 (0.033)	-0.262 (0.035)		-0.165 (0.070)	0.165 (0.065)
SF32672	S07F05	0.443 (0.072)	0.390 (0.235)	0.246 (0.054)		
SF32392	S07F06	0.557 (0.060)	-0.928 (0.196)	0.221 (0.054)		
SF32425	S07F07	0.863 (0.118)	0.782 (0.093)	0.223 (0.030)		
SF32257	S07F08	1.068 (0.155)	1.053 (0.080)	0.213 (0.024)		
SF32663	S07F09	0.590 (0.118)	1.558 (0.179)	0.209 (0.034)		
SF32660	S07F10	0.804 (0.143)	1.524 (0.132)	0.171 (0.024)		
SF32555	S07F11	1.176 (0.078)	0.962 (0.048)			
SF32122	S07F12	0.726 (0.055)	1.088 (0.078)			
SF32542	S08F01	1.300 (0.158)	0.690 (0.063)	0.250 (0.025)		
SF32645	S08F02	0.904 (0.128)	0.889 (0.088)	0.220 (0.029)		
SF32530D	S08F03D	0.534 (0.033)	0.469 (0.048)		0.718 (0.073)	-0.718 (0.085)
SF32007	S08F04	0.812 (0.053)	0.367 (0.048)			
SF32502	S08F05	0.888 (0.109)	0.935 (0.079)	0.147 (0.024)		
SF32679	S08F06	0.907 (0.071)	1.338 (0.078)			
SF32184	S08F07	0.568 (0.100)	1.336 (0.157)	0.181 (0.035)		
SF32394	S08F08	0.882 (0.117)	0.651 (0.092)	0.232 (0.032)		
SF32151	S08F09	1.493 (0.163)	0.770 (0.049)	0.179 (0.020)		
SF32651A	S08F10A	1.336 (0.078)	0.455 (0.033)			
SF32651B	S08F10B	1.140 (0.081)	1.140 (0.055)			
SF32665A	S08F11A	1.070 (0.071)	0.858 (0.047)			
SF32665B	S08F11B	1.123 (0.083)	1.256 (0.062)			
SF32665C	S08F11C	0.993 (0.077)	1.329 (0.072)			
SF32465	S11F01	0.929 (0.099)	-0.101 (0.104)	0.244 (0.040)		
SF32315	S11F02	0.931 (0.108)	0.525 (0.082)	0.208 (0.030)		
SF32306	S11F03	0.587 (0.025)	0.551 (0.038)		-1.188 (0.097)	1.188 (0.103)
SF32640	S11F04	0.664 (0.046)	0.063 (0.054)			
SF32579	S11F05	0.768 (0.138)	1.282 (0.119)	0.217 (0.030)		
SF32570	S11F06	0.944 (0.066)	0.949 (0.055)			

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SF32024	S11F07	0.871 (0.166)	1.396 (0.121)	0.248 (0.027)		
SF32272	S11F08	1.147 (0.092)	1.449 (0.072)			
SF32141	S11F09	1.422 (0.185)	1.088 (0.057)	0.167 (0.019)		
SF32060	S11F10	1.066 (0.062)	-0.034 (0.038)			
SF32463	S11F11	1.545 (0.165)	0.588 (0.050)	0.212 (0.022)		
SF32650D	S11F12D	0.686 (0.034)	0.549 (0.036)		-0.228 (0.066)	0.228 (0.073)
SF32514	S11F13	1.026 (0.161)	1.111 (0.085)	0.234 (0.025)		
SF32611	S12F01	0.842 (0.140)	1.378 (0.113)	0.182 (0.024)		
SF32614	S12F02	0.802 (0.051)	0.129 (0.047)			
SF32451	S12F03	0.644 (0.026)	0.221 (0.034)		-1.197 (0.095)	1.197 (0.096)
SF32156	S12F04	1.056 (0.139)	0.889 (0.074)	0.214 (0.025)		
SF32056	S12F05	0.934 (0.063)	0.777 (0.050)			
SF32087	S12F06	0.671 (0.106)	1.070 (0.120)	0.185 (0.033)		
SF32279	S12F07	0.718 (0.127)	1.491 (0.137)	0.165 (0.027)		
SF32238	S12F08	1.163 (0.124)	0.655 (0.059)	0.160 (0.023)		
SF32369	S12F09	0.617 (0.031)	0.673 (0.041)		-0.403 (0.075)	0.403 (0.085)
SF32160	S12F10	0.887 (0.126)	0.616 (0.103)	0.296 (0.034)		
SF32654	S12F11	1.001 (0.115)	0.778 (0.069)	0.154 (0.024)		
SF32126	S12F12	0.775 (0.053)	0.529 (0.052)			
SF32510	S12F13	0.882 (0.103)	-0.014 (0.116)	0.282 (0.041)		
SF32158	S12F14	0.943 (0.130)	0.626 (0.094)	0.290 (0.032)		
SF32574	S13F01	1.049 (0.086)	0.870 (0.047)	0.320 (0.016)		
SF32532	S13F02	0.760 (0.025)	-0.123 (0.025)			
SF32562	S13F03	0.677 (0.014)	0.441 (0.017)		-0.676 (0.038)	0.676 (0.040)
SF32422	S13F04	1.182 (0.064)	0.383 (0.036)	0.223 (0.015)		
SF32375	S13F05	0.589 (0.013)	0.907 (0.023)		-1.066 (0.048)	1.066 (0.054)
SF32714	S13F06	1.140 (0.063)	-0.134 (0.051)	0.301 (0.022)		
SF32704	S13F07	0.930 (0.033)	1.101 (0.031)			
SF32705A	S13F08A	0.953 (0.031)	0.748 (0.024)			
SF32705B	S13F08B	1.005 (0.030)	0.414 (0.020)			
SF32706A	S13F09A	0.953 (0.032)	0.914 (0.027)			
SF32706B	S13F09B	1.119 (0.038)	1.044 (0.025)			
SF32707	S13F10	1.558 (0.063)	1.446 (0.028)			
SF32115	S14F01	0.975 (0.052)	0.633 (0.034)	0.097 (0.014)		
SF32565	S14F02	0.838 (0.032)	1.231 (0.037)			
SF32403	S14F03	1.127 (0.090)	1.101 (0.040)	0.258 (0.013)		
SF32273	S14F04	0.679 (0.091)	1.786 (0.099)	0.233 (0.018)		
SF32019A	S14F05A	1.062 (0.043)	1.488 (0.038)			
SF32019B	S14F05B	1.150 (0.056)	1.891 (0.055)			
SF32516	S14F06	0.800 (0.026)	0.091 (0.023)			
SF32620	S14F07	0.840 (0.083)	1.667 (0.068)	0.139 (0.012)		
SF32693A	S14F08A	0.902 (0.029)	0.530 (0.023)			
SF32693B	S14F08B	1.088 (0.033)	0.281 (0.019)			
SF32695	S14F09	0.674 (0.019)	1.131 (0.027)		-0.315 (0.037)	0.315 (0.048)

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SF32697D	S14F10D	0.688 (0.018)	1.001 (0.023)		-0.234 (0.033)	0.234 (0.042)
SZ32115	S01_01	0.984 (0.085)	0.629 (0.056)	0.148 (0.022)		
SZ32565	S01_02	0.772 (0.043)	0.923 (0.049)			
SZ32403	S01_03	1.074 (0.124)	0.964 (0.062)	0.263 (0.021)		
SZ32273	S01_04	0.722 (0.131)	1.697 (0.138)	0.253 (0.024)		
SZ32019A	S01_05A	0.982 (0.055)	1.171 (0.048)			
SZ32019B	S01_05B	1.079 (0.072)	1.660 (0.070)			
SZ32516	S01_06	0.781 (0.038)	-0.414 (0.041)			
SZ32620	S01_07	0.678 (0.111)	1.651 (0.128)	0.189 (0.024)		
SZ32693A	S01_08A	0.935 (0.043)	0.102 (0.031)			
SZ32693B	S01_08B	1.079 (0.048)	-0.164 (0.030)			
SZ32695	S01_09	0.668 (0.027)	0.633 (0.030)		-0.165 (0.051)	0.165 (0.059)
SZ32697D	S01_10D	0.732 (0.029)	0.493 (0.027)		-0.040 (0.047)	0.040 (0.052)
SZ42009	S02_01	0.990 (0.135)	0.865 (0.083)	0.391 (0.025)		
SZ42313	S02_02	0.703 (0.036)	-0.694 (0.049)			
SZ42059	S02_03	0.855 (0.096)	0.795 (0.077)	0.243 (0.026)		
SZ42011	S02_04	0.790 (0.047)	1.270 (0.062)			
SZ42028	S02_05	0.837 (0.106)	1.273 (0.081)	0.186 (0.021)		
SZ42001	S02_06	0.993 (0.122)	1.367 (0.073)	0.165 (0.017)		
SZ42276	S02_07	0.714 (0.113)	1.085 (0.112)	0.321 (0.031)		
SZ42279	S02_08	0.389 (0.054)	0.477 (0.228)	0.224 (0.050)		
SZ42083	S02_09	0.793 (0.033)	1.073 (0.034)		-0.158 (0.046)	0.158 (0.061)
SZ42106	S02_10	1.035 (0.052)	0.859 (0.037)			
SZ42071	S02_11	0.875 (0.112)	1.107 (0.077)	0.234 (0.023)		
SZ42101	S02_12	0.831 (0.047)	1.069 (0.051)			
SZ42307	S02_13	0.562 (0.034)	0.561 (0.054)			
SZ42405	S02_14	0.945 (0.109)	0.714 (0.077)	0.309 (0.026)		
SZ42244A	S02_15A	1.076 (0.057)	1.014 (0.040)			
SZ42244B	S02_15B	0.909 (0.058)	1.492 (0.068)			
SZ42153	S02_16	0.806 (0.043)	0.395 (0.038)			
SZ22183	S03_01	1.365 (0.148)	1.018 (0.050)	0.258 (0.017)		
SZ22276	S03_02	0.753 (0.077)	0.338 (0.097)	0.224 (0.034)		
SZ22115	S03_03	0.929 (0.077)	0.058 (0.076)	0.217 (0.030)		
SZ22022	S03_04	0.672 (0.035)	0.091 (0.041)			
SZ22019	S03_05	0.877 (0.074)	-0.271 (0.097)	0.248 (0.037)		
SZ22002	S03_06	1.037 (0.093)	0.488 (0.060)	0.216 (0.024)		
SZ22294	S03_07	0.828 (0.084)	0.098 (0.106)	0.309 (0.036)		
SZ22244	S03_09	1.161 (0.059)	0.959 (0.036)			
SZ22150	S03_10	0.863 (0.087)	0.586 (0.075)	0.223 (0.027)		
SZ22042	S03_11	1.134 (0.093)	0.503 (0.051)	0.186 (0.021)		
SZ22289	S03_12	0.786 (0.033)	0.905 (0.031)		0.802 (0.038)	-0.802 (0.057)
SZ22069	S03_13	0.942 (0.047)	0.596 (0.035)			
SZ22268	S03_14	0.637 (0.038)	0.663 (0.050)			
SZ42013	S04_01	0.910 (0.069)	-0.786 (0.104)	0.231 (0.041)		

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SZ42006	S04_02	0.625 (0.078)	0.293 (0.148)	0.299 (0.043)		
SZ42310	S04_03	0.532 (0.021)	0.087 (0.032)		-0.277 (0.064)	0.277 (0.063)
SZ42052	S04_04	0.567 (0.026)	0.197 (0.033)		0.381 (0.057)	-0.381 (0.058)
SZ42054	S04_05	0.325 (0.037)	-0.590 (0.277)	0.207 (0.053)		
SZ42043	S04_06	0.418 (0.032)	1.165 (0.096)			
SZ42196	S04_07	0.767 (0.047)	1.355 (0.067)			
SZ42061	S04_08	0.603 (0.090)	1.429 (0.117)	0.179 (0.028)		
SZ42292	S04_09	0.420 (0.016)	0.728 (0.043)		-1.126 (0.089)	1.126 (0.098)
SZ42109	S04_10	1.372 (0.111)	0.131 (0.055)	0.295 (0.025)		
SZ42232A	S04_11A	0.775 (0.039)	0.380 (0.037)			
SZ42232B	S04_11B	1.120 (0.082)	1.854 (0.082)			
SZ42232C	S04_11C	1.527 (0.183)	1.747 (0.094)	0.330 (0.012)		
SZ42294	S04_12	1.405 (0.145)	0.955 (0.046)	0.241 (0.017)		
SZ42149	S04_13	0.541 (0.033)	0.256 (0.050)			
SZ42155	S04_14	0.819 (0.043)	0.668 (0.040)			
SZ42150	S04_15	0.755 (0.087)	0.580 (0.098)	0.267 (0.032)		
SZ22290	S05_01	1.116 (0.097)	0.382 (0.059)	0.239 (0.024)		
SZ22292	S05_02	0.672 (0.038)	0.628 (0.046)			
SZ22054	S05_03	1.100 (0.108)	0.589 (0.060)	0.265 (0.024)		
SZ22181	S05_04	1.055 (0.111)	0.800 (0.061)	0.254 (0.022)		
SZ22208	S05_05	1.145 (0.133)	1.048 (0.059)	0.254 (0.019)		
SZ22078	S05_06	1.138 (0.051)	0.173 (0.027)			
SZ22126	S05_07	0.583 (0.066)	0.399 (0.127)	0.207 (0.038)		
SZ22281	S05_08	0.608 (0.039)	1.176 (0.071)			
SZ32385	S05_09	0.893 (0.083)	0.093 (0.088)	0.267 (0.033)		
SZ32035	S05_10	1.080 (0.083)	0.349 (0.052)	0.160 (0.022)		
SZ32519	S05_11	0.667 (0.037)	0.505 (0.044)			
SZ32683	S05_12	1.024 (0.102)	0.808 (0.058)	0.206 (0.022)		
SZ32258	S05_13	0.885 (0.073)	-0.046 (0.082)	0.209 (0.032)		
SZ32120A	S05_14A	0.784 (0.049)	1.369 (0.067)			
SZ32120B	S05_14B	0.901 (0.052)	0.962 (0.045)			
SZ42304	S06_01	0.731 (0.059)	-0.032 (0.089)	0.160 (0.032)		
SZ42038	S06_02	0.836 (0.074)	0.435 (0.072)	0.161 (0.027)		
SZ42298	S06_03	0.897 (0.043)	0.304 (0.033)			
SZ42261	S06_04	0.785 (0.044)	0.960 (0.049)			
SZ42051A	S06_05A	0.777 (0.038)	-0.024 (0.037)			
SZ42051B	S06_05B	1.168 (0.057)	0.817 (0.032)			
SZ42076	S06_06	0.907 (0.046)	0.708 (0.037)			
SZ42404	S06_07	0.695 (0.032)	1.301 (0.045)		-0.079 (0.051)	0.079 (0.073)
SZ42306	S06_08	1.217 (0.148)	1.133 (0.060)	0.276 (0.018)		
SZ42403	S06_09	0.830 (0.044)	0.808 (0.043)			
SZ42272	S06_10	0.836 (0.082)	0.403 (0.083)	0.241 (0.030)		
SZ42100	S06_11	0.442 (0.022)	0.597 (0.042)		-0.035 (0.070)	0.035 (0.079)
SZ42238A	S06_12A	0.728 (0.078)	0.907 (0.079)	0.159 (0.025)		

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SZ42238B	S06_12B	0.703 (0.044)	1.335 (0.071)			
SZ42141	S06_13	0.750 (0.060)	-0.320 (0.100)	0.185 (0.036)		
SZ42215	S06_14	0.525 (0.090)	1.555 (0.149)	0.196 (0.033)		
SZ32606	S07_01	0.907 (0.073)	-0.671 (0.109)	0.263 (0.042)		
SZ32015	S07_02	0.690 (0.037)	0.528 (0.044)			
SZ32310D	S07_03D	0.503 (0.020)	-0.257 (0.035)		-0.230 (0.070)	0.230 (0.065)
SZ32680	S07_04	0.654 (0.026)	-0.453 (0.031)		0.089 (0.058)	-0.089 (0.049)
SZ32672	S07_05	0.380 (0.052)	0.096 (0.269)	0.252 (0.057)		
SZ32392	S07_06	0.433 (0.040)	-1.644 (0.243)	0.220 (0.056)		
SZ32425	S07_07	0.843 (0.098)	0.772 (0.081)	0.260 (0.027)		
SZ32257	S07_08	1.246 (0.144)	1.068 (0.056)	0.259 (0.018)		
SZ32663	S07_09	0.519 (0.093)	1.620 (0.159)	0.217 (0.034)		
SZ32660	S07_10	0.797 (0.115)	1.449 (0.096)	0.202 (0.022)		
SZ32555	S07_11	1.239 (0.058)	0.668 (0.029)			
SZ32122	S07_12	0.646 (0.040)	0.933 (0.058)			
SZ42053	S08_01	1.109 (0.089)	-0.082 (0.070)	0.265 (0.030)		
SZ42408	S08_02	0.761 (0.039)	0.471 (0.039)			
SZ42015	S08_03	0.854 (0.110)	0.896 (0.084)	0.294 (0.027)		
SZ42309	S08_04	0.340 (0.060)	1.364 (0.260)	0.220 (0.047)		
SZ42049A	S08_05A	0.948 (0.044)	-0.578 (0.037)			
SZ42049B	S08_05B	1.192 (0.052)	0.266 (0.026)			
SZ42182	S08_06	0.607 (0.056)	-0.377 (0.144)	0.212 (0.045)		
SZ42402	S08_07	0.762 (0.047)	1.344 (0.068)			
SZ42228A	S08_08A	1.213 (0.063)	1.079 (0.038)			
SZ42228B	S08_08B	1.121 (0.049)	-0.003 (0.028)			
SZ42228C	S08_08C	1.368 (0.061)	0.566 (0.025)			
SZ42126	S08_09	0.597 (0.076)	0.040 (0.184)	0.312 (0.051)		
SZ42210	S08_10	0.766 (0.132)	1.516 (0.116)	0.267 (0.025)		
SZ42176	S08_11	0.954 (0.047)	0.626 (0.035)			
SZ42211	S08_12	0.897 (0.042)	0.165 (0.032)			
SZ42135	S08_13	0.783 (0.039)	-0.013 (0.036)			
SZ42257	S08_14	0.747 (0.113)	1.172 (0.102)	0.287 (0.028)		
SZ32542	S09_01	1.199 (0.115)	0.674 (0.054)	0.262 (0.021)		
SZ32645	S09_02	0.950 (0.106)	0.796 (0.070)	0.261 (0.025)		
SZ32530D	S09_03D	0.540 (0.024)	0.321 (0.036)		0.886 (0.056)	-0.886 (0.062)
SZ32007	S09_04	0.780 (0.039)	0.333 (0.037)			
SZ32502	S09_05	1.016 (0.088)	0.661 (0.053)	0.160 (0.021)		
SZ32679	S09_06	0.791 (0.051)	1.505 (0.074)			
SZ32184	S09_07	0.448 (0.072)	1.154 (0.175)	0.211 (0.042)		
SZ32394	S09_08	0.812 (0.095)	0.693 (0.088)	0.275 (0.029)		
SZ32151	S09_09	1.195 (0.107)	0.787 (0.047)	0.180 (0.018)		
SZ32651A	S09_10A	1.071 (0.049)	0.365 (0.029)			
SZ32651B	S09_10B	0.949 (0.053)	1.147 (0.048)			
SZ32665A	S09_11A	0.935 (0.047)	0.621 (0.035)			

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SZ32665B	S09_11B	1.032 (0.057)	1.119 (0.044)			
SZ32665C	S09_11C	0.938 (0.055)	1.064 (0.047)			
SZ42073	S10_01	0.453 (0.045)	-1.667 (0.275)	0.257 (0.065)		
SZ42017	S10_02	0.829 (0.112)	1.318 (0.085)	0.203 (0.022)		
SZ42007	S10_03	1.166 (0.113)	0.827 (0.052)	0.204 (0.020)		
SZ42024	S10_04	1.002 (0.148)	1.455 (0.085)	0.248 (0.019)		
SZ42095	S10_05	0.971 (0.074)	-0.191 (0.075)	0.204 (0.031)		
SZ42022	S10_06	0.879 (0.044)	0.535 (0.035)			
SZ42063	S10_07	0.761 (0.059)	-1.799 (0.162)	0.235 (0.057)		
SZ42197	S10_08	0.986 (0.103)	0.850 (0.062)	0.217 (0.022)		
SZ42297	S10_09	0.575 (0.024)	1.296 (0.048)		-0.694 (0.069)	0.694 (0.087)
SZ42305	S10_10	0.507 (0.028)	1.302 (0.058)		0.350 (0.059)	-0.350 (0.088)
SZ42112	S10_11	0.439 (0.054)	0.274 (0.194)	0.213 (0.047)		
SZ42173	S10_12	0.460 (0.020)	-0.521 (0.045)		1.296 (0.080)	-1.296 (0.063)
SZ42407	S10_13	0.425 (0.030)	0.653 (0.071)			
SZ42278	S10_14	0.787 (0.043)	0.832 (0.045)			
SZ42274	S10_15	0.975 (0.144)	1.493 (0.087)	0.222 (0.019)		
SZ42317	S10_17	0.574 (0.022)	-0.014 (0.031)		-0.415 (0.065)	0.415 (0.063)
SZ32465	S11_01	0.664 (0.061)	-0.230 (0.128)	0.215 (0.042)		
SZ32315	S11_02	0.818 (0.088)	0.671 (0.080)	0.233 (0.028)		
SZ32306	S11_03	0.479 (0.016)	0.478 (0.034)		-1.261 (0.084)	1.261 (0.088)
SZ32640	S11_04	0.617 (0.034)	-0.067 (0.045)			
SZ32579	S11_05	0.784 (0.132)	1.479 (0.110)	0.277 (0.024)		
SZ32570	S11_06	0.914 (0.045)	0.589 (0.035)			
SZ32024	S11_07	0.835 (0.137)	1.511 (0.107)	0.270 (0.022)		
SZ32272	S11_08	0.972 (0.060)	1.476 (0.063)			
SZ32141	S11_09	1.466 (0.147)	1.059 (0.044)	0.179 (0.016)		
SZ32060	S11_10	1.063 (0.050)	-0.364 (0.034)			
SZ32463	S11_11	1.350 (0.108)	0.329 (0.048)	0.236 (0.022)		
SZ32650D	S11_12D	0.734 (0.029)	0.151 (0.026)		0.126 (0.046)	-0.126 (0.046)
SZ32514	S11_13	0.792 (0.108)	1.031 (0.092)	0.279 (0.028)		
SZ42042	S12_01	0.652 (0.061)	-0.524 (0.153)	0.246 (0.049)		
SZ42030	S12_02	0.822 (0.048)	1.269 (0.059)			
SZ42003	S12_03	0.529 (0.076)	0.915 (0.144)	0.229 (0.040)		
SZ42110	S12_04	0.592 (0.051)	-0.775 (0.158)	0.211 (0.048)		
SZ42222A	S12_05A	1.054 (0.060)	1.308 (0.050)			
SZ42222B	S12_05B	1.008 (0.053)	1.010 (0.041)			
SZ42222C	S12_05C	0.783 (0.072)	0.132 (0.094)	0.222 (0.034)		
SZ42065	S12_06	0.641 (0.058)	-1.115 (0.186)	0.264 (0.058)		
SZ42280	S12_07	1.277 (0.093)	0.289 (0.046)	0.175 (0.021)		
SZ42088	S12_08	0.620 (0.034)	-0.057 (0.045)			
SZ42218	S12_09	1.391 (0.128)	0.695 (0.046)	0.252 (0.019)		
SZ42104	S12_10	0.817 (0.046)	1.105 (0.052)			
SZ42064	S12_11	0.827 (0.045)	0.937 (0.046)			

( ) Standard errors appear in parentheses

Exhibit D.4 IRT Parameters for TIMSS Joint 2003-2007 Eighth Grade Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
SZ42273	S12_12	0.990 (0.045)	0.179 (0.030)			
SZ42301	S12_13	0.907 (0.043)	0.079 (0.032)			
SZ42312	S12_14	0.442 (0.047)	-0.398 (0.213)	0.217 (0.052)		
SZ42217	S12_15	1.432 (0.136)	0.742 (0.046)	0.258 (0.019)		
SZ42406	S12_16	1.023 (0.050)	0.680 (0.033)			
SZ32611	S13_01	0.860 (0.121)	1.351 (0.087)	0.230 (0.022)		
SZ32614	S13_02	0.791 (0.039)	-0.012 (0.036)			
SZ32451	S13_03	0.627 (0.020)	0.106 (0.027)		-1.157 (0.074)	1.157 (0.073)
SZ32156	S13_04	1.173 (0.109)	0.741 (0.052)	0.221 (0.020)		
SZ32056	S13_05	0.879 (0.044)	0.597 (0.036)			
SZ32087	S13_06	0.688 (0.082)	0.929 (0.092)	0.188 (0.029)		
SZ32279	S13_07	0.962 (0.129)	1.410 (0.079)	0.201 (0.019)		
SZ32238	S13_08	1.074 (0.090)	0.578 (0.052)	0.169 (0.021)		
SZ32369	S13_09	0.614 (0.025)	0.725 (0.033)		-0.168 (0.054)	0.168 (0.063)
SZ32160	S13_10	0.747 (0.086)	0.344 (0.116)	0.307 (0.037)		
SZ32654	S13_11	0.872 (0.083)	0.622 (0.069)	0.185 (0.026)		
SZ32126	S13_12	0.729 (0.039)	0.257 (0.039)			
SZ32510	S13_13	0.752 (0.066)	-0.395 (0.123)	0.238 (0.043)		
SZ32158	S13_14	0.835 (0.088)	0.388 (0.093)	0.280 (0.033)		
SZ42258	S14_01	0.766 (0.095)	1.085 (0.083)	0.201 (0.026)		
SZ42005	S14_02	0.358 (0.013)	0.285 (0.041)		-1.948 (0.114)	1.948 (0.116)
SZ42016	S14_03	0.747 (0.130)	2.034 (0.162)	0.142 (0.018)		
SZ42300A	S14_04A	1.092 (0.049)	0.109 (0.028)			
SZ42300B	S14_04B	0.484 (0.038)	1.811 (0.130)			
SZ42300C	S14_04C	0.916 (0.044)	0.267 (0.032)			
SZ42319	S14_05	1.109 (0.055)	0.888 (0.035)			
SZ42068	S14_06	0.974 (0.118)	1.093 (0.068)	0.229 (0.022)		
SZ42216	S14_07	0.817 (0.088)	0.383 (0.096)	0.273 (0.034)		
SZ42249	S14_08	0.981 (0.100)	0.668 (0.067)	0.246 (0.025)		
SZ42094	S14_09	0.849 (0.043)	0.615 (0.038)			
SZ42293A	S14_10A	0.895 (0.042)	-0.319 (0.036)			
SZ42293B	S14_10B	0.736 (0.057)	2.031 (0.119)			
SZ42195	S14_11	0.550 (0.041)	1.779 (0.116)			
SZ42400	S14_12	0.889 (0.051)	1.189 (0.052)			
SZ42164	S14_14	1.144 (0.097)	0.722 (0.047)	0.153 (0.019)		

( ) Standard errors appear in parentheses

Exhibit D.5 IRT Parameters for TIMSS 2007 Fourth Grade Mathematics - Number

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M031286	M01_01	1.268 (0.053)	0.358 (0.025)			
M031106	M01_02	1.066 (0.045)	0.255 (0.028)			
M031282	M01_03	0.999 (0.033)	0.865 (0.022)		-0.509 (0.045)	0.509 (0.051)
M031227	M01_04	1.138 (0.062)	1.310 (0.046)			
M031335	M01_05	1.170 (0.089)	0.146 (0.054)	0.231 (0.023)		
M031068	M01_06	1.303 (0.055)	0.389 (0.024)			
M031299	M01_07	1.455 (0.060)	0.109 (0.023)			
M031301	M01_08	1.228 (0.050)	-0.415 (0.029)			
M041014	M02_01	0.851 (0.064)	-0.588 (0.099)	0.197 (0.038)		
M041039	M02_02	1.012 (0.081)	0.137 (0.064)	0.223 (0.025)		
M041278	M02_03	0.701 (0.034)	0.338 (0.039)			
M041006	M02_04	1.079 (0.084)	0.618 (0.046)	0.140 (0.017)		
M041250	M02_05	1.128 (0.048)	0.175 (0.027)			
M041094	M02_06	1.284 (0.119)	0.919 (0.046)	0.208 (0.016)		
M031235	M03_01	0.858 (0.040)	0.567 (0.036)			
M031285	M03_02	0.818 (0.041)	0.902 (0.044)			
M031050	M03_03	1.243 (0.113)	0.717 (0.049)	0.249 (0.018)		
M031258	M03_04	1.146 (0.053)	0.811 (0.032)			
M031334	M03_05	1.483 (0.137)	0.899 (0.042)	0.237 (0.015)		
M031255	M03_06	1.147 (0.100)	0.370 (0.059)	0.286 (0.023)		
M041052	M04_01	0.820 (0.073)	-0.405 (0.119)	0.305 (0.040)		
M041056	M04_02	0.926 (0.041)	0.321 (0.031)			
M041069	M04_03	1.414 (0.112)	1.097 (0.037)	0.086 (0.010)		
M041076	M04_04	0.887 (0.040)	0.467 (0.034)			
M041281	M04_05	1.212 (0.071)	-0.199 (0.045)	0.104 (0.019)		
M031303	M05_01	1.415 (0.102)	-0.170 (0.053)	0.277 (0.024)		
M031309	M05_02	1.300 (0.052)	-0.167 (0.025)			
M031245	M05_03	1.727 (0.138)	1.043 (0.032)	0.106 (0.010)		
M031242A	M05_04A	0.945 (0.040)	-0.120 (0.031)			
M031247	M05_05	0.597 (0.026)	1.281 (0.046)		-0.187 (0.056)	0.187 (0.077)
M031173	M05_07	1.608 (0.096)	-0.078 (0.035)	0.125 (0.017)		
M041010	M06_01	1.041 (0.089)	-0.014 (0.077)	0.318 (0.029)		
M041098	M06_02	1.543 (0.131)	0.719 (0.039)	0.245 (0.016)		
M041064	M06_03	0.853 (0.037)	-0.292 (0.035)			
M041003	M06_04	0.907 (0.040)	0.163 (0.031)			
M041104	M06_05	1.102 (0.046)	0.062 (0.027)			
M041299	M06_06	1.343 (0.063)	0.916 (0.030)			
M031029	M07_01	1.252 (0.120)	0.656 (0.053)	0.317 (0.019)		
M031030	M07_02	0.819 (0.050)	1.662 (0.076)			
M031332	M07_03	1.087 (0.091)	0.170 (0.065)	0.285 (0.025)		
M031098	M07_04	1.367 (0.091)	0.329 (0.037)	0.140 (0.016)		
M031254	M07_05	1.181 (0.089)	0.342 (0.048)	0.192 (0.020)		
M031276	M07_07	1.444 (0.113)	0.382 (0.043)	0.251 (0.019)		
M031064	M07_08	1.028 (0.092)	0.747 (0.054)	0.192 (0.019)		

( ) Standard errors appear in parentheses

Exhibit D.5 IRT Parameters for TIMSS 2007 Fourth Grade Mathematics - Number (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M041291	M08_01	0.637 (0.030)	-0.750 (0.049)			
M041289	M08_02	1.072 (0.095)	0.536 (0.057)	0.252 (0.021)		
M041068	M08_03	1.186 (0.079)	0.755 (0.035)	0.061 (0.011)		
M041065A	M08_04A	1.896 (0.141)	0.798 (0.029)	0.143 (0.011)		
M041065B	M08_04B	1.280 (0.065)	1.152 (0.037)			
M041096	M08_05	0.948 (0.067)	0.409 (0.048)	0.098 (0.018)		
M041125	M08_06	1.193 (0.100)	0.850 (0.044)	0.157 (0.015)		
M031128	M09_01	0.559 (0.028)	-1.136 (0.064)			
M031016	M09_02	1.181 (0.055)	0.822 (0.031)			
M031183	M09_03	0.812 (0.031)	0.290 (0.025)		0.609 (0.038)	-0.609 (0.041)
M031187	M09_05	0.811 (0.065)	-0.458 (0.105)	0.218 (0.038)		
M031251	M09_06	1.493 (0.117)	0.637 (0.037)	0.203 (0.016)		
M031294	M09_07	1.271 (0.078)	0.055 (0.040)	0.113 (0.017)		
M031218	M09_09	1.453 (0.097)	0.162 (0.038)	0.163 (0.018)		
M041107	M10_01	0.997 (0.063)	-0.955 (0.081)	0.139 (0.034)		
M041011	M10_02	1.244 (0.086)	-0.054 (0.051)	0.204 (0.022)		
M041122	M10_03	0.471 (0.017)	0.660 (0.037)		-0.678 (0.071)	0.678 (0.079)
M041041	M10_04	1.061 (0.096)	0.359 (0.065)	0.299 (0.024)		
M041320	M10_05	1.565 (0.112)	0.557 (0.033)	0.166 (0.015)		
M041115A	M10_06A	0.933 (0.040)	-0.176 (0.032)			
M041115B	M10_06B	1.216 (0.051)	0.240 (0.025)			
M031210	M11_01	1.225 (0.107)	0.746 (0.046)	0.221 (0.017)		
M031009	M11_02	0.939 (0.043)	0.582 (0.033)			
M031252	M11_03	1.030 (0.073)	-0.045 (0.059)	0.164 (0.024)		
M031316	M11_04	0.690 (0.034)	-1.770 (0.072)			
M031317	M11_05	1.168 (0.090)	0.741 (0.041)	0.132 (0.015)		
M031079B	M11_06B	1.226 (0.049)	-0.503 (0.029)			
M031079C	M11_06C	0.858 (0.040)	0.587 (0.035)			
M031043	M11_08	1.290 (0.094)	0.382 (0.042)	0.180 (0.018)		
M041298	M12_01	1.110 (0.078)	-0.204 (0.062)	0.205 (0.026)		
M041007	M12_02	0.761 (0.073)	0.663 (0.073)	0.171 (0.025)		
M041280	M12_03	0.833 (0.082)	0.781 (0.068)	0.191 (0.023)		
M041059	M12_04	0.875 (0.040)	0.379 (0.033)			
M041046	M12_05	1.323 (0.093)	0.499 (0.037)	0.140 (0.016)		
M041048	M12_06	1.275 (0.113)	0.672 (0.048)	0.260 (0.019)		
M031346A	M13_01A	2.108 (0.087)	-0.202 (0.018)			
M031346B	M13_01B	2.312 (0.102)	0.550 (0.017)			
M031346C	M13_01C	1.796 (0.064)	0.367 (0.014)		0.437 (0.020)	-0.437 (0.022)
M031379	M13_02	1.267 (0.060)	0.917 (0.031)			
M031380	M13_03	1.135 (0.059)	1.227 (0.043)			
M031313	M13_05	0.627 (0.031)	-1.007 (0.057)			
M031185	M13_08	1.265 (0.100)	0.297 (0.050)	0.241 (0.022)		
M041004	M14_01	0.894 (0.063)	-1.138 (0.112)	0.193 (0.046)		
M041023	M14_02	1.369 (0.087)	-0.738 (0.057)	0.185 (0.028)		

( ) Standard errors appear in parentheses

Exhibit D.5 IRT Parameters for TIMSS 2007 Fourth Grade Mathematics - Number (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M041034	M14_03	0.780 (0.052)	-0.284 (0.072)	0.093 (0.026)		
M041087	M14_04	0.809 (0.037)	0.249 (0.034)			
M041124	M14_05	0.861 (0.038)	-0.015 (0.033)			

( ) Standard errors appear in parentheses

Exhibit D.6 **IRT Parameters for TIMSS 2007 Fourth Grade Mathematics - Geometric Shapes and Measures**

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M031271	M01_09	0.678 (0.033)	-1.317 (0.061)			
M041330	M02_07	0.706 (0.073)	0.235 (0.095)	0.183 (0.033)		
M041300A	M02_08A	1.877 (0.079)	0.224 (0.018)			
M041300B	M02_08B	2.587 (0.113)	0.173 (0.015)			
M041300C	M02_08C	1.933 (0.087)	0.452 (0.019)			
M041300D	M02_08D	2.204 (0.103)	0.576 (0.018)			
M041173	M02_09	0.807 (0.102)	1.174 (0.080)	0.155 (0.021)		
M031041	M03_07	0.919 (0.043)	0.292 (0.031)			
M031350A	M03_08A	1.507 (0.068)	0.492 (0.022)			
M031350B	M03_08B	1.523 (0.066)	0.175 (0.022)			
M031350C	M03_08C	1.144 (0.058)	0.729 (0.031)			
M031274	M03_09	0.941 (0.042)	-0.210 (0.033)			
M041164	M04_06	0.852 (0.071)	-0.297 (0.097)	0.204 (0.036)		
M041146	M04_07	0.883 (0.040)	-0.182 (0.034)			
M041152	M04_08	0.940 (0.093)	0.751 (0.058)	0.191 (0.022)		
M041258A	M04_09A	1.275 (0.054)	0.042 (0.025)			
M041258B	M04_09B	1.087 (0.051)	0.481 (0.028)			
M041131	M04_10	0.851 (0.118)	1.475 (0.093)	0.187 (0.019)		
M031219	M05_06	0.751 (0.099)	0.899 (0.090)	0.268 (0.028)		
M031085	M05_08	1.070 (0.120)	0.832 (0.061)	0.294 (0.022)		
M041329	M06_07	1.132 (0.106)	0.225 (0.065)	0.338 (0.025)		
M041143	M06_08	0.457 (0.014)	-0.144 (0.033)		-1.557 (0.091)	1.557 (0.089)
M041158	M06_09	0.930 (0.074)	-0.184 (0.078)	0.202 (0.030)		
M041328	M06_10	1.025 (0.044)	-0.071 (0.029)			
M041155	M06_11	1.050 (0.079)	0.276 (0.048)	0.138 (0.020)		
M041284	M06_12	0.949 (0.039)	0.725 (0.024)		0.378 (0.033)	-0.378 (0.043)
M031038	M07_06	0.993 (0.074)	-0.382 (0.076)	0.189 (0.031)		
M031006	M07_09	1.086 (0.093)	-0.281 (0.083)	0.328 (0.032)		
M031330	M07_10	0.893 (0.038)	-0.838 (0.041)			
M031351	M07_11	1.124 (0.085)	0.405 (0.044)	0.142 (0.019)		
M041135	M08_07	0.722 (0.078)	-0.313 (0.153)	0.328 (0.046)		
M041257	M08_08	0.853 (0.043)	0.642 (0.037)			
M041268	M08_09	1.693 (0.158)	0.974 (0.036)	0.179 (0.013)		
M041151	M08_10	1.033 (0.096)	0.143 (0.074)	0.318 (0.028)		
M041264	M08_11	1.138 (0.097)	0.399 (0.052)	0.228 (0.022)		
M031297	M09_08	1.020 (0.048)	0.511 (0.030)			
M031109	M09_10	0.991 (0.092)	0.169 (0.073)	0.288 (0.027)		
M031159	M09_11	1.183 (0.090)	-0.043 (0.057)	0.230 (0.025)		
M041160A	M10_07A	1.572 (0.066)	-0.760 (0.029)			
M041160B	M10_07B	1.873 (0.083)	-0.809 (0.027)			
M041327	M10_08	0.684 (0.035)	0.195 (0.039)			
M041148	M10_09	0.478 (0.023)	0.180 (0.036)		0.259 (0.065)	-0.259 (0.066)
M041265	M10_10	1.156 (0.097)	0.744 (0.042)	0.130 (0.016)		
M031004	M11_07	1.120 (0.104)	1.030 (0.048)	0.122 (0.015)		

( ) Standard errors appear in parentheses

Exhibit D.6 IRT Parameters for TIMSS 2007 Fourth Grade Mathematics - Geometric Shapes and Measures (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M031325	M11_09	0.889 (0.046)	0.818 (0.040)			
M031088	M11_10	0.984 (0.083)	-0.173 (0.082)	0.263 (0.032)		
M031093	M11_11	0.850 (0.098)	0.713 (0.076)	0.261 (0.027)		
M041169	M12_07	1.294 (0.094)	0.309 (0.041)	0.171 (0.018)		
M041333	M12_08	1.296 (0.103)	0.715 (0.037)	0.135 (0.015)		
M041262	M12_09	1.557 (0.138)	0.761 (0.037)	0.215 (0.015)		
M041267	M12_10	0.659 (0.037)	0.743 (0.049)			
M031083	M13_06	1.110 (0.085)	-0.158 (0.065)	0.231 (0.027)		
M031071	M13_07	1.130 (0.098)	0.734 (0.045)	0.155 (0.017)		
M041302A	M14_06A	1.158 (0.076)	-0.626 (0.066)	0.161 (0.029)		
M041302B	M14_06B	0.841 (0.037)	-0.199 (0.034)			
M041302C	M14_06C	1.495 (0.061)	-0.170 (0.023)			
M041254	M14_07	1.146 (0.099)	0.490 (0.049)	0.219 (0.020)		
M041153	M14_08	1.242 (0.095)	0.412 (0.042)	0.172 (0.018)		
M041132	M14_09	0.668 (0.082)	0.839 (0.088)	0.184 (0.029)		
M041165	M14_10	0.444 (0.017)	0.640 (0.038)		-0.942 (0.078)	0.942 (0.086)

( ) Standard errors appear in parentheses

Exhibit D.7 IRT Parameters for TIMSS 2007 Fourth Grade Mathematics - Data Display

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M031134	M01_10	0.721 (0.040)	0.892 (0.050)			
M031045	M01_11	1.658 (0.115)	-0.146 (0.042)	0.202 (0.020)		
M041274	M02_10	1.080 (0.048)	-0.020 (0.029)			
M041203	M02_11	1.052 (0.050)	0.236 (0.029)			
M031240	M03_10	0.794 (0.037)	-0.207 (0.038)			
M041275	M04_11	0.610 (0.018)	-0.318 (0.029)		-1.092 (0.074)	1.092 (0.070)
M041186	M04_12	1.440 (0.097)	0.145 (0.036)	0.130 (0.016)		
M041336	M04_13	1.062 (0.102)	0.938 (0.050)	0.147 (0.017)		
M031242B	M05_04B	1.608 (0.070)	0.351 (0.020)			
M031242C	M05_04C	2.263 (0.190)	0.430 (0.031)	0.332 (0.017)		
M031172	M05_09	1.569 (0.099)	-0.030 (0.035)	0.128 (0.017)		
M041335	M06_13	0.995 (0.076)	-0.608 (0.094)	0.208 (0.037)		
M041184	M06_14	1.166 (0.085)	-0.304 (0.065)	0.190 (0.027)		
M031135	M07_12	1.302 (0.090)	-0.321 (0.056)	0.184 (0.024)		
M041182	M08_12	1.017 (0.042)	-1.083 (0.041)			
M041200	M08_13	0.947 (0.032)	-0.013 (0.021)		-0.071 (0.040)	0.071 (0.037)
M031133	M09_12	1.172 (0.047)	-0.736 (0.034)			
M041175	M10_11	1.042 (0.079)	-0.724 (0.096)	0.228 (0.039)		
M041199	M10_12	1.631 (0.104)	-0.337 (0.042)	0.144 (0.020)		
M031155	M11_12	1.263 (0.109)	0.403 (0.047)	0.251 (0.020)		
M041177	M12_11	1.160 (0.084)	-0.379 (0.066)	0.196 (0.027)		
M041271	M12_12	1.186 (0.087)	-0.100 (0.055)	0.184 (0.023)		
M041276A	M12_13A	1.317 (0.058)	0.217 (0.024)			
M041276B	M12_13B	1.129 (0.059)	0.755 (0.032)			
M041174	M14_11	1.226 (0.049)	-0.508 (0.030)			
M041191	M14_12	1.031 (0.077)	-0.990 (0.109)	0.240 (0.046)		

( ) Standard errors appear in parentheses

Exhibit D.8 IRT Parameters for TIMSS 2007 Fourth Grade Mathematics - Knowing

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M031286	M01_01	1.272 (0.056)	0.360 (0.024)			
M031271	M01_09	0.684 (0.034)	-1.222 (0.061)			
M041014	M02_01	0.862 (0.067)	-0.538 (0.098)	0.189 (0.039)		
M041278	M02_03	0.952 (0.044)	0.321 (0.030)			
M041006	M02_04	1.027 (0.087)	0.672 (0.049)	0.153 (0.019)		
M041250	M02_05	1.528 (0.065)	0.198 (0.021)			
M041173	M02_09	1.091 (0.124)	1.236 (0.064)	0.190 (0.017)		
M041052	M04_01	0.941 (0.086)	-0.298 (0.100)	0.331 (0.036)		
M041056	M04_02	1.065 (0.048)	0.286 (0.027)			
M041069	M04_03	1.646 (0.126)	0.952 (0.032)	0.083 (0.001)		
M041076	M04_04	1.106 (0.050)	0.401 (0.027)			
M041164	M04_06	1.088 (0.092)	-0.131 (0.073)	0.290 (0.030)		
M041131	M04_10	0.728 (0.010)	1.465 (0.103)	0.171 (0.022)		
M041186	M04_12	1.055 (0.072)	0.061 (0.048)	0.102 (0.020)		
M031242B	M05_04B	1.226 (0.054)	0.338 (0.025)			
M031219	M05_06	0.655 (0.084)	0.788 (0.104)	0.235 (0.032)		
M031085	M05_08	0.944 (0.102)	0.740 (0.068)	0.268 (0.024)		
M041010	M06_01	1.270 (0.109)	0.040 (0.060)	0.343 (0.025)		
M041003	M06_04	1.036 (0.046)	0.133 (0.028)			
M041104	M06_05	1.104 (0.048)	0.036 (0.027)			
M041299	M06_06	1.438 (0.069)	0.833 (0.028)			
M041329	M06_07	1.092 (0.102)	0.224 (0.068)	0.333 (0.026)		
M041143	M06_08	0.478 (0.015)	-0.117 (0.031)		-1.478 (0.087)	1.478 (0.085)
M041335	M06_13	0.936 (0.066)	-0.724 (0.089)	0.157 (0.036)		
M031029	M07_01	1.412 (0.132)	0.586 (0.047)	0.313 (0.019)		
M031030	M07_02	0.952 (0.057)	1.476 (0.064)			
M031332	M07_03	1.128 (0.094)	0.128 (0.060)	0.269 (0.025)		
M031038	M07_06	0.933 (0.068)	-0.446 (0.079)	0.164 (0.033)		
M031276	M07_07	1.517 (0.124)	0.401 (0.041)	0.267 (0.019)		
M031006	M07_09	0.793 (0.071)	-0.529 (0.126)	0.257 (0.045)		
M031330	M07_10	0.856 (0.038)	-0.852 (0.043)			
M041291	M08_01	0.753 (0.035)	-0.617 (0.042)			
M041068	M08_03	1.542 (0.103)	0.699 (0.029)	0.075 (0.011)		
M041065A	M08_04A	1.596 (0.126)	0.794 (0.033)	0.144 (0.013)		
M041096	M08_05	1.095 (0.076)	0.394 (0.040)	0.099 (0.017)		
M041135	M08_07	0.664 (0.072)	-0.396 (0.170)	0.303 (0.051)		
M041257	M08_08	0.883 (0.044)	0.636 (0.036)			
M031128	M09_01	0.610 (0.030)	-1.031 (0.059)			
M031294	M09_07	1.503 (0.095)	0.109 (0.034)	0.132 (0.017)		
M031109	M09_10	0.992 (0.087)	0.124 (0.071)	0.261 (0.028)		
M031159	M09_11	1.309 (0.097)	-0.008 (0.050)	0.228 (0.023)		
M041011	M10_02	1.407 (0.098)	-0.025 (0.044)	0.207 (0.022)		
M041122	M10_03	0.511 (0.019)	0.619 (0.034)		-0.622 (0.065)	0.622 (0.072)
M041041	M10_04	1.212 (0.108)	0.346 (0.055)	0.303 (0.023)		

( ) Standard errors appear in parentheses

Exhibit D.8 IRT Parameters for TIMSS 2007 Fourth Grade Mathematics - Knowing (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M041320	M10_05	1.996 (0.139)	0.488 (0.026)	0.161 (0.014)		
M041160A	M10_07A	1.088 (0.046)	-0.887 (0.038)			
M041148	M10_09	0.443 (0.022)	0.199 (0.038)		0.250 (0.069)	-0.250 (0.071)
M041175	M10_11	1.054 (0.072)	-0.763 (0.080)	0.156 (0.035)		
M031210	M11_01	1.361 (0.113)	0.632 (0.041)	0.207 (0.018)		
M031316	M11_04	0.736 (0.036)	-1.628 (0.068)			
M031317	M11_05	1.216 (0.091)	0.647 (0.038)	0.118 (0.015)		
M031093	M11_11	0.721 (0.086)	0.726 (0.095)	0.247 (0.031)		
M041298	M12_01	1.693 (0.108)	-0.180 (0.037)	0.177 (0.020)		
M041007	M12_02	0.888 (0.083)	0.625 (0.062)	0.183 (0.023)		
M041280	M12_03	0.940 (0.088)	0.681 (0.058)	0.185 (0.022)		
M041059	M12_04	1.320 (0.057)	0.316 (0.023)			
M041046	M12_05	1.778 (0.113)	0.384 (0.026)	0.115 (0.013)		
M041169	M12_07	1.054 (0.079)	0.279 (0.049)	0.147 (0.021)		
M041276A	M12_13A	0.974 (0.045)	0.234 (0.029)			
M031083	M13_06	1.137 (0.088)	-0.120 (0.063)	0.238 (0.027)		
M031071	M13_07	1.146 (0.098)	0.748 (0.045)	0.158 (0.017)		
M041004	M14_01	1.131 (0.083)	-0.894 (0.089)	0.245 (0.039)		
M041023	M14_02	1.974 (0.133)	-0.541 (0.041)	0.224 (0.023)		
M041034	M14_03	1.044 (0.070)	-0.150 (0.055)	0.128 (0.023)		
M041087	M14_04	0.906 (0.042)	0.234 (0.031)			
M041302A	M14_06A	1.054 (0.074)	-0.590 (0.076)	0.188 (0.033)		
M041254	M14_07	1.066 (0.093)	0.487 (0.053)	0.212 (0.022)		
M041174	M14_11	1.134 (0.048)	-0.480 (0.031)			

( ) Standard errors appear in parentheses

Exhibit D.9 IRT Parameters for TIMSS 2007 Fourth Grade Mathematics - Applying

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M031299	M01_07	1.378 (0.058)	0.062 (0.023)			
M031301	M01_08	1.230 (0.050)	-0.441 (0.028)			
M031134	M01_10	0.670 (0.039)	0.899 (0.055)			
M041039	M02_02	0.928 (0.086)	0.185 (0.074)	0.252 (0.028)		
M041094	M02_06	1.164 (0.126)	0.958 (0.053)	0.225 (0.018)		
M041330	M02_07	0.877 (0.074)	0.182 (0.067)	0.160 (0.026)		
M041300A	M02_08A	1.369 (0.059)	0.259 (0.023)			
M041300B	M02_08B	1.491 (0.064)	0.190 (0.022)			
M041274	M02_10	0.884 (0.041)	-0.058 (0.033)			
M031050	M03_03	1.396 (0.134)	0.731 (0.044)	0.266 (0.018)		
M031334	M03_05	1.718 (0.170)	0.885 (0.037)	0.255 (0.015)		
M031255	M03_06	1.210 (0.115)	0.444 (0.056)	0.317 (0.022)		
M031041	M03_07	1.036 (0.047)	0.289 (0.028)			
M031350A	M03_08A	1.255 (0.058)	0.524 (0.026)			
M031350C	M03_08C	0.974 (0.050)	0.795 (0.037)			
M031274	M03_09	0.968 (0.043)	-0.201 (0.033)			
M031240	M03_10	0.782 (0.037)	-0.178 (0.038)			
M041281	M04_05	1.428 (0.089)	-0.083 (0.040)	0.138 (0.019)		
M041146	M04_07	0.822 (0.037)	-0.248 (0.036)			
M041152	M04_08	1.333 (0.112)	0.646 (0.041)	0.196 (0.017)		
M041275	M04_11	0.524 (0.016)	-0.315 (0.032)		-1.366 (0.085)	1.366 (0.081)
M031303	M05_01	1.674 (0.122)	-0.076 (0.044)	0.298 (0.022)		
M031309	M05_02	1.375 (0.056)	-0.130 (0.024)			
M031245	M05_03	2.095 (0.172)	0.971 (0.027)	0.113 (0.010)		
M031242A	M05_04A	1.040 (0.044)	-0.087 (0.028)			
M031173	M05_07	1.985 (0.119)	-0.002 (0.029)	0.141 (0.016)		
M031172	M05_09	1.626 (0.098)	-0.029 (0.034)	0.127 (0.017)		
M041098	M06_02	1.868 (0.162)	0.693 (0.033)	0.254 (0.015)		
M041064	M06_03	0.838 (0.037)	-0.283 (0.035)			
M041158	M06_09	1.061 (0.081)	-0.085 (0.066)	0.214 (0.027)		
M041328	M06_10	0.947 (0.042)	-0.029 (0.030)			
M041155	M06_11	1.183 (0.087)	0.343 (0.043)	0.150 (0.019)		
M031098	M07_04	1.970 (0.139)	0.421 (0.028)	0.191 (0.015)		
M031254	M07_05	1.481 (0.117)	0.424 (0.039)	0.233 (0.018)		
M031351	M07_11	1.046 (0.082)	0.411 (0.048)	0.141 (0.020)		
M041289	M08_02	1.171 (0.096)	0.400 (0.048)	0.216 (0.020)		
M041125	M08_06	1.385 (0.119)	0.801 (0.038)	0.166 (0.015)		
M041268	M08_09	1.870 (0.170)	0.939 (0.032)	0.179 (0.012)		
M041264	M08_11	1.053 (0.101)	0.503 (0.059)	0.267 (0.022)		
M041182	M08_12	1.015 (0.043)	-1.114 (0.041)			
M041200	M08_13	0.686 (0.023)	-0.089 (0.026)		-0.260 (0.053)	0.260 (0.050)
M031183	M09_03	0.869 (0.033)	0.268 (0.023)		0.568 (0.036)	-0.568 (0.038)
M031187	M09_05	0.919 (0.077)	-0.332 (0.094)	0.257 (0.035)		
M031251	M09_06	1.854 (0.146)	0.597 (0.031)	0.210 (0.015)		

( ) Standard errors appear in parentheses

Exhibit D.9 IRT Parameters for TIMSS 2007 Fourth Grade Mathematics - Applying (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M031297	M09_08	1.052 (0.049)	0.514 (0.030)			
M031218	M09_09	1.588 (0.112)	0.230 (0.036)	0.198 (0.018)		
M031133	M09_12	1.106 (0.046)	-0.770 (0.035)			
M041107	M10_01	1.090 (0.074)	-0.820 (0.082)	0.180 (0.036)		
M041115A	M10_06A	0.922 (0.040)	-0.172 (0.032)			
M041160B	M10_07B	1.299 (0.054)	-0.951 (0.034)			
M041327	M10_08	0.673 (0.034)	0.197 (0.040)			
M031009	M11_02	1.135 (0.052)	0.522 (0.027)			
M031252	M11_03	1.224 (0.086)	0.006 (0.050)	0.181 (0.022)		
M031079B	M11_06B	1.277 (0.051)	-0.466 (0.028)			
M031004	M11_07	1.147 (0.103)	0.993 (0.044)	0.113 (0.014)		
M031043	M11_08	1.538 (0.109)	0.360 (0.035)	0.181 (0.017)		
M031325	M11_09	0.859 (0.045)	0.839 (0.041)			
M031088	M11_10	1.067 (0.085)	-0.184 (0.073)	0.249 (0.030)		
M031155	M11_12	1.610 (0.135)	0.471 (0.039)	0.274 (0.019)		
M041333	M12_08	1.398 (0.108)	0.705 (0.034)	0.132 (0.014)		
M041262	M12_09	1.375 (0.129)	0.818 (0.042)	0.218 (0.016)		
M041267	M12_10	0.604 (0.035)	0.803 (0.055)			
M041177	M12_11	1.257 (0.087)	-0.301 (0.058)	0.206 (0.026)		
M031346A	M13_01A	1.169 (0.047)	-0.284 (0.028)			
M031313	M13_05	0.692 (0.032)	-0.957 (0.051)			
M041124	M14_05	0.981 (0.042)	-0.021 (0.029)			
M041153	M14_08	1.192 (0.093)	0.411 (0.044)	0.174 (0.019)		
M041132	M14_09	0.624 (0.077)	0.797 (0.095)	0.168 (0.031)		
M041165	M14_10	0.406 (0.015)	0.663 (0.042)		-1.061 (0.085)	1.061 (0.094)
M041191	M14_12	1.053 (0.084)	-0.874 (0.107)	0.284 (0.044)		

( ) Standard errors appear in parentheses

Exhibit D.10 IRT Parameters for TIMSS 2007 Fourth Grade Mathematics - Reasoning

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M031106	M01_02	0.871 (0.039)	0.263 (0.032)			
M031282	M01_03	0.913 (0.030)	0.860 (0.024)		-0.613 (0.049)	0.613 (0.055)
M031227	M01_04	1.291 (0.070)	1.209 (0.039)			
M031335	M01_05	1.365 (0.100)	0.157 (0.044)	0.235 (0.020)		
M031068	M01_06	1.474 (0.063)	0.372 (0.022)			
M031045	M01_11	1.335 (0.092)	-0.235 (0.050)	0.172 (0.023)		
M041300C	M02_08C	1.215 (0.056)	0.570 (0.027)			
M041300D	M02_08D	1.264 (0.060)	0.737 (0.029)			
M041203	M02_11	0.954 (0.045)	0.237 (0.032)			
M031235	M03_01	0.858 (0.041)	0.561 (0.035)			
M031285	M03_02	0.904 (0.045)	0.845 (0.039)			
M031258	M03_04	1.054 (0.051)	0.826 (0.034)			
M031350B	M03_08B	1.022 (0.046)	0.171 (0.029)			
M041258A	M04_09A	1.464 (0.062)	0.046 (0.023)			
M041258B	M04_09B	1.125 (0.051)	0.478 (0.027)			
M041336	M04_13	1.132 (0.111)	1.045 (0.050)	0.166 (0.017)		
M031242C	M05_04C	1.209 (0.110)	0.497 (0.056)	0.326 (0.022)		
M031247	M05_05	0.608 (0.027)	1.285 (0.046)		-0.210 (0.054)	0.210 (0.075)
M041284	M06_12	0.799 (0.034)	0.831 (0.029)		0.387 (0.038)	-0.387 (0.050)
M041184	M06_14	1.189 (0.083)	-0.338 (0.061)	0.174 (0.026)		
M031064	M07_08	1.179 (0.103)	0.768 (0.047)	0.205 (0.018)		
M031135	M07_12	1.370 (0.093)	-0.302 (0.051)	0.178 (0.023)		
M041065B	M08_04B	1.072 (0.058)	1.243 (0.046)			
M041151	M08_10	0.838 (0.078)	-0.003 (0.095)	0.262 (0.033)		
M031016	M09_02	1.195 (0.056)	0.815 (0.030)			
M041115B	M10_06B	1.102 (0.048)	0.221 (0.027)			
M041265	M10_10	0.982 (0.081)	0.752 (0.048)	0.114 (0.017)		
M041199	M10_12	1.485 (0.010)	-0.325 (0.047)	0.171 (0.022)		
M031079C	M11_06C	0.821 (0.040)	0.582 (0.037)			
M041048	M12_06	1.102 (0.103)	0.690 (0.055)	0.257 (0.020)		
M041271	M12_12	0.947 (0.071)	-0.152 (0.070)	0.158 (0.028)		
M041276B	M12_13B	0.771 (0.043)	0.926 (0.048)			
M031346B	M13_01B	2.957 (0.130)	0.535 (0.015)			
M031346C	M13_01C	1.636 (0.060)	0.386 (0.015)		0.415 (0.021)	-0.415 (0.024)
M031379	M13_02	1.526 (0.070)	0.860 (0.026)			
M031380	M13_03	1.277 (0.066)	1.158 (0.037)			
M031185	M13_08	1.227 (0.097)	0.321 (0.050)	0.244 (0.022)		
M041302B	M14_06B	0.637 (0.031)	-0.228 (0.043)			
M041302C	M14_06C	1.003 (0.042)	-0.220 (0.031)			

( ) Standard errors appear in parentheses

Exhibit D.11 IRT Parameters for TIMSS 2007 Fourth Grade Science - Life Science

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S031193	S01_04	0.772 (0.082)	-0.013 (0.106)	0.276 (0.037)		
S031264	S01_05	0.881 (0.068)	-0.485 (0.084)	0.158 (0.032)		
S031347	S01_06	0.830 (0.080)	-0.417 (0.115)	0.290 (0.040)		
S031346	S01_07	0.900 (0.059)	1.166 (0.059)			
S041007	S02_01	0.813 (0.083)	0.121 (0.090)	0.259 (0.034)		
S041164	S02_02	1.662 (0.138)	0.645 (0.031)	0.173 (0.016)		
S041018	S02_03	0.865 (0.035)	0.049 (0.021)		0.349 (0.038)	-0.349 (0.035)
S041160	S02_04	0.697 (0.118)	1.498 (0.131)	0.184 (0.025)		
S041042	S02_05	0.754 (0.065)	-0.972 (0.143)	0.211 (0.048)		
S031229	S03_01	1.673 (0.156)	0.580 (0.036)	0.268 (0.019)		
S031270	S03_02	0.690 (0.054)	1.566 (0.102)			
S031026	S03_03	0.680 (0.024)	0.034 (0.024)		-0.442 (0.052)	0.442 (0.050)
S031319	S03_04	1.558 (0.197)	1.090 (0.051)	0.224 (0.015)		
S041165	S04_01	0.780 (0.102)	0.786 (0.082)	0.250 (0.028)		
S041023	S04_02	1.297 (0.060)	0.297 (0.022)			
S041047	S04_03	0.569 (0.055)	-0.799 (0.174)	0.175 (0.050)		
S041001	S04_04	0.450 (0.024)	0.399 (0.039)		0.149 (0.067)	-0.149 (0.074)
S041029	S04_05	0.975 (0.045)	-0.481 (0.035)			
S041179	S04_08	0.960 (0.053)	0.639 (0.034)			
S031255	S05_01	1.024 (0.086)	-0.071 (0.068)	0.244 (0.029)		
S031240D	S05_02D	0.783 (0.030)	0.005 (0.025)		0.724 (0.042)	-0.724 (0.038)
S031239	S05_03	1.146 (0.130)	0.190 (0.081)	0.468 (0.029)		
S031235A	S05_04A	1.896 (0.084)	0.400 (0.017)			
S031235B	S05_04B	2.095 (0.094)	0.504 (0.016)			
S041003	S06_03	0.756 (0.041)	0.249 (0.035)			
S041224	S06_04	1.123 (0.046)	0.504 (0.018)		0.305 (0.027)	-0.305 (0.032)
S041163	S06_05	0.786 (0.120)	1.148 (0.095)	0.240 (0.026)		
S041039	S06_06	0.977 (0.048)	0.135 (0.028)			
S041014	S06_07	1.638 (0.256)	1.342 (0.070)	0.233 (0.014)		
S041181	S06_08	0.821 (0.041)	-0.214 (0.036)			
S041174	S06_09	1.017 (0.053)	0.549 (0.030)			
S031317	S07_01	1.022 (0.105)	0.053 (0.084)	0.387 (0.031)		
S031190	S07_02	1.267 (0.064)	0.602 (0.026)			
S031431	S07_03	0.911 (0.139)	1.474 (0.106)	0.163 (0.019)		
S031283	S07_04	0.685 (0.075)	-0.322 (0.146)	0.288 (0.045)		
S031426	S07_05	1.084 (0.087)	-0.008 (0.060)	0.222 (0.027)		
S041009	S08_01	1.018 (0.087)	-0.489 (0.091)	0.316 (0.035)		
S041223	S08_02	1.201 (0.108)	0.244 (0.055)	0.289 (0.025)		
S041026	S08_03	1.119 (0.108)	0.567 (0.050)	0.224 (0.022)		
S041177	S08_04	0.560 (0.030)	1.051 (0.048)		0.377 (0.049)	-0.377 (0.072)
S041183	S08_05	0.742 (0.028)	0.332 (0.027)		0.892 (0.041)	-0.892 (0.046)
S041008	S08_06	1.513 (0.135)	0.582 (0.037)	0.224 (0.019)		
S031340	S09_01	1.047 (0.119)	0.797 (0.057)	0.227 (0.022)		
S031236	S09_02	1.037 (0.078)	-0.545 (0.077)	0.212 (0.032)		

( ) Standard errors appear in parentheses

Exhibit D.11 IRT Parameters for TIMSS 2007 Fourth Grade Science - Life Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S031361	S09_04	0.988 (0.102)	0.487 (0.062)	0.244 (0.026)		
S031001	S09_05	1.352 (0.091)	-0.413 (0.053)	0.203 (0.026)		
S041178	S10_02	0.894 (0.111)	0.810 (0.070)	0.249 (0.025)		
S041182	S10_03	0.764 (0.044)	0.513 (0.039)			
S041180	S10_04	1.709 (0.129)	0.293 (0.034)	0.229 (0.020)		
S041013A	S10_06A	0.733 (0.043)	0.529 (0.040)			
S041013B	S10_06B	0.753 (0.050)	1.112 (0.064)			
S031254	S11_01	0.670 (0.099)	0.572 (0.120)	0.340 (0.036)		
S031266	S11_02	1.371 (0.098)	0.188 (0.038)	0.170 (0.020)		
S031233	S11_03	0.810 (0.041)	-0.176 (0.035)			
S031281	S11_07	1.189 (0.089)	-0.792 (0.080)	0.250 (0.034)		
S041027	S12_01	0.866 (0.040)	-1.301 (0.057)			
S041043	S12_02	0.729 (0.038)	-0.245 (0.039)			
S041006	S12_05	0.602 (0.028)	0.489 (0.030)		0.137 (0.048)	-0.137 (0.055)
S041301	S12_07	0.869 (0.050)	0.718 (0.039)			
S041033	S12_09	1.175 (0.064)	0.849 (0.035)			
S031356	S13_01	1.171 (0.102)	-1.016 (0.111)	0.381 (0.041)		
S031291	S13_02	1.523 (0.100)	-0.665 (0.053)	0.197 (0.027)		
S031230	S13_03	0.848 (0.068)	-1.047 (0.127)	0.213 (0.045)		
S031325	S13_04	0.833 (0.046)	0.449 (0.034)			
S031390D	S13_11D	1.149 (0.046)	0.254 (0.017)		0.255 (0.029)	-0.255 (0.029)
S041010	S14_01	1.467 (0.097)	-0.482 (0.051)	0.211 (0.026)		
S041034	S14_02	0.645 (0.065)	-0.404 (0.138)	0.208 (0.044)		
S041017	S14_03	0.856 (0.111)	0.915 (0.074)	0.232 (0.025)		
S041124	S14_04	1.327 (0.155)	0.895 (0.050)	0.258 (0.019)		
S041037	S14_06	0.625 (0.025)	0.241 (0.025)		-0.250 (0.051)	0.250 (0.053)
S041032	S14_11	1.185 (0.055)	-0.972 (0.043)			

( ) Standard errors appear in parentheses

Exhibit D.12 IRT Parameters for TIMSS 2007 Fourth Grade Science - Physical Science

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S031446A	S01_01A	1.730 (0.081)	0.575 (0.021)			
S031446B	S01_01B	1.334 (0.068)	0.740 (0.029)			
S031446C	S01_01C	0.823 (0.042)	0.014 (0.033)			
S031445A	S01_02A	2.717 (0.116)	0.388 (0.014)			
S031445B	S01_02B	1.370 (0.061)	-0.467 (0.028)			
S031447	S01_03	0.552 (0.029)	0.931 (0.044)		0.490 (0.050)	-0.490 (0.071)
S041079	S02_06	0.934 (0.069)	-0.780 (0.089)	0.163 (0.033)		
S041073	S02_07	0.681 (0.037)	0.064 (0.038)			
S041217	S02_08	0.927 (0.080)	0.265 (0.058)	0.169 (0.024)		
S041196	S02_09	0.645 (0.070)	-0.295 (0.141)	0.242 (0.044)		
S041211	S02_10	0.887 (0.044)	-0.041 (0.031)			
S041051	S02_11	0.994 (0.147)	1.145 (0.083)	0.272 (0.022)		
S031414A	S03_05A	3.273 (0.140)	-0.028 (0.012)			
S031414B	S03_05B	2.932 (0.125)	-0.060 (0.013)			
S031078	S03_06	0.727 (0.078)	0.185 (0.096)	0.231 (0.034)		
S031009	S03_07	0.727 (0.039)	-0.133 (0.038)			
S041054	S04_06	0.548 (0.058)	-1.317 (0.269)	0.265 (0.072)		
S041308	S04_07	0.829 (0.094)	0.624 (0.074)	0.237 (0.027)		
S041216	S04_11	0.668 (0.039)	0.399 (0.041)			
S041061	S04_12	0.897 (0.045)	-0.108 (0.032)			
S031205	S05_05	0.556 (0.074)	0.312 (0.144)	0.237 (0.043)		
S031399A	S05_06A	3.148 (0.133)	0.279 (0.012)			
S031399B	S05_06B	2.890 (0.122)	0.122 (0.013)			
S041117	S06_01	0.503 (0.046)	-2.274 (0.332)	0.265 (0.091)		
S041120	S06_02	0.729 (0.149)	1.423 (0.157)	0.410 (0.026)		
S041049	S06_10	1.021 (0.098)	0.514 (0.055)	0.212 (0.023)		
S041060	S06_12	0.945 (0.060)	1.160 (0.057)			
S031422	S07_06	0.934 (0.080)	-0.963 (0.120)	0.279 (0.042)		
S031427	S07_07	0.841 (0.072)	-0.020 (0.071)	0.167 (0.028)		
S031075	S07_08	0.577 (0.107)	0.928 (0.156)	0.383 (0.037)		
S041195	S08_08	0.977 (0.059)	1.005 (0.048)			
S041134A	S08_09A	1.709 (0.075)	0.386 (0.019)			
S041134B	S08_09B	1.694 (0.073)	0.156 (0.019)			
S041134C	S08_09C	0.974 (0.089)	0.514 (0.052)	0.165 (0.022)		
S041191	S08_10	0.662 (0.095)	0.938 (0.010)	0.230 (0.030)		
S031410	S09_07	0.529 (0.072)	0.087 (0.176)	0.256 (0.048)		
S031421	S09_08	0.634 (0.034)	-0.423 (0.047)			
S031298	S09_09	1.133 (0.198)	1.543 (0.115)	0.253 (0.017)		
S031076	S09_10	1.007 (0.053)	0.499 (0.030)			
S041311	S10_01	0.748 (0.061)	-2.003 (0.213)	0.252 (0.075)		
S041187	S10_05	1.084 (0.190)	1.632 (0.128)	0.209 (0.016)		
S041067	S10_07	0.988 (0.046)	-0.222 (0.031)			
S041305	S10_08	1.409 (0.127)	0.618 (0.038)	0.200 (0.018)		
S041048	S10_09	1.038 (0.051)	0.293 (0.027)			

( ) Standard errors appear in parentheses

Exhibit D.12 IRT Parameters for TIMSS 2007 Fourth Grade Science - Physical Science (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S041069	S10_11	1.235 (0.114)	0.333 (0.054)	0.296 (0.024)		
S031204	S11_04	0.760 (0.045)	0.572 (0.041)			
S031273	S11_05	1.784 (0.138)	0.345 (0.033)	0.235 (0.019)		
S031299	S11_06	0.723 (0.042)	0.517 (0.041)			
S031077	S11_08	1.145 (0.099)	-0.088 (0.068)	0.329 (0.029)		
S031311	S11_09	1.363 (0.108)	0.213 (0.045)	0.240 (0.023)		
S041050	S12_03	0.680 (0.085)	0.721 (0.088)	0.204 (0.029)		
S041070	S12_04	1.173 (0.094)	0.366 (0.044)	0.182 (0.021)		
S041052	S12_06	1.238 (0.102)	-0.133 (0.063)	0.332 (0.028)		
S041080	S12_08	0.812 (0.118)	1.394 (0.104)	0.169 (0.020)		
S041077	S12_11	1.152 (0.055)	0.344 (0.025)			
S031068	S13_05	1.198 (0.010)	0.486 (0.043)	0.173 (0.020)		
S031418	S13_06	1.200 (0.104)	0.664 (0.041)	0.145 (0.018)		
S031197D	S13_07D	0.645 (0.023)	-0.457 (0.031)		-0.377 (0.061)	0.377 (0.051)
S031371	S13_08	1.133 (0.105)	0.575 (0.048)	0.198 (0.021)		
S041186	S14_05	0.652 (0.044)	1.128 (0.069)			
S041119	S14_07	0.642 (0.084)	-0.141 (0.171)	0.383 (0.045)		
S041193	S14_09	0.494 (0.053)	-0.552 (0.197)	0.174 (0.053)		
S041068	S14_12	0.726 (0.041)	0.386 (0.038)			
S041303	S14_13	0.545 (0.080)	0.784 (0.126)	0.209 (0.038)		

( ) Standard errors appear in parentheses

Exhibit D.13 IRT Parameters for TIMSS 2007 Fourth Grade Science - Earth Science

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S031081	S01_08	0.668 (0.034)	-0.574 (0.050)			
S041089	S02_12	1.350 (0.108)	0.277 (0.045)	0.268 (0.022)		
S041156A	S02_13A	3.261 (0.222)	0.106 (0.021)	0.241 (0.018)		
S041156B	S02_13B	1.652 (0.076)	0.283 (0.019)			
S031401	S03_08	1.552 (0.158)	0.708 (0.042)	0.291 (0.019)		
S031384A	S03_09A	1.265 (0.057)	-0.985 (0.044)			
S031384B	S03_09B	1.329 (0.066)	-0.274 (0.029)			
S041087	S04_09	1.908 (0.153)	0.563 (0.030)	0.222 (0.017)		
S041205	S04_10	1.507 (0.127)	0.458 (0.041)	0.276 (0.020)		
S041202	S04_13	0.544 (0.020)	0.662 (0.031)		-1.164 (0.073)	1.164 (0.079)
S041215	S04_14	1.837 (0.193)	0.849 (0.035)	0.235 (0.017)		
S031393	S05_07	0.700 (0.033)	-1.347 (0.068)			
S031278	S05_08	0.566 (0.032)	-0.595 (0.059)			
S041208	S06_11	0.592 (0.128)	1.399 (0.176)	0.379 (0.030)		
S041201A	S06_13A	2.492 (0.106)	0.262 (0.014)			
S041201B	S06_13B	2.385 (0.105)	0.353 (0.015)			
S031047	S07_09	0.727 (0.039)	0.028 (0.037)			
S031387	S07_10	1.199 (0.166)	1.176 (0.065)	0.197 (0.019)		
S031396D	S07_11D	0.502 (0.018)	-0.912 (0.047)		-0.466 (0.085)	0.466 (0.067)
S041107	S08_11	0.555 (0.019)	-0.415 (0.034)		-0.847 (0.074)	0.847 (0.065)
S041113	S08_12	1.091 (0.056)	0.384 (0.026)			
S031391D	S09_03D	0.630 (0.028)	0.487 (0.028)		0.034 (0.048)	-0.034 (0.053)
S031275	S09_11	1.144 (0.169)	1.288 (0.077)	0.206 (0.018)		
S041110	S10_10	0.835 (0.042)	-0.147 (0.035)			
S041100	S10_12	1.563 (0.125)	0.539 (0.035)	0.213 (0.018)		
S041092	S10_13	1.205 (0.117)	0.464 (0.055)	0.326 (0.023)		
S031088D	S11_10D	0.613 (0.020)	0.283 (0.035)		1.299 (0.052)	-1.299 (0.056)
S031389	S11_11	1.003 (0.112)	0.881 (0.059)	0.204 (0.021)		
S041209	S12_12	1.163 (0.105)	0.637 (0.046)	0.199 (0.020)		
S041081	S12_13	0.653 (0.026)	0.552 (0.027)		-0.402 (0.051)	0.402 (0.056)
S041102	S12_14	1.448 (0.116)	0.203 (0.046)	0.305 (0.023)		
S031376	S13_09	1.068 (0.114)	0.854 (0.054)	0.206 (0.020)		
S031044	S13_10	1.082 (0.054)	0.211 (0.026)			
S041105	S14_08	1.336 (0.099)	0.209 (0.042)	0.235 (0.021)		
S041149D	S14_10D	0.685 (0.025)	0.781 (0.028)		-0.784 (0.057)	0.784 (0.064)

( ) Standard errors appear in parentheses

Exhibit D.14 IRT Parameters for TIMSS 2007 Fourth Grade Science - Knowing

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S041164	S02_02	1.360 (0.126)	0.718 (0.041)	0.194 (0.018)		
S041018	S02_03	0.746 (0.032)	0.008 (0.024)		0.342 (0.043)	-0.342 (0.040)
S041160	S02_04	0.663 (0.118)	1.568 (0.148)	0.190 (0.026)		
S041042	S02_05	0.723 (0.062)	-1.044 (0.147)	0.202 (0.048)		
S041196	S02_09	0.774 (0.075)	-0.191 (0.104)	0.246 (0.038)		
S041089	S02_12	1.087 (0.092)	0.150 (0.057)	0.219 (0.027)		
S031229	S03_01	1.836 (0.185)	0.683 (0.038)	0.316 (0.018)		
S031319	S03_04	1.415 (0.179)	1.125 (0.056)	0.222 (0.017)		
S031414A	S03_05A	3.337 (0.143)	0.003 (0.012)			
S031414B	S03_05B	3.218 (0.138)	-0.024 (0.012)			
S031384A	S03_09A	0.789 (0.041)	-1.169 (0.063)			
S041165	S04_01	0.734 (0.097)	0.809 (0.090)	0.239 (0.031)		
S041023	S04_02	1.163 (0.056)	0.336 (0.024)			
S041308	S04_07	1.159 (0.118)	0.599 (0.053)	0.258 (0.024)		
S041202	S04_13	0.596 (0.021)	0.618 (0.029)		-1.066 (0.067)	1.066 (0.072)
S031255	S05_01	1.058 (0.084)	-0.111 (0.064)	0.224 (0.029)		
S031240D	S05_02D	0.877 (0.033)	0.009 (0.023)		0.685 (0.038)	-0.685 (0.034)
S031239	S05_03	0.802 (0.093)	-0.054 (0.124)	0.397 (0.039)		
S031205	S05_05	0.832 (0.088)	0.322 (0.082)	0.250 (0.032)		
S031393	S05_07	0.919 (0.047)	-1.007 (0.054)			
S031278	S05_08	0.729 (0.041)	-0.428 (0.046)			
S041117	S06_01	0.610 (0.063)	-1.781 (0.288)	0.324 (0.079)		
S041120	S06_02	1.245 (0.237)	1.238 (0.092)	0.442 (0.019)		
S041003	S06_03	0.760 (0.042)	0.226 (0.034)			
S041224	S06_04	1.126 (0.046)	0.477 (0.018)		0.306 (0.027)	-0.306 (0.032)
S041014	S06_07	1.744 (0.246)	1.251 (0.055)	0.227 (0.014)		
S041181	S06_08	0.883 (0.044)	-0.215 (0.033)			
S041208	S06_11	0.682 (0.122)	1.077 (0.127)	0.350 (0.033)		
S031317	S07_01	0.953 (0.103)	0.055 (0.094)	0.392 (0.034)		
S031283	S07_04	0.732 (0.073)	-0.367 (0.126)	0.260 (0.043)		
S031422	S07_06	1.084 (0.087)	-0.743 (0.092)	0.284 (0.038)		
S031396D	S07_11D	0.619 (0.024)	-0.699 (0.038)		-0.386 (0.069)	0.386 (0.054)
S041009	S08_01	0.996 (0.083)	-0.522 (0.091)	0.294 (0.036)		
S041223	S08_02	1.248 (0.112)	0.251 (0.053)	0.292 (0.025)		
S041026	S08_03	0.997 (0.096)	0.536 (0.056)	0.203 (0.024)		
S041008	S08_06	1.374 (0.119)	0.550 (0.040)	0.205 (0.020)		
S041191	S08_10	0.875 (0.112)	0.862 (0.075)	0.246 (0.026)		
S041107	S08_11	0.589 (0.020)	-0.375 (0.030)		-0.836 (0.069)	0.836 (0.062)
S031340	S09_01	1.115 (0.122)	0.787 (0.055)	0.234 (0.021)		
S031236	S09_02	0.977 (0.071)	-0.636 (0.078)	0.174 (0.031)		
S031391D	S09_03D	0.647 (0.029)	0.457 (0.028)		0.011 (0.046)	-0.011 (0.052)
S031410	S09_07	0.637 (0.073)	0.024 (0.126)	0.238 (0.040)		
S031421	S09_08	0.682 (0.037)	-0.398 (0.044)			
S041311	S10_01	0.830 (0.077)	-1.665 (0.199)	0.331 (0.064)		

( ) Standard errors appear in parentheses

Exhibit D.14 IRT Parameters for TIMSS 2007 Fourth Grade Science - Knowing (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S041178	S10_02	1.013 (0.120)	0.797 (0.064)	0.263 (0.024)		
S041187	S10_05	1.220 (0.215)	1.499 (0.104)	0.208 (0.016)		
S041067	S10_07	1.113 (0.052)	-0.175 (0.027)			
S041110	S10_10	0.959 (0.047)	-0.114 (0.030)			
S041100	S10_12	1.506 (0.122)	0.466 (0.035)	0.192 (0.019)		
S041092	S10_13	1.127 (0.112)	0.366 (0.060)	0.295 (0.027)		
S031254	S11_01	0.631 (0.096)	0.588 (0.134)	0.335 (0.039)		
S031233	S11_03	0.850 (0.043)	-0.147 (0.033)			
S031299	S11_06	0.841 (0.047)	0.494 (0.035)			
S031281	S11_07	1.232 (0.088)	-0.770 (0.073)	0.225 (0.033)		
S031311	S11_09	1.393 (0.108)	0.184 (0.043)	0.216 (0.023)		
S031088D	S11_10D	0.773 (0.026)	0.272 (0.028)		1.064 (0.041)	-1.064 (0.045)
S031389	S11_11	1.125 (0.123)	0.820 (0.054)	0.208 (0.022)		
S041027	S12_01	0.952 (0.045)	-1.175 (0.053)			
S041043	S12_02	0.730 (0.039)	-0.225 (0.038)			
S041070	S12_04	1.093 (0.089)	0.354 (0.047)	0.164 (0.022)		
S041052	S12_06	0.950 (0.083)	-0.265 (0.086)	0.277 (0.035)		
S041209	S12_12	0.954 (0.098)	0.660 (0.059)	0.192 (0.024)		
S041081	S12_13	0.696 (0.027)	0.526 (0.025)		-0.386 (0.048)	0.386 (0.053)
S031356	S13_01	1.188 (0.102)	-0.967 (0.105)	0.381 (0.041)		
S031230	S13_03	0.992 (0.077)	-0.873 (0.099)	0.235 (0.039)		
S031197D	S13_07D	0.676 (0.024)	-0.396 (0.029)		-0.383 (0.057)	0.383 (0.049)
S031376	S13_09	1.019 (0.113)	0.843 (0.059)	0.200 (0.023)		
S041010	S14_01	1.528 (0.100)	-0.444 (0.049)	0.210 (0.027)		
S041034	S14_02	0.742 (0.069)	-0.310 (0.110)	0.218 (0.039)		
S041186	S14_05	0.793 (0.050)	0.988 (0.054)			
S041105	S14_08	1.199 (0.084)	0.045 (0.044)	0.144 (0.023)		
S041032	S14_11	1.341 (0.064)	-0.858 (0.038)			
S041068	S14_12	0.966 (0.051)	0.360 (0.029)			
S041303	S14_13	0.726 (0.085)	0.608 (0.085)	0.193 (0.032)		

( ) Standard errors appear in parentheses

Exhibit D.15 IRT Parameters for TIMSS 2007 Fourth Grade Science - Applying

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S031193	S01_04	0.641 (0.069)	-0.216 (0.130)	0.228 (0.040)		
S031264	S01_05	0.845 (0.067)	-0.531 (0.089)	0.167 (0.031)		
S031347	S01_06	0.672 (0.068)	-0.662 (0.153)	0.247 (0.045)		
S031081	S01_08	0.730 (0.038)	-0.513 (0.046)			
S041007	S02_01	0.749 (0.078)	0.085 (0.096)	0.249 (0.034)		
S041073	S02_07	0.740 (0.041)	0.126 (0.035)			
S041211	S02_10	1.060 (0.052)	0.042 (0.027)			
S041051	S02_11	1.218 (0.163)	1.028 (0.060)	0.264 (0.020)		
S041156A	S02_13A	2.273 (0.148)	0.061 (0.028)	0.211 (0.020)		
S041156B	S02_13B	1.674 (0.076)	0.284 (0.019)			
S031009	S03_07	0.833 (0.044)	-0.085 (0.034)			
S031401	S03_08	1.562 (0.146)	0.600 (0.039)	0.256 (0.020)		
S031384B	S03_09B	0.945 (0.048)	-0.392 (0.038)			
S041047	S04_03	0.524 (0.058)	-0.724 (0.210)	0.215 (0.055)		
S041001	S04_04	0.450 (0.024)	0.431 (0.039)		0.143 (0.067)	-0.143 (0.074)
S041054	S04_06	0.699 (0.074)	-0.911 (0.186)	0.343 (0.051)		
S041087	S04_09	1.454 (0.117)	0.489 (0.036)	0.185 (0.019)		
S041205	S04_10	1.052 (0.087)	0.267 (0.052)	0.189 (0.024)		
S041216	S04_11	0.869 (0.047)	0.359 (0.032)			
S041061	S04_12	1.067 (0.052)	-0.067 (0.028)			
S041215	S04_14	1.240 (0.131)	0.809 (0.047)	0.193 (0.020)		
S031235A	S05_04A	1.458 (0.068)	0.435 (0.021)			
S031235B	S05_04B	1.534 (0.072)	0.555 (0.021)			
S031399A	S05_06A	1.661 (0.074)	0.328 (0.018)			
S031399B	S05_06B	1.505 (0.069)	0.143 (0.020)			
S041163	S06_05	0.795 (0.133)	1.248 (0.107)	0.263 (0.026)		
S041039	S06_06	0.955 (0.048)	0.131 (0.028)			
S041174	S06_09	0.916 (0.051)	0.570 (0.033)			
S041060	S06_12	1.175 (0.070)	1.047 (0.043)			
S041201A	S06_13A	1.858 (0.082)	0.281 (0.017)			
S041201B	S06_13B	1.710 (0.080)	0.375 (0.019)			
S031427	S07_07	0.906 (0.081)	0.160 (0.068)	0.216 (0.027)		
S031075	S07_08	0.678 (0.121)	0.940 (0.128)	0.394 (0.033)		
S031387	S07_10	1.121 (0.147)	1.126 (0.064)	0.177 (0.020)		
S041177	S08_04	0.519 (0.029)	1.104 (0.053)		0.382 (0.053)	-0.382 (0.078)
S041183	S08_05	0.811 (0.030)	0.315 (0.026)		0.857 (0.038)	-0.857 (0.042)
S041195	S08_08	0.974 (0.060)	1.025 (0.048)			
S041113	S08_12	1.069 (0.055)	0.361 (0.027)			
S031001	S09_05	1.269 (0.091)	-0.369 (0.058)	0.236 (0.027)		
S031298	S09_09	1.145 (0.227)	1.528 (0.122)	0.255 (0.018)		
S031275	S09_11	1.044 (0.146)	1.221 (0.076)	0.184 (0.019)		
S041182	S10_03	0.816 (0.046)	0.501 (0.036)			
S041180	S10_04	1.581 (0.118)	0.291 (0.036)	0.224 (0.020)		
S041013A	S10_06A	0.744 (0.044)	0.529 (0.040)			

( ) Standard errors appear in parentheses

Exhibit D.15 IRT Parameters for TIMSS 2007 Fourth Grade Science - Applying (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S041013B	S10_06B	0.748 (0.051)	1.121 (0.065)			
S041069	S10_11	1.251 (0.117)	0.369 (0.053)	0.303 (0.024)		
S031204	S11_04	0.788 (0.047)	0.570 (0.039)			
S031273	S11_05	1.570 (0.120)	0.334 (0.036)	0.219 (0.020)		
S031077	S11_08	1.038 (0.093)	-0.109 (0.075)	0.318 (0.030)		
S041006	S12_05	0.576 (0.028)	0.503 (0.031)		0.132 (0.051)	-0.132 (0.058)
S041301	S12_07	0.833 (0.049)	0.740 (0.041)			
S041033	S12_09	1.233 (0.066)	0.834 (0.033)			
S041077	S12_11	1.006 (0.052)	0.379 (0.028)			
S041102	S12_14	1.161 (0.102)	0.144 (0.059)	0.285 (0.027)		
S031291	S13_02	1.615 (0.112)	-0.568 (0.052)	0.244 (0.026)		
S031325	S13_04	0.833 (0.046)	0.455 (0.034)			
S031418	S13_06	1.044 (0.097)	0.721 (0.047)	0.143 (0.019)		
S031371	S13_08	1.071 (0.102)	0.604 (0.050)	0.197 (0.022)		
S031390D	S13_11D	1.058 (0.043)	0.265 (0.018)		0.249 (0.031)	-0.249 (0.031)
S041017	S14_03	1.036 (0.128)	0.900 (0.062)	0.253 (0.022)		
S041124	S14_04	1.276 (0.160)	0.955 (0.054)	0.269 (0.019)		
S041119	S14_07	0.922 (0.102)	0.037 (0.098)	0.422 (0.032)		
S041193	S14_09	0.561 (0.058)	-0.490 (0.158)	0.177 (0.045)		

( ) Standard errors appear in parentheses

Exhibit D.16 IRT Parameters for TIMSS 2007 Fourth Grade Science - Reasoning

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S031446A	S01_01A	1.572 (0.074)	0.662 (0.023)			
S031446B	S01_01B	1.244 (0.063)	0.833 (0.030)			
S031446C	S01_01C	0.771 (0.037)	0.043 (0.035)			
S031445A	S01_02A	2.449 (0.107)	0.463 (0.015)			
S031445B	S01_02B	1.270 (0.051)	-0.502 (0.029)			
S031447	S01_03	0.539 (0.027)	1.002 (0.045)		0.513 (0.051)	-0.513 (0.073)
S031346	S01_07	0.761 (0.050)	1.331 (0.070)			
S041079	S02_06	0.926 (0.067)	-0.741 (0.099)	0.181 (0.041)		
S041217	S02_08	0.998 (0.093)	0.422 (0.059)	0.212 (0.024)		
S031270	S03_02	0.676 (0.049)	1.634 (0.097)			
S031026	S03_03	0.654 (0.021)	0.030 (0.025)		-0.424 (0.054)	0.424 (0.052)
S031078	S03_06	0.770 (0.081)	0.195 (0.099)	0.234 (0.035)		
S041029	S04_05	0.777 (0.035)	-0.593 (0.041)			
S041179	S04_08	0.902 (0.046)	0.695 (0.036)			
S041049	S06_10	1.143 (0.102)	0.508 (0.048)	0.203 (0.020)		
S031190	S07_02	1.074 (0.053)	0.630 (0.031)			
S031431	S07_03	0.755 (0.127)	1.647 (0.132)	0.159 (0.020)		
S031426	S07_05	0.861 (0.081)	-0.056 (0.094)	0.245 (0.034)		
S031047	S07_09	0.670 (0.034)	-0.011 (0.040)			
S041134A	S08_09A	1.554 (0.068)	0.454 (0.020)			
S041134B	S08_09B	1.459 (0.062)	0.193 (0.022)			
S041134C	S08_09C	0.910 (0.088)	0.582 (0.058)	0.165 (0.023)		
S031361	S09_04	0.762 (0.101)	0.716 (0.091)	0.274 (0.031)		
S031076	S09_10	0.836 (0.043)	0.586 (0.036)			
S041305	S10_08	1.136 (0.109)	0.724 (0.047)	0.197 (0.019)		
S041048	S10_09	0.992 (0.045)	0.337 (0.028)			
S031266	S11_02	1.424 (0.099)	0.216 (0.038)	0.167 (0.018)		
S041050	S12_03	0.724 (0.101)	0.901 (0.089)	0.246 (0.029)		
S041080	S12_08	0.866 (0.128)	1.432 (0.095)	0.176 (0.019)		
S031068	S13_05	1.541 (0.139)	0.660 (0.037)	0.242 (0.017)		
S031044	S13_10	0.800 (0.039)	0.208 (0.034)			
S041037	S14_06	0.521 (0.020)	0.273 (0.030)		-0.325 (0.061)	0.325 (0.063)
S041149D	S14_10D	0.481 (0.018)	0.977 (0.041)		-1.220 (0.080)	1.220 (0.090)

( ) Standard errors appear in parentheses

Exhibit D.17 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Number

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M022043	M01_01	0.643 (0.049)	-0.274 (0.125)	0.156 (0.045)		
M022046	M01_02	0.847 (0.031)	-0.266 (0.029)			
M022057	M01_06	0.542 (0.064)	0.304 (0.186)	0.285 (0.050)		
M022066	M01_09	1.164 (0.063)	0.370 (0.035)	0.083 (0.014)		
M022097	M03_01	1.121 (0.062)	0.092 (0.042)	0.112 (0.019)		
M022104	M03_03	0.993 (0.057)	-0.085 (0.053)	0.117 (0.024)		
M022106	M03_05	0.970 (0.038)	0.889 (0.033)			
M022110	M03_07	0.651 (0.028)	-0.421 (0.037)			
M022232	M01_10	0.642 (0.021)	1.707 (0.042)		-1.994 (0.109)	1.994 (0.121)
M022234B	M01_11B	0.823 (0.026)	1.392 (0.029)		-1.696 (0.097)	1.696 (0.103)
M032064	M09_03	1.557 (0.058)	0.815 (0.022)			
M032094	M09_01	1.302 (0.082)	0.158 (0.044)	0.227 (0.020)		
M032142	M05_01	1.197 (0.148)	1.335 (0.064)	0.406 (0.014)		
M032160	M07_03	2.167 (0.188)	1.394 (0.030)	0.139 (0.008)		
M032166	M13_01	0.918 (0.061)	0.155 (0.060)	0.155 (0.025)		
M032307	M03_09	1.385 (0.055)	1.108 (0.028)			
M032381	M07_01	1.057 (0.038)	0.463 (0.025)			
M032416	M07_02	1.209 (0.078)	0.985 (0.035)	0.092 (0.011)		
M032523	M03_10	1.893 (0.158)	1.267 (0.033)	0.168 (0.009)		
M032525	M03_13	0.927 (0.059)	0.180 (0.057)	0.135 (0.024)		
M032529	M07_12	1.927 (0.143)	1.104 (0.029)	0.160 (0.009)		
M032595	M13_08	1.148 (0.066)	0.346 (0.037)	0.096 (0.016)		
M032626	M13_07	0.665 (0.060)	0.538 (0.091)	0.166 (0.031)		
M032662	M09_02	1.769 (0.153)	1.442 (0.036)	0.126 (0.008)		
M032701	M03_11	1.289 (0.075)	-0.880 (0.062)	0.180 (0.037)		
M032704	M03_12	1.212 (0.071)	-0.093 (0.046)	0.164 (0.023)		
M032725	M11_02	0.930 (0.039)	1.114 (0.038)			
M032755	M05_06	0.853 (0.029)	1.499 (0.033)		-0.404 (0.046)	0.404 (0.062)
M042001	M04_01	0.764 (0.057)	-0.252 (0.104)	0.199 (0.041)		
M042002	M10_04	0.812 (0.034)	1.064 (0.041)			
M042003	M02_01	0.788 (0.060)	0.049 (0.091)	0.216 (0.034)		
M042015	M12_01	0.747 (0.056)	-0.248 (0.107)	0.199 (0.041)		
M042016	M10_03	0.961 (0.078)	0.760 (0.055)	0.216 (0.020)		
M042018	M02_03	0.965 (0.035)	0.597 (0.028)			
M042019	M08_03	0.916 (0.035)	0.687 (0.030)			
M042022	M04_02	0.954 (0.098)	1.105 (0.063)	0.283 (0.019)		
M042023	M08_04	1.161 (0.042)	0.617 (0.025)			
M042024	M10_02	1.548 (0.085)	0.284 (0.030)	0.137 (0.014)		
M042031	M06_02	1.426 (0.010)	0.698 (0.037)	0.224 (0.014)		
M042032	M06_01	0.921 (0.071)	-0.162 (0.092)	0.311 (0.035)		
M042039	M02_05	0.718 (0.049)	0.513 (0.062)	0.085 (0.022)		
M042041	M10_01	1.285 (0.081)	-0.192 (0.053)	0.257 (0.025)		
M042052	M14_04	1.133 (0.063)	0.049 (0.042)	0.115 (0.019)		
M042055	M02_04	1.263 (0.110)	0.973 (0.047)	0.289 (0.015)		

( ) Standard errors appear in parentheses

Exhibit D.17 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Number (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M042059	M06_04	0.894 (0.024)	0.253 (0.018)		-0.108 (0.033)	0.108 (0.035)
M042060	M08_02	1.383 (0.079)	0.317 (0.034)	0.145 (0.015)		
M042079	M02_02	1.106 (0.080)	-0.083 (0.069)	0.320 (0.028)		
M042081	M14_02	0.937 (0.038)	0.919 (0.034)			
M042114A	M12_04A	3.182 (0.122)	-0.002 (0.012)			
M042114B	M12_04B	3.456 (0.137)	0.247 (0.011)			
M042182	M14_01	1.105 (0.086)	0.327 (0.059)	0.310 (0.022)		
M042183	M08_01	0.753 (0.053)	-0.057 (0.087)	0.151 (0.034)		
M042186	M06_03	1.033 (0.037)	0.500 (0.026)			
M042194	M12_03	1.060 (0.038)	-0.319 (0.024)			
M042196	M12_02	0.899 (0.051)	0.255 (0.045)	0.071 (0.018)		
M042197	M08_05	1.180 (0.046)	0.997 (0.029)			
M042302A	M14_06A	1.429 (0.041)	0.498 (0.014)		0.003 (0.023)	-0.003 (0.026)
M042302B	M14_06B	1.597 (0.044)	0.600 (0.013)		-0.226 (0.026)	0.226 (0.028)
M042302C	M14_06C	0.642 (0.021)	1.440 (0.039)		-0.590 (0.054)	0.590 (0.071)
M042304A	M04_05A	1.552 (0.053)	-0.243 (0.018)			
M042304B	M04_05B	1.304 (0.042)	0.798 (0.017)		0.431 (0.022)	-0.431 (0.030)
M042304C	M04_05C	2.070 (0.081)	0.857 (0.019)			
M042304D	M04_05D	0.699 (0.017)	0.428 (0.021)		-1.091 (0.055)	1.091 (0.058)

( ) Standard errors appear in parentheses

Exhibit D.18 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Algebra

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M022050	M01_04	0.956 (0.084)	1.408 (0.057)	0.169 (0.013)		
M032047	M11_09	1.098 (0.134)	1.341 (0.072)	0.436 (0.015)		
M032163	M05_10	1.244 (0.099)	0.895 (0.043)	0.218 (0.014)		
M032198	M05_02	1.002 (0.071)	0.544 (0.050)	0.187 (0.019)		
M032273	M07_04	0.777 (0.058)	-0.141 (0.097)	0.222 (0.035)		
M032295	M11_05	1.202 (0.068)	-0.551 (0.055)	0.172 (0.027)		
M032352	M11_01	1.008 (0.095)	0.692 (0.069)	0.390 (0.020)		
M032419	M09_04	1.134 (0.096)	1.067 (0.048)	0.241 (0.014)		
M032424	M11_12	0.938 (0.067)	0.662 (0.051)	0.160 (0.018)		
M032477	M09_05	1.280 (0.095)	0.852 (0.041)	0.224 (0.014)		
M032538	M09_06	1.220 (0.042)	0.428 (0.023)			
M032540	M07_05	1.115 (0.088)	0.444 (0.058)	0.327 (0.020)		
M032640	M05_03	0.483 (0.017)	1.888 (0.060)		-0.943 (0.074)	0.943 (0.101)
M032673	M13_09	1.203 (0.088)	0.715 (0.043)	0.219 (0.016)		
M032683	M11_03	0.542 (0.015)	1.004 (0.031)		-1.394 (0.071)	1.394 (0.079)
M032698	M07_06	0.783 (0.060)	0.767 (0.061)	0.134 (0.020)		
M032738	M11_04	1.123 (0.064)	-0.212 (0.050)	0.151 (0.023)		
M032757	M13_03	0.500 (0.012)	-0.211 (0.026)		-1.889 (0.084)	1.889 (0.082)
M032760A	M13_04A	1.107 (0.030)	0.814 (0.017)		-0.898 (0.050)	0.898 (0.053)
M032760B	M13_04B	1.947 (0.083)	1.130 (0.022)			
M032760C	M13_04C	2.166 (0.106)	1.364 (0.024)			
M032761	M13_05	1.160 (0.041)	1.455 (0.026)		-0.181 (0.036)	0.181 (0.048)
M042049	M14_03	0.835 (0.062)	0.035 (0.085)	0.243 (0.031)		
M042050	M12_07	0.943 (0.037)	0.899 (0.033)			
M042066	M08_07	0.619 (0.026)	0.506 (0.040)			
M042067	M10_08	1.460 (0.139)	1.295 (0.045)	0.269 (0.012)		
M042074A	M12_08A	1.401 (0.051)	0.711 (0.023)			
M042074B	M12_08B	1.347 (0.051)	0.879 (0.025)			
M042074C	M12_08C	2.099 (0.093)	1.178 (0.022)			
M042076	M14_05	1.053 (0.076)	0.691 (0.048)	0.204 (0.017)		
M042077	M10_06	1.597 (0.106)	0.510 (0.034)	0.239 (0.014)		
M042080A	M08_11A	0.842 (0.032)	0.656 (0.033)			
M042080B	M08_11B	1.216 (0.055)	1.542 (0.041)			
M042082	M04_03	1.629 (0.107)	0.886 (0.029)	0.142 (0.010)		
M042086	M06_08	1.184 (0.044)	0.783 (0.027)			
M042088	M04_04	1.303 (0.089)	0.567 (0.040)	0.224 (0.015)		
M042093	M14_10	1.323 (0.060)	1.484 (0.036)			
M042100	M14_07	1.040 (0.069)	0.281 (0.053)	0.205 (0.021)		
M042103	M06_07	1.027 (0.044)	1.341 (0.040)			
M042109	M12_06	1.095 (0.107)	1.375 (0.056)	0.237 (0.014)		
M042112	M12_05	0.395 (0.055)	0.800 (0.241)	0.236 (0.054)		
M042198A	M10_05A	1.172 (0.041)	-0.579 (0.025)			
M042198B	M10_05B	1.118 (0.040)	0.495 (0.025)			
M042198C	M10_05C	1.804 (0.080)	1.256 (0.025)			

( ) Standard errors appear in parentheses

Exhibit D.18 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Algebra (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M042199	M02_06	1.555 (0.147)	0.756 (0.048)	0.450 (0.015)		
M042202	M14_08	1.151 (0.090)	0.776 (0.048)	0.258 (0.016)		
M042226	M06_06	1.515 (0.052)	0.370 (0.020)			
M042228	M06_09	0.668 (0.029)	0.897 (0.044)			
M042229A	M08_10A	1.970 (0.076)	0.839 (0.019)			
M042229B	M08_10B	2.167 (0.087)	0.903 (0.019)			
M042234	M08_06	1.227 (0.074)	0.426 (0.038)	0.154 (0.015)		
M042235	M10_07	1.476 (0.079)	0.372 (0.029)	0.111 (0.012)		
M042236	M06_05	1.338 (0.085)	0.341 (0.039)	0.214 (0.016)		
M042238	M04_08	1.167 (0.101)	1.365 (0.048)	0.171 (0.012)		
M042239	M04_07	1.685 (0.109)	1.076 (0.028)	0.086 (0.008)		
M042240	M14_09	1.133 (0.063)	0.297 (0.039)	0.109 (0.016)		
M042243	M08_08	1.550 (0.080)	0.551 (0.026)	0.072 (0.009)		
M042245	M06_10	2.023 (0.161)	1.268 (0.029)	0.132 (0.008)		
M042248	M08_09	1.416 (0.053)	0.880 (0.024)			
M042263	M02_08	0.799 (0.036)	1.530 (0.054)			
M042267	M04_06	1.301 (0.076)	0.764 (0.031)	0.088 (0.010)		
M042301A	M02_07A	1.061 (0.037)	0.083 (0.024)			
M042301B	M02_07B	2.358 (0.092)	0.693 (0.016)			
M042301C	M02_07C	2.621 (0.122)	1.151 (0.019)			

( ) Standard errors appear in parentheses

Exhibit D.19 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Geometry

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M022049	M01_03	0.804 (0.089)	0.704 (0.092)	0.364 (0.027)		
M022055	M01_05	1.429 (0.051)	0.673 (0.022)			
M022062	M01_08	1.099 (0.076)	0.813 (0.039)	0.119 (0.014)		
M022105	M03_04	0.735 (0.069)	0.998 (0.067)	0.148 (0.022)		
M022108	M03_06	0.946 (0.069)	0.326 (0.062)	0.207 (0.024)		
M022234A	M01_11A	0.858 (0.025)	0.945 (0.023)		-0.489 (0.043)	0.489 (0.050)
M022243	M01_12	1.246 (0.046)	0.592 (0.025)			
M032097	M07_07	1.588 (0.154)	1.383 (0.040)	0.192 (0.001)		
M032100	M09_09	0.988 (0.066)	0.560 (0.046)	0.130 (0.018)		
M032116	M09_08	1.447 (0.134)	1.044 (0.041)	0.294 (0.013)		
M032205	M05_09	0.645 (0.068)	0.455 (0.122)	0.252 (0.038)		
M032294	M07_10	1.036 (0.074)	0.163 (0.062)	0.252 (0.025)		
M032324	M09_07	1.340 (0.104)	0.932 (0.038)	0.196 (0.013)		
M032331	M11_06	2.522 (0.247)	1.356 (0.028)	0.204 (0.008)		
M032344	M05_04	1.391 (0.050)	0.657 (0.022)			
M032397	M09_12	1.285 (0.110)	1.015 (0.042)	0.228 (0.014)		
M032398	M11_10	2.144 (0.197)	1.055 (0.031)	0.313 (0.011)		
M032402	M09_10	0.892 (0.093)	0.936 (0.069)	0.293 (0.021)		
M032414	M07_09	1.173 (0.043)	0.699 (0.026)			
M032575	M07_08	2.346 (0.168)	0.752 (0.025)	0.224 (0.011)		
M032579	M03_14	1.289 (0.075)	-0.048 (0.044)	0.184 (0.021)		
M032623	M11_07	2.288 (0.160)	0.910 (0.023)	0.165 (0.009)		
M032679	M11_08	1.414 (0.109)	0.652 (0.041)	0.291 (0.016)		
M032691	M03_15	0.962 (0.035)	0.407 (0.027)			
M032692	M13_06	0.591 (0.016)	1.165 (0.032)		-1.386 (0.071)	1.386 (0.081)
M032734	M09_11	0.770 (0.029)	-0.183 (0.030)			
M032754	M05_05	0.679 (0.028)	-0.299 (0.034)			
M042036	M04_10	1.650 (0.114)	0.931 (0.029)	0.141 (0.001)		
M042120	M08_12	1.050 (0.076)	0.077 (0.067)	0.287 (0.026)		
M042130	M04_11	0.484 (0.024)	-0.022 (0.045)			
M042132	M12_10	1.659 (0.154)	1.306 (0.036)	0.202 (0.010)		
M042137	M02_10	1.299 (0.101)	0.823 (0.041)	0.229 (0.015)		
M042148	M02_11	0.901 (0.072)	0.040 (0.087)	0.300 (0.031)		
M042150	M10_09	0.946 (0.093)	1.134 (0.056)	0.213 (0.017)		
M042151	M12_09	0.965 (0.034)	0.060 (0.026)			
M042152	M06_13	0.924 (0.089)	1.014 (0.059)	0.227 (0.019)		
M042201	M06_12	1.322 (0.045)	0.219 (0.021)			
M042203	M08_13	1.695 (0.098)	0.319 (0.030)	0.172 (0.014)		
M042257	M12_11	0.817 (0.074)	0.954 (0.062)	0.171 (0.021)		
M042264	M08_14	0.833 (0.038)	1.365 (0.050)			
M042265	M02_09	0.977 (0.076)	0.690 (0.052)	0.191 (0.019)		
M042268	M14_12	1.533 (0.117)	1.148 (0.032)	0.125 (0.009)		
M042270	M06_11	0.924 (0.033)	0.069 (0.026)			
M042271	M14_11	1.057 (0.072)	0.547 (0.045)	0.156 (0.018)		

( ) Standard errors appear in parentheses

Exhibit D.19 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Geometry (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M042279	M04_09	1.061 (0.064)	0.262 (0.045)	0.131 (0.019)		
M042300A	M10_10A	1.780 (0.060)	0.201 (0.017)			
M042300B	M10_10B	1.851 (0.064)	0.408 (0.017)			

( ) Standard errors appear in parentheses

Exhibit D.20 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Data and Chance

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M022101	M03_02	1.308 (0.078)	0.007 (0.043)	0.195 (0.021)		
M022181	M03_08	1.115 (0.075)	-0.420 (0.073)	0.273 (0.034)		
M022257	M01_07	1.349 (0.110)	0.770 (0.041)	0.263 (0.015)		
M032132	M09_14	0.723 (0.056)	0.560 (0.070)	0.120 (0.025)		
M032507	M11_11	1.527 (0.140)	1.276 (0.039)	0.176 (0.010)		
M032637A	M07_13A	0.986 (0.036)	-0.199 (0.026)			
M032637B	M07_13B	2.177 (0.081)	-0.040 (0.016)			
M032637C	M07_13C	2.014 (0.076)	0.443 (0.017)			
M032681A	M11_13A	0.657 (0.028)	-0.296 (0.036)			
M032681B	M11_13B	0.674 (0.031)	0.949 (0.046)			
M032681C	M11_13C	1.572 (0.058)	0.634 (0.020)			
M032688	M07_11	0.835 (0.035)	0.955 (0.038)			
M032695	M09_13	0.568 (0.015)	0.065 (0.024)		-0.954 (0.059)	0.954 (0.059)
M032721	M13_02	0.809 (0.092)	1.438 (0.079)	0.246 (0.019)		
M032753A	M05_07A	3.229 (0.123)	0.823 (0.001)		0.127 (0.015)	-0.127 (0.016)
M032753B	M05_07B	3.621 (0.152)	0.945 (0.001)		0.183 (0.013)	-0.183 (0.016)
M032753C	M05_07C	1.487 (0.054)	0.582 (0.021)			
M032756	M05_08	0.823 (0.033)	0.479 (0.032)			
M042158	M12_12	0.638 (0.067)	0.418 (0.129)	0.272 (0.039)		
M042159	M14_13	0.640 (0.028)	-0.491 (0.038)			
M042164	M14_14	1.505 (0.057)	0.747 (0.022)			
M042167	M14_15	1.385 (0.057)	0.926 (0.026)			
M042169A	M10_12A	1.247 (0.044)	0.461 (0.022)			
M042169B	M10_12B	0.655 (0.033)	1.362 (0.061)			
M042169C	M10_12C	1.505 (0.079)	1.547 (0.041)			
M042177	M06_16	1.058 (0.081)	0.521 (0.053)	0.229 (0.021)		
M042179	M06_15	0.980 (0.082)	0.726 (0.056)	0.229 (0.021)		
M042207	M06_17	0.408 (0.010)	0.226 (0.030)		-3.173 (0.125)	3.173 (0.126)
M042220	M02_14	0.817 (0.021)	0.603 (0.020)		-1.231 (0.060)	1.231 (0.063)
M042222	M04_13	1.041 (0.067)	0.494 (0.041)	0.097 (0.016)		
M042224	M08_16	0.966 (0.035)	0.161 (0.026)			
M042250	M02_13	1.427 (0.050)	-0.495 (0.021)			
M042252	M12_13	1.552 (0.115)	0.893 (0.032)	0.165 (0.012)		
M042254	M02_12	0.822 (0.057)	-0.968 (0.126)	0.217 (0.056)		
M042255	M08_15	0.694 (0.053)	0.001 (0.010)	0.153 (0.037)		
M042260	M10_11	1.046 (0.080)	0.264 (0.062)	0.268 (0.025)		
M042261	M12_14	1.039 (0.081)	0.467 (0.056)	0.237 (0.022)		
M042269	M06_14	0.851 (0.086)	0.645 (0.084)	0.336 (0.027)		
M042303A	M04_12A	0.882 (0.034)	0.459 (0.030)			
M042303B	M04_12B	0.370 (0.016)	0.935 (0.051)		-0.037 (0.067)	0.037 (0.084)

( ) Standard errors appear in parentheses

Exhibit D.21 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Knowing

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M022043	M01_01	0.614 (0.048)	-0.347 (0.130)	0.151 (0.044)		
M022050	M01_04	1.112 (0.084)	1.025 (0.041)	0.128 (0.013)		
M022066	M01_09	1.982 (0.104)	0.344 (0.022)	0.106 (0.010)		
M022097	M03_01	1.111 (0.061)	0.077 (0.040)	0.104 (0.017)		
M022101	M03_02	0.963 (0.060)	-0.094 (0.059)	0.154 (0.025)		
M022104	M03_03	1.080 (0.061)	-0.065 (0.047)	0.126 (0.021)		
M022105	M03_04	0.679 (0.064)	0.979 (0.070)	0.132 (0.023)		
M022110	M03_07	0.614 (0.026)	-0.460 (0.039)			
M032094	M09_01	1.350 (0.088)	0.212 (0.042)	0.252 (0.018)		
M032132	M09_14	0.659 (0.052)	0.508 (0.074)	0.108 (0.026)		
M032166	M13_01	1.151 (0.074)	0.232 (0.045)	0.190 (0.019)		
M032198	M05_02	1.407 (0.103)	0.635 (0.037)	0.247 (0.015)		
M032295	M11_05	1.746 (0.110)	-0.242 (0.040)	0.300 (0.020)		
M032397	M09_12	1.335 (0.109)	0.980 (0.040)	0.225 (0.013)		
M032416	M07_02	1.584 (0.098)	0.879 (0.027)	0.010 (0.009)		
M032419	M09_04	1.441 (0.115)	0.914 (0.037)	0.243 (0.013)		
M032477	M09_05	1.839 (0.131)	0.747 (0.029)	0.236 (0.012)		
M032525	M03_13	1.159 (0.076)	0.309 (0.044)	0.196 (0.019)		
M032538	M09_06	1.380 (0.048)	0.380 (0.020)			
M032540	M07_05	1.462 (0.117)	0.503 (0.043)	0.366 (0.017)		
M032626	M13_07	1.023 (0.077)	0.560 (0.050)	0.209 (0.019)		
M032637B	M07_13B	0.965 (0.037)	-0.231 (0.028)			
M032673	M13_09	1.699 (0.115)	0.609 (0.030)	0.221 (0.013)		
M032679	M11_08	1.132 (0.081)	0.480 (0.046)	0.225 (0.018)		
M032681A	M11_13A	0.574 (0.026)	-0.338 (0.041)			
M032683	M11_03	0.676 (0.018)	0.876 (0.025)		-1.078 (0.058)	1.078 (0.064)
M032698	M07_06	1.012 (0.075)	0.710 (0.045)	0.158 (0.017)		
M032725	M11_02	1.313 (0.053)	0.954 (0.027)			
M032734	M09_11	0.762 (0.029)	-0.182 (0.031)			
M032738	M11_04	1.634 (0.097)	-0.005 (0.035)	0.234 (0.018)		
M032753C	M05_07C	0.677 (0.031)	0.828 (0.044)			
M042001	M04_01	0.771 (0.058)	-0.263 (0.101)	0.214 (0.037)		
M042003	M02_01	0.789 (0.058)	-0.051 (0.084)	0.184 (0.032)		
M042015	M12_01	1.387 (0.088)	-0.023 (0.045)	0.265 (0.021)		
M042019	M08_03	0.875 (0.034)	0.639 (0.032)			
M042022	M04_02	1.084 (0.097)	0.910 (0.050)	0.262 (0.016)		
M042024	M10_02	1.540 (0.091)	0.372 (0.030)	0.172 (0.014)		
M042032	M06_01	0.955 (0.072)	-0.195 (0.083)	0.298 (0.032)		
M042049	M14_03	1.024 (0.075)	0.043 (0.064)	0.270 (0.025)		
M042050	M12_07	1.286 (0.049)	0.768 (0.024)			
M042052	M14_04	1.864 (0.099)	0.090 (0.025)	0.142 (0.013)		
M042059	M06_04	0.667 (0.019)	0.274 (0.022)		-0.313 (0.043)	0.313 (0.045)
M042060	M08_02	1.371 (0.083)	0.333 (0.034)	0.174 (0.015)		
M042076	M14_05	1.252 (0.083)	0.501 (0.038)	0.184 (0.016)		

( ) Standard errors appear in parentheses

Exhibit D.21 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Knowing (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M042077	M10_06	1.544 (0.100)	0.458 (0.033)	0.223 (0.015)		
M042079	M02_02	1.362 (0.104)	0.109 (0.054)	0.393 (0.021)		
M042080A	M08_11A	0.831 (0.033)	0.582 (0.032)			
M042080B	M08_11B	1.294 (0.060)	1.361 (0.037)			
M042081	M14_02	1.059 (0.042)	0.844 (0.030)			
M042082	M04_03	2.079 (0.146)	0.877 (0.025)	0.174 (0.001)		
M042088	M04_04	1.416 (0.092)	0.462 (0.035)	0.206 (0.015)		
M042100	M14_07	1.309 (0.091)	0.336 (0.042)	0.259 (0.018)		
M042103	M06_07	1.102 (0.049)	1.251 (0.037)			
M042114A	M12_04A	1.110 (0.039)	0.012 (0.023)			
M042120	M08_12	1.007 (0.078)	0.095 (0.068)	0.303 (0.025)		
M042148	M02_11	1.083 (0.084)	0.137 (0.064)	0.328 (0.024)		
M042150	M10_09	0.981 (0.088)	1.064 (0.052)	0.202 (0.016)		
M042152	M06_13	0.711 (0.075)	1.041 (0.076)	0.197 (0.024)		
M042158	M12_12	0.803 (0.083)	0.614 (0.087)	0.341 (0.027)		
M042169A	M10_12A	1.157 (0.042)	0.464 (0.023)			
M042169B	M10_12B	0.403 (0.028)	1.892 (0.126)			
M042183	M08_01	0.716 (0.054)	-0.093 (0.094)	0.165 (0.034)		
M042194	M12_03	1.127 (0.039)	-0.290 (0.024)			
M042196	M12_02	1.093 (0.063)	0.324 (0.038)	0.101 (0.016)		
M042198A	M10_05A	1.039 (0.036)	-0.580 (0.027)			
M042199	M02_06	1.589 (0.010)	0.122 (0.036)	0.258 (0.017)		
M042222	M04_13	0.913 (0.063)	0.578 (0.046)	0.103 (0.017)		
M042224	M08_16	0.914 (0.034)	0.091 (0.027)			
M042226	M06_06	1.700 (0.060)	0.381 (0.017)			
M042229B	M08_10B	2.007 (0.079)	0.813 (0.019)			
M042234	M08_06	1.625 (0.099)	0.414 (0.030)	0.185 (0.013)		
M042235	M10_07	1.870 (0.100)	0.374 (0.023)	0.119 (0.011)		
M042236	M06_05	1.836 (0.118)	0.409 (0.029)	0.242 (0.014)		
M042239	M04_07	1.791 (0.121)	1.030 (0.027)	0.102 (0.008)		
M042243	M08_08	2.050 (0.109)	0.497 (0.021)	0.090 (0.009)		
M042248	M08_09	1.542 (0.059)	0.767 (0.022)			
M042250	M02_13	1.052 (0.037)	-0.549 (0.027)			
M042260	M10_11	0.982 (0.079)	0.335 (0.065)	0.295 (0.024)		
M042261	M12_14	0.847 (0.068)	0.387 (0.068)	0.201 (0.025)		
M042301A	M02_07A	0.446 (0.023)	0.109 (0.048)			
M042303A	M04_12A	1.114 (0.041)	0.442 (0.024)			

( ) Standard errors appear in parentheses

Exhibit D.22 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Applying

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M022046	M01_02	0.831 (0.031)	-0.296 (0.029)			
M022055	M01_05	1.303 (0.048)	0.709 (0.024)			
M022057	M01_06	0.553 (0.066)	0.335 (0.178)	0.300 (0.048)		
M022062	M01_08	1.011 (0.074)	0.889 (0.043)	0.127 (0.015)		
M022106	M03_05	0.984 (0.038)	0.885 (0.032)			
M022108	M03_06	0.976 (0.074)	0.430 (0.060)	0.241 (0.023)		
M022181	M03_08	1.103 (0.071)	-0.495 (0.071)	0.242 (0.034)		
M022232	M01_10	0.639 (0.021)	1.685 (0.042)		-2.003 (0.109)	2.003 (0.121)
M022234A	M01_11A	0.959 (0.028)	0.940 (0.021)		-0.391 (0.039)	0.391 (0.045)
M022234B	M01_11B	1.029 (0.034)	1.300 (0.024)		-1.270 (0.078)	1.270 (0.083)
M022243	M01_12	1.120 (0.043)	0.639 (0.027)			
M022257	M01_07	1.561 (0.118)	0.754 (0.036)	0.265 (0.014)		
M032047	M11_09	1.449 (0.172)	1.243 (0.056)	0.445 (0.013)		
M032064	M09_03	1.366 (0.051)	0.871 (0.024)			
M032097	M07_07	1.310 (0.126)	1.490 (0.048)	0.186 (0.011)		
M032100	M09_09	0.969 (0.064)	0.586 (0.046)	0.126 (0.018)		
M032116	M09_08	1.348 (0.116)	1.019 (0.043)	0.278 (0.014)		
M032142	M05_01	1.751 (0.202)	1.263 (0.047)	0.418 (0.012)		
M032160	M07_03	2.161 (0.187)	1.405 (0.030)	0.139 (0.008)		
M032163	M05_10	1.590 (0.128)	0.927 (0.037)	0.249 (0.013)		
M032205	M05_09	0.572 (0.064)	0.486 (0.145)	0.248 (0.042)		
M032273	M07_04	1.037 (0.073)	0.003 (0.066)	0.265 (0.027)		
M032294	M07_10	0.918 (0.066)	0.125 (0.070)	0.217 (0.028)		
M032307	M03_09	1.449 (0.058)	1.099 (0.027)			
M032331	M11_06	2.251 (0.211)	1.375 (0.031)	0.199 (0.009)		
M032344	M05_04	1.112 (0.041)	0.720 (0.027)			
M032352	M11_01	1.612 (0.136)	0.618 (0.042)	0.400 (0.015)		
M032414	M07_09	1.102 (0.041)	0.761 (0.027)			
M032507	M11_11	1.916 (0.162)	1.286 (0.032)	0.185 (0.009)		
M032523	M03_10	2.145 (0.178)	1.258 (0.030)	0.170 (0.009)		
M032529	M07_12	1.774 (0.134)	1.144 (0.031)	0.163 (0.001)		
M032575	M07_08	2.291 (0.161)	0.764 (0.025)	0.217 (0.011)		
M032579	M03_14	1.168 (0.073)	0.003 (0.051)	0.203 (0.023)		
M032595	M13_08	1.340 (0.078)	0.378 (0.033)	0.122 (0.014)		
M032623	M11_07	1.858 (0.124)	0.896 (0.027)	0.153 (0.001)		
M032637A	M07_13A	0.841 (0.032)	-0.250 (0.029)			
M032637C	M07_13C	1.117 (0.043)	0.441 (0.026)			
M032681B	M11_13B	0.647 (0.030)	0.980 (0.048)			
M032681C	M11_13C	1.202 (0.045)	0.704 (0.025)			
M032688	M07_11	0.848 (0.035)	0.922 (0.036)			
M032691	M03_15	0.863 (0.033)	0.441 (0.030)			
M032695	M09_13	0.542 (0.014)	0.053 (0.025)		-1.011 (0.061)	1.011 (0.062)
M032701	M03_11	1.239 (0.075)	-0.878 (0.068)	0.208 (0.038)		
M032704	M03_12	1.251 (0.075)	-0.054 (0.046)	0.188 (0.022)		

( ) Standard errors appear in parentheses

Exhibit D.22 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Applying (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M032754	M05_05	0.728 (0.029)	-0.282 (0.033)			
M042016	M10_03	0.879 (0.073)	0.754 (0.060)	0.202 (0.021)		
M042018	M02_03	0.933 (0.035)	0.581 (0.029)			
M042023	M08_04	1.125 (0.041)	0.652 (0.025)			
M042031	M06_02	1.521 (0.113)	0.762 (0.037)	0.253 (0.013)		
M042039	M02_05	0.720 (0.047)	0.462 (0.060)	0.077 (0.021)		
M042041	M10_01	1.322 (0.096)	0.048 (0.055)	0.360 (0.023)		
M042055	M02_04	1.272 (0.115)	1.009 (0.048)	0.302 (0.015)		
M042067	M10_08	1.406 (0.146)	1.339 (0.048)	0.284 (0.012)		
M042086	M06_08	1.116 (0.042)	0.776 (0.028)			
M042093	M14_10	1.435 (0.064)	1.303 (0.032)			
M042109	M12_06	1.274 (0.121)	1.302 (0.048)	0.245 (0.013)		
M042112	M12_05	0.501 (0.080)	1.192 (0.169)	0.333 (0.040)		
M042114B	M12_04B	1.239 (0.043)	0.303 (0.022)			
M042130	M04_11	0.477 (0.024)	0.011 (0.045)			
M042137	M02_10	1.178 (0.093)	0.876 (0.045)	0.232 (0.015)		
M042151	M12_09	0.839 (0.031)	0.071 (0.028)			
M042159	M14_13	0.559 (0.025)	-0.589 (0.044)			
M042169C	M10_12C	0.786 (0.047)	2.085 (0.088)			
M042177	M06_16	1.128 (0.077)	0.368 (0.048)	0.208 (0.020)		
M042179	M06_15	0.898 (0.068)	0.520 (0.060)	0.184 (0.022)		
M042182	M14_01	1.232 (0.096)	0.358 (0.052)	0.330 (0.020)		
M042201	M06_12	1.252 (0.043)	0.228 (0.022)			
M042202	M14_08	1.360 (0.105)	0.676 (0.040)	0.267 (0.015)		
M042203	M08_13	1.526 (0.088)	0.347 (0.031)	0.158 (0.014)		
M042207	M06_17	0.441 (0.011)	0.134 (0.028)		-2.890 (0.116)	2.890 (0.116)
M042220	M02_14	0.692 (0.017)	0.654 (0.023)		-1.508 (0.070)	1.508 (0.074)
M042229A	M08_10A	1.415 (0.053)	0.903 (0.024)			
M042238	M04_08	1.151 (0.111)	1.416 (0.052)	0.188 (0.012)		
M042240	M14_09	1.114 (0.069)	0.304 (0.042)	0.140 (0.018)		
M042245	M06_10	1.510 (0.120)	1.324 (0.037)	0.133 (0.009)		
M042252	M12_13	1.289 (0.092)	0.913 (0.037)	0.157 (0.013)		
M042254	M02_12	0.624 (0.047)	-1.141 (0.189)	0.225 (0.068)		
M042255	M08_15	0.678 (0.052)	0.034 (0.103)	0.161 (0.037)		
M042267	M04_06	1.271 (0.088)	0.912 (0.036)	0.137 (0.012)		
M042270	M06_11	0.909 (0.033)	0.073 (0.027)			
M042271	M14_11	1.056 (0.064)	0.399 (0.040)	0.104 (0.016)		
M042300A	M10_10A	1.481 (0.050)	0.230 (0.019)			
M042300B	M10_10B	1.568 (0.055)	0.440 (0.019)			
M042302A	M14_06A	1.186 (0.033)	0.502 (0.016)		-0.074 (0.027)	0.074 (0.030)
M042302B	M14_06B	1.197 (0.032)	0.614 (0.016)		-0.417 (0.033)	0.417 (0.036)
M042304A	M04_05A	1.448 (0.049)	-0.250 (0.020)			
M042304B	M04_05B	1.059 (0.035)	0.849 (0.021)		0.420 (0.026)	-0.420 (0.036)
M042304C	M04_05C	1.652 (0.064)	0.915 (0.022)			

( ) Standard errors appear in parentheses

Exhibit D.23 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Reasoning

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M022049	M01_03	0.623 (0.072)	0.614 (0.138)	0.316 (0.038)		
M032324	M09_07	1.058 (0.086)	0.983 (0.048)	0.184 (0.016)		
M032381	M07_01	1.154 (0.040)	0.463 (0.024)			
M032398	M11_10	0.982 (0.104)	1.212 (0.065)	0.295 (0.017)		
M032402	M09_10	0.919 (0.094)	0.982 (0.069)	0.303 (0.021)		
M032424	M11_12	0.832 (0.064)	0.640 (0.062)	0.155 (0.023)		
M032640	M05_03	0.510 (0.018)	1.779 (0.057)		-0.925 (0.070)	0.925 (0.096)
M032662	M09_02	1.449 (0.136)	1.573 (0.046)	0.130 (0.009)		
M032692	M13_06	0.561 (0.016)	1.224 (0.034)		-1.493 (0.075)	1.493 (0.085)
M032721	M13_02	0.803 (0.101)	1.484 (0.081)	0.253 (0.020)		
M032753A	M05_07A	2.749 (0.100)	0.861 (0.011)		0.125 (0.016)	-0.125 (0.019)
M032753B	M05_07B	3.058 (0.122)	0.998 (0.011)		0.202 (0.015)	-0.202 (0.018)
M032755	M05_06	0.856 (0.029)	1.487 (0.034)		-0.398 (0.046)	0.398 (0.063)
M032756	M05_08	0.755 (0.031)	0.521 (0.034)			
M032757	M13_03	0.576 (0.014)	-0.170 (0.023)		-1.623 (0.073)	1.623 (0.071)
M032760A	M13_04A	1.257 (0.033)	0.762 (0.016)		-0.769 (0.044)	0.769 (0.047)
M032760B	M13_04B	2.167 (0.092)	1.057 (0.019)			
M032760C	M13_04C	2.405 (0.118)	1.273 (0.021)			
M032761	M13_05	1.224 (0.042)	1.366 (0.024)		-0.186 (0.034)	0.186 (0.045)
M042002	M10_04	0.626 (0.030)	1.248 (0.057)			
M042036	M04_10	1.056 (0.083)	1.022 (0.046)	0.139 (0.014)		
M042066	M08_07	0.760 (0.030)	0.439 (0.033)			
M042074A	M12_08A	1.761 (0.063)	0.619 (0.019)			
M042074B	M12_08B	1.702 (0.063)	0.771 (0.021)			
M042074C	M12_08C	2.275 (0.101)	1.103 (0.020)			
M042132	M12_10	1.447 (0.151)	1.491 (0.049)	0.221 (0.011)		
M042164	M14_14	1.181 (0.044)	0.836 (0.027)			
M042167	M14_15	1.121 (0.045)	1.027 (0.032)			
M042186	M06_03	0.821 (0.032)	0.611 (0.032)			
M042197	M08_05	1.213 (0.046)	0.981 (0.029)			
M042198B	M10_05B	1.651 (0.057)	0.408 (0.018)			
M042198C	M10_05C	2.970 (0.143)	1.093 (0.017)			
M042228	M06_09	0.698 (0.030)	0.936 (0.042)			
M042257	M12_11	0.646 (0.071)	1.158 (0.088)	0.181 (0.027)		
M042263	M02_08	0.973 (0.042)	1.320 (0.042)			
M042264	M08_14	0.761 (0.036)	1.471 (0.056)			
M042265	M02_09	0.753 (0.069)	0.805 (0.075)	0.189 (0.026)		
M042268	M14_12	1.190 (0.108)	1.440 (0.048)	0.147 (0.011)		

( ) Standard errors appear in parentheses

Exhibit D.23 IRT Parameters for TIMSS 2007 Eighth Grade Mathematics - Reasoning (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
M042269	M06_14	0.756 (0.072)	0.476 (0.101)	0.284 (0.033)		
M042279	M04_09	1.106 (0.073)	0.308 (0.047)	0.171 (0.020)		
M042301B	M02_07B	1.713 (0.062)	0.682 (0.020)			
M042301C	M02_07C	4.022 (0.214)	1.043 (0.014)			
M042302C	M14_06C	0.488 (0.017)	1.720 (0.054)		-0.907 (0.069)	0.907 (0.093)
M042303B	M04_12B	0.391 (0.017)	0.889 (0.048)		-0.019 (0.063)	0.019 (0.079)
M042304D	M04_05D	0.487 (0.012)	0.430 (0.028)		-1.752 (0.078)	1.752 (0.082)

( ) Standard errors appear in parentheses

Exhibit D.24 IRT Parameters for TIMSS 2007 Eighth Grade Science - Biology

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S022115	S03_03	1.229 (0.090)	0.228 (0.047)	0.265 (0.022)		
S022126	S05_07	0.656 (0.068)	0.434 (0.095)	0.206 (0.033)		
S022150	S03_10	1.079 (0.090)	0.543 (0.048)	0.224 (0.021)		
S022289	S03_12	0.904 (0.033)	0.873 (0.024)		0.661 (0.027)	-0.661 (0.044)
S032007	S09_04	0.910 (0.039)	0.363 (0.027)			
S032015	S07_02	0.889 (0.039)	0.516 (0.029)			
S032035	S05_10	1.349 (0.090)	0.402 (0.035)	0.171 (0.017)		
S032087	S13_06	0.856 (0.090)	0.885 (0.063)	0.205 (0.023)		
S032122	S07_12	0.750 (0.040)	0.891 (0.045)			
S032126	S13_12	0.792 (0.036)	0.293 (0.030)			
S032258	S05_13	1.109 (0.077)	0.059 (0.052)	0.224 (0.025)		
S032306	S11_03	0.552 (0.016)	0.592 (0.026)		-1.202 (0.062)	1.202 (0.067)
S032310D	S07_03D	0.630 (0.021)	-0.094 (0.023)		-0.118 (0.045)	0.118 (0.042)
S032315	S11_02	0.977 (0.093)	0.711 (0.055)	0.238 (0.022)		
S032385	S05_09	0.925 (0.077)	0.058 (0.075)	0.293 (0.030)		
S032451	S13_03	0.652 (0.017)	0.174 (0.020)		-1.171 (0.057)	1.171 (0.058)
S032465	S11_01	0.821 (0.065)	0.013 (0.077)	0.219 (0.031)		
S032514	S11_13	0.754 (0.102)	1.051 (0.086)	0.277 (0.026)		
S032530D	S09_03D	0.578 (0.022)	0.379 (0.028)		0.753 (0.042)	-0.753 (0.049)
S032542	S09_01	1.185 (0.106)	0.674 (0.046)	0.252 (0.019)		
S032606	S07_01	1.226 (0.093)	-0.314 (0.069)	0.406 (0.030)		
S032611	S13_01	0.846 (0.110)	1.281 (0.079)	0.206 (0.020)		
S032614	S13_02	0.865 (0.037)	0.046 (0.027)			
S032620	S01_07	0.830 (0.117)	1.461 (0.094)	0.179 (0.019)		
S032640	S11_04	0.653 (0.031)	-0.026 (0.034)			
S032645	S09_02	1.042 (0.106)	0.756 (0.055)	0.271 (0.022)		
S032665A	S09_11A	1.126 (0.048)	0.631 (0.026)			
S032665B	S09_11B	1.751 (0.075)	0.876 (0.023)			
S032665C	S09_11C	1.560 (0.070)	0.833 (0.024)			
S032693A	S01_08A	1.177 (0.046)	0.247 (0.021)			
S032693B	S01_08B	1.290 (0.049)	0.009 (0.020)			
S032695	S01_09	0.837 (0.028)	0.632 (0.021)		-0.139 (0.034)	0.139 (0.041)
S032697D	S01_10D	0.935 (0.031)	0.523 (0.019)		-0.020 (0.031)	0.020 (0.036)
S042001	S02_06	1.393 (0.143)	1.164 (0.045)	0.171 (0.012)		
S042003	S12_03	0.533 (0.073)	0.844 (0.123)	0.217 (0.037)		
S042005	S14_02	0.397 (0.012)	0.341 (0.031)		-1.812 (0.085)	1.812 (0.088)
S042006	S04_02	0.764 (0.085)	0.394 (0.098)	0.345 (0.032)		
S042007	S10_03	1.145 (0.101)	0.772 (0.044)	0.201 (0.018)		
S042009	S02_01	1.199 (0.131)	0.785 (0.054)	0.361 (0.019)		
S042011	S02_04	0.999 (0.050)	1.156 (0.045)			
S042013	S04_01	1.216 (0.080)	-0.428 (0.061)	0.294 (0.029)		
S042015	S08_03	1.075 (0.106)	0.755 (0.052)	0.266 (0.020)		
S042016	S14_03	0.775 (0.114)	1.728 (0.123)	0.126 (0.017)		
S042017	S10_02	0.957 (0.108)	1.139 (0.061)	0.194 (0.018)		

( ) Standard errors appear in parentheses

Exhibit D.24 IRT Parameters for TIMSS 2007 Eighth Grade Science - Biology (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S042022	S10_06	0.999 (0.043)	0.542 (0.027)			
S042024	S10_04	0.982 (0.146)	1.425 (0.087)	0.257 (0.018)		
S042028	S02_05	0.868 (0.110)	1.276 (0.076)	0.194 (0.019)		
S042030	S12_02	0.910 (0.046)	1.117 (0.046)			
S042038	S06_02	1.149 (0.081)	0.467 (0.040)	0.145 (0.019)		
S042042	S12_01	0.686 (0.065)	-0.376 (0.137)	0.297 (0.044)		
S042043	S04_06	0.415 (0.029)	1.136 (0.087)			
S042049A	S08_05A	1.172 (0.045)	-0.294 (0.023)			
S042049B	S08_05B	1.431 (0.053)	0.344 (0.018)			
S042051A	S06_05A	0.972 (0.040)	0.176 (0.024)			
S042051B	S06_05B	1.415 (0.059)	0.796 (0.024)			
S042052	S04_04	0.700 (0.027)	0.249 (0.022)		0.300 (0.038)	-0.300 (0.040)
S042053	S08_01	1.327 (0.090)	0.034 (0.046)	0.278 (0.023)		
S042054	S04_05	0.385 (0.054)	-0.162 (0.311)	0.256 (0.068)		
S042059	S02_03	1.213 (0.118)	0.826 (0.046)	0.260 (0.018)		
S042222A	S12_05A	2.074 (0.091)	0.951 (0.021)			
S042222B	S12_05B	1.952 (0.079)	0.765 (0.019)			
S042222C	S12_05C	0.928 (0.077)	0.252 (0.065)	0.251 (0.027)		
S042258	S14_01	0.918 (0.099)	0.947 (0.060)	0.213 (0.021)		
S042261	S06_04	0.890 (0.043)	0.870 (0.038)			
S042297	S10_09	0.671 (0.023)	1.179 (0.036)		-0.707 (0.052)	0.707 (0.066)
S042298	S06_03	1.196 (0.047)	0.401 (0.022)			
S042300A	S14_04A	1.595 (0.058)	0.155 (0.017)			
S042300B	S14_04B	0.651 (0.041)	1.493 (0.083)			
S042300C	S14_04C	1.483 (0.055)	0.288 (0.018)			
S042304	S06_01	0.927 (0.066)	0.152 (0.057)	0.173 (0.025)		
S042309	S08_04	0.438 (0.087)	1.444 (0.190)	0.266 (0.043)		
S042310	S04_03	0.648 (0.021)	0.144 (0.021)		-0.234 (0.043)	0.234 (0.044)
S042313	S02_02	0.826 (0.035)	-0.499 (0.033)			
S042319	S14_05	1.300 (0.056)	0.890 (0.029)			
S042408	S08_02	0.830 (0.037)	0.434 (0.030)			

( ) Standard errors appear in parentheses

Exhibit D.25 IRT Parameters for TIMSS 2007 Eighth Grade Science - Chemistry

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S022181	S05_04	1.002 (0.010)	0.945 (0.056)	0.252 (0.019)		
S022183	S03_01	1.176 (0.112)	1.046 (0.047)	0.227 (0.015)		
S022208	S05_05	1.027 (0.103)	1.029 (0.054)	0.240 (0.018)		
S022276	S03_02	0.773 (0.076)	0.503 (0.086)	0.253 (0.030)		
S032056	S13_05	0.869 (0.036)	0.674 (0.032)			
S032156	S13_04	1.249 (0.108)	0.869 (0.042)	0.223 (0.016)		
S032502	S09_05	1.148 (0.098)	0.887 (0.044)	0.199 (0.016)		
S032565	S01_02	0.740 (0.035)	1.046 (0.047)			
S032570	S11_06	0.968 (0.039)	0.682 (0.030)			
S032579	S11_05	1.141 (0.136)	1.291 (0.062)	0.292 (0.015)		
S032672	S07_05	0.377 (0.064)	0.379 (0.371)	0.331 (0.072)		
S032679	S09_06	0.643 (0.038)	1.802 (0.089)			
S032680	S07_04	0.580 (0.020)	-0.330 (0.026)		0.030 (0.049)	-0.030 (0.045)
S032683	S05_12	0.991 (0.084)	0.845 (0.048)	0.165 (0.018)		
S042063	S10_07	0.804 (0.059)	-1.553 (0.173)	0.266 (0.077)		
S042064	S12_11	0.785 (0.036)	0.953 (0.042)			
S042065	S12_06	0.730 (0.061)	-0.834 (0.165)	0.281 (0.061)		
S042068	S14_06	1.001 (0.101)	1.105 (0.055)	0.215 (0.017)		
S042071	S02_11	1.124 (0.112)	1.047 (0.051)	0.245 (0.016)		
S042073	S10_01	0.616 (0.057)	-1.110 (0.243)	0.319 (0.078)		
S042076	S06_06	0.988 (0.039)	0.667 (0.029)			
S042083	S02_09	0.837 (0.028)	1.097 (0.029)		-0.141 (0.037)	0.141 (0.050)
S042088	S12_08	0.696 (0.030)	0.087 (0.033)			
S042094	S14_09	0.907 (0.037)	0.697 (0.032)			
S042095	S10_05	1.005 (0.070)	-0.074 (0.067)	0.222 (0.029)		
S042100	S06_11	0.409 (0.017)	0.747 (0.041)		-0.083 (0.061)	0.083 (0.073)
S042101	S02_12	0.956 (0.043)	1.068 (0.039)			
S042104	S12_10	0.887 (0.040)	1.101 (0.042)			
S042106	S02_10	1.028 (0.044)	0.954 (0.034)			
S042109	S04_10	1.256 (0.087)	0.168 (0.048)	0.258 (0.022)		
S042110	S12_04	0.658 (0.056)	-0.588 (0.160)	0.229 (0.056)		
S042112	S10_11	0.420 (0.058)	0.388 (0.250)	0.229 (0.061)		
S042228A	S08_08A	0.978 (0.046)	1.294 (0.046)			
S042228B	S08_08B	2.449 (0.086)	0.145 (0.013)			
S042228C	S08_08C	3.519 (0.146)	0.583 (0.012)			
S042232A	S04_11A	0.818 (0.034)	0.513 (0.031)			
S042232B	S04_11B	1.195 (0.069)	1.825 (0.064)			
S042232C	S04_11C	1.691 (0.278)	1.771 (0.078)	0.332 (0.011)		
S042305	S10_10	0.546 (0.024)	1.306 (0.047)		0.363 (0.045)	-0.363 (0.071)
S042306	S06_08	1.363 (0.133)	1.060 (0.044)	0.265 (0.014)		
S042400	S14_12	0.934 (0.044)	1.240 (0.045)			

( ) Standard errors appear in parentheses

Exhibit D.26 IRT Parameters for TIMSS 2007 Eighth Grade Science - Physics

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S022002	S03_06	1.023 (0.081)	0.494 (0.051)	0.199 (0.021)		
S022019	S03_05	0.904 (0.072)	-0.128 (0.087)	0.263 (0.035)		
S022022	S03_04	0.837 (0.034)	0.199 (0.028)			
S022042	S03_11	1.367 (0.094)	0.553 (0.034)	0.171 (0.015)		
S022054	S05_03	0.997 (0.086)	0.536 (0.056)	0.243 (0.022)		
S022069	S03_13	0.999 (0.042)	0.673 (0.029)			
S022268	S03_14	0.672 (0.033)	0.783 (0.044)			
S022281	S05_08	0.574 (0.032)	1.254 (0.069)			
S022292	S05_02	0.753 (0.033)	0.614 (0.036)			
S032024	S11_07	0.902 (0.121)	1.468 (0.083)	0.243 (0.017)		
S032141	S11_09	1.580 (0.138)	1.045 (0.036)	0.187 (0.012)		
S032158	S13_14	0.845 (0.084)	0.467 (0.081)	0.290 (0.029)		
S032184	S09_07	0.688 (0.100)	1.238 (0.098)	0.270 (0.027)		
S032238	S13_08	1.281 (0.092)	0.600 (0.037)	0.176 (0.016)		
S032257	S07_08	1.203 (0.128)	1.176 (0.052)	0.237 (0.015)		
S032272	S11_08	0.993 (0.051)	1.490 (0.056)			
S032273	S01_04	0.702 (0.113)	1.644 (0.117)	0.249 (0.022)		
S032279	S13_07	1.127 (0.123)	1.332 (0.059)	0.198 (0.014)		
S032369	S13_09	0.630 (0.022)	0.750 (0.029)		-0.091 (0.043)	0.091 (0.052)
S032392	S07_06	0.496 (0.047)	-1.225 (0.285)	0.257 (0.083)		
S032394	S09_08	1.108 (0.109)	0.804 (0.053)	0.301 (0.019)		
S032403	S01_03	1.168 (0.114)	0.984 (0.048)	0.251 (0.016)		
S032425	S07_07	0.877 (0.092)	0.827 (0.066)	0.262 (0.023)		
S042061	S04_08	0.709 (0.095)	1.424 (0.090)	0.190 (0.022)		
S042173	S10_12	0.565 (0.019)	-0.359 (0.030)		1.109 (0.052)	-1.109 (0.044)
S042176	S08_11	0.990 (0.040)	0.648 (0.029)			
S042182	S08_06	0.698 (0.064)	-0.140 (0.128)	0.243 (0.045)		
S042195	S14_11	0.609 (0.037)	1.688 (0.089)			
S042196	S04_07	0.811 (0.042)	1.388 (0.058)			
S042197	S10_08	0.914 (0.093)	0.981 (0.059)	0.212 (0.020)		
S042210	S08_10	0.931 (0.128)	1.423 (0.082)	0.277 (0.018)		
S042211	S08_12	1.005 (0.038)	0.200 (0.024)			
S042216	S14_07	1.001 (0.091)	0.460 (0.063)	0.287 (0.024)		
S042217	S12_15	1.372 (0.113)	0.733 (0.039)	0.245 (0.016)		
S042218	S12_09	1.435 (0.113)	0.679 (0.037)	0.245 (0.015)		
S042238A	S06_12A	0.920 (0.085)	0.916 (0.053)	0.169 (0.019)		
S042238B	S06_12B	0.692 (0.037)	1.361 (0.064)			
S042244A	S02_15A	1.202 (0.053)	1.030 (0.033)			
S042244B	S02_15B	1.028 (0.054)	1.453 (0.054)			
S042249	S14_08	1.029 (0.088)	0.674 (0.050)	0.218 (0.020)		
S042272	S06_10	0.889 (0.077)	0.425 (0.067)	0.228 (0.026)		
S042273	S12_12	1.088 (0.040)	0.196 (0.023)			
S042274	S10_15	1.177 (0.136)	1.356 (0.061)	0.229 (0.014)		
S042276	S02_07	1.156 (0.159)	1.273 (0.070)	0.387 (0.016)		

( ) Standard errors appear in parentheses

Exhibit D.26 IRT Parameters for TIMSS 2007 Eighth Grade Science - Physics (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S042278	S10_14	0.751 (0.035)	0.892 (0.043)			
S042279	S02_08	0.474 (0.085)	1.025 (0.196)	0.327 (0.047)		
S042280	S12_07	1.385 (0.097)	0.441 (0.037)	0.214 (0.017)		
S042292	S04_09	0.488 (0.016)	0.767 (0.034)		-0.843 (0.062)	0.843 (0.071)
S042293A	S14_10A	0.979 (0.036)	-0.248 (0.026)			
S042293B	S14_10B	0.757 (0.049)	2.042 (0.105)			
S042294	S04_12	1.434 (0.133)	0.981 (0.040)	0.251 (0.014)		
S042402	S08_07	0.778 (0.039)	1.296 (0.056)			
S042403	S06_09	0.894 (0.039)	0.805 (0.035)			
S042407	S10_13	0.419 (0.026)	0.674 (0.063)			

( ) Standard errors appear in parentheses

Exhibit D.27 IRT Parameters for TIMSS 2007 Eighth Grade Science - Earth Science

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S022078	S05_06	1.065 (0.039)	0.272 (0.024)			
S022244	S03_09	1.256 (0.054)	1.012 (0.030)			
S022290	S05_01	1.080 (0.087)	0.459 (0.051)	0.254 (0.021)		
S022294	S03_07	0.863 (0.085)	0.248 (0.091)	0.358 (0.030)		
S032019A	S01_05A	1.647 (0.073)	1.059 (0.025)			
S032019B	S01_05B	1.813 (0.097)	1.405 (0.034)			
S032060	S11_10	1.119 (0.040)	-0.278 (0.025)			
S032115	S01_01	0.775 (0.064)	0.645 (0.057)	0.108 (0.021)		
S032120A	S05_14A	1.052 (0.051)	1.222 (0.041)			
S032120B	S05_14B	1.025 (0.048)	0.949 (0.034)			
S032151	S09_09	1.091 (0.084)	0.725 (0.040)	0.152 (0.016)		
S032160	S13_10	0.882 (0.091)	0.492 (0.080)	0.346 (0.027)		
S032463	S11_11	1.261 (0.088)	0.360 (0.041)	0.220 (0.018)		
S032510	S13_13	0.910 (0.074)	-0.141 (0.091)	0.303 (0.034)		
S032516	S01_06	0.723 (0.029)	-0.372 (0.034)			
S032519	S05_11	0.801 (0.034)	0.583 (0.033)			
S032555	S07_11	1.215 (0.049)	0.775 (0.026)			
S032650D	S11_12D	0.718 (0.024)	0.253 (0.022)		0.164 (0.038)	-0.164 (0.039)
S032651A	S09_10A	1.391 (0.051)	0.424 (0.020)			
S032651B	S09_10B	1.349 (0.060)	1.036 (0.029)			
S032654	S13_11	1.309 (0.107)	0.749 (0.039)	0.237 (0.016)		
S032660	S07_10	1.126 (0.125)	1.329 (0.058)	0.200 (0.014)		
S032663	S07_09	0.742 (0.112)	1.466 (0.098)	0.265 (0.022)		
S042126	S08_09	0.752 (0.106)	0.609 (0.120)	0.437 (0.032)		
S042135	S08_13	0.739 (0.030)	-0.030 (0.031)			
S042141	S06_13	0.775 (0.064)	-0.142 (0.102)	0.227 (0.037)		
S042149	S04_13	0.590 (0.027)	0.258 (0.038)			
S042150	S04_15	0.875 (0.093)	0.780 (0.068)	0.286 (0.023)		
S042153	S02_16	0.750 (0.035)	0.585 (0.036)			
S042155	S04_14	0.947 (0.039)	0.690 (0.030)			
S042164	S14_14	1.015 (0.077)	0.727 (0.042)	0.133 (0.016)		
S042215	S06_14	0.638 (0.093)	1.413 (0.099)	0.207 (0.025)		
S042257	S08_14	0.861 (0.109)	1.123 (0.072)	0.291 (0.021)		
S042301	S12_13	0.930 (0.035)	0.175 (0.026)			
S042307	S02_13	0.668 (0.030)	0.539 (0.038)			
S042312	S12_14	0.467 (0.058)	-0.218 (0.271)	0.275 (0.067)		
S042317	S10_17	0.532 (0.017)	0.036 (0.027)		-0.393 (0.056)	0.393 (0.055)
S042404	S06_07	0.680 (0.027)	1.407 (0.043)		-0.227 (0.045)	0.227 (0.067)
S042405	S02_14	0.856 (0.100)	0.882 (0.072)	0.309 (0.023)		
S042406	S12_16	1.145 (0.046)	0.745 (0.026)			

( ) Standard errors appear in parentheses

Exhibit D.28 IRT Parameters for TIMSS 2007 Eighth Grade Science - Knowing

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S022126	S05_07	0.727 (0.071)	0.431 (0.081)	0.217 (0.030)		
S022181	S05_04	1.279 (0.114)	0.724 (0.041)	0.237 (0.018)		
S022183	S03_01	2.151 (0.178)	0.784 (0.026)	0.231 (0.012)		
S022208	S05_05	1.347 (0.130)	0.853 (0.041)	0.246 (0.016)		
S022276	S03_02	0.827 (0.073)	0.352 (0.068)	0.216 (0.027)		
S022290	S05_01	1.426 (0.104)	0.384 (0.038)	0.256 (0.019)		
S022294	S03_07	1.219 (0.101)	0.220 (0.056)	0.358 (0.024)		
S032007	S09_04	1.065 (0.044)	0.317 (0.023)			
S032015	S07_02	1.058 (0.044)	0.476 (0.025)			
S032024	S11_07	1.347 (0.179)	1.267 (0.060)	0.265 (0.014)		
S032035	S05_10	1.300 (0.084)	0.363 (0.034)	0.153 (0.017)		
S032087	S13_06	0.990 (0.094)	0.789 (0.052)	0.200 (0.020)		
S032115	S01_01	1.256 (0.085)	0.537 (0.034)	0.127 (0.016)		
S032122	S07_12	0.725 (0.040)	0.914 (0.048)			
S032126	S13_12	0.857 (0.039)	0.290 (0.028)			
S032151	S09_09	1.605 (0.118)	0.646 (0.029)	0.177 (0.014)		
S032158	S13_14	1.087 (0.097)	0.448 (0.056)	0.301 (0.024)		
S032160	S13_10	1.050 (0.102)	0.443 (0.064)	0.352 (0.025)		
S032257	S07_08	1.276 (0.151)	1.118 (0.052)	0.242 (0.016)		
S032258	S05_13	1.078 (0.076)	0.060 (0.053)	0.228 (0.025)		
S032273	S01_04	0.849 (0.134)	1.445 (0.102)	0.256 (0.020)		
S032310D	S07_03D	0.714 (0.024)	-0.072 (0.020)		-0.068 (0.040)	0.068 (0.038)
S032385	S05_09	1.031 (0.079)	0.030 (0.063)	0.282 (0.027)		
S032403	S01_03	1.227 (0.119)	0.857 (0.044)	0.235 (0.017)		
S032425	S07_07	1.229 (0.121)	0.765 (0.047)	0.286 (0.019)		
S032463	S11_11	1.532 (0.101)	0.316 (0.033)	0.216 (0.018)		
S032465	S11_01	1.028 (0.079)	0.138 (0.059)	0.266 (0.026)		
S032510	S13_13	1.041 (0.079)	-0.116 (0.068)	0.302 (0.029)		
S032514	S11_13	1.031 (0.122)	0.942 (0.061)	0.293 (0.021)		
S032519	S05_11	0.744 (0.037)	0.577 (0.036)			
S032530D	S09_03D	0.609 (0.023)	0.343 (0.027)		0.739 (0.040)	-0.739 (0.046)
S032606	S07_01	1.350 (0.101)	-0.262 (0.061)	0.414 (0.028)		
S032611	S13_01	1.020 (0.124)	1.184 (0.064)	0.216 (0.018)		
S032614	S13_02	0.901 (0.038)	0.049 (0.026)			
S032640	S11_04	0.648 (0.032)	-0.017 (0.034)			
S032645	S09_02	1.188 (0.111)	0.679 (0.047)	0.269 (0.020)		
S032660	S07_10	1.209 (0.144)	1.220 (0.057)	0.203 (0.015)		
S032672	S07_05	0.587 (0.079)	0.374 (0.153)	0.355 (0.042)		
S042001	S02_06	1.176 (0.123)	1.153 (0.052)	0.153 (0.014)		
S042003	S12_03	0.713 (0.088)	0.847 (0.086)	0.256 (0.028)		
S042006	S04_02	0.738 (0.077)	0.278 (0.099)	0.310 (0.033)		
S042009	S02_01	1.054 (0.100)	0.569 (0.056)	0.295 (0.022)		
S042013	S04_01	1.123 (0.076)	-0.475 (0.066)	0.277 (0.030)		
S042015	S08_03	1.288 (0.120)	0.718 (0.043)	0.275 (0.018)		

( ) Standard errors appear in parentheses

Exhibit D.28 IRT Parameters for TIMSS 2007 Eighth Grade Science - Knowing (Continued)

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S042016	S14_03	1.054 (0.135)	1.488 (0.081)	0.138 (0.014)		
S042024	S10_04	1.350 (0.181)	1.257 (0.060)	0.270 (0.014)		
S042028	S02_05	1.026 (0.105)	1.027 (0.053)	0.170 (0.017)		
S042038	S06_02	1.126 (0.087)	0.526 (0.043)	0.171 (0.020)		
S042042	S12_01	0.922 (0.081)	-0.119 (0.089)	0.369 (0.032)		
S042054	S04_05	0.461 (0.062)	-0.028 (0.234)	0.284 (0.057)		
S042059	S02_03	1.076 (0.095)	0.682 (0.047)	0.211 (0.020)		
S042065	S12_06	1.095 (0.098)	-0.198 (0.085)	0.474 (0.030)		
S042068	S14_06	1.146 (0.115)	0.955 (0.048)	0.209 (0.017)		
S042071	S02_11	1.259 (0.122)	0.860 (0.043)	0.229 (0.017)		
S042073	S10_01	0.741 (0.084)	-0.465 (0.171)	0.505 (0.043)		
S042076	S06_06	1.161 (0.048)	0.590 (0.025)			
S042095	S10_05	1.201 (0.075)	-0.062 (0.046)	0.207 (0.023)		
S042100	S06_11	0.480 (0.021)	0.649 (0.035)		-0.064 (0.052)	0.064 (0.062)
S042109	S04_10	2.032 (0.131)	0.202 (0.029)	0.278 (0.017)		
S042112	S10_11	0.537 (0.065)	0.326 (0.143)	0.234 (0.043)		
S042126	S08_09	0.866 (0.113)	0.539 (0.098)	0.436 (0.030)		
S042135	S08_13	0.907 (0.038)	0.001 (0.026)			
S042141	S06_13	0.933 (0.069)	-0.087 (0.068)	0.230 (0.029)		
S042149	S04_13	0.558 (0.030)	0.259 (0.040)			
S042150	S04_15	0.993 (0.089)	0.544 (0.055)	0.242 (0.023)		
S042153	S02_16	0.741 (0.038)	0.559 (0.037)			
S042164	S14_14	1.678 (0.111)	0.587 (0.027)	0.146 (0.013)		
S042182	S08_06	0.938 (0.085)	0.155 (0.076)	0.345 (0.029)		
S042210	S08_10	0.971 (0.141)	1.295 (0.081)	0.273 (0.019)		
S042215	S06_14	0.775 (0.096)	1.140 (0.077)	0.197 (0.023)		
S042217	S12_15	1.781 (0.145)	0.667 (0.031)	0.254 (0.015)		
S042249	S14_08	0.996 (0.084)	0.603 (0.050)	0.194 (0.021)		
S042257	S08_14	1.260 (0.144)	0.945 (0.051)	0.309 (0.018)		
S042261	S06_04	0.981 (0.046)	0.817 (0.034)			
S042272	S06_10	1.156 (0.096)	0.474 (0.048)	0.264 (0.022)		
S042274	S10_15	0.988 (0.123)	1.250 (0.069)	0.199 (0.018)		
S042279	S02_08	0.464 (0.072)	0.675 (0.178)	0.268 (0.047)		
S042293A	S14_10A	1.029 (0.041)	-0.195 (0.025)			
S042301	S12_13	1.041 (0.042)	0.159 (0.023)			
S042306	S06_08	1.373 (0.144)	0.954 (0.043)	0.260 (0.016)		
S042307	S02_13	0.716 (0.036)	0.501 (0.036)			
S042312	S12_14	0.563 (0.064)	-0.160 (0.173)	0.283 (0.049)		
S042317	S10_17	0.663 (0.022)	0.066 (0.022)		-0.298 (0.045)	0.298 (0.044)

( ) Standard errors appear in parentheses

Exhibit D.29 IRT Parameters for TIMSS 2007 Eighth Grade Science - Applying

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S022002	S03_06	1.186 (0.095)	0.535 (0.044)	0.227 (0.020)		
S022019	S03_05	1.026 (0.077)	-0.060 (0.067)	0.278 (0.029)		
S022054	S05_03	1.035 (0.092)	0.539 (0.054)	0.249 (0.022)		
S022069	S03_13	1.146 (0.048)	0.624 (0.026)			
S022078	S05_06	1.297 (0.048)	0.236 (0.020)			
S022150	S03_10	0.960 (0.082)	0.520 (0.055)	0.206 (0.023)		
S022281	S05_08	0.659 (0.037)	1.118 (0.060)			
S032019A	S01_05A	1.249 (0.059)	1.081 (0.036)			
S032019B	S01_05B	1.383 (0.078)	1.442 (0.050)			
S032056	S13_05	0.974 (0.042)	0.596 (0.029)			
S032060	S11_10	1.211 (0.046)	-0.267 (0.023)			
S032120A	S05_14A	1.037 (0.052)	1.160 (0.044)			
S032120B	S05_14B	1.099 (0.053)	0.853 (0.033)			
S032141	S11_09	1.836 (0.169)	0.978 (0.032)	0.198 (0.012)		
S032184	S09_07	0.582 (0.085)	1.069 (0.112)	0.220 (0.034)		
S032238	S13_08	1.385 (0.103)	0.598 (0.035)	0.189 (0.016)		
S032279	S13_07	1.120 (0.132)	1.270 (0.061)	0.195 (0.015)		
S032306	S11_03	0.594 (0.017)	0.542 (0.024)		-1.086 (0.057)	1.086 (0.062)
S032369	S13_09	0.730 (0.026)	0.679 (0.025)		-0.067 (0.037)	0.067 (0.045)
S032392	S07_06	0.498 (0.050)	-1.206 (0.275)	0.269 (0.074)		
S032394	S09_08	1.146 (0.111)	0.712 (0.051)	0.289 (0.020)		
S032451	S13_03	0.638 (0.017)	0.167 (0.021)		-1.206 (0.059)	1.206 (0.059)
S032502	S09_05	1.401 (0.101)	0.643 (0.032)	0.161 (0.015)		
S032516	S01_06	0.836 (0.035)	-0.330 (0.030)			
S032542	S09_01	1.474 (0.122)	0.641 (0.037)	0.261 (0.017)		
S032570	S11_06	1.132 (0.046)	0.597 (0.026)			
S032579	S11_05	0.987 (0.130)	1.185 (0.070)	0.271 (0.019)		
S032650D	S11_12D	0.821 (0.028)	0.199 (0.019)		0.143 (0.033)	-0.143 (0.035)
S032651A	S09_10A	1.341 (0.051)	0.417 (0.020)			
S032663	S07_09	0.630 (0.108)	1.483 (0.126)	0.253 (0.027)		
S032679	S09_06	0.906 (0.049)	1.358 (0.057)			
S032683	S05_12	1.272 (0.103)	0.726 (0.037)	0.173 (0.016)		
S032697D	S01_10D	0.818 (0.028)	0.521 (0.021)		-0.060 (0.035)	0.060 (0.040)
S042005	S14_02	0.365 (0.011)	0.349 (0.034)		-1.992 (0.092)	1.992 (0.096)
S042007	S10_03	1.423 (0.118)	0.721 (0.036)	0.214 (0.016)		
S042011	S02_04	1.016 (0.050)	1.129 (0.043)			
S042017	S10_02	1.036 (0.115)	1.110 (0.057)	0.202 (0.017)		
S042030	S12_02	0.904 (0.045)	1.118 (0.046)			
S042043	S04_06	0.440 (0.029)	1.077 (0.080)			
S042049A	S08_05A	1.056 (0.040)	-0.345 (0.025)			
S042051A	S06_05A	0.941 (0.038)	0.156 (0.025)			
S042052	S04_04	0.638 (0.025)	0.254 (0.024)		0.295 (0.041)	-0.295 (0.044)
S042053	S08_01	1.433 (0.099)	0.057 (0.044)	0.301 (0.022)		
S042061	S04_08	0.679 (0.092)	1.319 (0.092)	0.169 (0.024)		

( ) Standard errors appear in parentheses

Exhibit D.29 IRT Parameters for TIMSS 2007 Eighth Grade Science - Applying (Continued)

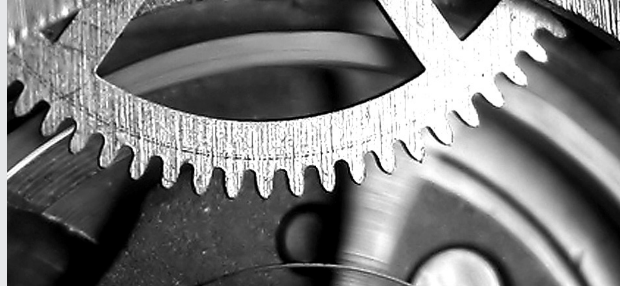
Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S042063	S10_07	0.853 (0.069)	-1.356 (0.159)	0.346 (0.060)		
S042064	S12_11	0.990 (0.044)	0.797 (0.033)			
S042088	S12_08	0.718 (0.033)	0.084 (0.031)			
S042094	S14_09	0.986 (0.043)	0.632 (0.029)			
S042101	S02_12	1.074 (0.049)	0.945 (0.035)			
S042106	S02_10	1.290 (0.054)	0.801 (0.027)			
S042110	S12_04	0.692 (0.059)	-0.523 (0.133)	0.244 (0.047)		
S042155	S04_14	1.006 (0.043)	0.628 (0.028)			
S042173	S10_12	0.499 (0.018)	-0.387 (0.033)		1.171 (0.058)	-1.171 (0.049)
S042195	S14_11	0.749 (0.044)	1.418 (0.070)			
S042196	S04_07	0.959 (0.050)	1.217 (0.049)			
S042211	S08_12	1.030 (0.041)	0.182 (0.023)			
S042216	S14_07	1.054 (0.107)	0.574 (0.062)	0.330 (0.024)		
S042218	S12_09	1.618 (0.129)	0.639 (0.033)	0.250 (0.015)		
S042222A	S12_05A	1.332 (0.063)	1.125 (0.035)			
S042222B	S12_05B	1.283 (0.056)	0.892 (0.029)			
S042222C	S12_05C	0.981 (0.084)	0.300 (0.064)	0.275 (0.026)		
S042228C	S08_08C	1.496 (0.058)	0.616 (0.021)			
S042244A	S02_15A	1.262 (0.057)	0.957 (0.031)			
S042244B	S02_15B	1.146 (0.062)	1.321 (0.049)			
S042258	S14_01	0.892 (0.094)	0.912 (0.060)	0.202 (0.022)		
S042273	S12_12	1.114 (0.043)	0.199 (0.022)			
S042276	S02_07	1.070 (0.158)	1.176 (0.074)	0.375 (0.018)		
S042278	S10_14	0.946 (0.042)	0.751 (0.033)			
S042292	S04_09	0.551 (0.018)	0.699 (0.030)		-0.737 (0.055)	0.737 (0.062)
S042293B	S14_10B	0.855 (0.057)	1.821 (0.094)			
S042298	S06_03	1.080 (0.043)	0.401 (0.024)			
S042300A	S14_04A	1.589 (0.057)	0.146 (0.017)			
S042300B	S14_04B	0.681 (0.041)	1.440 (0.076)			
S042300C	S14_04C	1.439 (0.053)	0.284 (0.018)			
S042305	S10_10	0.590 (0.027)	1.169 (0.044)		0.308 (0.042)	-0.308 (0.065)
S042309	S08_04	0.439 (0.085)	1.376 (0.188)	0.259 (0.044)		
S042313	S02_02	0.799 (0.034)	-0.534 (0.034)			
S042400	S14_12	1.085 (0.052)	1.075 (0.039)			
S042402	S08_07	0.862 (0.044)	1.162 (0.050)			
S042404	S06_07	0.844 (0.032)	1.189 (0.034)		-0.146 (0.037)	0.146 (0.054)
S042405	S02_14	1.237 (0.120)	0.686 (0.048)	0.309 (0.019)		
S042406	S12_16	1.129 (0.048)	0.686 (0.027)			
S042407	S10_13	0.419 (0.027)	0.661 (0.064)			
S042408	S08_02	0.773 (0.035)	0.433 (0.032)			

( ) Standard errors appear in parentheses

Exhibit D.30 IRT Parameters for TIMSS 2007 Eighth Grade Science - Reasoning

Item		Slope (a_j)	Location (b_j)	Guessing (c_j)	Step 1 (d_{j1})	Step 2 (d_{j2})
S022022	S03_04	0.693 (0.029)	0.216 (0.034)			
S022042	S03_11	1.273 (0.092)	0.656 (0.038)	0.194 (0.015)		
S022115	S03_03	0.937 (0.072)	0.158 (0.069)	0.239 (0.027)		
S022244	S03_09	1.072 (0.046)	1.095 (0.035)			
S022268	S03_14	0.690 (0.032)	0.783 (0.042)			
S022289	S03_12	0.744 (0.026)	1.024 (0.029)		0.815 (0.033)	-0.815 (0.054)
S022292	S05_02	0.634 (0.029)	0.737 (0.043)			
S032156	S13_04	1.188 (0.099)	0.901 (0.043)	0.216 (0.015)		
S032272	S11_08	0.810 (0.045)	1.742 (0.072)			
S032315	S11_02	0.871 (0.088)	0.869 (0.065)	0.255 (0.022)		
S032555	S07_11	1.297 (0.050)	0.779 (0.024)			
S032565	S01_02	0.769 (0.035)	1.049 (0.044)			
S032620	S01_07	0.772 (0.110)	1.747 (0.104)	0.195 (0.017)		
S032651B	S09_10B	0.768 (0.040)	1.437 (0.059)			
S032654	S13_11	0.903 (0.085)	0.863 (0.059)	0.227 (0.021)		
S032665A	S09_11A	0.921 (0.038)	0.750 (0.032)			
S032665B	S09_11B	1.709 (0.073)	1.006 (0.023)			
S032665C	S09_11C	1.461 (0.064)	0.962 (0.026)			
S032680	S07_04	0.640 (0.020)	-0.318 (0.024)		0.106 (0.045)	-0.106 (0.041)
S032693A	S01_08A	1.033 (0.037)	0.267 (0.025)			
S032693B	S01_08B	1.147 (0.040)	-0.021 (0.023)			
S032695	S01_09	0.701 (0.023)	0.750 (0.026)		-0.145 (0.041)	0.145 (0.049)
S042022	S10_06	0.852 (0.034)	0.616 (0.032)			
S042049B	S08_05B	1.051 (0.039)	0.410 (0.025)			
S042051B	S06_05B	1.154 (0.047)	0.955 (0.030)			
S042083	S02_09	0.640 (0.022)	1.272 (0.038)		-0.300 (0.047)	0.300 (0.064)
S042104	S12_10	0.847 (0.039)	1.166 (0.044)			
S042176	S08_11	0.875 (0.036)	0.719 (0.033)			
S042197	S10_08	0.979 (0.102)	1.138 (0.057)	0.241 (0.017)		
S042228A	S08_08A	1.207 (0.054)	1.196 (0.035)			
S042228B	S08_08B	1.066 (0.038)	0.171 (0.024)			
S042232A	S04_11A	0.979 (0.037)	0.488 (0.027)			
S042232B	S04_11B	1.261 (0.074)	1.813 (0.059)			
S042232C	S04_11C	1.637 (0.280)	1.786 (0.078)	0.329 (0.011)		
S042238A	S06_12A	0.789 (0.076)	0.998 (0.062)	0.159 (0.021)		
S042238B	S06_12B	0.672 (0.035)	1.430 (0.064)			
S042280	S12_07	1.453 (0.086)	0.335 (0.032)	0.155 (0.015)		
S042294	S04_12	1.173 (0.116)	1.133 (0.049)	0.249 (0.015)		
S042297	S10_09	0.531 (0.018)	1.439 (0.046)		-0.903 (0.065)	0.903 (0.083)
S042304	S06_01	0.833 (0.065)	0.211 (0.075)	0.196 (0.028)		
S042310	S04_03	0.498 (0.016)	0.167 (0.027)		-0.347 (0.055)	0.347 (0.057)
S042319	S14_05	1.107 (0.047)	1.057 (0.033)			
S042403	S06_09	0.800 (0.035)	0.895 (0.039)			

( ) Standard errors appear in parentheses



Appendix E

Summary Statistics and Standard Errors for Proficiency in Mathematics and Science Content Domains and Cognitive Domains in the Fourth and Eighth Grades

Exhibit E.1 Summary Statistics and Standard Errors for Proficiency in Number in the Fourth Grade

Country	Sample Size	Number			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	390.710	91.852	4.986	5.025
Armenia	4,079	521.780	87.024	3.898	4.005
Australia	4,108	496.443	95.241	3.474	3.669
Austria	4,859	502.395	68.213	1.900	2.188
Chinese Taipei	4,131	581.167	74.717	1.752	1.897
Colombia	4,801	359.708	88.768	4.334	4.339
Czech Republic	4,235	481.901	69.955	2.676	2.830
Denmark	3,519	508.740	74.195	2.657	2.910
El Salvador	4,166	317.087	100.003	3.680	3.928
England	4,316	530.860	96.382	3.163	3.228
Georgia	4,108	464.430	82.897	3.660	3.753
Germany	5,200	521.146	68.537	2.131	2.215
Hong Kong SAR	3,791	606.196	75.313	3.678	3.815
Hungary	4,048	509.829	94.144	3.618	3.653
Iran, Islamic Rep. of	3,833	398.383	84.475	3.532	3.552
Italy	4,470	505.231	78.532	3.098	3.154
Japan	4,487	560.979	84.174	2.144	2.166
Kazakhstan	3,990	555.743	81.597	6.378	6.594
Kuwait	3,803	320.618	96.746	3.224	3.497
Latvia	3,908	535.657	74.911	2.014	2.061
Lithuania	3,980	533.016	75.455	2.232	2.255
Morocco	3,894	353.393	96.588	4.536	4.726
Netherlands	3,349	534.734	63.514	1.934	2.221
New Zealand	4,940	477.823	101.486	2.573	2.654
Norway	4,108	460.983	83.254	2.530	2.832
Qatar	7,019	291.850	97.307	1.012	1.176
Russian Federation	4,464	546.403	76.920	4.311	4.395
Scotland	3,929	480.932	88.783	2.455	2.622
Singapore	5,041	610.549	92.680	4.081	4.252
Slovak Republic	4,963	495.171	83.286	3.860	3.927
Slovenia	4,351	484.568	79.168	1.777	1.858
Sweden	4,676	489.971	72.729	2.523	2.549
Tunisia	4,134	352.482	104.420	3.971	4.467
Ukraine	4,292	479.915	81.032	2.625	2.900
United States	7,896	524.176	85.120	2.706	2.744
Yemen	+	+	+	+	+
Benchmarking Participants					
Alberta, Canada	4,037	489.445	77.420	3.106	3.275
British Columbia, Canada	4,153	492.799	80.609	2.659	2.784
Dubai, UAE	3,064	444.234	93.006	1.877	1.998
Massachusetts, US	1,747	571.317	75.766	3.707	3.972
Minnesota, US	1,846	545.599	89.029	6.135	6.164
Ontario, Canada	3,496	489.115	81.921	3.274	3.600
Quebec, Canada	3,885	510.571	73.222	2.883	2.998

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.2 Summary Statistics and Standard Errors for Proficiency in Geometric Shapes and Measures in the Fourth Grade

Country	Sample Size	Geometric Shapes and Measures			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	382.820	86.105	4.224	4.497
Armenia	4,079	483.326	101.062	4.386	4.715
Australia	4,108	535.784	77.443	2.928	3.056
Austria	4,859	509.013	67.306	1.997	2.432
Chinese Taipei	4,131	555.737	74.380	1.764	2.182
Colombia	4,801	361.299	92.218	4.372	4.831
Czech Republic	4,235	494.165	71.299	2.667	2.778
Denmark	3,519	543.509	69.047	2.276	2.627
El Salvador	4,166	333.015	95.837	3.921	4.257
England	4,316	547.933	78.425	2.682	2.734
Georgia	4,108	415.187	95.200	4.646	4.840
Germany	5,200	528.447	64.609	1.995	2.031
Hong Kong SAR	3,791	598.855	62.912	2.987	3.076
Hungary	4,048	509.736	84.784	3.224	3.284
Iran, Islamic Rep. of	3,833	428.549	79.180	3.111	3.288
Italy	4,470	509.252	77.373	2.981	2.999
Japan	4,487	565.781	74.586	1.819	2.227
Kazakhstan	3,990	542.194	95.291	7.297	7.379
Kuwait	3,803	316.495	100.567	3.186	3.632
Latvia	3,908	532.106	66.083	1.887	2.621
Lithuania	3,980	517.941	71.088	2.107	2.360
Morocco	3,894	364.809	90.152	3.939	4.328
Netherlands	3,349	522.475	58.439	2.028	2.294
New Zealand	4,940	502.115	73.425	2.056	2.269
Norway	4,108	489.711	69.428	2.485	2.998
Qatar	7,019	296.299	97.234	0.982	1.430
Russian Federation	4,464	538.337	86.560	4.858	5.088
Scotland	3,929	503.380	72.160	1.998	2.575
Singapore	5,041	570.192	79.775	3.529	3.650
Slovak Republic	4,963	499.479	81.899	3.995	4.259
Slovenia	4,351	522.223	61.401	1.543	1.756
Sweden	4,676	508.331	62.197	2.253	2.303
Tunisia	4,134	334.383	107.635	4.044	4.464
Ukraine	4,292	457.210	86.113	2.685	2.824
United States	7,896	522.443	72.769	2.356	2.498
Yemen	+	+	+	+	+
Benchmarking Participants					
Alberta, Canada	4,037	512.261	61.869	2.590	2.899
British Columbia, Canada	4,153	509.942	68.954	2.659	2.941
Dubai, UAE	3,064	440.213	90.378	2.025	2.814
Massachusetts, US	1,747	564.368	71.818	3.782	4.078
Minnesota, US	1,846	556.240	76.126	4.944	5.258
Ontario, Canada	3,496	530.411	67.449	2.861	3.047
Quebec, Canada	3,885	525.078	68.312	3.156	3.245

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.3 Summary Statistics and Standard Errors for Proficiency in Data Display in the Fourth Grade

Country	Sample Size	Data Display			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	361.401	93.498	4.998	5.157
Armenia	4,079	457.833	95.256	4.162	4.270
Australia	4,108	533.623	72.290	2.735	3.120
Austria	4,859	508.047	69.124	1.904	2.625
Chinese Taipei	4,131	566.549	56.620	1.145	1.956
Colombia	4,801	363.156	106.033	5.208	5.895
Czech Republic	4,235	493.047	83.361	2.902	3.328
Denmark	3,519	528.758	73.073	2.913	3.384
El Salvador	4,166	367.255	87.502	3.235	3.481
England	4,316	546.677	68.286	2.069	2.473
Georgia	4,108	414.263	88.411	4.289	4.582
Germany	5,200	533.510	77.809	2.162	3.077
Hong Kong SAR	3,791	585.210	53.397	2.507	2.651
Hungary	4,048	504.083	90.908	3.354	3.518
Iran, Islamic Rep. of	3,833	399.807	88.222	3.531	4.044
Italy	4,470	506.384	74.663	2.870	3.407
Japan	4,487	578.285	62.783	1.758	2.840
Kazakhstan	3,990	521.620	68.564	5.628	5.775
Kuwait	3,803	317.517	110.496	3.689	4.661
Latvia	3,908	535.591	71.313	1.897	3.019
Lithuania	3,980	530.352	74.260	2.089	2.873
Morocco	3,894	315.555	107.776	4.958	6.068
Netherlands	3,349	542.810	60.553	1.919	2.316
New Zealand	4,940	512.823	72.667	2.034	2.623
Norway	4,108	486.719	75.823	2.358	2.611
Qatar	7,019	326.030	81.249	0.826	1.559
Russian Federation	4,464	530.068	85.168	4.778	4.930
Scotland	3,929	515.785	69.519	2.025	2.186
Singapore	5,041	583.245	70.772	3.021	3.187
Slovak Republic	4,963	492.059	83.359	4.035	4.224
Slovenia	4,351	517.565	72.306	1.894	2.494
Sweden	4,676	529.264	75.256	2.419	2.667
Tunisia	4,134	307.304	113.710	4.707	4.833
Ukraine	4,292	462.083	84.490	2.811	3.248
United States	7,896	543.264	61.271	1.971	2.442
Yemen	+	+	+	+	+
Benchmarking Participants					
Alberta, Canada	4,037	537.239	66.787	2.844	3.675
British Columbia, Canada	4,153	531.238	65.087	2.329	2.783
Dubai, UAE	3,064	461.020	87.313	2.093	2.669
Massachusetts, US	1,747	571.013	72.148	3.242	4.025
Minnesota, US	1,846	557.245	71.466	4.340	4.753
Ontario, Canada	3,496	543.821	62.522	2.562	3.425
Quebec, Canada	3,885	527.072	66.798	2.563	3.621

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.4 Summary Statistics and Standard Errors for Proficiency in Mathematics Knowing in the Fourth Grade

Country	Sample Size	Mathematics Knowing			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	383.954	97.298	5.071	5.412
Armenia	4,079	517.582	93.651	4.652	4.805
Australia	4,108	509.048	90.528	4.014	4.213
Austria	4,859	504.774	61.294	1.688	1.958
Chinese Taipei	4,131	583.997	65.080	1.572	1.732
Colombia	4,801	359.523	85.599	4.680	5.152
Czech Republic	4,235	472.999	57.774	2.277	2.410
Denmark	3,519	512.785	67.318	2.539	2.731
El Salvador	4,166	312.465	95.251	3.590	4.055
England	4,316	544.165	91.465	3.187	3.558
Georgia	4,108	450.243	84.188	3.983	4.027
Germany	5,200	514.427	62.999	1.964	2.021
Hong Kong SAR	3,791	616.614	67.527	3.265	3.509
Hungary	4,048	510.725	85.763	3.237	3.424
Iran, Islamic Rep. of	3,833	409.878	83.634	3.464	3.556
Italy	4,470	514.285	75.088	2.972	3.206
Japan	4,487	564.626	70.549	1.997	2.139
Kazakhstan	3,990	558.627	87.289	7.115	7.251
Kuwait	3,803	326.117	103.480	3.673	4.576
Latvia	3,908	529.768	64.743	2.119	2.223
Lithuania	3,980	520.059	72.534	2.279	2.776
Morocco	3,894	353.727	97.815	4.531	4.825
Netherlands	3,349	525.442	56.532	2.007	2.190
New Zealand	4,940	481.568	92.462	2.479	2.544
Norway	4,108	460.501	74.656	2.507	2.901
Qatar	7,019	292.677	101.483	1.013	1.262
Russian Federation	4,464	538.031	71.323	4.327	4.456
Scotland	3,929	488.930	77.790	2.518	2.617
Singapore	5,041	620.492	85.177	3.840	3.969
Slovak Republic	4,963	492.270	74.647	3.864	3.921
Slovenia	4,351	497.220	60.760	1.393	1.792
Sweden	4,676	482.016	65.016	2.376	2.519
Tunisia	4,134	343.236	104.713	4.587	4.885
Ukraine	4,292	471.933	77.729	2.813	3.019
United States	7,896	541.183	76.644	2.495	2.563
Yemen	+	+	+	+	+
Benchmarking Participants					
Alberta, Canada	4,037	494.241	64.418	3.063	3.101
British Columbia, Canada	4,153	497.723	69.406	2.432	2.485
Dubai, UAE	3,064	456.974	88.156	1.948	2.114
Massachusetts, US	1,747	580.920	70.258	3.900	4.115
Minnesota, US	1,846	565.311	80.275	5.809	6.209
Ontario, Canada	3,496	497.835	70.653	2.937	3.189
Quebec, Canada	3,885	517.442	65.525	2.927	3.189

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.5 Summary Statistics and Standard Errors for Proficiency in Mathematics Applying in the Fourth Grade

Country	Sample Size	Mathematics Applying			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	375.570	88.789	4.855	5.167
Armenia	4,079	492.816	87.513	4.055	4.070
Australia	4,108	522.855	79.031	3.435	3.525
Austria	4,859	507.039	66.527	1.752	1.833
Chinese Taipei	4,131	569.405	64.839	1.590	1.728
Colombia	4,801	357.434	90.074	4.858	5.116
Czech Republic	4,235	496.138	72.967	2.632	2.683
Denmark	3,519	527.709	68.163	2.435	2.502
El Salvador	4,166	338.630	88.134	3.477	3.683
England	4,316	540.486	78.698	2.769	3.050
Georgia	4,108	433.190	96.246	4.486	4.502
Germany	5,200	530.861	64.415	2.061	2.204
Hong Kong SAR	3,791	599.495	61.237	3.158	3.420
Hungary	4,048	507.369	84.116	3.204	3.485
Iran, Islamic Rep. of	3,833	404.619	82.521	3.464	3.696
Italy	4,470	500.987	71.789	2.865	2.922
Japan	4,487	565.605	73.033	1.893	1.975
Kazakhstan	3,990	547.383	88.363	7.092	7.151
Kuwait	3,803	304.992	103.463	3.619	4.142
Latvia	3,908	540.093	68.358	2.178	2.543
Lithuania	3,980	539.176	73.854	2.207	2.446
Morocco	3,894	345.679	95.791	4.699	4.724
Netherlands	3,349	539.908	57.984	1.899	1.995
New Zealand	4,940	495.241	78.599	2.153	2.335
Norway	4,108	479.069	70.045	2.432	2.826
Qatar	7,019	296.431	90.045	0.907	1.159
Russian Federation	4,464	546.737	86.500	4.762	4.802
Scotland	3,929	499.679	74.092	2.103	2.350
Singapore	5,041	589.533	77.304	3.435	3.676
Slovak Republic	4,963	497.976	82.411	3.999	4.035
Slovenia	4,351	503.754	68.019	1.564	1.906
Sweden	4,676	508.404	60.104	2.086	2.168
Tunisia	4,134	329.002	111.229	4.776	4.843
Ukraine	4,292	466.441	86.659	2.927	3.146
United States	7,896	523.904	73.506	2.385	2.610
Yemen	+	+	+	+	+
Benchmarking Participants					
Alberta, Canada	4,037	505.432	64.432	2.818	2.867
British Columbia, Canada	4,153	505.221	67.309	2.562	2.566
Dubai, UAE	3,064	440.841	85.015	1.550	1.705
Massachusetts, US	1,747	565.787	68.838	3.242	3.476
Minnesota, US	1,846	547.871	79.260	5.359	5.458
Ontario, Canada	3,496	514.840	64.560	2.746	3.076
Quebec, Canada	3,885	517.289	63.871	2.592	2.753

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.6 Summary Statistics and Standard Errors for Proficiency in Mathematics Reasoning in the Fourth Grade

Country	Sample Size	Mathematics Reasoning			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	386.715	93.274	4.713	4.750
Armenia	4,079	489.302	95.625	4.535	4.693
Australia	4,108	516.265	73.607	3.085	3.391
Austria	4,859	506.257	70.190	1.730	2.113
Chinese Taipei	4,131	565.766	71.783	1.662	1.862
Colombia	4,801	372.023	97.979	4.514	4.877
Czech Republic	4,235	492.634	79.388	3.105	3.408
Denmark	3,519	524.397	71.383	1.946	2.088
El Salvador	4,166	355.889	93.231	3.388	4.003
England	4,316	537.381	76.988	2.629	3.096
Georgia	4,108	437.371	85.694	3.784	4.181
Germany	5,200	528.289	68.439	2.001	2.546
Hong Kong SAR	3,791	588.933	70.283	3.160	3.493
Hungary	4,048	509.030	95.572	3.667	3.821
Iran, Islamic Rep. of	3,833	410.055	83.351	3.312	3.776
Italy	4,470	509.412	78.255	2.782	3.078
Japan	4,487	563.031	78.224	2.017	2.092
Kazakhstan	3,990	538.775	74.152	6.025	6.091
Kuwait	+	+	+	+	+
Latvia	3,908	537.314	73.699	2.298	2.458
Lithuania	3,980	526.013	75.850	2.237	2.549
Morocco	+	+	+	+	+
Netherlands	3,349	534.152	67.055	2.197	2.356
New Zealand	4,940	503.208	81.234	2.317	2.808
Norway	4,108	488.764	76.252	2.444	2.687
Qatar	+	+	+	+	+
Russian Federation	4,464	540.366	81.098	4.745	4.827
Scotland	3,929	497.093	78.729	2.149	2.207
Singapore	5,041	577.657	83.469	3.745	3.778
Slovak Republic	4,963	499.403	87.103	3.908	3.982
Slovenia	4,351	504.970	77.080	1.529	2.054
Sweden	4,676	519.077	67.844	2.307	2.503
Tunisia	+	+	+	+	+
Ukraine	4,292	474.257	81.309	2.796	3.198
United States	7,896	523.211	67.120	2.134	2.194
Yemen	+	+	+	+	+
Benchmarking Participants					
Alberta, Canada	4,037	519.312	65.103	2.645	3.056
British Columbia, Canada	4,153	516.418	68.246	2.148	2.341
Dubai, UAE	3,064	445.695	92.340	1.981	2.876
Massachusetts, US	1,747	564.656	66.253	3.025	3.237
Minnesota, US	1,846	542.604	66.303	4.839	5.051
Ontario, Canada	3,496	525.702	65.325	2.470	2.641
Quebec, Canada	3,885	522.984	68.436	2.609	2.969

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.7 Summary Statistics and Standard Errors for Proficiency in Life Science in the Fourth Grade

Country	Sample Size	Life Science			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	351.168	105.043	6.019	6.207
Armenia	4,079	489.098	117.828	5.486	5.861
Australia	4,108	528.201	71.898	2.978	3.384
Austria	4,859	525.823	66.658	1.828	2.027
Chinese Taipei	4,131	540.879	67.619	1.930	2.149
Colombia	4,801	408.445	91.678	4.992	5.207
Czech Republic	4,235	519.564	68.287	2.564	2.898
Denmark	3,519	526.957	70.161	2.162	2.445
El Salvador	4,166	409.749	89.399	3.025	3.555
England	4,316	532.264	77.132	2.592	2.692
Georgia	4,108	426.944	75.857	3.444	3.496
Germany	5,200	528.906	66.312	1.909	2.009
Hong Kong SAR	3,791	532.105	60.694	3.134	3.491
Hungary	4,048	547.744	75.764	2.780	2.850
Iran, Islamic Rep. of	3,833	442.484	90.677	4.041	4.378
Italy	4,470	548.981	74.988	2.835	2.962
Japan	4,487	530.357	67.275	1.649	1.991
Kazakhstan	3,990	527.568	62.753	4.725	5.026
Kuwait	3,803	352.820	138.928	4.381	4.929
Latvia	3,908	535.055	58.841	1.703	2.118
Lithuania	3,980	516.282	59.387	1.588	1.801
Morocco	3,894	291.706	127.103	5.861	6.766
Netherlands	3,349	535.822	54.755	2.090	2.222
New Zealand	4,940	506.500	86.256	2.422	2.505
Norway	4,108	486.638	66.617	2.391	2.536
Qatar	7,019	291.105	130.979	1.313	1.433
Russian Federation	4,464	538.699	67.811	4.098	4.108
Scotland	3,929	503.754	74.422	2.035	2.172
Singapore	5,041	582.189	95.356	3.829	4.083
Slovak Republic	4,963	531.768	78.620	4.022	4.041
Slovenia	4,351	510.508	65.511	1.585	2.175
Sweden	4,676	530.822	63.010	2.343	2.529
Tunisia	4,134	323.427	136.110	5.344	5.565
Ukraine	4,292	481.915	72.161	2.260	2.478
United States	7,896	539.622	81.025	2.444	2.533
Yemen	+	+	+	+	+
Benchmarking Participants					
Alberta, Canada	4,037	541.112	70.962	3.418	3.693
British Columbia, Canada	4,153	538.147	67.623	2.348	2.766
Dubai, UAE	3,064	457.495	103.945	2.496	2.842
Massachusetts, US	1,747	568.138	70.411	3.306	3.491
Minnesota, US	1,846	544.836	75.295	5.787	6.102
Ontario, Canada	3,496	535.132	76.964	3.464	3.709
Quebec, Canada	3,885	522.286	63.146	2.523	2.659

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.8 Summary Statistics and Standard Errors for Proficiency in Physical Science in the Fourth Grade

Country	Sample Size	Physical Science			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	377.110	107.678	5.223	5.313
Armenia	4,079	492.365	107.232	4.933	5.081
Australia	4,108	522.334	75.069	3.004	3.081
Austria	4,859	513.647	78.231	2.164	2.404
Chinese Taipei	4,131	559.332	85.425	2.192	2.549
Colombia	4,801	411.234	100.703	4.832	4.880
Czech Republic	4,235	510.822	68.129	2.447	2.794
Denmark	3,519	502.328	80.656	2.373	2.484
El Salvador	4,166	391.749	98.094	3.023	3.780
England	4,316	542.745	72.053	2.535	2.657
Georgia	4,108	413.514	88.050	3.693	4.019
Germany	5,200	523.734	78.175	2.259	2.544
Hong Kong SAR	3,791	558.063	66.349	3.225	3.484
Hungary	4,048	529.224	78.967	2.996	3.262
Iran, Islamic Rep. of	3,833	453.703	101.188	4.144	4.186
Italy	4,470	520.880	77.479	2.926	3.056
Japan	4,487	564.382	65.851	1.706	2.340
Kazakhstan	3,990	527.612	76.391	5.709	5.802
Kuwait	3,803	345.348	127.324	3.982	5.152
Latvia	3,908	543.956	64.517	2.187	2.407
Lithuania	3,980	514.069	61.281	1.393	1.447
Morocco	3,894	323.801	136.165	5.151	5.510
Netherlands	3,349	502.685	58.112	1.804	2.309
New Zealand	4,940	498.330	85.728	2.262	2.495
Norway	4,108	468.714	69.474	2.105	2.661
Qatar	7,019	303.044	152.671	1.577	2.055
Russian Federation	4,464	547.384	80.496	4.426	4.551
Scotland	3,929	499.399	68.495	1.833	1.897
Singapore	5,041	585.055	92.126	3.505	3.895
Slovak Republic	4,963	512.651	84.012	4.375	4.571
Slovenia	4,351	529.645	69.225	1.484	1.644
Sweden	4,676	508.103	75.978	2.633	2.656
Tunisia	4,134	339.809	152.881	6.076	6.419
Ukraine	4,292	474.817	83.439	2.479	2.723
United States	7,896	534.158	77.372	2.221	2.350
Yemen	+	+	+	+	+
Benchmarking Participants					
Alberta, Canada	4,037	534.519	68.421	2.938	3.125
British Columbia, Canada	4,153	530.812	68.136	2.438	2.618
Dubai, UAE	3,064	466.808	105.525	2.222	2.820
Massachusetts, US	1,747	560.149	62.025	3.615	4.391
Minnesota, US	1,846	544.543	75.348	5.328	5.351
Ontario, Canada	3,496	534.935	71.870	2.725	2.942
Quebec, Canada	3,885	513.218	64.679	2.308	2.577

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.9 Summary Statistics and Standard Errors for Proficiency in Earth Science in the Fourth Grade

Country	Sample Size	Earth Science			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	365.160	109.492	5.574	5.748
Armenia	4,079	478.865	115.865	5.272	5.536
Australia	4,108	534.144	72.157	3.006	3.194
Austria	4,859	532.239	82.279	1.922	1.939
Chinese Taipei	4,131	553.276	74.995	1.699	1.947
Colombia	4,801	401.055	106.686	5.079	5.578
Czech Republic	4,235	517.598	72.376	2.356	2.607
Denmark	3,519	521.832	71.991	2.627	2.716
El Salvador	4,166	393.319	102.009	3.017	3.335
England	4,316	538.446	73.486	2.449	2.949
Georgia	4,108	431.778	104.063	4.719	4.977
Germany	5,200	523.912	78.809	2.255	2.370
Hong Kong SAR	3,791	559.544	65.760	2.912	3.150
Hungary	4,048	516.726	87.028	3.025	3.466
Iran, Islamic Rep. of	3,833	433.364	98.274	3.825	4.106
Italy	4,470	526.031	80.692	2.870	3.043
Japan	4,487	528.608	71.574	1.859	2.713
Kazakhstan	3,990	534.056	71.102	4.947	5.160
Kuwait	3,803	362.660	120.570	3.660	3.835
Latvia	3,908	535.805	68.202	2.100	2.201
Lithuania	3,980	510.584	68.521	2.024	2.455
Morocco	3,894	292.501	140.623	6.019	6.213
Netherlands	3,349	523.548	72.288	2.235	2.537
New Zealand	4,940	514.787	78.006	2.019	2.647
Norway	4,108	496.638	80.938	2.614	2.932
Qatar	7,019	304.917	127.406	1.287	2.218
Russian Federation	4,464	536.337	76.222	4.242	4.302
Scotland	3,929	507.849	69.763	1.802	2.519
Singapore	5,041	553.520	81.208	3.230	3.317
Slovak Republic	4,963	530.287	79.911	4.171	4.781
Slovenia	4,351	517.011	79.206	2.040	2.502
Sweden	4,676	534.880	75.213	2.455	2.710
Tunisia	4,134	325.029	142.181	5.208	5.789
Ukraine	4,292	473.776	84.301	2.777	3.087
United States	7,896	533.463	76.848	2.362	2.586
Yemen	+	+	+	+	+
Benchmarking Participants					
Alberta, Canada	4,037	544.019	67.736	3.065	3.343
British Columbia, Canada	4,153	536.806	64.506	2.280	2.659
Dubai, UAE	3,064	470.760	95.846	2.182	2.569
Massachusetts, US	1,747	557.924	72.579	4.139	4.434
Minnesota, US	1,846	547.013	76.697	5.305	5.784
Ontario, Canada	3,496	529.540	67.655	2.885	3.175
Quebec, Canada	3,885	523.219	61.114	2.349	2.589

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.10 Summary Statistics and Standard Errors for Proficiency in Science Knowing in the Fourth Grade

Country	Sample Size	Science Knowing			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	350.300	102.061	5.574	5.771
Armenia	4,079	486.136	110.880	5.225	5.232
Australia	4,108	529.092	74.426	3.090	3.119
Austria	4,859	529.244	72.189	1.992	2.013
Chinese Taipei	4,131	536.348	75.403	1.894	2.483
Colombia	4,801	409.078	99.597	5.250	5.453
Czech Republic	4,235	519.559	68.346	2.381	2.664
Denmark	3,519	515.932	75.522	2.725	2.929
El Salvador	4,166	410.218	95.333	3.294	3.910
England	4,316	542.960	80.349	2.690	2.912
Georgia	4,108	434.337	81.268	3.769	3.839
Germany	5,200	527.486	71.734	2.016	2.151
Hong Kong SAR	3,791	545.536	64.624	2.957	3.249
Hungary	4,048	539.670	77.225	2.884	2.978
Iran, Islamic Rep. of	3,833	437.377	96.592	4.032	4.330
Italy	4,470	529.934	82.853	3.374	3.892
Japan	4,487	528.293	66.863	1.909	2.201
Kazakhstan	3,990	533.687	78.255	5.770	5.836
Kuwait	3,803	360.002	120.660	3.620	3.879
Latvia	3,908	539.561	61.317	1.962	2.165
Lithuania	3,980	510.729	57.054	1.485	1.708
Morocco	3,894	290.725	130.830	5.626	5.836
Netherlands	3,349	517.683	57.943	1.815	2.479
New Zealand	4,940	510.533	82.562	2.157	2.530
Norway	4,108	484.920	68.773	2.331	2.423
Qatar	7,019	303.548	140.056	1.473	2.260
Russian Federation	4,464	541.670	73.846	4.643	4.838
Scotland	3,929	510.710	69.556	1.865	2.045
Singapore	5,041	587.014	101.504	3.925	4.079
Slovak Republic	4,963	527.369	81.409	4.224	4.393
Slovenia	4,351	511.313	68.595	1.593	1.648
Sweden	4,676	525.595	68.973	2.459	2.499
Tunisia	4,134	315.512	144.012	5.553	5.861
Ukraine	4,292	476.024	74.547	2.305	2.439
United States	7,896	541.317	78.928	2.198	2.274
Yemen	+	+	+	+	+
Benchmarking Participants					
Alberta, Canada	4,037	548.730	71.785	3.232	3.486
British Columbia, Canada	4,153	539.490	69.737	2.342	2.536
Dubai, UAE	3,064	462.769	107.432	2.087	2.557
Massachusetts, US	1,747	566.167	69.880	3.862	4.427
Minnesota, US	1,846	549.634	78.041	5.646	5.853
Ontario, Canada	3,496	538.367	75.323	3.221	3.397
Quebec, Canada	3,885	516.166	64.219	2.417	2.756

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.11 Summary Statistics and Standard Errors for Proficiency in Science Applying in the Fourth Grade

Country	Sample Size	Science Applying			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	379.021	99.361	5.280	5.684
Armenia	4,079	486.720	121.406	5.580	5.604
Australia	4,108	523.066	77.132	3.020	3.312
Austria	4,859	526.030	76.186	2.018	2.156
Chinese Taipei	4,131	556.064	71.070	1.803	2.098
Colombia	4,801	403.567	97.254	5.155	5.400
Czech Republic	4,235	515.871	72.146	2.746	3.114
Denmark	3,519	515.208	72.422	2.263	2.556
El Salvador	4,166	393.469	92.780	3.322	3.569
England	4,316	536.118	76.289	2.443	2.653
Georgia	4,108	423.639	89.445	4.057	4.105
Germany	5,200	526.309	74.281	2.159	2.177
Hong Kong SAR	3,791	549.416	56.871	2.751	2.950
Hungary	4,048	530.669	80.881	3.126	3.198
Iran, Islamic Rep. of	3,833	450.520	89.407	3.927	4.256
Italy	4,470	539.142	74.620	2.981	3.073
Japan	4,487	542.351	64.050	1.775	2.745
Kazakhstan	3,990	535.537	66.630	4.786	4.937
Kuwait	3,803	337.972	140.437	4.100	4.268
Latvia	3,908	534.649	64.100	2.149	2.425
Lithuania	3,980	515.060	65.114	1.812	2.755
Morocco	3,894	311.121	130.922	5.938	6.280
Netherlands	3,349	525.241	59.405	2.018	2.225
New Zealand	4,940	499.577	88.891	2.275	2.391
Norway	4,108	478.350	76.833	2.490	2.790
Qatar	7,019	283.039	140.934	1.432	2.664
Russian Federation	4,464	546.307	72.631	4.564	4.706
Scotland	3,929	493.812	76.592	2.064	2.425
Singapore	5,041	578.640	81.501	3.377	3.699
Slovak Republic	4,963	526.801	82.573	4.303	4.422
Slovenia	4,351	525.100	73.863	1.887	2.130
Sweden	4,676	520.714	72.083	2.499	2.852
Tunisia	4,134	329.075	146.437	5.772	6.266
Ukraine	4,292	477.285	77.649	2.636	3.152
United States	7,896	532.990	82.756	2.565	2.786
Yemen	+	+	+	+	+
Benchmarking Participants					
Alberta, Canada	4,037	534.949	71.335	3.499	3.712
British Columbia, Canada	4,153	533.176	68.513	2.263	2.385
Dubai, UAE	3,064	463.196	101.442	1.942	2.626
Massachusetts, US	1,747	563.121	68.675	4.140	4.360
Minnesota, US	1,846	544.233	81.236	5.795	5.890
Ontario, Canada	3,496	528.056	76.712	3.169	3.417
Quebec, Canada	3,885	515.121	62.083	2.501	2.650

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.12 Summary Statistics and Standard Errors for Proficiency in Science Reasoning in the Fourth Grade

Country	Sample Size	Science Reasoning			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	4,223	357.397	107.884	4.919	5.769
Armenia	4,079	484.175	116.357	4.985	5.258
Australia	4,108	530.259	79.407	3.116	3.378
Austria	4,859	513.256	77.111	2.218	2.338
Chinese Taipei	4,131	570.802	87.174	2.063	2.383
Colombia	4,801	409.056	97.454	4.827	5.123
Czech Republic	4,235	510.027	75.305	2.812	2.901
Denmark	3,519	525.434	80.010	2.865	3.791
El Salvador	4,166	376.398	99.483	3.481	4.034
England	4,316	537.252	70.511	2.363	2.689
Georgia	4,108	388.091	98.600	4.507	4.943
Germany	5,200	525.005	82.058	2.225	2.275
Hong Kong SAR	3,791	561.105	78.734	3.632	4.363
Hungary	4,048	529.055	94.505	3.534	3.691
Iran, Islamic Rep. of	3,833	436.072	97.377	3.999	4.350
Italy	4,470	525.686	80.614	2.995	3.781
Japan	4,487	567.424	61.462	1.729	2.081
Kazakhstan	3,990	519.380	76.992	5.241	5.324
Kuwait	3,803	330.941	132.494	4.042	5.358
Latvia	3,908	550.769	71.556	2.274	2.747
Lithuania	3,980	524.135	71.919	1.987	2.373
Morocco	3,894	317.712	119.143	4.931	5.350
Netherlands	3,349	525.421	55.642	1.915	2.252
New Zealand	4,940	505.200	85.759	2.044	2.908
Norway	4,108	480.024	70.618	2.097	3.202
Qatar	7,019	292.623	119.315	1.243	2.858
Russian Federation	4,464	542.295	87.213	4.611	4.629
Scotland	3,929	500.674	76.907	1.824	2.186
Singapore	5,041	567.752	87.578	3.214	3.707
Slovak Republic	4,963	513.444	92.232	4.665	4.878
Slovenia	4,351	527.440	75.537	1.555	1.794
Sweden	4,676	527.296	70.950	2.614	3.489
Tunisia	4,134	349.024	126.868	4.869	5.344
Ukraine	4,292	478.145	88.814	2.757	2.963
United States	7,896	534.873	78.012	2.236	2.641
Yemen	+	+	+	+	+
Benchmarking Participants					
Alberta, Canada	4,037	536.740	74.222	3.658	4.401
British Columbia, Canada	4,153	535.881	72.688	2.390	2.708
Dubai, UAE	3,064	461.739	100.757	2.452	2.602
Massachusetts, US	1,747	568.581	68.198	4.268	6.205
Minnesota, US	1,846	548.842	77.528	5.648	6.358
Ontario, Canada	3,496	541.019	75.233	2.966	3.140
Quebec, Canada	3,885	527.518	67.942	2.616	3.255

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.13 Summary Statistics and Standard Errors for Proficiency in Number in the Eighth Grade

Country	Sample Size	Number			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	403.213	65.685	1.460	1.712
Armenia	4,689	492.486	89.825	3.100	3.124
Australia	4,069	503.343	83.587	3.625	3.695
Bahrain	4,230	387.993	83.246	1.362	2.000
Bosnia and Herzegovina	4,220	450.777	74.211	2.519	2.975
Botswana	4,208	366.481	86.941	2.417	2.927
Bulgaria	4,019	457.930	97.873	4.453	4.665
Chinese Taipei	4,046	576.956	103.754	4.103	4.240
Colombia	4,873	369.308	97.213	3.369	3.453
Cyprus	4,399	464.280	88.895	1.470	1.649
Czech Republic	4,845	511.045	78.659	2.361	2.488
Egypt	6,582	392.659	95.445	2.969	3.138
El Salvador	4,063	354.709	81.048	2.555	2.954
England	4,025	509.720	85.300	4.782	4.977
Georgia	4,178	420.808	87.158	5.442	5.552
Ghana	5,294	309.799	96.730	3.880	3.950
Hong Kong SAR	3,470	566.891	95.544	5.604	5.638
Hungary	4,111	516.915	87.947	3.334	3.599
Indonesia	4,203	399.181	87.214	3.419	3.722
Iran, Islamic Rep. of	3,981	394.931	91.360	3.822	3.878
Israel	3,294	468.888	92.556	3.144	3.241
Italy	4,408	477.509	75.709	2.600	2.751
Japan	4,312	551.397	91.314	2.152	2.304
Jordan	5,251	416.473	100.425	3.998	4.293
Korea, Rep. of	4,240	583.341	96.344	2.346	2.365
Kuwait	4,091	346.661	87.770	2.343	3.127
Lebanon	3,786	454.436	71.255	3.305	3.351
Lithuania	3,991	506.007	80.087	2.386	2.659
Malaysia	4,466	490.504	85.116	4.951	5.135
Malta	4,670	495.912	96.272	1.008	1.319
Morocco	3,060	389.488	83.023	2.960	3.385
Norway	4,627	487.511	68.064	1.873	1.967
Oman	4,752	362.824	88.413	2.411	2.732
Palestinian Nat'l Auth.	4,378	365.552	102.227	3.187	3.241
Qatar	7,184	334.344	87.241	0.754	1.566
Romania	4,198	456.990	93.722	3.423	3.492
Russian Federation	4,472	506.639	82.304	3.423	3.776
Saudi Arabia	4,243	309.411	96.547	2.603	3.280
Scotland	4,070	488.883	81.803	3.537	3.691
Serbia	4,045	477.765	86.099	2.617	2.888
Singapore	4,599	597.225	94.007	3.482	3.519
Slovenia	4,043	502.038	76.281	1.737	2.257
Sweden	5,215	506.844	67.118	1.707	1.760
Syrian Arab Republic	4,650	393.023	84.237	3.132	3.431
Thailand	5,412	443.654	95.546	4.727	4.805
Tunisia	4,080	425.031	70.800	2.283	2.583
Turkey	4,498	429.167	101.776	3.872	4.048
Ukraine	4,424	459.864	86.046	3.281	3.659
United States	7,377	510.165	80.628	2.726	2.739
Benchmarking Participants					
Basque Country, Spain	2,296	508.930	66.526	2.630	2.852
British Columbia, Canada	4,256	520.053	78.464	3.098	3.186
Dubai, UAE	3,195	458.321	97.499	2.463	3.164
Massachusetts, US	1,897	548.299	86.883	5.101	5.185
Minnesota, US	1,777	536.754	71.257	4.118	4.289
Ontario, Canada	3,448	524.789	77.030	3.780	3.959
Quebec, Canada	3,956	534.112	71.183	3.087	3.407

Exhibit E.14 Summary Statistics and Standard Errors for Proficiency in Algebra in the Eighth Grade

Country	Sample Size	Algebra			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	349.470	83.997	1.936	2.434
Armenia	4,689	531.532	82.046	2.327	2.498
Australia	4,069	470.888	80.289	3.557	3.707
Bahrain	4,230	403.175	93.720	1.622	1.828
Bosnia and Herzegovina	4,220	475.055	78.642	2.798	3.196
Botswana	4,208	393.943	70.298	1.757	2.169
Bulgaria	4,019	475.942	105.529	5.042	5.101
Chinese Taipei	4,046	617.380	128.261	5.276	5.439
Colombia	4,873	390.247	78.267	2.763	3.123
Cyprus	4,399	467.927	84.365	1.537	2.001
Czech Republic	4,845	483.684	73.813	2.266	2.352
Egypt	6,582	409.360	94.309	3.275	3.299
El Salvador	4,063	331.123	82.233	2.742	3.702
England	4,025	491.763	83.552	4.473	4.621
Georgia	4,178	421.189	107.125	6.517	6.595
Ghana	5,294	358.077	87.397	3.439	3.605
Hong Kong SAR	3,470	565.497	96.965	5.524	5.570
Hungary	4,111	503.227	84.179	3.445	3.578
Indonesia	4,203	405.263	87.852	3.219	3.469
Iran, Islamic Rep. of	3,981	408.052	84.800	3.848	3.918
Israel	3,294	469.599	96.168	3.690	3.930
Italy	4,408	460.466	79.610	3.110	3.242
Japan	4,312	559.081	92.795	2.419	2.540
Jordan	5,251	447.939	97.242	3.985	4.079
Korea, Rep. of	4,240	596.223	110.273	2.907	3.008
Kuwait	4,091	354.173	87.375	2.391	2.992
Lebanon	3,786	464.788	75.426	3.173	3.201
Lithuania	3,991	482.548	87.700	2.623	2.728
Malaysia	4,466	454.060	74.394	4.273	4.294
Malta	4,670	473.396	81.683	0.906	1.442
Morocco	3,060	362.339	103.047	3.383	3.958
Norway	4,627	425.378	70.588	2.176	2.773
Oman	4,752	391.466	102.415	3.002	3.178
Palestinian Nat'l Auth.	4,378	382.288	96.950	3.000	3.354
Qatar	7,184	312.067	96.987	0.929	1.488
Romania	4,198	478.360	105.002	4.385	4.603
Russian Federation	4,472	518.395	91.525	4.458	4.531
Saudi Arabia	4,243	343.915	79.755	1.980	2.797
Scotland	4,070	466.847	80.959	3.530	3.710
Serbia	4,045	500.177	94.192	3.073	3.247
Singapore	4,599	579.148	93.721	3.592	3.661
Slovenia	4,043	488.090	74.954	2.218	2.361
Sweden	5,215	456.398	76.887	2.235	2.431
Syrian Arab Republic	4,650	405.619	89.820	3.635	3.694
Thailand	5,412	433.389	95.006	4.998	5.040
Tunisia	4,080	423.151	64.624	2.437	2.608
Turkey	4,498	440.248	114.887	4.879	5.139
Ukraine	4,424	463.975	95.091	3.605	3.868
United States	7,377	500.601	74.262	2.517	2.727
Benchmarking Participants					
Basque Country, Spain	2,296	484.884	71.626	2.931	3.100
British Columbia, Canada	4,256	489.306	73.565	2.910	3.074
Dubai, UAE	3,195	474.535	94.549	2.196	2.423
Massachusetts, US	1,897	537.901	79.814	4.678	4.846
Minnesota, US	1,777	515.119	69.180	4.546	4.659
Ontario, Canada	3,448	489.601	71.559	3.151	3.659
Quebec, Canada	3,956	504.690	72.950	3.101	3.301

Exhibit E.15 Summary Statistics and Standard Errors for Proficiency in Geometry in the Eighth Grade

Country	Sample Size	Geometry			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	432.138	63.313	1.439	2.098
Armenia	4,689	492.845	90.514	4.051	4.122
Australia	4,069	487.438	74.741	3.271	3.628
Bahrain	4,230	412.270	86.990	1.448	2.113
Bosnia and Herzegovina	4,220	450.897	79.179	2.806	3.475
Botswana	4,208	324.546	99.348	2.661	3.180
Bulgaria	4,019	468.231	98.455	4.421	5.047
Chinese Taipei	4,046	591.966	101.853	4.084	4.620
Colombia	4,873	371.346	76.549	3.214	3.299
Cyprus	4,399	457.676	96.585	1.990	2.698
Czech Republic	4,845	497.622	76.540	2.238	2.725
Egypt	6,582	406.314	100.057	3.319	3.408
El Salvador	4,063	317.663	87.204	3.212	3.680
England	4,025	510.117	77.803	4.289	4.408
Georgia	4,178	408.607	104.577	6.489	6.710
Ghana	5,294	274.516	109.317	4.472	4.859
Hong Kong SAR	3,470	569.904	85.246	5.375	5.472
Hungary	4,111	507.594	87.700	3.547	3.627
Indonesia	4,203	394.579	98.354	3.929	4.450
Iran, Islamic Rep. of	3,981	422.674	90.126	4.104	4.382
Israel	3,294	436.045	97.892	3.901	4.277
Italy	4,408	489.591	78.356	2.977	3.054
Japan	4,312	572.856	71.616	1.906	2.158
Jordan	5,251	435.601	96.429	3.746	3.852
Korea, Rep. of	4,240	586.590	83.479	2.067	2.327
Kuwait	4,091	384.645	81.606	2.659	2.841
Lebanon	3,786	462.127	75.199	3.506	3.955
Lithuania	3,991	506.922	79.745	2.235	2.647
Malaysia	4,466	476.893	88.244	5.520	5.553
Malta	4,670	495.116	87.215	0.804	1.118
Morocco	3,060	396.372	90.518	3.301	3.643
Norway	4,627	458.710	76.513	2.264	2.285
Oman	4,752	387.457	95.476	2.658	3.032
Palestinian Nat'l Auth.	4,378	388.168	99.800	3.101	3.779
Qatar	7,184	301.457	102.787	1.081	1.804
Romania	4,198	466.435	97.862	3.900	4.027
Russian Federation	4,472	509.631	76.112	3.669	4.069
Saudi Arabia	4,243	358.898	82.147	2.302	2.608
Scotland	4,070	485.452	74.903	3.271	3.854
Serbia	4,045	485.742	90.015	3.416	3.599
Singapore	4,599	578.350	85.182	3.332	3.369
Slovenia	4,043	499.461	70.114	2.140	2.402
Sweden	5,215	471.686	80.670	2.329	2.517
Syrian Arab Republic	4,650	417.187	85.441	3.160	3.437
Thailand	5,412	441.933	95.640	5.075	5.322
Tunisia	4,080	436.791	68.914	2.282	2.588
Turkey	4,498	411.109	113.026	4.807	5.091
Ukraine	4,424	467.221	82.808	3.252	3.567
United States	7,377	479.943	70.712	2.399	2.525
Benchmarking Participants					
Basque Country, Spain	2,296	476.328	76.259	3.376	3.716
British Columbia, Canada	4,256	486.936	75.056	3.054	3.718
Dubai, UAE	3,195	450.873	88.844	2.409	3.422
Massachusetts, US	1,897	519.231	74.315	4.198	4.297
Minnesota, US	1,777	505.065	67.359	4.057	4.372
Ontario, Canada	3,448	507.977	70.981	3.923	4.215
Quebec, Canada	3,956	523.001	69.729	3.164	3.269

Exhibit E.16 Summary Statistics and Standard Errors for Proficiency in Data and Chance in the Eighth Grade

Country	Sample Size	Data and Chance			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	371.143	65.189	1.425	1.717
Armenia	4,689	426.761	108.793	3.863	3.925
Australia	4,069	525.295	84.705	3.084	3.209
Bahrain	4,230	418.222	70.031	1.299	2.131
Bosnia and Herzegovina	4,220	437.493	85.816	2.234	2.313
Botswana	4,208	383.835	74.550	1.548	2.605
Bulgaria	4,019	440.206	110.674	4.391	4.733
Chinese Taipei	4,046	565.603	88.894	3.367	3.646
Colombia	4,873	405.115	84.356	3.625	3.834
Cyprus	4,399	464.303	90.753	1.427	1.649
Czech Republic	4,845	511.667	81.205	2.569	2.799
Egypt	6,582	383.997	87.310	2.574	3.149
El Salvador	4,063	361.784	80.658	2.345	2.994
England	4,025	547.261	94.618	4.837	4.970
Georgia	4,178	372.737	106.358	3.807	4.256
Ghana	5,294	320.558	86.422	3.180	3.647
Hong Kong SAR	3,470	548.909	80.549	4.710	4.730
Hungary	4,111	523.621	82.712	3.143	3.254
Indonesia	4,203	402.427	89.834	3.165	3.580
Iran, Islamic Rep. of	3,981	414.506	76.862	3.241	3.466
Israel	3,294	465.406	114.125	4.203	4.382
Italy	4,408	490.601	83.617	2.944	3.128
Japan	4,312	573.325	74.262	1.646	2.192
Jordan	5,251	424.877	94.222	3.547	3.781
Korea, Rep. of	4,240	579.661	75.758	1.715	1.968
Kuwait	4,091	365.996	77.967	2.180	3.538
Lebanon	3,786	407.275	85.230	4.076	4.437
Lithuania	3,991	523.385	80.025	1.871	2.273
Malaysia	4,466	468.836	71.798	3.874	4.093
Malta	4,670	486.961	102.254	0.980	1.413
Morocco	3,060	371.141	92.504	3.294	3.382
Norway	4,627	504.982	94.770	2.391	2.481
Oman	4,752	389.389	87.227	2.468	3.015
Palestinian Nat'l Auth.	4,378	370.585	93.907	2.721	2.915
Qatar	7,184	305.240	97.657	0.875	1.649
Romania	4,198	428.768	99.453	3.591	3.688
Russian Federation	4,472	487.139	85.842	3.481	3.837
Saudi Arabia	4,243	348.391	74.424	1.749	2.161
Scotland	4,070	516.512	82.023	3.334	3.451
Serbia	4,045	458.224	93.973	2.733	3.033
Singapore	4,599	573.958	96.242	3.525	3.866
Slovenia	4,043	510.962	74.122	2.068	2.317
Sweden	5,215	525.690	96.760	2.775	3.044
Syrian Arab Republic	4,650	387.241	77.129	2.449	2.660
Thailand	5,412	452.884	79.393	3.592	4.082
Tunisia	4,080	411.176	73.373	2.171	2.327
Turkey	4,498	444.830	98.557	3.966	4.361
Ukraine	4,424	457.571	87.490	3.209	3.548
United States	7,377	530.647	86.819	2.792	2.817
Benchmarking Participants					
Basque Country, Spain	2,296	503.734	80.762	2.859	3.661
British Columbia, Canada	4,256	529.151	81.799	3.129	3.241
Dubai, UAE	3,195	456.834	99.907	2.503	3.205
Massachusetts, US	1,897	568.805	95.351	4.770	5.170
Minnesota, US	1,777	560.223	85.620	4.894	5.357
Ontario, Canada	3,448	543.461	83.528	3.840	4.244
Quebec, Canada	3,956	533.170	73.853	2.887	2.979

Exhibit E.17 Summary Statistics and Standard Errors for Proficiency in Mathematics Knowing in the Eighth Grade

Country	Sample Size	Mathematics Knowing			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	370.844	66.511	1.714	1.886
Armenia	4,689	506.739	76.632	3.079	3.136
Australia	4,069	487.476	70.703	3.308	3.335
Bahrain	4,230	394.660	88.193	1.511	1.734
Bosnia and Herzegovina	4,220	477.982	75.876	2.646	2.908
Botswana	4,208	376.467	76.079	1.935	2.067
Bulgaria	4,019	476.974	98.191	4.612	4.662
Chinese Taipei	4,046	593.722	104.249	4.371	4.531
Colombia	4,873	364.183	79.311	3.304	3.442
Cyprus	4,399	468.458	77.953	1.531	1.600
Czech Republic	4,845	502.414	68.269	2.402	2.478
Egypt	6,582	392.096	100.852	3.553	3.614
El Salvador	4,063	335.589	78.854	2.801	3.141
England	4,025	503.328	70.176	3.922	4.041
Georgia	4,178	426.850	98.228	5.739	5.812
Ghana	5,294	313.231	105.121	4.539	4.600
Hong Kong SAR	3,470	573.639	87.202	5.312	5.351
Hungary	4,111	518.288	80.307	3.229	3.279
Indonesia	4,203	396.619	96.328	3.808	3.978
Iran, Islamic Rep. of	3,981	403.309	83.560	3.748	4.077
Israel	3,294	473.284	89.123	3.453	3.730
Italy	4,408	476.043	71.452	2.967	2.999
Japan	4,312	560.004	77.224	1.828	2.249
Jordan	5,251	431.751	101.723	4.158	4.224
Korea, Rep. of	4,240	596.328	90.930	2.360	2.542
Kuwait	4,091	347.026	85.732	2.612	3.060
Lebanon	3,786	464.065	74.588	3.798	3.879
Lithuania	3,991	507.604	80.388	2.410	2.483
Malaysia	4,466	476.671	75.581	4.777	4.817
Malta	4,670	490.438	86.102	0.822	1.610
Morocco	3,060	364.908	100.584	3.922	4.400
Norway	4,627	458.171	51.440	1.624	1.839
Oman	4,752	372.075	101.219	3.328	3.468
Palestinian Nat'l Auth.	4,378	365.218	107.992	3.633	3.798
Qatar	7,184	306.934	99.277	1.057	1.438
Romania	4,198	470.063	100.190	4.068	4.173
Russian Federation	4,472	521.147	82.054	3.857	3.890
Saudi Arabia	4,243	307.727	90.522	2.499	2.582
Scotland	4,070	480.900	67.774	3.242	3.256
Serbia	4,045	500.096	84.365	2.950	3.232
Singapore	4,599	581.458	81.192	3.281	3.407
Slovenia	4,043	499.738	68.504	2.014	2.197
Sweden	5,215	478.244	54.384	1.676	2.019
Syrian Arab Republic	4,650	393.144	88.910	3.559	4.172
Thailand	5,412	436.002	86.827	4.786	4.823
Tunisia	4,080	420.619	66.436	2.409	2.594
Turkey	4,498	439.242	108.735	4.512	4.842
Ukraine	4,424	471.266	89.895	3.401	3.493
United States	7,377	513.980	68.426	2.515	2.562
Benchmarking Participants					
Basque Country, Spain	2,296	501.394	59.738	2.724	2.946
British Columbia, Canada	4,256	504.409	62.500	2.704	2.921
Dubai, UAE	3,195	469.000	90.988	2.037	2.310
Massachusetts, US	1,897	546.382	75.201	4.356	4.493
Minnesota, US	1,777	532.183	61.751	4.310	4.610
Ontario, Canada	3,448	505.146	59.571	2.937	3.219
Quebec, Canada	3,956	519.509	61.070	2.626	2.710

Exhibit E.18 Summary Statistics and Standard Errors for Proficiency in Mathematics Applying in the Eighth Grade

Country	Sample Size	Mathematics Applying			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	411.850	61.428	1.508	1.987
Armenia	4,689	492.722	90.860	3.752	3.831
Australia	4,069	499.940	78.553	3.355	3.442
Bahrain	4,230	402.648	78.090	1.565	1.918
Bosnia and Herzegovina	4,220	440.313	81.577	2.599	2.610
Botswana	4,208	351.095	82.065	2.184	2.632
Bulgaria	4,019	457.872	103.823	4.622	4.766
Chinese Taipei	4,046	592.168	101.796	4.061	4.177
Colombia	4,873	383.943	80.751	3.255	3.737
Cyprus	4,399	465.074	92.791	1.772	1.811
Czech Republic	4,845	504.302	75.591	2.610	2.655
Egypt	6,582	393.278	101.745	3.316	3.581
El Salvador	4,063	346.656	72.373	2.519	3.283
England	4,025	514.286	83.674	4.764	4.900
Georgia	4,178	401.063	102.317	5.317	5.516
Ghana	5,294	296.698	95.290	4.047	4.177
Hong Kong SAR	3,470	568.634	92.272	5.696	5.907
Hungary	4,111	513.358	84.032	3.080	3.113
Indonesia	4,203	398.328	88.119	3.626	3.691
Iran, Islamic Rep. of	3,981	401.572	88.816	3.945	4.167
Israel	3,294	455.903	102.738	3.948	4.051
Italy	4,408	482.989	74.280	2.768	2.852
Japan	4,312	565.043	82.926	2.103	2.224
Jordan	5,251	422.237	101.562	3.979	4.118
Korea, Rep. of	4,240	595.252	93.479	2.596	2.759
Kuwait	4,091	361.052	80.175	2.350	2.650
Lebanon	3,786	448.024	74.532	3.930	4.557
Lithuania	3,991	511.386	77.605	2.302	2.397
Malaysia	4,466	478.063	79.645	4.917	4.949
Malta	4,670	492.207	92.596	0.870	1.032
Morocco	3,060	389.406	79.269	3.201	3.262
Norway	4,627	477.205	68.119	1.914	2.182
Oman	4,752	367.525	96.152	2.921	3.045
Palestinian Nat'l Auth.	4,378	370.805	97.757	3.243	3.445
Qatar	7,184	304.612	96.383	1.084	1.448
Romania	4,198	462.047	96.758	3.911	3.961
Russian Federation	4,472	509.614	81.491	3.705	3.746
Saudi Arabia	4,243	335.245	81.751	2.167	2.303
Scotland	4,070	488.989	79.925	3.658	3.710
Serbia	4,045	478.465	93.081	3.292	3.347
Singapore	4,599	593.030	90.292	3.629	3.642
Slovenia	4,043	502.984	70.096	1.884	1.992
Sweden	5,215	497.100	71.588	1.962	2.015
Syrian Arab Republic	4,650	400.914	82.532	3.175	3.396
Thailand	5,412	446.403	88.081	4.626	4.665
Tunisia	4,080	423.358	70.815	2.410	2.430
Turkey	4,498	424.923	107.101	4.494	4.524
Ukraine	4,424	463.747	88.024	3.478	3.541
United States	7,377	502.647	79.417	2.796	2.864
Benchmarking Participants					
Basque Country, Spain	2,296	494.655	70.643	2.889	3.016
British Columbia, Canada	4,256	509.211	72.667	2.996	3.054
Dubai, UAE	3,195	455.519	94.610	2.477	2.946
Massachusetts, US	1,897	542.410	81.408	4.294	4.350
Minnesota, US	1,777	529.510	72.998	4.615	4.782
Ontario, Canada	3,448	518.254	71.353	3.467	3.662
Quebec, Canada	3,956	529.014	67.904	3.065	3.132

Exhibit E.19 Summary Statistics and Standard Errors for Proficiency in Mathematics Reasoning in the Eighth Grade

Country	Sample Size	Mathematics Reasoning			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	+	+	+	+	+
Armenia	4,689	489.443	101.065	3.420	3.846
Australia	4,069	501.731	79.206	3.222	3.339
Bahrain	4,230	413.316	87.897	1.731	2.054
Bosnia and Herzegovina	4,220	452.214	85.979	2.714	2.907
Botswana	+	+	+	+	+
Bulgaria	4,019	455.005	108.868	4.395	4.662
Chinese Taipei	4,046	591.419	108.883	4.030	4.103
Colombia	4,873	415.715	81.491	3.021	3.296
Cyprus	4,399	460.853	97.840	1.882	2.064
Czech Republic	4,845	499.812	77.289	2.479	2.557
Egypt	6,582	396.500	93.357	3.025	3.379
El Salvador	+	+	+	+	+
England	4,025	517.611	83.229	4.078	4.282
Georgia	4,178	389.302	109.117	5.466	5.770
Ghana	+	+	+	+	+
Hong Kong SAR	3,470	556.982	99.812	5.523	5.565
Hungary	4,111	512.643	88.453	3.222	3.246
Indonesia	4,203	405.061	89.524	3.285	3.332
Iran, Islamic Rep. of	3,981	426.610	82.926	3.449	3.524
Israel	3,294	462.453	99.520	3.717	4.135
Italy	4,408	483.461	80.110	2.730	2.797
Japan	4,312	567.803	92.726	2.357	2.422
Jordan	5,251	440.361	98.357	3.366	3.569
Korea, Rep. of	4,240	579.023	92.697	2.246	2.279
Kuwait	+	+	+	+	+
Lebanon	3,786	429.411	91.268	3.839	3.998
Lithuania	3,991	485.758	84.651	2.297	2.517
Malaysia	4,466	467.818	70.375	3.738	3.762
Malta	4,670	474.746	88.159	0.903	1.300
Morocco	3,060	383.314	90.609	3.327	3.536
Norway	4,627	475.382	74.860	2.094	2.251
Oman	4,752	397.110	94.743	3.095	3.331
Palestinian Nat'l Auth.	4,378	381.309	101.337	3.156	3.502
Qatar	+	+	+	+	+
Romania	4,198	448.586	114.037	4.423	4.551
Russian Federation	4,472	496.771	91.278	3.621	3.646
Saudi Arabia	+	+	+	+	+
Scotland	4,070	495.310	81.146	3.299	3.329
Serbia	4,045	473.780	94.677	3.275	3.290
Singapore	4,599	578.684	102.821	3.949	4.110
Slovenia	4,043	495.811	79.172	2.361	2.526
Sweden	5,215	490.118	84.140	2.327	2.553
Syrian Arab Republic	4,650	395.736	91.083	3.002	3.366
Thailand	5,412	456.242	87.175	4.243	4.359
Tunisia	4,080	425.476	66.299	1.813	2.315
Turkey	4,498	440.722	107.095	4.046	4.202
Ukraine	4,424	444.739	96.352	3.374	3.751
United States	7,377	504.727	72.137	2.305	2.376
Benchmarking Participants					
Basque Country, Spain	2,296	496.030	79.086	2.834	3.523
British Columbia, Canada	4,256	510.224	73.148	2.842	3.254
Dubai, UAE	3,195	464.645	98.991	2.508	2.768
Massachusetts, US	1,897	543.099	76.272	3.992	4.076
Minnesota, US	1,777	523.210	65.210	3.851	4.190
Ontario, Canada	3,448	521.371	68.086	3.005	3.182
Quebec, Canada	3,956	523.897	72.623	2.905	2.955

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Exhibit E.20 Summary Statistics and Standard Errors for Proficiency in Biology in the Eighth Grade

Country	Sample Size	Biology			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	411.430	72.948	1.567	1.862
Armenia	4,689	490.431	100.798	5.784	5.861
Australia	4,069	518.317	82.614	3.321	3.445
Bahrain	4,230	473.409	87.819	1.428	1.959
Bosnia and Herzegovina	4,220	464.044	80.476	2.821	3.036
Botswana	4,208	358.618	104.987	2.657	2.916
Bulgaria	3,079	466.611	106.516	5.863	6.039
Chinese Taipei	4,046	548.647	85.270	3.187	3.360
Colombia	4,873	433.958	79.182	3.581	3.679
Cyprus	4,399	446.668	85.590	1.623	1.919
Czech Republic	4,845	530.870	71.570	1.776	2.144
Egypt	6,582	406.477	96.651	3.248	3.409
El Salvador	4,063	398.456	71.553	2.621	2.985
England	4,025	540.886	81.899	4.118	4.366
Georgia	4,178	422.847	85.200	3.587	3.901
Ghana	5,294	303.775	113.587	4.832	4.950
Hong Kong SAR	3,470	527.126	75.317	4.464	4.586
Hungary	4,111	533.751	73.743	2.588	2.679
Indonesia	4,203	428.014	80.487	2.983	3.065
Iran, Islamic Rep. of	3,981	448.702	80.731	3.422	3.613
Israel	3,294	472.183	101.219	4.093	4.180
Italy	4,408	502.223	78.248	2.747	2.989
Japan	4,312	552.537	72.191	1.487	1.856
Jordan	5,251	478.266	90.727	3.659	3.798
Korea, Rep. of	4,240	547.726	68.962	1.653	1.900
Kuwait	4,091	419.483	87.726	2.360	2.556
Lebanon	3,786	405.258	102.444	5.912	6.159
Lithuania	3,991	526.537	81.353	2.280	2.336
Malaysia	4,466	469.150	86.476	5.726	5.772
Malta	4,670	452.838	115.155	1.137	1.654
Morocco	3,060	394.732	87.445	3.060	3.478
Norway	4,627	486.819	73.027	2.089	2.304
Oman	4,752	413.596	100.727	2.881	3.092
Palestinian Nat'l Auth.	4,378	401.647	110.088	3.560	4.064
Qatar	7,184	318.023	128.337	1.177	1.679
Romania	4,198	459.087	88.963	2.976	3.213
Russian Federation	4,472	524.926	75.652	3.421	3.627
Saudi Arabia	4,243	407.371	84.417	1.990	2.360
Scotland	4,070	495.077	80.248	3.111	3.183
Serbia	4,045	473.684	84.149	3.059	3.187
Singapore	4,599	563.815	98.546	4.043	4.166
Slovenia	4,043	529.764	73.156	1.983	2.322
Sweden	5,215	514.736	77.031	2.260	2.444
Syrian Arab Republic	4,650	459.486	79.105	2.535	2.699
Thailand	5,412	478.423	85.576	4.301	4.481
Tunisia	4,080	451.728	65.236	2.083	2.205
Turkey	4,498	461.963	87.950	3.319	3.429
Ukraine	4,424	476.718	83.970	3.179	3.444
United States	7,377	529.868	81.680	2.678	2.819
Benchmarking Participants					
Basque Country, Spain	2,296	497.946	69.446	2.605	2.885
British Columbia, Canada	4,256	534.656	74.834	2.962	3.225
Dubai, UAE	3,195	484.856	95.941	3.138	3.381
Massachusetts, US	1,897	562.593	80.184	3.751	4.295
Minnesota, US	1,777	554.671	78.064	4.705	5.191
Ontario, Canada	3,448	537.318	72.060	3.547	3.780
Quebec, Canada	3,956	512.685	68.427	2.555	2.934

Exhibit E.21 Summary Statistics and Standard Errors for Proficiency in Chemistry in the Eighth Grade

Country	Sample Size	Chemistry			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	414.109	74.088	1.433	1.743
Armenia	4,689	478.354	116.937	6.071	6.317
Australia	4,069	504.714	75.396	3.404	3.558
Bahrain	4,230	468.175	86.380	1.384	2.364
Bosnia and Herzegovina	4,220	467.819	83.049	2.762	2.884
Botswana	4,208	370.987	92.341	2.050	2.434
Bulgaria	3,079	472.343	111.215	5.982	6.080
Chinese Taipei	4,046	573.171	103.757	4.191	4.212
Colombia	4,873	419.856	72.871	2.777	3.099
Cyprus	4,399	452.227	93.468	1.884	2.487
Czech Republic	4,845	535.355	73.565	1.831	2.738
Egypt	6,582	413.440	106.480	3.633	3.993
El Salvador	4,063	377.048	78.582	2.780	3.193
England	4,025	533.938	79.375	3.870	3.953
Georgia	4,178	417.718	100.547	4.239	4.563
Ghana	5,294	342.001	109.599	4.714	4.869
Hong Kong SAR	3,470	517.423	76.092	4.419	4.559
Hungary	4,111	536.414	82.423	3.259	3.481
Indonesia	4,203	420.762	74.336	3.049	3.434
Iran, Islamic Rep. of	3,981	462.683	81.535	3.415	3.487
Israel	3,294	467.486	101.090	4.053	4.589
Italy	4,408	480.747	76.367	2.631	2.860
Japan	4,312	551.399	69.341	1.723	1.901
Jordan	5,251	490.813	97.699	3.976	4.109
Korea, Rep. of	4,240	535.794	74.303	1.820	2.418
Kuwait	4,091	417.713	99.056	2.818	3.832
Lebanon	3,786	446.595	103.279	5.201	5.496
Lithuania	3,991	506.767	80.249	2.121	2.255
Malaysia	4,466	478.982	81.571	4.905	5.017
Malta	4,670	460.864	120.742	1.248	2.071
Morocco	3,060	415.509	88.269	2.547	2.953
Norway	4,627	482.767	64.208	1.818	2.203
Oman	4,752	416.252	102.440	3.075	3.648
Palestinian Nat'l Auth.	4,378	413.447	111.994	3.640	4.189
Qatar	7,184	322.241	130.065	1.216	1.755
Romania	4,198	463.286	93.158	3.225	4.031
Russian Federation	4,472	534.625	84.102	3.545	3.725
Saudi Arabia	4,243	390.013	83.524	2.238	2.486
Scotland	4,070	496.917	73.582	2.975	3.201
Serbia	4,045	466.928	88.046	2.832	3.689
Singapore	4,599	560.309	101.580	3.989	4.092
Slovenia	4,043	539.191	76.370	2.158	2.461
Sweden	5,215	499.101	79.179	2.280	2.414
Syrian Arab Republic	4,650	449.754	76.518	2.689	2.914
Thailand	5,412	461.885	78.893	3.936	4.057
Tunisia	4,080	458.467	58.563	1.962	2.451
Turkey	4,498	434.869	107.005	4.375	5.182
Ukraine	4,424	490.053	87.050	3.260	3.327
United States	7,377	510.377	76.256	2.583	2.678
Benchmarking Participants					
Basque Country, Spain	2,296	472.465	81.704	3.070	3.451
British Columbia, Canada	4,256	504.771	70.445	2.632	2.704
Dubai, UAE	3,195	493.025	93.213	2.980	3.521
Massachusetts, US	1,897	540.384	85.388	4.278	4.576
Minnesota, US	1,777	518.823	71.032	4.451	4.906
Ontario, Canada	3,448	504.585	66.301	2.880	3.433
Quebec, Canada	3,956	496.649	74.481	2.920	3.129

Exhibit E.22 Summary Statistics and Standard Errors for Proficiency in Physics in the Eighth Grade

Country	Sample Size	Physics			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	396.862	73.498	1.488	2.230
Armenia	4,689	502.885	95.673	5.507	5.646
Australia	4,069	507.678	77.241	3.481	4.161
Bahrain	4,230	465.722	77.714	1.267	1.501
Bosnia and Herzegovina	4,220	463.175	80.931	2.850	3.084
Botswana	4,208	350.664	106.845	2.609	3.175
Bulgaria	3,079	465.966	104.758	5.455	5.570
Chinese Taipei	4,046	554.298	86.404	3.554	3.715
Colombia	4,873	407.098	80.119	3.392	3.515
Cyprus	4,399	457.610	87.081	1.700	2.776
Czech Republic	4,845	536.993	72.164	2.024	2.102
Egypt	6,582	413.456	92.283	3.165	3.344
El Salvador	4,063	380.492	78.774	2.531	3.506
England	4,025	544.929	77.188	3.945	4.020
Georgia	4,178	415.747	97.494	5.414	5.769
Ghana	5,294	276.337	128.229	5.299	5.784
Hong Kong SAR	3,470	528.069	85.741	4.803	4.837
Hungary	4,111	540.895	81.198	2.757	3.244
Indonesia	4,203	431.945	79.078	2.744	3.110
Iran, Islamic Rep. of	3,981	470.424	78.103	3.271	3.626
Israel	3,294	471.573	91.306	4.272	4.591
Italy	4,408	489.273	71.622	2.656	3.104
Japan	4,312	558.393	80.956	1.843	1.949
Jordan	5,251	478.720	94.368	3.961	4.167
Korea, Rep. of	4,240	571.294	81.906	2.111	2.396
Kuwait	4,091	437.951	82.738	2.186	2.781
Lebanon	3,786	430.556	87.732	4.718	5.066
Lithuania	3,991	505.432	77.407	2.349	2.928
Malaysia	4,466	483.533	84.956	5.556	5.750
Malta	4,670	469.770	95.868	1.010	1.721
Morocco	3,060	404.737	91.998	2.779	3.065
Norway	4,627	475.433	75.369	2.167	3.036
Oman	4,752	443.145	90.056	2.525	2.876
Palestinian Nat'l Auth.	4,378	413.955	108.922	3.575	3.680
Qatar	7,184	346.654	119.773	1.142	2.121
Romania	4,198	458.146	90.599	3.399	3.420
Russian Federation	4,472	519.164	82.714	3.820	4.037
Saudi Arabia	4,243	408.168	75.743	2.086	2.338
Scotland	4,070	493.710	76.894	3.077	3.658
Serbia	4,045	467.411	82.308	2.761	2.958
Singapore	4,599	575.449	95.548	3.830	3.873
Slovenia	4,043	524.498	68.753	1.857	2.049
Sweden	5,215	506.429	78.249	2.390	2.709
Syrian Arab Republic	4,650	446.777	77.995	2.674	2.728
Thailand	5,412	457.650	81.668	4.177	4.205
Tunisia	4,080	431.845	67.053	2.177	2.452
Turkey	4,498	445.421	96.560	3.570	4.255
Ukraine	4,424	492.437	85.387	3.086	3.853
United States	7,377	502.535	78.145	2.667	2.715
Benchmarking Participants					
Basque Country, Spain	2,296	493.473	75.724	2.805	3.398
British Columbia, Canada	4,256	516.525	66.363	2.470	2.838
Dubai, UAE	3,195	489.464	86.993	2.682	3.377
Massachusetts, US	1,897	534.908	82.874	3.994	4.993
Minnesota, US	1,777	514.166	78.197	4.573	4.775
Ontario, Canada	3,448	520.368	71.353	3.611	4.141
Quebec, Canada	3,956	491.809	73.613	3.235	3.378

Exhibit E.23 Summary Statistics and Standard Errors for Proficiency in Earth Science in the Eighth Grade

Country	Sample Size	Earth Science			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	412.796	68.924	1.366	1.613
Armenia	4,689	474.650	107.329	5.598	5.793
Australia	4,069	519.312	81.101	3.454	3.788
Bahrain	4,230	464.857	89.542	1.520	2.391
Bosnia and Herzegovina	4,220	468.871	89.283	2.711	3.391
Botswana	4,208	360.519	113.144	2.817	4.002
Bulgaria	3,079	479.591	99.011	5.157	5.506
Chinese Taipei	4,046	545.213	75.453	2.675	2.887
Colombia	4,873	406.843	85.572	3.727	3.896
Cyprus	4,399	457.279	89.772	1.735	2.339
Czech Republic	4,845	533.953	71.631	1.886	1.954
Egypt	6,582	426.151	91.260	3.085	3.777
El Salvador	4,063	400.126	75.963	2.686	2.878
England	4,025	529.172	88.953	4.115	4.317
Georgia	4,178	424.998	89.518	4.018	4.085
Ghana	5,294	294.128	118.803	4.856	5.803
Hong Kong SAR	3,470	532.207	76.924	4.291	4.470
Hungary	4,111	531.353	79.850	2.827	2.858
Indonesia	4,203	441.672	82.663	2.780	3.257
Iran, Islamic Rep. of	3,981	475.855	80.522	3.327	3.700
Israel	3,294	462.441	94.488	3.743	4.131
Italy	4,408	502.708	84.245	2.932	3.102
Japan	4,312	532.502	71.331	1.827	2.533
Jordan	5,251	483.612	92.164	3.407	3.648
Korea, Rep. of	4,240	538.328	66.992	1.777	2.242
Kuwait	4,091	410.236	90.804	2.466	3.022
Lebanon	3,786	388.837	109.626	6.018	6.417
Lithuania	3,991	514.958	82.715	2.315	2.488
Malaysia	4,466	462.588	86.571	5.221	5.383
Malta	4,670	455.837	116.148	1.275	1.511
Morocco	3,060	397.438	95.754	3.192	3.785
Norway	4,627	502.211	77.527	2.104	2.508
Oman	4,752	438.881	82.056	2.462	2.526
Palestinian Nat'l Auth.	4,378	408.464	107.161	3.335	3.662
Qatar	7,184	312.066	129.557	1.111	1.910
Romania	4,198	470.621	90.471	3.166	3.303
Russian Federation	4,472	524.673	74.842	3.246	3.442
Saudi Arabia	4,243	423.283	77.436	2.017	2.336
Scotland	4,070	497.670	79.373	3.106	3.218
Serbia	4,045	466.122	96.218	3.051	3.806
Singapore	4,599	540.873	99.802	3.969	4.081
Slovenia	4,043	541.958	73.026	2.037	2.194
Sweden	5,215	510.048	78.251	2.247	3.037
Syrian Arab Republic	4,650	448.438	80.802	2.635	3.153
Thailand	5,412	488.455	72.414	3.615	3.814
Tunisia	4,080	447.474	59.713	1.635	1.757
Turkey	4,498	466.463	79.124	2.816	3.299
Ukraine	4,424	482.202	87.414	3.300	3.980
United States	7,377	524.590	86.044	2.865	3.144
Benchmarking Participants					
Basque Country, Spain	2,296	513.950	77.870	2.684	2.846
British Columbia, Canada	4,256	530.098	71.038	2.451	2.716
Dubai, UAE	3,195	490.057	97.635	3.022	3.238
Massachusetts, US	1,897	559.889	79.092	3.834	4.008
Minnesota, US	1,777	544.823	80.421	4.950	5.493
Ontario, Canada	3,448	529.797	75.905	3.846	4.254
Quebec, Canada	3,956	513.010	69.203	2.685	3.520

Exhibit E.24 Summary Statistics and Standard Errors for Proficiency in Science Knowing in the Eighth Grade

Country	Sample Size	Science Knowing			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	409.186	65.742	1.702	1.870
Armenia	4,689	493.085	107.973	6.344	6.368
Australia	4,069	500.713	78.735	3.059	3.141
Bahrain	4,230	468.628	87.495	1.673	2.125
Bosnia and Herzegovina	4,220	486.390	84.046	3.336	3.730
Botswana	4,208	360.946	105.022	2.535	2.919
Bulgaria	3,079	489.363	101.573	5.502	5.791
Chinese Taipei	4,046	565.153	82.307	3.268	3.476
Colombia	4,873	418.171	84.158	3.561	4.032
Cyprus	4,399	438.318	91.179	1.847	2.646
Czech Republic	4,845	532.635	67.775	1.826	2.145
Egypt	6,582	434.028	109.141	3.572	3.850
El Salvador	4,063	394.203	75.679	2.762	3.215
England	4,025	530.489	90.229	4.846	4.865
Georgia	4,178	440.029	94.796	4.866	5.063
Ghana	5,294	316.284	128.276	5.553	5.667
Hong Kong SAR	3,470	532.049	73.653	4.330	4.458
Hungary	4,111	524.221	78.650	2.795	3.013
Indonesia	4,203	425.760	80.244	3.367	3.644
Iran, Islamic Rep. of	3,981	468.327	83.879	3.421	3.866
Israel	3,294	456.255	98.014	4.409	4.952
Italy	4,408	494.016	84.589	3.113	3.325
Japan	4,312	534.228	69.843	1.760	2.187
Jordan	5,251	490.870	102.758	4.447	4.495
Korea, Rep. of	4,240	543.011	70.167	1.769	2.000
Kuwait	4,091	429.852	83.911	2.191	2.463
Lebanon	3,786	402.934	103.952	5.843	5.897
Lithuania	3,991	513.080	74.437	2.271	2.372
Malaysia	4,466	458.040	93.350	6.280	6.465
Malta	4,670	436.427	122.080	1.107	1.458
Morocco	3,060	395.606	91.346	2.894	3.087
Norway	4,627	485.588	69.188	1.816	1.998
Oman	4,752	427.930	95.531	3.188	3.451
Palestinian Nat'l Auth.	4,378	406.797	108.824	3.437	3.514
Qatar	7,184	325.087	137.764	1.129	1.685
Romania	4,198	450.799	94.568	3.748	4.175
Russian Federation	4,472	534.451	81.793	4.228	4.278
Saudi Arabia	4,243	417.283	73.282	1.883	2.080
Scotland	4,070	480.256	84.088	3.382	3.892
Serbia	4,045	484.833	85.338	2.775	2.791
Singapore	4,599	553.763	102.084	4.357	4.481
Slovenia	4,043	532.999	70.456	1.982	2.020
Sweden	5,215	504.812	74.021	2.246	2.311
Syrian Arab Republic	4,650	474.413	76.487	2.661	2.900
Thailand	5,412	472.633	82.115	4.324	4.389
Tunisia	4,080	440.594	59.411	1.833	1.977
Turkey	4,498	461.816	89.808	3.535	3.598
Ukraine	4,424	476.931	90.556	3.532	3.786
United States	7,377	512.086	81.806	2.681	2.935
Benchmarking Participants					
Basque Country, Spain	2,296	490.312	71.772	2.541	2.993
British Columbia, Canada	4,256	515.644	67.225	2.537	2.881
Dubai, UAE	3,195	494.502	93.191	2.601	3.292
Massachusetts, US	1,897	544.684	86.127	4.049	4.223
Minnesota, US	1,777	526.369	78.492	4.388	4.753
Ontario, Canada	3,448	509.976	66.808	3.161	3.288
Quebec, Canada	3,956	495.377	67.413	2.661	2.860

Exhibit E.25 Summary Statistics and Standard Errors for Proficiency in Science Applying in the Eighth Grade

Country	Sample Size	Science Applying			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	410.244	66.093	1.459	2.351
Armenia	4,689	502.170	94.350	5.297	5.391
Australia	4,069	510.155	74.564	3.054	3.235
Bahrain	4,230	467.918	85.645	1.452	2.065
Bosnia and Herzegovina	4,220	462.670	82.231	2.827	2.836
Botswana	4,208	357.728	104.364	2.692	3.179
Bulgaria	3,079	471.326	106.401	5.992	6.096
Chinese Taipei	4,046	560.383	85.883	3.212	3.366
Colombia	4,873	416.759	75.072	2.919	3.110
Cyprus	4,399	455.741	85.644	1.757	2.030
Czech Republic	4,845	539.189	71.061	1.776	1.923
Egypt	6,582	403.804	98.746	3.311	3.559
El Salvador	4,063	388.373	69.781	2.540	3.190
England	4,025	537.608	78.562	3.946	4.033
Georgia	4,178	421.593	87.586	4.375	4.493
Ghana	5,294	290.627	120.794	5.293	5.463
Hong Kong SAR	3,470	522.296	77.685	4.666	4.947
Hungary	4,111	549.037	76.925	2.818	2.999
Indonesia	4,203	424.720	74.416	2.866	3.138
Iran, Islamic Rep. of	3,981	454.306	82.885	3.620	3.804
Israel	3,294	471.616	93.752	3.795	4.154
Italy	4,408	498.472	73.543	2.698	2.940
Japan	4,312	554.963	75.182	1.789	1.991
Jordan	5,251	484.994	95.254	3.862	4.054
Korea, Rep. of	4,240	547.011	72.215	1.797	2.027
Kuwait	4,091	416.617	93.595	2.665	2.927
Lebanon	3,786	422.312	98.051	5.599	5.845
Lithuania	3,991	512.371	76.455	2.148	2.169
Malaysia	4,466	473.153	88.192	5.834	5.937
Malta	4,670	462.412	107.108	1.155	1.585
Morocco	3,060	399.627	86.350	2.629	3.261
Norway	4,627	486.105	72.455	2.173	2.281
Oman	4,752	422.694	98.620	2.858	3.222
Palestinian Nat'l Auth.	4,378	412.121	113.102	3.561	3.994
Qatar	7,184	321.629	130.663	1.132	1.460
Romania	4,198	470.162	88.553	3.355	3.453
Russian Federation	4,472	526.664	76.683	3.579	3.805
Saudi Arabia	4,243	403.439	84.964	2.207	2.684
Scotland	4,070	494.676	74.080	2.985	3.096
Serbia	4,045	469.422	86.803	3.091	3.557
Singapore	4,599	567.350	99.632	4.140	4.228
Slovenia	4,043	533.346	69.845	2.015	2.216
Sweden	5,215	508.802	77.676	2.478	2.707
Syrian Arab Republic	4,650	444.838	80.503	2.741	2.976
Thailand	5,412	472.417	79.783	3.943	4.102
Tunisia	4,080	444.816	61.411	1.831	2.321
Turkey	4,498	449.598	91.421	3.487	3.578
Ukraine	4,424	487.585	81.181	3.441	3.716
United States	7,377	515.971	76.561	2.595	2.698
Benchmarking Participants					
Basque Country, Spain	2,296	499.431	73.679	2.716	2.853
British Columbia, Canada	4,256	521.104	67.084	2.566	2.805
Dubai, UAE	3,195	488.786	95.068	2.899	3.099
Massachusetts, US	1,897	549.933	75.770	3.781	4.038
Minnesota, US	1,777	534.183	69.870	4.595	4.758
Ontario, Canada	3,448	522.379	68.196	3.512	3.610
Quebec, Canada	3,956	500.449	67.609	2.820	3.060

Exhibit E.26 Summary Statistics and Standard Errors for Proficiency in Science Reasoning in the Eighth Grade

Country	Sample Size	Science Reasoning			
		Mean Proficiency	Standard Deviation	Jackknife Sampling Error	Overall Standard Error
Algeria	5,447	413.540	68.997	1.409	1.928
Armenia	4,689	459.389	112.703	6.412	6.514
Australia	4,069	530.345	79.887	3.358	3.592
Bahrain	4,230	468.610	82.629	1.320	2.025
Bosnia and Herzegovina	4,220	451.571	80.650	2.733	3.141
Botswana	4,208	362.466	96.541	2.470	2.702
Bulgaria	3,079	447.991	109.910	5.902	6.130
Chinese Taipei	4,046	541.305	88.476	3.354	3.464
Colombia	4,873	427.727	73.632	2.665	2.720
Cyprus	4,399	459.684	87.212	2.020	2.298
Czech Republic	4,845	533.983	74.488	2.122	2.252
Egypt	6,582	395.435	99.841	3.140	3.363
El Salvador	4,063	383.509	76.626	2.883	3.439
England	4,025	546.697	79.071	3.886	4.044
Georgia	4,178	394.435	94.628	4.081	4.629
Ghana	+	+	+	+	+
Hong Kong SAR	3,470	533.291	81.433	4.937	5.030
Hungary	4,111	530.089	78.112	2.811	3.020
Indonesia	4,203	438.362	81.145	3.049	3.168
Iran, Islamic Rep. of	3,981	461.781	74.920	3.356	3.781
Israel	3,294	481.011	100.133	4.058	4.199
Italy	4,408	492.898	71.356	2.425	2.559
Japan	4,312	559.853	75.781	1.693	2.007
Jordan	5,251	470.960	90.083	3.727	4.074
Korea, Rep. of	4,240	558.313	72.099	1.795	2.020
Kuwait	4,091	410.657	96.748	2.884	2.946
Lebanon	3,786	420.328	97.329	5.450	5.631
Lithuania	3,991	526.859	83.130	2.413	2.502
Malaysia	4,466	487.034	73.308	4.707	4.907
Malta	4,670	473.415	105.021	1.150	1.387
Morocco	3,060	412.750	86.006	2.171	3.008
Norway	4,627	491.329	72.443	2.089	2.789
Oman	4,752	427.730	89.050	2.577	3.502
Palestinian Nat'l Auth.	4,378	395.555	108.150	3.373	3.772
Qatar	+	+	+	+	+
Romania	4,198	459.503	86.852	3.453	3.500
Russian Federation	4,472	520.344	73.089	3.489	3.661
Saudi Arabia	4,243	395.392	83.935	2.101	2.485
Scotland	4,070	510.920	78.275	3.403	3.573
Serbia	4,045	454.779	93.555	3.328	3.546
Singapore	4,599	564.035	92.978	3.890	4.073
Slovenia	4,043	537.703	74.532	2.062	2.197
Sweden	5,215	516.909	75.519	2.275	2.593
Syrian Arab Republic	4,650	439.948	74.266	2.414	2.674
Thailand	5,412	473.006	78.704	3.768	3.963
Tunisia	4,080	458.137	65.298	2.121	2.918
Turkey	4,498	462.431	91.416	3.212	3.387
Ukraine	4,424	487.823	83.155	3.428	3.947
United States	7,377	528.955	76.327	2.712	2.855
Benchmarking Participants					
Basque Country, Spain	2,296	498.520	74.384	3.140	3.331
British Columbia, Canada	4,256	534.525	73.648	2.857	2.965
Dubai, UAE	3,195	482.678	88.845	2.692	3.308
Massachusetts, US	1,897	563.638	74.608	3.503	3.959
Minnesota, US	1,777	545.250	71.875	5.058	5.321
Ontario, Canada	3,448	542.385	74.527	3.944	3.990
Quebec, Canada	3,956	523.443	66.544	2.954	3.119

Note: A plus sign (+) indicates average achievement could not be accurately estimated.

Appendix F

Item Descriptions Developed During the TIMSS 2007 Benchmarking

Fourth Grade – Mathematics

Items at Low International Benchmark (400)

Number

M08_01	Subtracts a three-digit number from another three-digit number.
M09_01	Adds a four-digit and a three-digit whole number.
M10_01	Identifies the number sentence that models a word problem involving subtraction.
M11_04	Finds the missing number in a number sentence involving multiplication.
M13_05	Solves a word problem involving addition of three-digit whole numbers.
M14_01	Identifies a four-digit number given in words.

Geometric Shapes and Measures

M01_09	Identifies two triangles with the same size and shape in a complex figure.
M07_09	Recognizes the inverse relationship between the size of a unit shown in a figure and the number of units required to cover an area.
M08_07	Identifies a pair of parallel lines.
M10_07A	Given the position, gives the informal coordinates of the position.
M10_07B	Given the informal coordinates, determines the position.

Data Display

M08_12	Completes a table from given information by counting.
M09_12	Completes a bar graph that represents a table of data.
M14_12	Identifies the largest increase shown in a bar graph.

Items at Intermediate International Benchmark (475)

Number

M01_08	Solves a measurement word problem involving subtraction of two-digit numbers.
M02_01	Identifies a set of two-digit numbers ordered from largest to smallest.
M02_05	Subtracts a number with one decimal place from another with one decimal place.
M04_01	Identifies a three-digit number described in units, 10s, and 100s.
M05_01	Identifies the appropriate operation to solve a word problem involving multiplication.
M05_04A	Extends entries in two tables according to numerical rules described in a situation.
M06_01	Identifies the value of a digit in a four-digit number.
M09_05	Selects the expression that represents a situation involving addition.
M10_06A	Extends a given geometrical pattern to determine a specified term.
M11_03	Generalizes from the first several terms of a numeric sequence to select another number that is also in the sequence.
M11_06	Extends a numeric sequence based on a geometric pattern.
M12_01	Identifies the rectangular model for a unit fraction.
M13_01A	Selects appropriate information and uses it to solve a simple proportion problem.
M14_02	Solves a word problem involving multiplication of one-digit numbers.
M14_03	Identifies multiples of a given number.

Geometric Shapes and Measures

M04_06	Identifies an object with its line of symmetry shown.
M04_07	Draws a rectangle given two adjacent sides.
M06_08	Writes the names of three familiar geometrical shapes.
M07_06	Recognizes that area does not change when the parts of a figure are rearranged.
M07_10	Recognizes the triangles in a set of polygons.
M09_10	Orders four angles by size.
M09_11	Identifies a pattern generated by quarter turns clockwise.
M10_08	Draws the line of symmetry on a symmetrical polygon.
M11_10	Locates a point on an informal coordinate grid and identifies the moves to get there.
M13_06	Identifies a three-dimensional object given the pictorial representation of its faces.
M14_06A	Identifies the shape made by connecting specified dots on a circle.
M14_06B	Draws a specified geometrical shape by connecting dots on a circle.
M14_06C	Draws a specified geometrical shape by connecting dots on a circle.

Data Display

M01_11	Identifies the pie chart that matches the information shown in a table.
M03_10	Completes a two-by-two table to summarize information.
M04_11	Completes a bar graph from information given in a table.
M05_09	Uses information to identify the number of symbols needed to complete a pictograph when the symbol represents more than one.
M06_13	Identifies the bar graph that shows a given piece of information.
M06_14	Identifies the bar chart that matches the information shown in a pie chart.
M07_12	Identifies the pie chart that matches a given bar graph.
M10_11	Identifies information from a pie chart.
M10_12	Identifies the bar chart that matches the information shown in a table.
M12_11	Interprets a bar chart to solve a problem.
M14_11	Interprets information in a table to solve a problem.

Items at High International Benchmark (550)

Number

M01_01	Solves a word problem involving division of a three-digit number by a one-digit number.
M01_02	Determines the missing digit to give a specified difference in a three-digit subtraction problem.
M01_05	Solves a multistep word problem involving time and temperature.
M01_06	Solves a multistep word problem involving duration of time.
M01_07	Solves a word problem involving conversion of metric units of capacity.
M02_02	Identifies the operation needed to solve a problem involving division.
M02_03	Multiplies 2 two-digit numbers.
M03_06	Identifies a number that satisfies a number sentence involving division.
M04_04	Solves a word problem involving addition of two fractions with the same denominator.
M04_05	Identifies the operation needed to solve a problem involving division.
M05_02	Solves a word problem involving division of a three-digit number by a one-digit number.
M05_07	Solves a multistep word problem involving addition and multiplication of whole numbers.
M06_03	Shades half of a geometrical figure divided into four equal parts.
M06_04	Given five different digits, determines the smallest possible three-digit number.
M06_05	Writes a number between two consecutive whole numbers.
M07_01	Identifies the difference between two fractions with the same denominator.
M07_03	Selects the two-place decimal closest to a given whole number.
M07_04	Identifies the next term in a sequence of whole numbers formed by doubling.

M07_05	Identifies a number sentence that represents a situation involving division.
M07_07	Identifies the value of an unlabelled mark on a circular scale.
M08_02	Identifies the whole number closest to a given multiple of a hundred.
M09_06	Identifies the two-step rule for a linear relationship between the first and second numbers in a set of ordered pairs of numbers.
M09_07	Identifies the value of an unlabelled mark on a circular scale.
M09_09	Identifies the appropriate operation to solve a word problem involving division.
M10_02	Identifies the number that is a hundred more than a given four-digit number.
M10_04	Identifies appropriately rounded numbers in a multiplication situation.
M10_05	Identifies equivalent familiar fractions in a context.
M10_06B	Extends a given geometrical pattern to find the value of a specified term.
M11_02	Solves a multistep word problem involving halving, doubling, and adding.
M11_08	Solves a word problem involving addition of time and conversion between hours and minutes.
M12_02	Uses knowledge of place value to solve a problem involving a five-digit number.
M12_04	Writes a fraction that represents a subset of a set of objects.
M12_05	Identifies the largest of a set of unit fractions.
M13_08	Solves a word problem involving measures and proportional reasoning.
M14_04	Adds 2 two-place decimals.
M14_05	Follows a rule to complete a table.

Geometric Shapes and Measures

M02_07	Determines the perimeter of a rectangle given its dimensions.
M02_08A	Uses two specified geometric tiles to make a four-sided figure.
M02_08B	Uses two specified geometric tiles to make a six-sided figure.
M02_08C	Uses two specified geometric tiles to make a different six-sided figure from one made previously.
M03_07	Determines the number of nonstandard units of area needed to cover a figure.
M03_08B	On a map drawn to scale, positions a building within a range of distance from a specified point.
M03_09	Given a figure and the line of symmetry on a grid, draws the reflection.
M04_09A	States a property that two shapes have in common.
M04_09B	States a property that two shapes do not have in common.
M06_07	Given a set of angles, identifies the right angle.
M06_09	Determines the number of cubes in a stack with some hidden.
M06_10	Given the line of reflection, draws the reflection of a given figure.
M06_11	Identifies the distance around a square given the length of one side.
M08_10	Identifies a net of a cube.
M08_11	Identifies the area of a right triangle drawn on a grid.
M11_09	Draws an angle greater than 90°.
M11_11	Identifies the figure in which a line of symmetry is shown.
M12_07	Identifies a pair of shapes which are not mirror images of each other.
M14_07	Identifies the number of edges of a solid shown in a picture.
M14_08	Determines the perimeter of a figure made of squares.

Data Display

M02_10	Completes the scale so that a bar graph shows information given in a table.
M02_11	Completes a bar graph to show a specified comparison.
M04_12	Reads a part symbol on a pictograph when the symbol represents more than one.
M05_04B	Reads and interprets data from two tables to answer a question.
M05_04C	Draws conclusions from data in two tables.
M08_13	Completes a bar graph from information given in a tally chart.
M11_12	Interprets data from a bar graph to solve a problem.
M12_12	Recognizes the bar graph labeled appropriately to show given information.
M12_13A	Finds totals and decides which one is the least.

Items at Advanced International Benchmark (625)

Number

M01_03	Selects the appropriate information and uses it to solve a multistep word problem involving whole numbers.
M01_04	Writes two-step rule for a linear relationship between pairs of numbers.
M02_04	Identifies the fraction that is equivalent to the shaded fraction of a rectangle.
M02_06	Solves a two-step word problem involving two-place decimals.
M03_01	Solves a multistep word problem involving divisibility.
M03_02	Solves a problem involving proportional reasoning.
M03_03	Solves a multistep measurement problem involving multiplication and subtraction.
M03_04	Writes a rule for a multiplicative relationship between the first and second numbers in a set of ordered pairs of numbers.
M03_05	Identifies the two-step rule used to describe the relationship between adjacent terms in a sequence of numbers.
M04_02	Given a unit fraction, shows that fraction of a given set of objects.

M04_03	Identifies a fraction equal to a given fraction.
M06_02	Solves a word problem involving division and rounding up the remainder.
M06_06	Adds two familiar unit fractions to solve a word problem.
M07_08	Solves a multistep problem involving conversion between hours and minutes.
M08_03	Identifies the smallest number from a set of one- and two-place decimals.
M08_04A	Identifies the circular representation of a nonunit fraction.
M08_04B	Explains why a chosen circular representation shows a given nonunit fraction.
M08_05	Identifies the missing first number in a number sentence involving subtraction.
M08_06	Identifies the two-step rule that relates the numbers in two columns of a table.
M09_02	Identifies all the numbers in a given interval ending in a given string of digits.
M09_03	Halves the amounts in a recipe involving whole numbers and fractions.
M10_03	Finds all the factors of a multifactor number less than 20.
M11_01	Given a unit fraction, identifies a larger fraction with a different denominator.
M11_05	Identifies the number that satisfies a number sentence involving addition of two terms on each side.
M11_06C	Generalizes from the first several terms of a numeric sequence to find the tenth term.
M12_03	Estimates the quotient of a four-digit number divided by a two-digit number.
M12_06	Solves a word problem involving proportional reasoning.
M13_01B	Selects appropriate information and uses it to solve a proportion problem.
M13_01C	Selects appropriate information and uses it to solve a multistep problem involving proportions.
M13_02	Selects appropriate information and uses it to solve a proportion problem.

Geometric Shapes and Measures

M02_08D	Uses three specified geometric tiles to make a seven-sided figure.
M02_09	Identifies a shape rotated by a 90° turn.
M03_08A	On a map drawn to scale, positions a park at a given distance from a specified point.
M03_08C	On a map drawn to scale, positions a building halfway between two specified points.
M04_08	Calculates the area of a rectangle.
M05_06	Recognizes that the area does not change when a figure is cut into parts and rearranged.
M05_08	Uses properties of rectangles and triangles to solve a problem.
M07_11	Recognizes the net of a triangular prism.
M08_08	Uses knowledge about properties of rectangles to classify statements as true or false.
M08_09	Solves a multistep word problem involving perimeter.
M09_08	Determines the area of a figure made up of squares and half squares on a grid.
M10_09	Uses knowledge of two common solids to classify statements about them as true or false.
M10_10	Matches a solid to its net.
M12_08	Determines the number of cubes in a given rectangular box.
M12_09	Identifies the area of an isosceles triangle drawn on a grid.
M12_10	Draws a line through a given point perpendicular to a given line.
M13_07	Identifies the position of a shape after a half-turn rotation.
M14_09	Estimates the length of a curved line in nonstandard units.

Data Display

M01_10	Organizes data and completes a tally chart to represent it.
M04_13	Uses data from two different graph types to solve a problem.
M12_13B	Draws and justifies a conclusion from data given in a table.

Items Above the Advanced International Benchmark (625)

Number

M05_03	Identifies the number that satisfies a number sentence involving division of two terms on each side.
M05_05	Solves a multistep problem to find one of the two unknown values.
M07_02	Subtracts a one-place decimal from a two-place decimal presented horizontally.
M13_03	Selects the appropriate information and uses it to solve a multistep problem involving two proportions.

Geometric Shapes and Measures

M04_10	Estimates a height using a nonstandard unit.
M06_12	Classifies polygons according to two given properties they either have or do not have.
M11_07	Estimates the length of a curved line next to the middle of a ruler.
M14_10	Draws all four lines of symmetry in a nonstandard shape.

Fourth Grade – Science

Items at Low International Benchmark (400)

Life Science

S13_03	Recognizes that wings are common to bird, bats, and butterflies.
S11_07	Recognizes that birds sit on their eggs to keep the eggs warm.
S13_01	Recognizes wolf as a predator.
S14_11	States one effect the sun can have on unprotected skin.
S07_04	Recognizes from diagrams of animals which animal is most likely to live in a desert.
S12_01	Completes a table by matching diagrams of animals to their ecosystems.
S02_05	Recognizes that the lung is the body organ most harmed by smoking.

Physical Science

S04_06	Recognizes that an iron object is most likely to be heavier than a wood or styrofoam object of the same shape and size.
S10_01	From a diagram, recognizes which thermometer reading shows the hottest water.
S14_09	Recognizes that the vibrations that produce sound in a guitar start with the strings.
S06_01	Identifies wind as the cause of movement in a sail boat.
S12_06	Identifies ice as the solid form of water.
S07_06	Recognizes that iron nails rust.
S14_07	From a diagram, identifies the direction of the force of Earth's gravity.

Earth Science

S03_09A	States the names of two seasons.

Items at Intermediate International Benchmark (475)

Life Science

S05_01	Recognizes that snakes shed their outer covering as they grow larger.
S04_05	In the context of an investigation of plant growth, describes a treatment that can cause one plant to grow better than another.
S12_02	Describes one way people can protect their teeth from decay, in addition to brushing.
S02_01	From a diagram, distinguishes non-living things from living things.
S08_01	Recognizes the stomach as an organ where digestion takes place.
S14_01	Recognizes that the function of seeds is to produce new plants.
S11_03	From pictures of animals, pairs each animal with its distinguishing biological characteristics (skeleton, milk production, number of legs).
S09_05	Recognizes that tadpoles hatch from frogs' eggs.
S01_05	Recognizes that a person's hair type can be predicted by his/her parents' hair type.

S05_03	Recognizes from diagram of birds which bird is most likely to eat mammals.
S09_02	Recognizes which foot structure belongs to a bird that lives in a pond.
S13_02	Recognizes that fat layers help keep a walrus warm.
S14_02	Recognizes that the body needs more oxygen during exercise.
S07_01	Recognizes that trees make their own food using sunlight.
S01_06	Interprets from a food chain that snakes eat voles.
S04_03	Recognizes that fruits and vegetables are the best source of vitamins and minerals.
S06_08	Describes how influenza can be passed from person to person.

Physical Science

S02_06	In the context of an investigation, recognizes that a floating body is lighter than bodies of the same shape and size that sink.
S03_07	From a diagram showing a person blowing into water using a straw, explains why bubbles rise to the top.
S09_08	From a list of common materials, indicates which of them will burn.
S01_02B	Given a diagram showing that the color of a white shirt appears to be different under different colored light bulbs, infers its color under blue light.
S11_08	Recognizes that an iron nail can complete an electrical circuit and allow a bulb to glow.
S10_07	Identifies electricity as the energy source for three household objects shown in a diagram.
S13_07	States two things that electricity can be used for in daily life.
S03_05A	States one way water in ice form is used by humans.
S03_05B	States one way water in liquid form is used by humans.

Earth Science

S05_07	Explains why people should not drink water directly from oceans and seas.
S07_11	States two different things human use wood for.
S01_08	Orders diagrams showing ribbons on poles by decreasing wind strength.
S03_09B	States one difference between two previously named seasons.
S05_08	States one difference between the sun and the moon.
S08_11	States two planets other than Earth that orbit the sun.

Items at High International Benchmark (550)

Life Science

S11_02	Recognizes that if the only remaining Siberian tigers are female, they will not be able to reproduce, and will die out.
S02_03	Complete a diagram showing the life cycle of a moth.
S04_02	Describes one way that seeds from a plant are dispersed.
S10_04	From a diagram of a food chain, identifies a predator-prey relationship.
S04_08	From information shown in a table, completes a food chain.
S06_06	Explains why traveling by bicycle is better for the environment than traveling by motorbike.
S14_06	From a list of human activities, identifies which have positive and which have negative effects on the environment.
S09_04	Recognizes that differences in light brightness cause eyes in one picture to look different from the eyes in a second picture.
S10_02	Recognizes that plants make food using energy from the sun.
S01_04	Recognizes that the teeth of monkeys are most like the teeth of humans.
S07_05	Recognizes from a picture types of seed that are scattered by wind.
S08_02	Recognizes from a diagram the part of a flowering plant that produces seeds.

S06_03	From a picture of a pond ecosystem, identifies three living and three nonliving things.
S05_04A	States one physical feature or behavior of fish that distinguishes them from sea mammals.
S03_03	Using knowledge of teeth, identifies and explains which of two skulls shows an animal that ate plants and an animal that ate meat.

Physical Science

S02_08	Identifies a method of separating a mixture of iron filings and sand.
S05_05	Recognizes that the hotter the water the more sugar will dissolve.
S08_09B	In the context of an investigation, states that candy dissolves faster in hot water than in cold.
S04_12	Completes a table by identifying examples of matter that exist as solid, liquid, or gas at room temperature.
S02_10	Explains why water disappears from a dish of water left in the sun.
S05_06A	Describes how a liquid can be turned into a gas.
S05_06B	Describes how a liquid can be turned into a solid.
S03_06	From a diagram showing a metal ruler heated at one end, recognizes the direction of heat transfer.
S11_05	Recognizes that metal conducts heat better than wood.
S06_10	From a table of properties of two materials, determine the identity of each.
S10_09	Given a diagram of three objects of different materials ordered by volume, justifies that objects with more volume do not necessarily weigh more.
S13_05	From a table showing the results of an experiment, identifies what was being studied in the experiment.
S08_09A	In the context of an investigation, explains that candy dissolves faster when it is crushed into small pieces.
S11_04	Recognizes that fine salt dissolves faster in water than coarse salt and explains why.
S13_08	Recognizes that heat needs to be supplied for melting and boiling but not for freezing.

S01_01C	From an investigation of the effect of different colored light on the apparent color of a shirt, infers the color of an unknown light bulb.
S10_11	From a diagram, recognizes the direction of a shadow.
S12_04	Recognizes what causes a shadow to be formed.
S01_02A	Describes the results of an investigation involving white shirt seen under different colored light bulbs.
S02_07	From a diagram of an electric circuit, states why an unbroken bulb does not light up.
S09_10	From a diagram showing two magnets on carts with the magnet poles marked, describes what happens to the carts when they are moved close together and let go.
S12_11	Completes the labeling of the poles on magnets shown in a diagram.
S04_07	Recognizes an example of an object moving because of the force of gravity.
S07_07	From a diagram showing three powders, recognizes those likely to be mixtures.
S09_07	Recognizes that salt water is a mixture.
S02_09	Identifies an object that runs only on electricity.
S07_08	Given a set of diagrams, recognizes that ice melts most slowly in a closed container.
S11_09	Recognizes that gravity causes an object to fall to the ground.

Earth Science

S02_12	Recognizes that most of Earth's surface is covered by water.
S10_13	Identifies that water that has had its salt removed so that it can be used as drinking water is most likely to come from the sea.
S07_09	Explains that early morning moisture can be due to condensation.
S06_13A	Describes one advantage of farming near a river.
S12_14	From a table showing temperature and cloud cover, identifies the place where it is most likely to snow.
S14_08	Recognizes that parts of animals that have hardened into rock are the best evidence that there were many kinds of animals on Earth that no longer exist today.

S04_10	Recognizes that a mountain side rock layer containing shellfish fossils was once part of a sea floor.
S10_10	States one form of energy Earth receives from the sun.
S13_10	Identifies the Earth, moon, and sun from a diagram.
S02_13A	From a table showing planetary distance from the sun, identifies the planet closest to the sun.
S02_13B	From a table showing planetary distance from the sun, identifies the planet most likely to have the lowest average surface temperature.

Items at Advanced International Benchmark (625)

Life Science

S05_04	States one physical feature or behavior of sea mammals that distinguishes them from fish.
S14_03	Recognizes examples of animals that take care of their young.
S07_02	Explains that the last surviving member of a species of a turtle cannot reproduce and gives a reason.
S06_09	Describes how migration increases the survival of birds.
S14_04	Recognizes an advantage to monarch butterflies of being poisonous to birds.
S10_03	States one thing a person can do to avoid catching flu from an infected person.
S12_09	Describes one physical change that can take place in a mammal as the weather gets cold.
S03_04	Recognizes that the energy needed to heal a cut comes from food.
S13_11	Describes two human activities that can lead to the extinction of animals.
S13_04	States one thing can cause the temperature of the human body to be higher than normal.
S02_02	Recognizes which animal has six legs.
S03_01	Recognizes a group of animals that are all mammals.
S04_01	From a diagram, recognizes an animal that has a skeleton on the outside of its body.
S08_06	Recognizes an animal that is classified as a mammal.
S08_03	Identifies the body covering that protects a reptile.

S10_06A	From a diagram of a tiger skull, identifies a function of the canines.
S11_01	Recognizes from a list of animals that humans have a young form that looks most like the adult form.
S09_01	Recognizes from a list of foods that cheese is the best source of calcium.
S01_07	Evaluates and supports argument for the need for a balanced diet.
S12_07	Explains why people should drink a lot of liquid every day.

Physical Science

S02_11	Recognizes that, regardless of their size, ice cubes float in water.
S04_11	Given a jar containing balls of the same volume but made of different metals, names one property that can be used to separate the balls into different groups.
S14_12	Names a source of energy other than coal, oil, or natural gas that is used to produce electricity.
S01_01A	Describes the results of an investigation involving a red shirt seen under different colored light bulbs.
S01_01B	From an investigation of the effect of different colored light on the apparent color of a shirt, concludes that the shirt looks different under different lights.
S11_06	Names one thing that shows that sunlight is made up of different colors.
S12_03	Using information in a table about physical properties of familiar items, identifies another item whose physical properties match those of one of the items in the table.
S13_06	Recognizes the diagram that best shows how ice floats in water.
S14_05	Labels the freezing point of water on a diagram of a thermometer.
S08_09C	In the context of an investigation, recognizes that more water in a solution makes a drink less sweet.
S14_13	Recognizes that burning results in new substances.
S08_10	From a list of familiar materials, recognizes the best conductor of heat.

S08_08	Given two electric circuits diagrams showing different battery configurations, explain which circuit will allow a bulb to light.
S06_02	Distinguishes objects that produce their own light from those that do not.
S10_08	From diagrams providing partial information about the weights of four cubes, draws a conclusion about the relative weight of one of the cubes.

Earth Science

S11_11	Recognizes a soil change due to natural causes.
S13_09	Recognizes that soil rich in decaying plants and animals makes plants grow.
S12_13	States two things that make up Earth's crust.
S04_09	Recognizes the pie chart that shows the proportions of land and water on Earth.
S04_14	Recognizes a common characteristic of different types of desert.
S16_13B	Describes one disadvantage of farming near a river.
S04_13	Provides an example of a natural resource, other than water, and describes its use.
S14_10	In the context of an investigation, explains why water does not fill a glass inverted in water but does fill it when the glass is tilted.
S10_12	Recognizes that the direction a river flows depends on the slope of the land.
S03_08	Recognizes that the moon is visible because it reflects the light from the sun.
S12_12	Recognizes how long it takes for Earth to orbit the sun.
S06_11	Recognizes how long it takes for Earth to rotate on its axis.
S08_12	From a diagram showing a shadow at different times of the day, explains why the shadow changed.

Items Above the Advanced International Benchmark (625)

Life Science

S04_04	States two characteristics that distinguish between living and nonliving things.
S12_05	States two characteristics that living things share, other than a need for water.
S06_05	Identifies a group of animals that contains only reptiles.
S05_02	States two reasons why humans need a skeleton.
S10_06B	From a diagram of a rat skull, identifies a function of the incisors.
S06_07	Identifies one function of fruit.
S06_04	From a diagram of a flowering plant, identifies numbered parts and states a function of each part.
S03_02	Predicts whether different types of plants can reproduce, and justifies the choice.
S07_03	Evaluates and explains the best experimental setup for investigating effect of salt on seaweed.
S02_04	Recognizes where plants get the energy to make food.
S08_04	Recognizes which living things make their own food.
S08_05	States one thing necessary to maintain good physical health and explains why.

Physical Science

S06_12	From a series of diagrams, identifies an unknown material as a gas based on its behavior in a closed container and justifies the answer.
S10_05	Recognizes a description of condensation.
S09_09	Determines changes in temperature when a hot object is put into cold water.
S01_03	Predicts and explains the color of a blue shirt under a blue light.
S12_08	Draws a conclusion about the relative weight of two objects made of different materials that both sink in water.

Earth Science

S11_10	Describes activities that require air.
S09_03	Describes two things people can do to avoid wasting water.
S07_10	Recognizes that fossils are evidence that land was once covered by the sea.
S09_11	Relates day and night on Earth to rotation on its axis.

Eighth Grade – Mathematics

Items at Low International Benchmark (400)

Number

M02_02	Multiplies a decimal by a power of ten.
M03_07	Multiplies a two-place decimal by a three-place decimal.
M03_11	Solves a word problem involving a proportion with unit ratio.
M04_01	Given a number in the millions in words recognizes the numeral.
M06_01	Given a three-place decimal recognizes the equivalent fraction.

Algebra

M10_05A	Finds the next term in a simple number pattern.

Data and Chance

M02_12	Given a table of values, selects the graph that could represent the given data.
M03_08	Given a table of values for two variables, selects the line graph that could represent the given data.
M14_13	Uses information in a table to complete a bar graph.

Items at Intermediate International Benchmark (475)

Number

M01_01	Identifies a circular model of a fraction that best approximates a given rectangular model of the same fraction.
M01_02	Solves a word problem by adding numbers with up to three decimal places.
M01_06	In a word problem selects the approximate quantity remaining after an amount is decreased by a given percent.
M02_01	Identifies a set of five-digit numbers ordered from largest to smallest.
M03_01	Reads the value indicated by an unlabeled tick mark on a circular scale.
M03_03	Selects the smallest fraction from a set of familiar fractions.
M03_12	Solves a word problem about distance and time by finding the missing term in a proportion.
M03_13	Identifies the integer that gives a specified result when divided by a given negative integer.
M04_05A	Completes a table by solving a simple word problem.
M08_01	Recognizes the power of 10 of the divisor in a division of decimals.
M10_01	In a word problem, given a unit fraction of a measure identifies the whole measure.
M12_01	Knows simple exponential notation.
M12_03	Uses knowledge of the whole being 100 percent to solve a simple word problem.
M12_04A	Completes a table of equivalent proportions.
M13_01	Rounds two-place decimals to whole numbers.
M14_01	Solves a word problem involving multiplication of a fraction and a decimal.
M14_04	Identifies equivalent ratios.

Algebra

M07_04	Selects the rule expressed in words that generates successive terms in a given number pattern.
M07_05	Solves a linear equation in one variable.

M11_04	Knows the meaning of a simple algebraic expression involving multiplication and addition.
M11_05	Identifies the algebraic expression that represents a situation, involving addition and multiplication.
M13_03	Extends number patterns derived from a sequence of geometric shapes.
M14_03	Recognizes the distributive property in evaluating an algebraic expression.
M14_07	Identifies the ordered pair of numbers that satisfies a linear equation.

Geometry

M02_11	Given its coordinates, determines that a point is in the second quadrant of the Cartesian plane.
M03_14	Determines the measure of the missing angle in a right triangle.
M04_11	Draws a triangle on a grid with twice the area of a given rectangle.
M05_05	Solves a word problem by comparing distances on a map drawn to scale with a given distance.
M07_10	Uses properties of an isosceles triangle to identify the coordinates of a point on a grid.
M09_11	Given a net of a three-dimensional object, completes a two-dimensional drawing of it from a specific viewpoint.
M10_10A	Given instructions, locates points on polar grid.

Data and Chance

M02_13	Reads a bar graph to identify quantities which satisfy a given condition.
M03_02	Recognizes that the probability of an outcome of a single event is inversely related to the number of elements in the population of events.
M07_13A	Identifies the straight line graph modeling a situation described in words.
M07_13B	Interprets two straight line graphs and uses their intersection to solve a problem.
M08_15	Given a table of percentages, selects the pie chart that could represent the given data.

M08_16	Interpolates from a line graph to provide an estimated value.
M10_11	Given a situation, judges the chance of an outcome as likely.
M11_13A	Selects the appropriate line on a graph and reads information from it.
M12_14	Given a situation, judges the chance of an outcome as unlikely.

Items at High International Benchmark (550)

Number

M01_09	Adds three fractions with different denominators which are less than 10.
M02_03	Uses knowledge of negative integers to produce the largest sum.
M02_05	Reduces an amount by a given percentage.
M04_02	Identifies the prime factorization of a number.
M04_05D	Combines the information from intermediate solutions to solve a problem involving time.
M05_01	Identifies equivalent ratios in a problem setting.
M06_02	Selects the numerator of a fraction to make two fractions equivalent when one denominator is not a multiple of the other.
M06_03	Continues a pattern of number sentences involving subtraction of negative integers.
M06_04	Given the part and the whole can express the part as a percentage and vice versa.
M07_01	Solves a word problem by determining a number between two given numbers that is divisible by only one of two other given numbers.
M08_02	Recognizes the fraction equivalent to a percentage.
M08_03	Approximates the sum of 5 three-digit numbers to the nearest 100.
M08_04	Identifies the larger of two fractions with different numerators and different denominators and explains why it is larger.
M08_05	Writes a rule for a multiplicative number pattern involving negative numbers.

M09_01	Identifies the decimal number that is equivalent to the sum of two fractions whose denominators are powers of ten.
M10_02	Identifies the decimal number represented by a point between two consecutive whole numbers on a number line with only the whole numbers labeled.
M10_03	Uses the law of exponents to express a product.
M12_02	Reads the value indicated by a minor unlabeled tick mark on a circular scale, when the previous major tick mark also is unlabeled.
M12_04B	Finds the unknown term in a proportion in a given situation.
M13_07	Identifies the prime factors of a given number.
M13_08	Uses percentages given in a pie chart to solve a problem.
M14_02	Uses knowledge of decimal place value to express a given sum as a decimal number.
M14_06A	Selects and combines information from two sources to solve a multistep word problem.
M14_06B	Selects and combines information from two sources to solve a multistep word problem.

Algebra

M02_06	Recognizes the simplification of an algebraic expression.
M02_07A	Continues a pattern involving the sum of interior angles of polygons based on triangles.
M04_04	Identifies the expression that represents a multiplicative situation.
M04_06	Solves a linear equation given in a word problem.
M05_02	Recognizes the product of two algebraic terms in one variable that involves exponents.
M05_10	Identifies the linear equation represented by a set of ordered pairs given in a table.
M06_05	Recognizes the collection of algebraic terms involving exponents.
M06_06	Evaluates an algebraic expression in two unknowns.
M06_08	Uses the value of a given algebraic expression to find the value of a related algebraic expression.
M07_06	Identifies an algebraic expression to model a situation.

M08_06	Solves a simple linear equation in one variable with a mixed number solution.
M08_07	Finds a missing term in a nonarithmetic and nongeometric number sequence.
M08_08	Identifes the linear equation satisfied by two given values.
M08_09	Solves a proportion expressed algebraically.
M08_11A	Adds two algebraic expressions and simplifies.
M09_05	Identifies the algebraic expression that represents a situation involving the sum of a constant term and a product.
M09_06	Uses a formula to determine the value of one variable given the value of the other.
M10_05B	Finds a specific term in a simple number pattern.
M10_06	Uses the distributive law to identify an algebraic expression equivalent to a given one.
M10_07	Determines the solution to a pair of simultaneous equations.
M11_01	Solves a word problem by using patterns in a two-column table to determine the number in the second column that would correspond to a number midway between two entries in the first column.
M11_12	Identifies the quantity that satisfies two inequalities represented by balances in a problem situation.
M12_05	Identifies the equation of a line that passes through points shown on a graph.
M12_07	Finds the value of an algebraic expression involving parentheses and negative terms.
M12_08A	Finds a specific term in a pattern presented numerically and geometrically.
M13_09	Given an interval containing a number, determines the interval containing the sum of that number and a whole number.
M14_05	Identifies the algebraic expression that represents a fraction of a variable.
M14_08	Identifies the equation that models a situation given in a word problem.
M14_09	Identifies values of two variables each satisfying a simple inequality.

Geometry

M01_03	Identifies a three-dimensional object after rotation.
M01_05	Finds the perimeter of a square, given its area is a square number.
M02_09	Identifies a net of a cube.
M03_06	Uses knowledge of a straight angle to find the measure of an angle.
M03_15	Uses properties of angles to draw and label a figure.
M04_09	Identifies how a three-dimensional object would look from a given viewpoint.
M05_04	Calculates the volume of a rectangular prism by using appropriate measure from its nets.
M05_09	Uses the properties of a triangle and regular hexagon to find the measure of an angle.
M06_11	Uses properties of triangles to draw a triangle of given dimensions on a grid.
M06_12	Given the volume and two dimensions of a rectangular solid, finds the other dimension.
M07_08	Calculates the area of an irregular figure formed by two rectangles.
M08_12	Identifies a true statement based on the properties of parallel and perpendicular lines.
M08_13	Uses the angle properties of triangles and rectangles to find a missing angle.
M09_09	Determines the number of cubes needed to fill a hole in a given shape.
M09_10	Identifies the justification that a triangle is a right triangle using the Pythagorean theorem.
M09_12	Identifies the transformations used to produce a sequence of figures.
M10_10B	Determines the measure of an angle drawn on a polar grid.
M11_08	Visualizes the unfolded shape of a figure shown on a folded piece of paper and uses properties of triangles to identify the shape.
M11_10	Applies properties of interior and exterior angles of a triangle to find an unknown angle in overlapping triangles.

M12_09	Draws a symmetrical shape given half of it and one of its lines of symmetry.
M12_11	Identifies two shapes that make a square.
M14_11	Given a cube made of unit cubes, uses the properties of a cube to identify the number of remaining unit cubes.

Data and Chance

M01_07	In a word problem, when given the possible number of outcomes and the probability of successful outcomes, solves for the number of successful outcomes.
M02_14	Uses the information in a pie chart showing percentages to draw a bar chart.
M04_12A	Calculates and compares the means of two sets of numbers given their totals.
M04_13	Given a word problem, determines the most likely outcome.
M05_07C	Draws conclusions from data in a table to meet given conditions.
M05_08	Compares and integrates several sets of data to determine which meet given conditions.
M06_14	Determines which of a set of statements involving averages must be true.
M06_15	Determines the probability of two of three possible outcomes.
M06_16	Uses data given as percentages to predict the outcome of a future event.
M06_17	Constructs and labels a pie chart representing a given situation.
M07_11	Uses experimental data and an understanding of probability to draw the spinner that could have produced the given data.
M07_13C	Reads values from two straight line graphs to solve a problem.
M09_13	Constructs and labels a pie chart representing a given situation.
M09_14	Identifies the statement that best describes the relative likelihood of two events.
M10_12A	Calculates the mean of a set of numbers.

M11_13C	Selects the appropriate line on a graph and determines the interval where the greatest change occurs.
M12_12	Reads values from two line graphs to solve a problem.
M12_13	Identifies a possible word representation for a part of a speed-time graph.
M14_14	Explains why a conclusion drawn from a given bar graph is incorrect.

Items at Advanced International Benchmark (625)

Number

M01_11B	Given the dimensions of two rectangles, expresses the ratio of their areas.
M02_04	Given the total number and the ratio of the two parts, identifies the value of one part.
M03_05	Given the total number and the ratio of the two parts, finds the value of one part.
M03_09	Selects appropriate data to solve a problem involving operations with fractions that have different denominators.
M03_10	Solves a word problem involving multiplication of a proper fraction and an improper fraction.
M04_05B	Given an average speed and distance, finds the duration and uses it to solve a problem.
M04_05C	Given an average speed and distance, finds the duration and uses it to solve a problem.
M07_02	Identifies a procedure for subtracting fractions with different denominators.
M07_03	Given the total number and the ratio of the two parts, identifies the value of one part.
M07_12	Given the original and reduced prices, finds the percentage of the reduction.
M09_02	Given two points on a number line representing unspecified fractions, identifies the point that represents their product.
M09_03	Solves a problem involving a fraction of a whole number of currency units.
M10_04	Arranges four given digits to obtain the greatest product of 2 two-digit numbers.
M11_02	Converts a mixed number to a decimal rounded to two places.

Algebra

M01_04	Solves a linear inequality involving a fraction.
M02_07B	Finds a specific term in a number pattern involving the sum of interior angles of polygons based on triangles.
M02_07C	Expresses the general term algebraically in a number pattern involving the sum of interior angles of polygons based on triangles.
M02_08	Solves a word problem that can be expressed as two linear equations with two variables.
M04_03	Evaluates an algebraic expression involving parentheses and negative terms.
M04_07	Simplifies an algebraic expression involving parentheses and negative terms.
M04_08	Given the equation of a straight line identifies a point on it.
M05_03	Extends a number pattern presented geometrically to solve a problem.
M06_09	Finds the missing term in a nonstandard number pattern.
M06_10	Identifies the linear equation that is satisfied by two ordered pairs.
M08_10A	Writes an equation to model a situation involving perimeter.
M08_10B	Solves a linear equation.
M08_11B	Subtracts one algebraic expression from another and simplifies.
M09_04	Identifies a diagram that models addition of two like algebraic terms.
M10_05C	Expresses the general term algebraically in a simple number pattern.
M10_08	Given the length of the sides of a rectangle in terms of a variable, identifies the algebraic expression for its area.
M11_03	Adds three simple algebraic rational expressions with different numerical denominators.
M11_09	Identifies the sum of three consecutive whole numbers given the middle number in general terms.
M12_06	Identifies the equation that models a situation involving distance, speed, and time.
M12_08B	Explains how to find a specific term in a pattern presented numerically and geometrically.

M12_08C	Expresses the general term algebraically in a pattern presented numerically and geometrically.
M13_04A	Extends a number pattern presented geometrically and numerically to solve a problem.
M13_04B	Extends a number pattern presented geometrically and numerically to solve a problem.
M13_04C	Extends a number pattern presented geometrically and numerically to solve a problem.
M13_05	Expresses the general term algebraically for two related number patterns.
M14_10	Uses a given formula to solve a word problem.

Geometry

M01_08	Uses properties of congruent triangles and the sum of the angles of a triangle to find the measure of an angle.
M01_11A	Uses computation with fractions to find the length and width of a rectangle and draws and labels that rectangle on a grid.
M01_12	Finds the area of a triangle inscribed in a square with known dimensions.
M02_10	Uses properties of parallel lines and triangles to find the measure of an angle sum.
M03_04	Identifies the image of a triangle under a rotation about a point in the plane.
M04_10	Uses properties of isosceles and right triangles to find the measure of an angle.
M06_13	Identifies the image of a triangle under a rotation about a point in the plane.
M07_09	Solves a problem involving angle bisectors and angles at a point on a straight line.
M08_14	Uses properties of similar triangles to identify equal angles.
M09_07	Uses information about the lengths of segments on a line to determine the distance between their midpoints.
M09_08	Finds the perimeter of a square, given its area is a square number.
M10_09	Identifies the polygon that has a line of symmetry.

M11_06	Uses knowledge of time, clocks, and angles to solve a problem.
M11_07	Determines the area of a trapezoid inscribed in a rectangle.
M12_10	Uses the Pythagorean theorem in finding the perimeter of a trapezoid.
M13_06	Uses knowledge of interior angles of a triangle to determine the angle sum of a given polygon.
M14_12	Uses Pythagorean theorem in finding the area of a triangle.

Data and Chance

M04_12B	Determines the truth of statements made about data shown in a scattergraph.
M05_07A	Completes a table by interpreting several timetables to identify times that meet a given set of conditions.
M05_07B	Derives information from given timetables to complete a table for a specified journey and check that it meets given conditions.
M10_12B	Finds the median of a set of numbers.
M11_11	Given a spinner, identifies the expected frequency of a particular outcome.
M11_13B	Interprets information from a line graph to determine an average.
M13_02	Solves a problem involving extrapolation of the data shown in a double bar graph.
M14_15	Uses understanding of average to solve a problem.

Items Above the Advanced International Benchmark (625)

Number

M01_10	Estimates the total time in minutes for an event made up of a series of events, each given in minutes and seconds.
M05_06	Calculates total costs for each of two groups given different unit costs and discounts.
M14_06C	Compares information from two sources and explains the result.

Algebra

M06_07	Solves an inequality.

Geometry

M07_07	Uses knowledge of the area of a circle and of average rate to solve a problem.

Data and Chance

M10_12C	Uses understanding of median and mean to solve a word problem.

Eighth Grade – Science

Items at Low International Benchmark (400)

Biology

S07_01	Identifies the circulatory system from a list of its components.
S04_01	Recognizes the cells that conduct messages.

Chemistry

S10_07	Recognizes the material that would complete an electric circuit.
S12_06	Recognizes the material that best conducts heat and electricity.
S10_01	Recognizes the chemical formula of carbon dioxide.

Physics

S07_06	Given the definition of work, identifies a diagram that shows that work is being done.
S08_06	Recognizes the form of energy in a compressed spring.

Items at Intermediate International Benchmark (475)

Biology

S08_05A	Based on data in a table, describes the changes in the population of two organisms over time.
S01_08	Based on a completed food web, predicts and explains what is most likely to happen to a predator population when its prey population is reduced.
S02_02	Explains why exposure to influenza does not necessarily lead to infection.

S08_01	Recognizes how vaccination helps prevent illnesses.
S11_01	Recognizes which cells destroy bacteria that enter the body.
S11_04	States why exercise is important for good health.
S13_02	Explains that an acquired characteristic cannot be passed onto the next generation.
S05_09	Recognizes a characteristic that is found only in mammals.
S04_02	From a diagram, identifies an organ of the digestive system.
S12_01	Recognizes an organism that is a producer.
S04_05	Recognizes a disease caused by a virus.

Chemistry

S10_05	Recognizes from a description of indicator color changes that neutralization has occurred.
S10_11	Recognizes a chemical process involving energy absorption.
S07_05	Identifies vinegar as an acidic solution.
S12_04	In the context of an investigation, identifies the condition under which nails would rust most.

Physics

S03_05	Applies knowledge that sound requires a medium to travel through by contrasting a situation on Earth to a situation on the moon.
S14_10A	Given a diagram showing a ball being thrown upwards, states the force that causes the ball to fall.

Earth Science

S06_13	Recognizes where active volcanoes are most likely to be found.
S01_06	Predicts a long-term effect of cutting down trees on the environment.
S08_13	Matches each of four processes that take place in the water cycle with descriptions of the processes.
S11_10	Given a starting point, orders the processes involved in the water cycle.
S13_13	Identifies paper from a list of common materials as the one that breaks down fastest.
S12_14	Recognizes what is caused by Earth rotating in its axis.

Items at High International Benchmark (550)

Biology

S02_01	Recognizes digestion from a description of the process.
S09_04	States one function of the uterus.
S14_02	Classifies animals in a list into two groups on the basis of a physical or behavioral characteristic and states the characteristic used.
S11_03	Compares two diagrams showing a pair of eyes and recognizes that more light results in smaller pupils.
S03_10	Recognizes the hierarchy of organization in living organisms (cell, tissue, organ, and organism).
S07_02	States one structure that is found in plant cells but not in animal cells.
S07_03	Given that chlorophyll is needed for photosynthesis, states two other factors that are needed.
S10_06	Given a graphical representation of the results of an investigation into the effects of light intensity and carbon dioxide concentration on the rate of photosynthesis, describes the relationship between carbon dioxide concentration and rate of photosynthesis.
S05_10	Recognizes that comparing genes can determine whether two people are related.
S04_03	In the context of an investigation comparing the growth of plants from genetically identical seeds under different conditions, predicts which plants will grow tallest and justifies the answer.
S13_03	Explains that camouflage helps snails avoid predators.
S01_08	Completes the food web of an ocean ecosystem based on information given in a table that lists a number of species and how they obtain their energy.
S14_04A	Indicates in a table which gas is released into the air and which gas is removed from the air during animal respiration.

S06_05A	From a graph showing the population changes over time of two organisms, identifies the time when the population of one of the organisms is at its highest.
S11_02	Interprets a graph showing a sudden drop in the size of a population of an organism and recognizes that loss of food supply is most likely to have caused this sudden drop.
S14_04C	Indicates in a table which gas is released into the air and which gas is removed from the air during photosynthesis.
S08_05B	Based on data in a table showing population changes over time, concludes that there is a population decline and gives an explanation for this decline.
S08_02	Applies knowledge of ecosystems to explain why birds of prey cannot survive in an environment without plants.
S06_03	Applies knowledge of competition to explain the importance of removing weeds from a field where crops are sown.
S13_12	States how a volcanic eruption impacts the environment.
S02_03	Recognizes the food that contains the highest percentage of protein.
S05_13	Recognizes the type of food that should be avoided by a person without a gall bladder.
S06_01	Interprets a graph showing changes in pulse rates before, during, and after exercise and recognizes what can be concluded from the graph.
S05_07	Recognizes the main function of chlorophyll.
S03_03	Applies knowledge of the processes of photosynthesis and respiration to identify gases used up and given off by plants and animals in a forest ecosystem pictured in a diagram.
S12_05C	Recognizes an advantage for a species of butterfly to resemble another species of butterfly that is toxic to birds.

Chemistry

S08_08B	In the context of an investigation about the gold content of jewelry, selects information from a table of properties of gold alloys to complete a table relating the density of alloys to number of carats and percentage of gold in each piece of jewelry.
S08_08C	In the context of an investigation about the gold content of jewelry, uses previously selected information and follows an example to calculate the mass of gold in jewelry.
S07_04	Interprets data in a table of physical properties to identify iron, water, and oxygen.
S04_11A	In the context of an investigation of density, interprets a table summarizing the methods used for measuring mass by four groups and explains why their results differed.
S11_06	Identifies a property of metals and describes how this property can be used to determine whether an unknown substance is a metal or nonmetal.
S06_06	Given the chemical formula for sulfuric acid, completes a table to show the number of atoms of each element in a molecule of the acid.
S12_08	In the context of an investigation, identifies which of two solutions is more dilute and justifies the selection.
S04_10	Recognizes that oxygen is necessary for burning.
S13_05	Explains what causes a balloon to inflate when sodium bicarbonate in the balloon mixes with vinegar.
S13_04	Recognizes the graph that most likely shows the effect of temperature on the solubility of sugar in water.
S03_02	Given a report of an experiment, distinguishes an observation from a prediction, conclusion, theory, or hypothesis.

Physics

S03_06	Based on a diagram demonstrating an investigation of thermal conductivity, recognizes that metal conducts heat faster than glass, wood, or plastic.
S06_10	Recognizes that molecules of a liquid slow down as the liquid cools.

S13_14	Recognizes that gas molecules move faster when temperature increases.
S12_07	Given a table showing speed of sound through different media, identifies the state of each medium and uses this information to recognize a conclusion that can be drawn from the table.
S03_11	Interprets data presented in a nonlinear distance vs. time graph.
S12_09	Recognizes why a helium balloon rises into the air.
S08_12	States the forces acting on students sitting on a wall.
S12_12	Explains why lightning is seen before thunder is heard during an electrical storm.
S03_04	Completes a table showing the relation between voltage and current.
S09_08	Identifies conduction as the process by which heat is transferred along a metal rod.
S05_03	Recognizes why the height of an alcohol column in a thermometer changes with increasing and decreasing temperature.
S14_07	Recognizes the pathway of light for an object to be seen.
S02_08	Recognizes how sound waves with large amplitude differ in energy and loudness from sound waves with smaller amplitude.
S14_08	Recognizes the object most likely to be used as a lever.

Earth Science

S09_10	Interprets a contour map to recognize a topographical representation of a mountain top.
S02_13	Describes how soil is formed.
S10_17	Explains how water evaporated from the sea ends up as rain on land.
S02_14	From a diagram showing the relative location of different towns and information about weather conditions in these towns, recognizes a prediction about future weather conditions.
S12_13	Describes what causes earthquakes.
S04_13	Describes one way groundwater can become polluted.

S04_14	Describes how trees can reduce soil erosion.
S05_11	Predicts one effect a new dam could have on wildlife.
S05_01	Recognizes the definition of an Earth year.
S05_06	Applies knowledge of the relative distances of the sun and the moon from Earth to explain why light from the moon reaches Earth in less time.
S04_15	Recognizes a nonrenewable resource.
S14_14	Recognizes a consequence of the gravitational pull of the moon on Earth.
S09_09	Recognizes the major cause of tides.
S08_09	Recognizes that carbon dioxide is increasing in Earth's atmosphere.
S03_07	Given a diagram of Earth's water cycle, recognizes the sun as the source of energy for the water cycle.
S11_11	Recognizes which soil change is due to a natural cause rather than human activity.
S13_10	Recognizes the main difference between planets and moons.

Items at Advanced International Benchmark (625)

Biology

S06_04	States a life function of a paramecium, other than taking in nutrients to produce energy.
S12_05B	In the context of an observation of butterflies and plants, identifies a developing stage in the life cycle of an organism and describes what takes place during that stage.
S14_05	Using the equipment and materials shown in a diagram, describes an investigation to find out how fertilizer affects the growth of plants.
S12_05A	In the context of an observation of butterflies and plants, identifies the growth stage in the life cycle of an organism and describes what takes place during that stage.
S02_05	Recognizes that a zygote is formed immediately after fertilization.
S01_09	From diagrams showing organisms that live in the intertidal zone, selects one organism, and explains how a physical feature or behavior helps the organism to survive low tide.

S01_10	States two conditions that are found at the bottom of the ocean that make it difficult for most organisms to live there.
S04_06	Completes a diagram to show the direction of the energy flow in a food web.
S09_11	Based on demographic and other information about two countries, predicts how their population will change over time.
S09_11	Given a table showing demographic data and data on grain production and oil consumption for two countries, predicts how a change in population in each country will affect land use over the next 10 years.
S09_11	Given a table showing demographic data and data on grain production and oil consumption for two countries, predicts how a change in population in each country will affect pollution over the next 10 years.
S04_04	Describes two environmental problems likely to occur when a city doubles in population over a short time.
S07_12	States one reason why the human population increased rapidly over the last 200 years.
S06_05B	Interprets a graph showing the population changes over time of two organisms and describes how the changes in population sizes are related.
S13_06	Recognizes that vaccines provide the body with long-term immunity.
S10_02	Recognizes the function of a labeled part of a plant cell.
S13_01	Recognizes that the purpose of cellular respiration is to provide energy for cell activities.
S09_01	Identifies food source as a criterion for classifying animals into two groups.
S12_03	Recognizes an organism in which oxygen and carbon dioxide are exchanged between air and blood through the skin.
S10_03	Recognizes an organ in a frog that has a function similar to that of lungs.
S14_01	Recognizes a function shared by the lungs, skin, and kidneys.
S08_03	Recognizes a function of the cell membrane.

S06_02	Recognizes that the first organisms that appeared on Earth lived in water.
S09_02	Recognizes that organisms that are producers use energy from the sun to make food.
S11_13	Recognizes that the increase in algal growth in a lake is most likely due to fertilizer runoff.

Chemistry

S12_11	Applies knowledge of density to explain why oil floats on water.
S08_08A	In the context of an investigation about the gold content of jewelry, describes the measurements to be taken using a graduated cylinder and water to find the volume of the jewelry.
S01_02	Based on an incomplete table comparing pure water and salt water, explains that addition of salt increases the density.
S05_12	Recognizes electrical conductivity as the criterion used for classifying materials into two groups.
S06_08	Recognizes the definition of a compound.
S14_09	Applies knowledge of expansion of water during freezing to explain why a bottle full of water cracked when it was left in a freezer.
S02_12	Explains that a chemical change in milk caused litmus paper to turn from blue to pink.
S06_11	Describes two things that might be observed as a chemical reaction takes place.
S12_10	Applies knowledge of conservation of mass during a neutralization reaction to explain what happens to mass when new substances are formed.
S05_05	Recognizes an example of a physical change.
S02_10	Applies knowledge of conservation of mass during a chemical reaction to explain what happens to mass when a new substance is formed.
S03_01	From a list of gases, identifies oxygen as the gas that causes rust formation.
S02_11	Recognizes a model showing the configuration of subatomic particles in an atom.

S11_05	Recognizes the concept map that best represents the particulate structure of matter going from molecules to atoms to subatomic particles (protons, neutrons, and electrons).
S09_05	Recognizes which diagram best represents the structure of water molecules.
S05_04	Recognizes that when sugar is dissolved in water, the sugar molecules continue to exist, but in solution.

Physics

S08_11	Describes how to distinguish between fresh water and salt water, using two hot plates but no thermometer.
S03_14	From a description of an experiment investigating the effect of dissolved salt on the freezing point of water, identifies the problem under investigation or states a conclusion based on prior knowledge.
S03_13	Applies knowledge of phase change and the boiling point of water to explain that the temperature of water does not exceed its boiling point despite the addition of heat.
S10_12	Identifies the characteristics or properties that change or remain the same as a liquid changes into a gas.
S05_08	Applies the principle of conservation of mass during phase change to explain why the mass of water remains unchanged after it is frozen.
S06_12B	In the context of an investigation into the relative efficiency of two heat sources, identifies a variable that was controlled.
S06_09	Given two metal bars, one of which is a magnet, describes how to use the magnet to determine if the other metal bar is a magnet.
S04_09	From a diagram showing three magnets, explain why two of them are touching and why the third remains separated.
S11_09	Recognizes that the force of gravity acts on a person regardless of position and movement.
S02_15A	In the context of an investigation about lifting blocks to build a pyramid, identifies the parts of an Egyptian lever, based on a model of the lever.

S05_02	Demonstrates an understanding that the surface of a liquid remains horizontal by drawing the level of the liquid on a frame-of-reference diagram depicting a tilted U-shaped container.
S10_14	On a diagram of a person looking through a periscope, draws the path and direction of a light ray through the periscope.
S07_08	Recognizes that plucking a guitar string harder causes the volume to increase but does not affect the pitch.
S13_09	Predicts the effect of removing air on the propagation of sound.
S13_07	Recognizes that when brought from a mountain top to a valley, a closed empty plastic bottle collapses because the air pressure in the valley is higher than on the mountain top.
S01_03	Recognizes that particles of a liquid move more slowly and are closer together than particles of a gas.
S07_07	Recognizes that mass is conserved during thermal expansion.
S06_12A	Recognizes where to place a thermometer in a liquid to take a reading while conducting an investigation.
S13_08	Recognizes that railway tracks are laid down with gaps between lengths to allow expansion on hot days.
S02_07	Recognizes that the color of an object is the same as the color of the light waves that are reflected by the object.
S12_15	Recognizes that a shadow is shortest when the sun is overhead.
S09_07	Interprets a circuit diagram to recognize that the current flowing through two bulbs is the same.
S10_08	From a description of an investigation about magnets, recognizes how the strength of a magnet is defined.
S04_12	From a diagram showing different liquids layered in a beaker, recognizes an accurate statement about relative densities.

Earth Science

S09_10	Draws on a contour map the path and direction of a river flowing from a mountain to a bay.

S11_12	Describes changes in atmospheric conditions that occur with increasing elevation.
S01_05	Identifies and explains a physical process that can cause weathering of rocks.
S12_16	Draws an arrow on a map to show the direction a river flows and explains why it flows in this direction.
S03_09	States that sulfur dioxide produced by burning coal combines with water vapor in the atmosphere to form acid rain.
S05_14	Describes how science and technology can be used to address global warming caused by increased levels of carbon dioxide in the atmosphere.
S02_16	Provides a reason why recycling household materials is important.
S07_11	Interprets data in a table to describe the effect of amount of fertilizer on the yield of rice.
S01_01	Recognizes the percentage of total water on Earth that is fresh water.
S13_11	Given a diagram showing weather conditions at different elevations on a mountain, identifies the most likely location of a jungle.
S07_09	Relates the tilt of Earth's axis as it orbits the sun to the seasons.
S08_14	Recognizes what causes the moon to appear to change shape.

Items Above the Advanced International Benchmark (625)

Biology

S02_06	Recognizes the likely classification of an animal with scales that uses only its lungs to exchange gases.
S02_04	Recognizes that the average body temperature of people living in hot climates is the same as those living in cold climates and provides a justification.
S03_12	Provides an explanation of why the heart beats faster during exercise.
S14_03	Recognizes which organelle produces energy for the cell.

S10_04	Recognizes an equation that summarizes the process of respiration.
S09_03	States two conditions needed for germination of seeds.
S12_02	Recognizes and describes an example of asexual reproduction.
S10_09	Designs an investigation to test a hypothesis about whether red and green peppers are produced by the same type of pepper plant.
S08_04	Recognizes an explanation for a change over time in a physical characteristic of an organism.
S14_04B	Indicates in a table which gas is released into the air and which gas is removed from the air during plant respiration.
S01_07	Recognizes the graph showing increasing rate of human population growth over the last 200 years.

Chemistry

S04_11B	In the context of an investigation of density, explains why two approaches to measuring the volume of an empty can gave different results.
S04_11C	As part of an investigation of density of a metal can, interprets a table of mass, volume, and density to identify the method that determined the density of the metal of the can.
S14_12	Explains why ice will stay frozen in a wooden container longer than in a metal container.
S10_10	Classifies items as elements, compounds, or mixtures.
S14_06	Recognizes air as a mixture.
S02_09	Describes the steps used to separate salt from a mixture of salt, sand, and leaves, and provides a reason for each step.
S09_06	States one thing that could be observed that shows energy has been released during a chemical reaction.

Physics

S04_08	Recognizes a diagrammatic representation of the particles in a metal after heating.
S08_10	Recognizes that mass is conserved and volume increases as water freezes.

S11_07	Recognizes a sequence of energy conversions that takes place in a battery-operated flashlight.
S08_07	Interprets a diagram and describes the direction of heat flow in metals.
S10_13	Explains why an unwrapped block of ice will melt faster than a block of ice wrapped in newspaper.
S11_08	Interprets a diagram showing air and water in a sphere attached to a U-tube and explains that heating the air in the sphere can cause the water level in the open tube to rise.
S10_15	Recognizes that light travels fastest through a vacuum.
S04_07	Describes an advantage of using parallel rather than series electrical circuits in homes.
S14_11	Applies Ohm's law to calculate resistance from current and voltage.
S01_04	Recognizes that an iron nail becomes magnetized when current flows through a wire coiled around the nail.
S14_10B	Given a diagram showing a ball being thrown upwards, falling to the ground and bouncing, explains why the ball will not bounce to the height from which it fell.
S02_15B	As part of an investigation about lifting blocks to build a pyramid, uses information shown in a diagram of a lever and applies a given formula to calculate the force needed to lift a block.

Earth Science

S07_10	Recognizes that most fresh water on Earth is located in the polar ice caps.
S01_05	Identifies and explains a chemical process that can cause weathering of rocks.
S06_07	Applies knowledge of condensation to explain why a liquid appeared on the outside of a pitcher of cold water.
S05_14	Describes how science and technology can be used to address oil spills in the oceans.
S06_14	Recognizes a diagrammatic representation of the sun, moon, and Earth during an eclipse of the moon.

TYPOGRAPHY: Set in Meridien, Minion, and Myriad.
PRODUCTION EDITOR: Debra R. Berger
PHOTOGRAPH: Felix Tchvertkin
COVER DESIGN: Mario A. Pita
BOOK DESIGN: Sue Farrell and Ruthanne Ryan
LAYOUT & PRODUCTION: Sue Farrell, Mario A. Pita, Ruthanne Ryan, Jennifer A. Moher, and Bobby Wong

TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

ISBN 1-889938-50-5

BOSTON COLLEGE

TIMSS 2007

timssandpirls.bc.edu