

**User Guide for the TIMSS International Database**  
**Final Year of Secondary School**

**(Population 3)**  
**Data Collected in 1995-96**

**Edited by**

Eugenio J. Gonzalez  
Teresa A. Smith  
Heiko Sibbers

**with contributions by**

Ray Adams  
Jean Dumais  
Pierre Foy  
Dirk Hastedt  
Dana L. Kelly  
Greg Macaskill  
Michael O. Martin  
Ina V.S. Mullis  
Knut Schwippert  
Ce Shen

**June 1998**



TIMSS International Study Center  
Boston College  
Chestnut Hill, MA, USA

© 1998 International Association for the Evaluation of Educational Achievement (IEA).

User Guide for the TIMSS International Database – Final Year of Secondary School  
1995 Assessment / Edited by Eugenio J. Gonzalez, Teresa A. Smith, and Heiko Sibberns

To obtain additional copies of the TIMSS International Database and User Guide contact the IEA.

International Association for the Evaluation of Educational Achievement  
The IEA Secretariat  
Herengracht 487  
1017 BT Amsterdam  
The Netherlands

Tel: +31 20 625 36 25  
Fax: +31 20 420 71 36  
email: [Department@IEA.nl](mailto:Department@IEA.nl)

For more information about TIMSS contact the TIMSS International Study Center

TIMSS International Study Center  
Campion Hall 323- CSTEELP  
School of Education  
Boston College  
Chestnut Hill, MA 02467  
United States

email: [timss@bc.edu](mailto:timss@bc.edu)  
<http://www.csteep.bc.edu/timss>

Funding for the international coordination of TIMSS is provided by the U.S. National Center for Education Statistics, the U.S. National Science Foundation, the IEA, and the Canadian government. Each participating country provides funding for the national implementation of TIMSS.

Boston College is an equal opportunity, affirmative action employer.

Printed and bound in the United States.

---

# Contents

<b>Chapter 1: Overview of TIMSS and the Database User Guide.....</b>	<b>1-1</b>
1.1 Overview of the International Database.....	1-1
1.2 Overview of TIMSS.....	1-2
1.3 TIMSS International Reports.....	1-5
1.4 Contents of the Database – Population 3.....	1-5
1.5 Contents of the User Guide.....	1-6
1.6 Management and Operations of TIMSS.....	1-7
1.7 Additional Resources.....	1-8
<b>Chapter 2: TIMSS Instruments and Booklet Design.....</b>	<b>2-1</b>
2.1 TIMSS Mathematics and Science Tests for Final-Year Students.....	2-1
2.1.1 Mathematics and Science Content.....	2-1
2.1.2 The TIMSS Items.....	2-3
2.1.3 TIMSS Test Design.....	2-6
2.1.4 Calculator Policy.....	2-9
2.2 Questionnaires.....	2-9
2.3 Release Status for TIMSS Test Items and Background Questionnaires – Final Year Assessment...2-10	
<b>Chapter 3: Sampling and Sampling Weights.....</b>	<b>3-1</b>
3.1 The Target Population.....	3-1
3.2 Coverage of the TIMSS Population.....	3-1
3.3 TIMSS Coverage Index.....	3-4
3.4 Sample Design.....	3-7
3.5 School Sampling.....	3-9
3.6 Student Sampling.....	3-9
3.7 Response Rates.....	3-10
3.7.1 School-Level Response Rates.....	3-10
3.7.2 Student-Level Response Rates.....	3-10
3.7.3 Overall Response Rates.....	3-11
3.8 Compliance with Sampling Guidelines.....	3-13
3.9 Sampling Weights.....	3-16
3.10 Sampling and Weight Variables Included in the Student Data Files.....	3-19
3.11 Selecting Students for Analyses.....	3-25
3.12 Weight Variables Included in the School Data Files.....	3-27
<b>Chapter 4: Data Collection, Materials Processing, Scoring, and Database Creation.....</b>	<b>4-1</b>
4.1 Data Collection and Field Administration.....	4-1
4.2 Free-Response Scoring.....	4-2
4.3 Data Entry.....	4-6
4.4 Database Creation.....	4-7
4.5 Instrument Deviations and National Adaptations.....	4-7

4.5.1 Cognitive Items .....	4-7
4.5.2 Background Questionnaire Items .....	4-9
<b>Chapter 5: TIMSS Scaling Procedures.....</b>	<b>5-1</b>
5.1 The TIMSS Scaling Model .....	5-1
5.2 The Unidimensional Random Coefficients Model .....	5-2
5.3 The Multidimensional Random Coefficients Multinomial Logit Model .....	5-3
5.4 The Population Model .....	5-4
5.5 Estimation.....	5-5
5.6 Latent Estimation and Prediction .....	5-6
5.7 Drawing Plausible Values.....	5-6
5.8 Scaling Steps .....	5-7
5.9 Standardizing the International Scale Scores.....	5-7
<b>Chapter 6: Student Achievement Scores.....</b>	<b>6-1</b>
6.1 Achievement Scores in the Student Files .....	6-1
6.1.1 Raw Scores.....	6-1
6.1.2 National Rasch Scores .....	6-2
6.1.3 International Expected a Posteriori (EAP) Scores.....	6-2
6.1.4 Plausible Values.....	6-3
6.2 Additional Scores Included in the Database .....	6-4
<b>Chapter 7: Content and Format of Database Files .....</b>	<b>7-1</b>
7.1 Introduction .....	7-1
7.2 Data Files.....	7-2
7.2.1 Background Files .....	7-3
7.2.1.1 Student Background File .....	7-3
7.2.1.2 School Background File.....	7-4
7.2.1.3 Identification Variables .....	7-4
7.2.1.4 Achievement Scores .....	7-6
7.2.1.5 Tracking Variables.....	7-8
7.2.1.6 International Background Variables .....	7-9
7.2.1.7 Variables Derived from Student Background Data.....	7-11
7.2.1.8 Sampling and Weight Variables .....	7-13
7.2.2 Written Assessment Files .....	7-15
7.2.2.1 Cognitive Item Variable Names.....	7-15
7.2.2.2 Cognitive Item Response Code Values.....	7-16
7.2.2.3 Analysis By Mathematics and Science Literacy, Physics, and Advanced Mathematics Content Area Reporting Categories.....	7-16
7.2.2.4 Release Status of TIMSS Test Items.....	7-20
7.2.2.5 Other Variables in the Written Assessment Files .....	7-20
7.2.3 Coding Reliability Files.....	7-20
7.2.4 Missing Codes in the International Data Files.....	7-22
7.3 Codebook Files .....	7-25
7.3.1 Accessing the Codebook Files.....	7-25
7.3.2 Using the Codebooks.....	7-26

7.4	Program Files.....	7-29
7.5	Data Almanacs.....	7-30
7.6	Test-Curriculum Matching Analysis Data Files .....	7-34
<b>Chapter 8: Estimating Sampling and Imputation Variance.....</b>		<b>8-1</b>
8.1	Estimating Sampling Variance .....	8-1
8.1.1	Computing Sampling Variance Using the JRR Method .....	8-1
8.1.2	Construction of Sampling Zones for Sampling Variance Estimation .....	8-2
8.1.3	Computing the JRR Replicate Weights .....	8-3
8.2	Estimating Imputation Variance .....	8-4
8.3	Combining Sampling and Imputation Variance .....	8-4
<b>Chapter 9: Performing Analyses with the TIMSS Data: Some Examples .....</b>		<b>9-1</b>
9.1	Contents of the CDs.....	9-1
9.2	Creating SAS Data Sets and SPSS System Files.....	9-3
9.3	Computing JRR Replicate Weights, Sampling, and Imputation Variance for the Achievement Scores Using SPSS and SAS.....	9-6
9.3.1	SAS Macro for Computing Mean Plausible Values and Percents with Corresponding Standard Errors (JACKPV.SAS).....	9-7
9.3.2	SPSS Macro for Computing Mean and Percents with Corresponding Standard Errors (JACKPV.SPS).....	9-18
9.4	Computing JRR Replicate Weights and Sampling Variance for Variables other than the Plausible Value Scores Using SPSS and SAS.....	9-26
9.4.1	SAS Macro for Computing Mean and Percents with Corresponding Standard Errors (JACKGEN.SAS).....	9-26
9.4.2	SPSS Macro for Computing Mean and Percents with Corresponding Standard Errors (JACKGEN.SPS).....	9-34
9.5	Performing Analyses with Student-Level Variables Using Plausible Values.....	9-40
9.6	Performing Analyses with Student-Level Variables Not Using Plausible Values.....	9-47
9.7	Performing Analyses with School-Level Variables.....	9-52
9.8	Scoring the Items.....	9-58

## References Cited in Text

## Acknowledgments

# Tables and Figures

Table 1.1	Countries Participating in TIMSS at Population 3 (Data Included in Database).....	1-2
Figure 2.1	The Major Categories of the TIMSS Curriculum Frameworks.....	2-2
Table 2.1	Mathematics and Science Literacy, Advanced Mathematics, and Physics Content Categories Final Year of Secondary School.....	2-3
Table 2.2	Distribution of Mathematics and Science Literacy Items by Reporting Category .....	2-5
Table 2.3	Distribution of Advanced Mathematics Items by Content Category .....	2-5
Table 2.4	Distribution of Physics Items by Content Category .....	2-6
Table 2.5	Assignment of Item Clusters to Test Booklets .....	2-7
Table 2.6	Number of Test Items per Booklet by Content Category .....	2-8
Figure 3.1	Relationship Between the Desired Populations and Exclusions .....	3-2
Table 3.1	Coverage of TIMSS Target Population .....	3-3
Table 3.2	TIMSS Coverage Indices (TCIs) .....	3-6
Table 3.3	TIMSS Coverage Indices (TCIs) for Advanced Mathematics and Physics.....	3-7
Table 3.4	Assignment of Test Booklet According to Student Classification .....	3-8
Table 3.5	Participation Rates – Mathematics and Science Literacy.....	3-11
Table 3.6	Participation Rates – Advanced Mathematics.....	3-12
Table 3.7	Participation Rates – Physics .....	3-13
Figure 3.2	Countries Grouped for Reporting Achievement According to Their Compliance with Guidelines for Sample Implementation and Participation Rates – Mathematics and Science Literacy.....	3-14
Figure 3.3	Countries Grouped for Reporting Achievement According to Their Compliance with Guidelines for Sample Implementation and Participation Rates Advanced Mathematics .....	3-15
Figure 3.4	Countries Grouped for Reporting Achievement According to Their Compliance with Guidelines for Sample Implementation and Participation Rates – Physics.....	3-16
Table 3.8	Sample Information for Students in their Final Year of Secondary School.....	3-17
Table 3.9	Sample Information for Students in their Final Year of Secondary School Having Taken Advanced Mathematics .....	3-18
Table 3.10	Sample Information for Students in their Final Year of Secondary School Having Taken Physics .....	3-18
Table 3.11	Versions of the Sampling Weights.....	3-22
Table 3.12	Participation Status Variables for the Student Data Files.....	3-26
Figure 4.1	Example Coding Guide for Short-Answer Science Literacy Item.....	4-3
Figure 4.2	Example Coding Guide for Extended-Response Advanced Mathematics Item.....	4-4
Table 4.1	TIMSS Within-Country Free-Response Coding Reliability Data .....	4-6
Table 4.2	List of Deleted Cognitive Items – Population 3 .....	4-8
Table 4.3	List of Recodes Made to Free-Response Item Codes .....	4-9
Table 6.1	Descriptive Statistics for Mathematics and Science Literacy Plausible Values .....	6-5
Table 6.2	Descriptive Statistics for Mathematics Literacy Plausible Values.....	6-6
Table 6.3	Descriptive Statistics for Science Literacy Plausible Values.....	6-7
Table 6.4	Descriptive Statistics for Advanced Mathematics Plausible Values.....	6-8
Table 6.5	Descriptive Statistics for Numbers, Equations, and Functions Plausible Values.....	6-9
Table 6.6	Descriptive Statistics for Calculus Plausible Values.....	6-10

Table 6.7	Descriptive Statistics for Geometry Plausible Values .....	6-11
Table 6.8	Descriptive Statistics for Physics Plausible Values.....	6-12
Table 6.9	Descriptive Statistics for Mechanics Plausible Values.....	6-13
Table 6.10	Descriptive Statistics for Electricity and Magnetism Plausible Values.....	6-14
Table 6.11	Descriptive Statistics for Heat Plausible Values.....	6-15
Table 6.12	Descriptive Statistics for Wave Phenomena Plausible Values.....	6-16
Table 6.13	Descriptive Statistics for Modern Physics Plausible Values.....	6-17
Table 6.14	Descriptive Statistics for Reasoning and Social Utility Plausible Values.....	6-18
Table 7.1	TIMSS Population 3 Data Files.....	7-2
Table 7.2	Country Identification and Inclusion Status in Population 3 Data Files.....	7-3
Table 7.3	Definitions of Stratification Options (IDSTRAT) for Population 3 Countries.....	7-5
Table 7.4	Background Questionnaire Item Field Location Format Conventions for Population 3.....	7-10
Table 7.5	International Background Variable Naming Conventions for Population 3.....	7-11
Table 7.6	International Report Table Location Reference Definition for Derived Variables in Population 3.....	7-12
Table 7.7	Variable Name Definitions for the Written Assessment Items in Population 3.....	7-15
Table 7.8	Classification of Population 3 Items by Mathematics and Science Literacy Reporting Categories.....	7-17
Table 7.9	Classification of Population 3 Items by Advanced Mathematics Content Area Reporting Categories.....	7-18
Table 7.10	Classification of Population 3 Items by Physics Content Area Reporting Categories.....	7-19
Table 7.11	Population 3 Codebook Files.....	7-25
Table 7.12	File Structure of Machine-Readable Codebook Files.....	7-26
Figure 7.1	Example Printout of the Student Background Codebook.....	7-27
Table 7.13	Population 3 Program Files.....	7-29
Table 7.14	Data Almanac Files for Population 3.....	7-30
Figure 7.2	Example Data Almanac Display for Categorical Student Background Variable.....	7-31
Figure 7.3	Example Data Almanac Display for a Numerical School Background Variable.....	7-33
Table 8.1	Sampling and Imputation Variance for the Mathematics and Science Literacy Scale Plausible Values.....	8-5
Table 8.2	Sampling and Imputation Variance for the Mathematics Literacy Scale Plausible Values.....	8-6
Table 8.3	Sampling and Imputation Variance for the Science Literacy Scale Plausible Values.....	8-7
Table 8.4	Sampling and Imputation Variance for the Advanced Mathematics Scale Plausible Values.....	8-8
Table 8.5	Sampling and Imputation Variance for the Numbers, Equations, and Functions Scale Plausible Values.....	8-8
Table 8.6	Sampling and Imputation Variance for the Calculus Scale Plausible Values.....	8-9
Table 8.7	Sampling and Imputation Variance for the Geometry Scale Plausible Values.....	8-9
Table 8.8	Sampling and Imputation Variance for the Physics Scale Plausible Values.....	8-10
Table 8.9	Sampling and Imputation Variance for the Mechanics Scale Plausible Values.....	8-10
Table 8.10	Sampling and Imputation Variance for the Electricity and Magnetism Scale Plausible Values.....	8-11
Table 8.11	Sampling and Imputation Variance for the Heat Scale Plausible Values.....	8-12
Table 8.12	Sampling and Imputation Variance for the Wave Phenomena Scale Plausible Values.....	8-13

Table 8.13	Sampling and Imputation Variance for the Modern Physics Scale Plausible Values .....	8-14
Table 8.14	Sampling and Imputation Variance for the Reasoning and Social Utility Scale Plausible Values .....	8-15
Table 9.1	Three-letter Extension Used to Identify the Files Contained in the CD .....	9-2
Figure 9.1	Extract of the SAS Control Code for Creating a Student Background SAS Dataset.....	9-5
Figure 9.2	Extract of the SPSS Control Code for Creating a Student Background SPSS System File.....	9-6
Figure 9.3	SAS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKPV.SAS).....	9-8
Table 9.2	Number of Replicate Weights Needed for Computing the JRR Error Variance Estimate for Each Sample .....	9-12
Table 9.3	Values for the Parameter ROOTPV .....	9-13
Figure 9.4	SAS Control Code and Output File for Using the Macro JACKPV.SAS.....	9-16
Figure 9.5	Sample Table for Analysis Using Plausible Values Taken From the TIMSS International Report "Mathematics and Science Achievement in the Final Year of Secondary School" .....	9-17
Figure 9.6	SPSS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKPV.SPS).....	9-18
Figure 9.7	SPSS Control Code and Extract of Output File for Using the Macro JACKPV.SPS.....	9-25
Figure 9.8	SAS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKGEN.SAS).....	9-27
Figure 9.9	SAS Control Code and Extract of Output File for Using the Macro JACKGEN.SAS.....	9-32
Figure 9.10	Sample Table for Analysis of Continuous Variables Taken From the TIMSS International Report "Mathematics and Science Achievement in the Final Year of Secondary School" .....	9-33
Figure 9.11	SPSS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKGEN.SPS).....	9-34
Figure 9.12	SPSS Control Code and Extract of Output File for Using the Macro JACK.SPS.....	9-40
Figure 9.13	Sample Table for Analysis of Student-Level Variables and Plausible Values Taken From the TIMSS International Report "Mathematics and Science Achievement in the Final Year of Secondary School" .....	9-43
Figure 9.14	SAS Control Statements for Performing Analyses with Student-level Variables Using Plausible Values (EXAMPLE1.SAS).....	9-44
Figure 9.15	SPSS Control Statements for Performing Analyses with Student-level Variables Using Plausible Values (EXAMPLE1.SPS).....	9-45
Figure 9.16	Extract of SAS Computer Output for Performing Analyses with Student-Level Variables Using Plausible Values.....	9-46
Figure 9.17	Extract of SPSS Computer Output for Performing Analyses with Student-Level Variables Using Plausible Values.....	9-47
Figure 9.18	SAS Control Statements for Performing Analyses with Student-Level Variables Not Using Plausible Values (EXAMPLE2.SAS).....	9-49
Figure 9.19	SPSS Control Statements for Performing Analyses with Student-Level Variables Not Using Plausible Values (EXAMPLE2.SPS).....	9-50
Figure 9.20	Extract of SAS Computer Output for Performing Analyses with Student-Level Variables Not Using Plausible Values.....	9-51
Figure 9.21	Extract of SPSS Computer Output for Performing Analyses with Student-Level Variables Not Using Plausible Values.....	9-52



---

Figure 9.22 SAS Control Statements for Performing Analyses with School-Level Variables Using Plausible Values (EXAMPLE3.SAS) .....	9-54
Figure 9.23 SPSS Control Statements for Performing Analyses with School-Level Variables Using Plausible Values (EXAMPLE3.SPS) .....	9-55
Figure 9.24 Extract of SAS Computer Output for Performing Analyses with School-Level Variables Using Plausible Values .....	9-56
Figure 9.25 Extract of SPSS Computer Output for Performing Analyses with School-Level Variables Using Plausible Values .....	9-57
Table 9.4 Definitions of Response Codes for the Multiple-Choice Items in the Written Assessment Data Files .....	9-58
Table 9.5 Definitions of Response Codes for the Open-Ended Items in the Written Assessment .....	9-58
Figure 9.26 SAS Control Code Used to Convert Cognitive Item Response Codes to Correctness-Score Levels (CSASCORE.SAS) .....	9-60
Figure 9.27 SPSS Control Code Used to Convert Cognitive Item Response Codes to Correctness-Score Levels (CSASCORE.SPS) .....	9-62



---

# Chapter 1

## Overview of TIMSS and the Database User Guide

### 1.1 Overview of the International Database

This User Guide accompanies the TIMSS International Database for the Final Year of Secondary School (TIMSS Population 3)<sup>1</sup>. The database, provided on compact disk, contains achievement data (mathematics and science literacy, advanced mathematics, and physics) and student and school background data collected in 24 countries in 1995-96. Table 1.1 lists the countries for which achievement data are included in the International Database. Each of these countries gave the IEA permission to release its national data.

The TIMSS International Database for the Final Year of Secondary School contains the following, for each country for which internationally comparable data are available.

- Mathematics literacy, science literacy, advanced mathematics, and physics proficiency scale scores
- Subscale scores for advanced mathematics and physics content areas
- Students' responses to cognitive items
- Student background questionnaire data
- School background questionnaire data
- Test-curriculum matching analysis data for advanced mathematics and physics
- Sampling weights
- International codebooks
- SPSS and SAS control statement files
- Data almanacs

Given the size and complexity of TIMSS and the psychometric innovations employed, the TIMSS database is enormous and extremely complex. There are more than 80 files on the disk, containing data and documentation. Every effort has been made to organize the database and provide adequate documentation so that researchers can access the database for secondary analysis. Reading this User Guide is the first step in using the TIMSS database. This guide describes TIMSS, including the data collection instruments, sample design, and data collection procedures; documents the content and format of the data files in the international database; and provides example analyses. Appropriate use of the various files and variables, as well as special considerations arising from the complex design are described. There are three supplements to this User Guide containing copies of the TIMSS international background questionnaires, documentation of national adaptations of the international background questionnaire items, and documentation of derived variables reported in the international report.

This chapter of the User Guide provides an overview of TIMSS, briefly describes the contents of the database, and describes the contents of this User Guide.

<sup>1</sup> The TIMSS International Database for the Primary and Middle School Years and an accompanying User Guide were released in 1997.

**Table 1.1****Countries Participating in TIMSS at Population 3 (Data Included in Database)**

<b>Mathematics and Science Literacy</b>	<b>Advanced Mathematics</b>	<b>Physics</b>
<ul style="list-style-type: none"> <li>• Australia</li> <li>• Austria</li> <li>• Canada</li> <li>• Cyprus</li> <li>• Czech Republic</li> <li>• Denmark</li> <li>• France</li> <li>• Germany</li> <li>• Hungary</li> <li>• Iceland</li> <li>• Israel</li> <li>• Italy</li> <li>• Lithuania</li> <li>• Netherlands</li> <li>• New Zealand</li> <li>• Norway</li> <li>• Russian Federation</li> <li>• Slovenia</li> <li>• South Africa</li> <li>• Sweden</li> <li>• Switzerland</li> <li>• United States</li> </ul>	<ul style="list-style-type: none"> <li>• Australia</li> <li>• Austria</li> <li>• Canada</li> <li>• Cyprus</li> <li>• Czech Republic</li> <li>• Denmark</li> <li>• France</li> <li>• Germany</li> <li>• Greece</li> <li>• Israel</li> <li>• Italy</li> <li>• Lithuania</li> <li>• Russian Federation</li> <li>• Slovenia</li> <li>• Sweden</li> <li>• Switzerland</li> <li>• United States</li> </ul>	<ul style="list-style-type: none"> <li>• Australia</li> <li>• Austria</li> <li>• Canada</li> <li>• Cyprus</li> <li>• Czech Republic</li> <li>• Denmark</li> <li>• France</li> <li>• Germany</li> <li>• Greece</li> <li>• Israel</li> <li>• Italy</li> <li>• Latvia</li> <li>• Norway</li> <li>• Russian Federation</li> <li>• Slovenia</li> <li>• Sweden</li> <li>• Switzerland</li> <li>• United States</li> </ul>

## 1.2 Overview of TIMSS

The Third International Mathematics and Science Study (TIMSS) was conducted in 1995 across more than 40 countries. TIMSS represents the continuation of a long series of studies conducted by the International Association for the Evaluation of Educational Achievement (IEA). Since its inception in 1959, the IEA has sponsored more than 15 studies of cross-national achievement in curricular areas such as mathematics, science, language, civics, and reading. The IEA conducted its First International Mathematics Study (FIMS) in 1964, and the Second International Mathematics Study (SIMS) in 1980-82. The First and Second International Science Studies (FISS and SISS) were carried out in 1970-71 and 1983-84, respectively. Since the subjects of mathematics and science are related in many respects and since there is broad interest in many countries in students'

abilities in both mathematics and science, the third studies were conducted together as an integrated effort.

The number of participating countries, the number of grades tested, and the simultaneous assessment of mathematics and science has resulted in TIMSS becoming the largest, most complex IEA study to date and the largest international study of educational achievement ever undertaken. Traditionally, IEA studies have systematically worked toward gaining a more in-depth understanding of how various factors contribute to the overall outcomes of schooling. Particular emphasis has been given to refining our understanding of students' opportunity to learn as this opportunity becomes successively defined and implemented by curricular and instructional practices.

Continuing the approach of previous IEA studies, TIMSS addressed three conceptual levels of curriculum. The intended curriculum is composed of the mathematics and science instructional and learning goals as defined at the system level. The implemented curriculum is the mathematics and science curriculum as interpreted by teachers and made available to students. The attained curriculum is the mathematics and science content that students have learned and their attitudes towards these subjects. To aid in interpretation and comparison of results, TIMSS also collected extensive information about the social and cultural contexts for learning, many of which are related to variation among educational systems.

Nearly 50 countries participated in one or more of the various components of the TIMSS data collection effort, including the curriculum analysis. To gather information about the intended curriculum, mathematics and science specialists within each participating country worked section by section through curriculum guides, textbooks, and other curricular materials to categorize aspects of these materials in accordance with detailed specifications derived from the TIMSS mathematics and science curriculum frameworks (Robitaille et al., 1993). Initial results from this component of TIMSS can be found in two companion volumes: *Many Visions, Many Aims: A Cross-National Investigation of Curricular Intentions in School Mathematics* (Schmidt et al., 1997a) and *Many Visions, Many Aims: A Cross-National Investigation of Curricular Intentions in School Science* (Schmidt et al., 1997b).<sup>2</sup>

To collect data about how the curriculum is implemented in classrooms, TIMSS administered a broad array of questionnaires, which also collected information about the social and cultural contexts for learning. Questionnaires were administered at the country level about decision-making and organizational features within the educational systems. The students who were tested answered questions pertaining to their attitudes towards mathematics and science, classroom activities, home background, and out-of-school activities. The mathematics and science teachers of sampled students responded to questions about teaching emphasis on the topics in the curriculum frameworks, instructional practices, textbook use, professional training and education, and their views on mathematics and science. The heads of schools responded to questions about school staffing and resources, mathematics and science course offerings, and support for teachers. In addition, a volume was compiled that presents descriptions of the educational systems of the participating countries (Robitaille, 1997).

To measure the attained curriculum, TIMSS tested more than half a million students in mathematics and science at three separate populations.

<sup>2</sup> The data from the TIMSS curriculum analysis are not included in the international database.

- **Population 1.** Students enrolled in the two adjacent grades that contained the largest proportion of 9-year-old students at the time of testing – third- and fourth-grade students in most countries.
- **Population 2.** Students enrolled in the two adjacent grades that contained the largest proportion of 13-year-old students at the time of testing – seventh- and eighth-grade students in most countries.
- **Population 3.** Students in their final year of secondary education. As an additional option, countries could test two special subgroups of these students: students having taken advanced courses in mathematics and students having taken courses in physics.

Countries participating in the study were required to administer tests to the students in the two grades at Population 2 but could choose whether or not to participate at the other levels. In many countries, subsamples of students in the upper grades of Populations 1 and 2 also participated in a performance assessment. The data for Populations 1 and 2 were released in September 1997 on compact disk and are available on the World Wide Web. The International Database for the Primary and Middle School Years is accompanied by the *User Guide for the TIMSS International Database for the Primary and Middle School Years* (Gonzalez and Smith, 1997).

In each country, a National Research Coordinator was responsible for conducting TIMSS in accordance with the international procedures established by the TIMSS International Study Center at Boston College. This included selecting a representative sample of schools and students for each population, translating the data collection instruments into the language(s) of testing, assembling the data collection instruments, sending them to the sampled schools, and arranging for data collection in the schools. In each school sampled for TIMSS, a School Coordinator and a Test Administrator administered the assessment instruments and followed security procedures. After the testing session, the School Coordinator returned the testing materials to the national research center. At that time, the National Research Coordinator arranged for scoring the open-ended responses and arranged to have the test and questionnaire responses entered into data files. These data files were then submitted to the IEA Data Processing Center for international processing. For each task, manuals documenting the international procedures were provided, together with various forms used to document the implementation of the tasks. In addition, international training sessions were held several times a year for National Research Coordinators and their staff members.

### 1.3 TIMSS International Reports

The TIMSS International Database for the Final Year of Secondary School contains the data that were summarized in a report published by the TIMSS International Study Center at Boston College in February 1998:

- *Mathematics and Science Achievement in the Final Year of Secondary School: IEA's Third International Mathematics and Science Study* (Mullis et al., 1998)

The International Database for the Primary and Middle School Years contains the data that were published in 1996 and 1997 in a series of reports prepared by the TIMSS International Study Center at Boston College:

- *Mathematics Achievement in the Primary School Years: IEA's Third International Mathematics and Science Study* (Mullis et al., 1997)
- *Science Achievement in the Primary School Years: IEA's Third International Mathematics and Science Study* (Martin et al., 1997)
- *Mathematics Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study* (Beaton et al., 1996a)
- *Science Achievement in the Middle School Years: IEA's Third International Mathematics and Science Study* (Beaton et al., 1996b)
- *Performance Assessment in IEA's Third International Mathematics and Science Study* (Harmon et al., 1997)

### 1.4 Contents of the Database – Population 3

The TIMSS International Database for the Final Year of Secondary School, provided on compact disk, includes more than 1200 variables in more than 80 files.

**Data Files.** These files include the written assessment data for the mathematics and science literacy test, advanced mathematics test, and physics test; background questionnaire data; coding reliability data; and sampling weights.

**Codebook Files.** The codebook files contain all information related to the structure of the data files as well as the source, format, descriptive labels, and response option codes for all variables.

**Program Files.** These files include programs that allow the user to convert the raw data files into SAS data sets or SPSS system files, estimate sampling variance using the jackknife repeated replication method, and convert item response codes to score values.

**Data Almanacs.** The data almanacs are text files that display weighted summary statistics for each participating country for each variable in the background questionnaires as well as derived variables.

**Test-Curriculum Matching Analysis Files.** These files contain data collected for the TIMSS test-curriculum matching analysis for advanced mathematics and physics.

These files are further described in Chapter 7. Each variable in the TIMSS database is designated by an alphanumeric variable name. Throughout this guide, these variables and the appropriate use of them in conducting analyses are described.

## **1.5 Contents of the User Guide**

Given the size and complexity of the TIMSS International Database, a description of its contents is also complicated. It is recommended that the user read through this guide to understand the study and get a sense of the structure and contents of the database, prior to trying to use the files contained on the CD. During this first reading, there may be particular sections that the user can skim and other sections that the user may want to read more carefully. Nonetheless, a preliminary reading (before actually opening up the files and trying to use them) would help the user better understand the complexities of the study and the International Database. When using the files, the user will need to follow certain sections of this guide more carefully than others and refer to the supplements to the guide. The contents of each chapter and the supplements are summarized below.

### **Chapter 2: TIMSS Instruments and Booklet Design**

This chapter describes the content and organization of the TIMSS tests for the students in their final year of secondary school, and the student and school background questionnaires. The TIMSS item release policy also is described.

### **Chapter 3: Sampling and Sampling Weights**

This chapter describes the sampling design for TIMSS, the use of sampling weights to obtain proper population estimates, and the weight variables included in the data files.

### **Chapter 4: Data Collection, Materials Processing, Scoring, and Database Creation**

This chapter describes the data collection and field administration procedures used in TIMSS, the scoring of the free-response items, data entry procedures, and the creation of the International Database, including the data verification and database restructuring.

### **Chapter 5: TIMSS Scaling Procedures**

This chapter provides an overview of the scaling methodology used by TIMSS, including a description of the scaling model, plausible values technology, the international calibration sample, and standardization of the international scale scores.

### **Chapter 6: Student Achievement Scores**

This chapter describes the student-level achievement scores that are available in the International Database, including how they were derived and used by TIMSS, and how they can be used by secondary analysts.



## **Chapter 7: Content and Format of Database Files**

This chapter provides detailed descriptions of the TIMSS data files, codebook files, data access programs, and data almanacs provided in the TIMSS database.

## **Chapter 8: Estimating Sampling Variance**

This chapter describes the jackknife repeated replication procedure for estimating sampling variance and the estimation of imputation variance.

## **Chapter 9: Performing Analyses with the TIMSS Data: Some Examples**

This chapter provides example programs in SPSS and SAS for conducting analyses on the TIMSS data, including merging data files and using the jackknife repeated replication procedure to estimate standard errors.

## **Supplement 1 - International Versions of the Background Questionnaires**

This supplement contains the international versions of the student and school background questionnaires and a table that maps each question to a variable in the database.

## **Supplement 2 - Documentation of National Adaptations of the International Background Questionnaire Items**

This supplement contains documentation of national adaptations of the international versions of the student and school questionnaire items. This documentation provides users with a guide to the availability of internationally comparable data for secondary analyses.

## **Supplement 3 - Documentation of Derived Variables Based on Student Background Questionnaire Items**

The TIMSS international report included a number of variables derived from questions in the student questionnaire. These derived variables are included in the database and are documented in this supplement to the User Guide.

## **1.6 Management and Operations of TIMSS**

TIMSS is directed by the International Study Center at Boston College in the United States. The TIMSS International Study Center was responsible for supervising all aspects of the design and implementation of the study at the international level, including development and design of the study, data collection instruments, and operational procedures; data analysis; reporting the international results; and quality assurance.

Several important TIMSS functions, including test and questionnaire development, translation checking, sampling consultations, data processing, and data analysis, were conducted by centers around the world, under the direction of the TIMSS International Study Center. The IEA Data

Processing Center (DPC), located in Hamburg, Germany, was responsible for checking and processing all TIMSS data and for constructing the international database. The DPC played a major role in developing and documenting the TIMSS field operations procedures. Statistics Canada, located in Ottawa, Canada, was responsible for advising National Research Coordinators (NRCs) on their sampling plans, for monitoring progress in all aspects of sampling, and for the computation of sampling weights. The Australian Council for Educational Research (ACER), located in Melbourne, Australia, has participated in the development of the achievement tests, has conducted psychometric analyses of field trial data, and was responsible for the development of scaling software and for scaling the achievement test data. The International Coordinating Center (ICC), in Vancouver, Canada, was responsible for international project coordination prior to the establishment of the International Study Center in August 1993. Since then, the ICC has provided support to the International Study Center, and in particular has managed translation verification in the achievement test development process; and has published several monographs in the TIMSS monograph series. As Sampling Referee, Keith Rust of Westat, Inc. (United States), worked with Statistics Canada and the NRCs to ensure that sampling plans met the TIMSS standards, and advised the International Study Director on all matters relating to sampling.

TIMSS was conducted in each country by the TIMSS National Research Coordinator (NRC) and the national research center. NRCs and their staff members were responsible for carrying out the TIMSS data collection, scoring, and data entry, and contributing to the study design and development, and the analysis plans.

The Acknowledgments section contains information about the management and operations of TIMSS, the National Research Coordinators, and the TIMSS advisory committees.

## 1.7 Additional Resources

Although this User Guide is intended to provide secondary analysts with sufficient information to conduct analyses on the TIMSS data, some users may want additional information about TIMSS. Further documentation on the study design implementation, and analysis can be found in these publications:

- *TIMSS: Quality Assurance in Data Collection* (Martin and Mullis, 1996)
- *TIMSS Technical Report, Volume I: Design and Development* (Martin and Kelly, 1996)
- *TIMSS Technical Report, Volume II: Implementation and Analysis, Primary and Middle School Years* (Martin and Kelly, 1997)
- *TIMSS Technical Report, Volume III: Implementation and Analysis, Final Year of Secondary School* (Martin and Kelly, 1998)

---

# Chapter 2

## TIMSS Instruments and Booklet Design

TIMSS used several types of instruments to collect data about students and schools – achievement tests, student background questionnaires, and school background questionnaires. This chapter describes the design and content of each of these instruments.

### 2.1 The TIMSS Tests for Final-Year Students

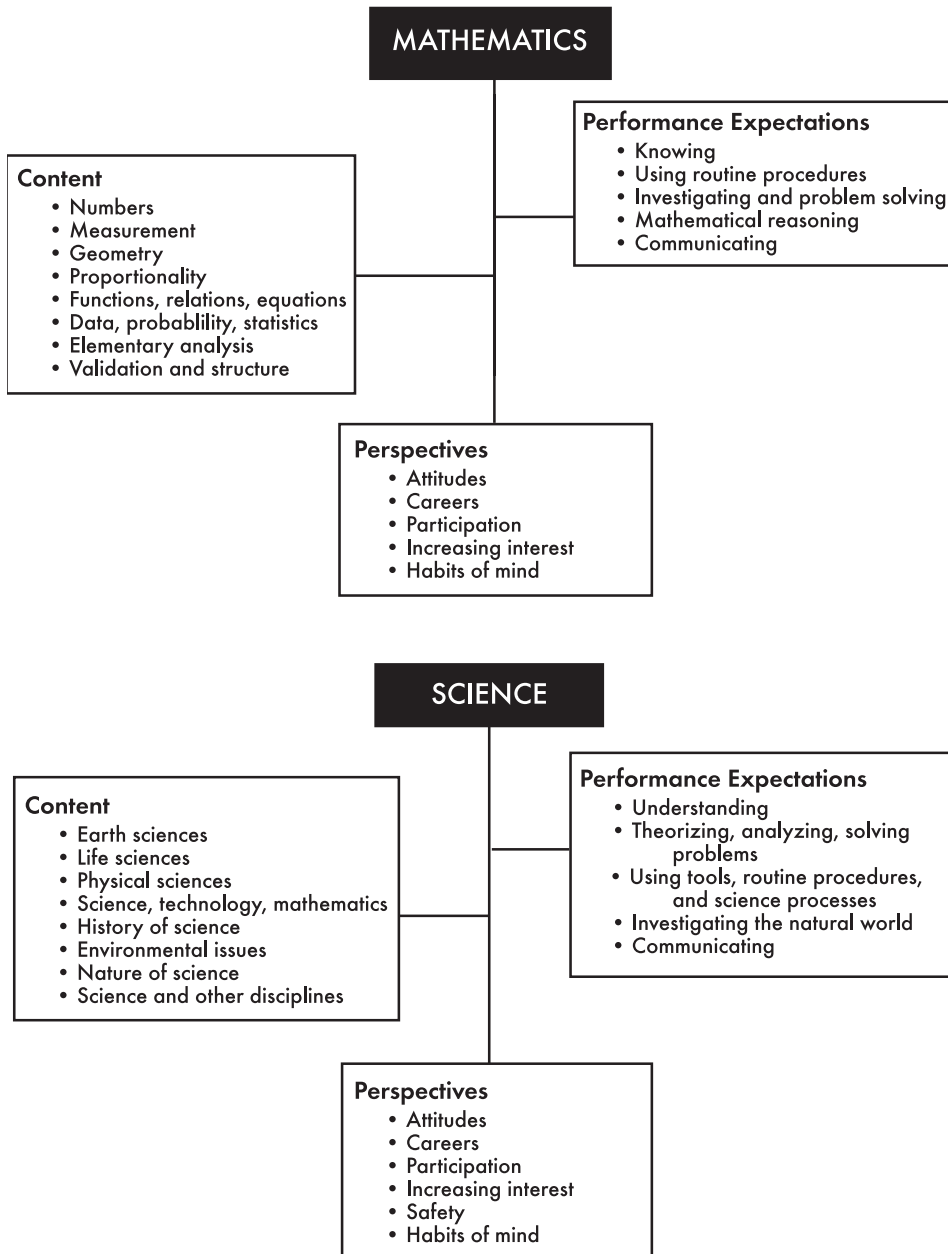
There were three separate tests for the three populations of students in their final year of secondary school tested by TIMSS: the mathematics and science literacy test for all students in their final year of secondary school, the advanced mathematics test for students having taken advanced mathematics; and the physics test for students having taken physics. The mathematics and science literacy test was designed to test students' general mathematics and scientific knowledge and understanding of mathematics and scientific principles. In order to examine how well students understand advanced mathematics concepts and can apply knowledge to solve problems, the advanced mathematics test was developed for students in their final year of secondary school having taken advanced mathematics. The physics test was developed for students in their final year of secondary school having taken physics, in order to examine how well they understand and can apply physics principles and concepts.

#### 2.1.1 Mathematics and Science Content

The TIMSS *Curriculum Frameworks for Mathematics and Science* (Robitaille et al., 1993) contain three dimensions – subject-matter content, performance expectations, and perspectives. *Subject-matter content* refers to the content of the mathematics or science test item under consideration. *Performance expectations* describe, in a non-hierarchical way, the kinds of performance or behavior that a given test item might elicit from students. The *perspectives* aspect focuses on the development of students' attitudes, interests, and motivations in mathematics and science.

As shown in Figure 2.1, each of the three aspects is partitioned into a number of categories. These major categories in turn were partitioned into subcategories specifying the content, performance expectations, and perspectives in more detail. For example, for each of the content categories there are up to six more specific subcategories.

**Figure 2.1**  
**The Major Categories of the TIMSS Curriculum Frameworks**



The two dimensions of the TIMSS frameworks used in developing the TIMSS tests were subject-matter content and performance expectations. During test development, each item was coded as to the content and performance expectations with which it is associated. The TIMSS item classification system permits an item to draw on multiple content areas and to involve more than one performance expectation, so that an item may have several content and performance codes. However, in constructing the tests, only the principal code was used for each of the two dimensions. For example, an item may be coded for content as “uncertainty and probability” (principal code) and “proportionality problem” (secondary code). When that item was selected for the test, only the principal code was considered.

Because of limitations in resources for data collection, a number of the detailed categories in the frameworks were combined into a few content categories for which scale scores were produced. In the analysis, each item in the TIMSS tests was included in one reporting category based on its principal content code. Table 2.1 presents the content categories for the mathematics and science literacy, advanced mathematics, and physics tests used in the international reports.

**Table 2.1**  
**Mathematics and Science Literacy, Advanced Mathematics, and Physics Content Categories – Final Year of Secondary School**

<b>Mathematics and Science Literacy Test</b>	Mathematics Literacy* Science Literacy*
<b>Advanced Mathematics Test</b>	Numbers and equations Calculus Geometry Probability and Statistics <sup>†</sup> Validation and Structure <sup>†</sup>
<b>Physics Test</b>	Mechanics Electricity and Magnetism Heat Wave Phenomena Modern Physics: Particle, Quantum and Astrophysics, and Relativity

\*The mathematics and science literacy test included items measuring reasoning and social utility; these are subsumed under the Mathematics Literacy and Science Literacy content categories.

<sup>†</sup>Probability and Statistics and Validation and Structure were not scaled separately. However, the overall Advanced Mathematics scale includes items in these categories.

### 2.1.2 The TIMSS Items

The task of putting together the achievement item pools for the different TIMSS tests was immense, and took more than three years to complete. Developing the TIMSS achievement tests necessitated building international consensus among NRCs, their national committees, mathematics and science experts, and measurement specialists. All NRCs worked to ensure that the items used in the tests were appropriate for their students and reflected their countries’ curricula.

Different types of achievement items were included in the item pools for TIMSS. The *multiple-choice* items consisted of a stem and either four or five answer choices. In the instructions at the front of the test booklets, students were encouraged to choose “the answer [they] think is best” when they were unsure. The instructions do not suggest or imply that students should guess if they do not know the answer. In the *free-response* items, students were asked to construct their own responses to the test questions by writing or drawing their answers. These included short-answer items and items where students were asked to provide extended responses. The free-response items were scored using the two-digit coding system developed by TIMSS (see Chapter 4).

For the mathematics and science literacy test, there were 76 items (44 mathematics literacy and 32 science literacy), including 52 multiple-choice items, 17 short-answer items, and 7 extended-response items. The distribution of items across the reporting categories is shown in Table 2.2.

For the advanced mathematics test, there were 65 items, including 47 multiple-choice items, 10 short-answer items, and 8 extended-response items. The distribution of items across the reporting categories is shown in Table 2.3.

For the physics test, there were 65 items, including 42 multiple-choice items, 15 short-answer items, and 8 extended-response items. The distribution of items across the reporting categories is shown in Table 2.4.

The classification of items into each mathematics and science literacy, advanced mathematics, and physics reporting category are shown in Tables 7.8, 7.9, and 7.10 in Chapter 7.

**Table 2.2**  
**Distribution of Mathematics and Science Literacy Items by Reporting Category**

Reporting Category	Percentage of Items	Number of Items	Number of Multiple-Choice Items	Number of Short-Answer Items	Number of Extended-Response Items	Number of Score Points <sup>1</sup>
Mathematics Literacy	58%	44	34	8	2	53
Science Literacy	42%	32	18	9	5	43
<b>Total</b>	100%	76	52	17	7	96

<sup>1</sup>In scoring the tests correct answers to most items were worth one point. However, responses to some constructed-response items were evaluated for partial credit with a fully correct answer awarded up to two or three points. In addition, some items had two parts. Thus, the number of score points exceeds the number of items in the test.

**Table 2.3**  
**Distribution of Advanced Mathematics Items by Content Category**

Content Category	Percentage of Items	Number of Items	Number of Multiple-Choice Items	Number of Short-Answer Items	Number of Extended-Response Items	Number of Score Points <sup>1</sup>
Numbers & Equations	26%	17	13	2	2	22
Calculus	23%	15	12	2	1	19
Geometry	35%	23	15	4	4	29
*Probability and Statistics	11%	7	5	2	0	8
*Validation and Structure	5%	3	2	0	1	4
<b>Total</b>	100%	65	47	10	8	82

<sup>1</sup>In scoring the tests correct answers to most items were worth one point. However, responses to some constructed-response items were evaluated for partial credit with a fully correct answer awarded up to two or three points. In addition, some items had two parts. Thus, the number of score points exceeds the number of items in the test.

\*Probability and Statistics and Validation and Structure were not scaled separately. However, the overall advanced mathematics scale includes those 10 items.

**Table 2.4**  
**Distribution of Physics Items by Content Category**

Content Category	Percentage of Items	Number of Items	Number of Multiple-Choice Items	Number of Short-Answer Items	Number of Extended-Response Items	Number of Score Points <sup>1</sup>
Mechanics	25%	16	11	4	1	19
Electricity and Magnetism	25%	16	10	3	3	21
Heat	14%	9	6	3	0	12
Wave Phenomena	15%	10	6	3	1	12
Modern Physics: Particle, Quantum and Astrophysics, and Relativity	22%	14	9	2	3	17
<b>Total</b>	100%	65	42	15	8	81

<sup>1</sup>In scoring the tests correct answers to most items were worth one point. However, responses to some constructed-response items were evaluated for partial credit with a fully correct answer awarded up to two points. In addition, some items had two parts. Thus, the number of score points exceeds the number of items in the test.

Because the percentages are rounded to the nearest whole number, the total may appear inconsistent.

### 2.1.3 TIMSS Test Design

The assessment of the final-year students was accomplished through a complex design that included four types of test booklets (nine booklets in total) that were distributed to students based on their academic preparation. The four types of test booklets below were intended to yield proficiency estimates in mathematics and science literacy, advanced mathematics, and physics:

- Two literacy booklets (booklets 1A and 1B) containing mathematics and science literacy items
- Three physics booklets (booklets 2A, 2B, and 2C) containing physics items only
- Three mathematics booklets (booklets 3A, 3B, and 3C) containing advanced mathematics items only
- One mathematics and science/advanced mathematics/physics booklet (booklet 4) containing items in physics, advanced mathematics, and mathematics and science literacy.



The TIMSS test design included 12 mutually exclusive clusters of items distributed among the four types of test booklets in a systematic fashion. The 12 clusters are labeled A through L. Each cluster could appear in more than one test booklet and, in a few cases, in different positions within the booklets. The items within a cluster always appear in the same order and position.<sup>1</sup>

Table 2.5 shows the assignment of clusters to test booklets, and the order in which the clusters appear in each booklet. Table 2.6 shows the number of items in each booklet by content category for each of the three tests.

**Table 2.5**  
**Assignment of Item Clusters to Test Booklets**

Cluster Label	Booklet*								
	1A	1B	2A	2B	2C	3A	3B	3C	4
Mathematics and Science Literacy	A	1	2						1
	B	2	1						
	C	3							
	D		3						
Physics	E		1	1	1				2
	F		2						
	G			2					
	H					2			
Advanced Mathematics	I					1	1	1	3
	J					2			
	K						2		
	L							2	

\*Number in cell indicates position of item cluster within the test booklet.

<sup>1</sup> The TIMSS test design is fully documented in Adams and Gonzalez (1996)

**Table 2.6**  
**Number of Test Items per Booklet by Content Category**

		Booklet								
		1A	1B	2A	2B	2C	3A	3B	3C	4
Mathematics and Science Literacy	Mathematics Literacy	32	31							6
	Science Literacy	27	24							6
Physics	Mechanics			6	8	6				2
	Electricity and Magnetism			7	7	8				3
	Heat			3	4	4				1
	Wave Phenomena			5	4	5				2
	Modern Physics			6	6	6				2
Advanced Mathematics	Numbers, Equations, and Functions						7	8	8	3
	Calculus						7	7	5	2
	Geometry						10	10	9	3
	Probability and Statistics						3	2	4	1
	Validation and Structure						2	1	2	1
<b>Total Number of Items</b>		59	55	27	29	29	29	28	28	32

To facilitate booklet rotation and ensure proper achievement estimates, students were classified as to their preparation in mathematics and physics. Each student was characterized as having taken advanced mathematics (M) or not (O), and as having taken physics (P) or not (O). This two-way classification yielded four mutually exclusive and exhaustive categories of students:

- OO Students having studied neither advanced mathematics nor physics
- OP Students having studied physics but not advanced mathematics
- MO Students having studied advanced mathematics but not physics
- MP Students having studied both advanced mathematics and physics

The nine test booklets were rotated among students based on this classification scheme (OO, OP, MO, MP), so that each student completed one 90-minute test booklet. Students classified as OO received either booklet 1A or 1B, the two booklets containing items related to mathematics and science literacy. Students classified as OP received either booklet 1A or 1B, or one of the three booklets containing physics material (2A, 2B, or 2C). Students classified as MO received either booklet 1A or 1B, or one of the three booklets containing advanced mathematics material (3A, 3B, or 3C). Students classified as MP also received one booklet, although in this case it could have been any one of the booklets (1A, 1B, 2A, 2B, 2C, 3A, 3B, 3C, 4).

### 2.1.4 Calculator Policy

Since calculator use by students in upper secondary school is very common in many countries, final-year students were given the option of using a calculator when completing the TIMSS test. At the end of the test booklet students completed a survey with questions about the frequency with which they used a calculator during the test and the type of calculator used. The frequency of calculator use on the TIMSS test is reported in the TIMSS international report for the final year of secondary school (Mullis et al., 1998) and responses to the calculator survey questions are included in the international database.

## 2.2 Questionnaires

The *student questionnaire* was a separate instrument from the test booklets. According to TIMSS test administration procedures, students were to be given a break after completing their test booklets. Subsequent to the break, test administrators were to distribute the student questionnaires. In many countries, the student questionnaires were distributed in a packet with the test booklets, not separately. Nevertheless, the quality assurance monitoring revealed no indication of any problems with the distribution of the student questionnaires (Martin and Mullis, 1996). The test administration script prescribed 20 minutes to administer the student questionnaire, but often more time was required (about 20 minutes extra).

The *student questionnaire* asked about students' demographics and home environment, including academic activities outside of school, people living in the home, parental education, books in the home, possessions in the home, and the importance students' mothers, peers, and friends placed on different aspects of education. Students also were queried about their attitudes towards mathematics and science and their future plans. The final sections of the questionnaires asked about programs of study, the most advanced mathematics and science courses taken, and classroom experiences in mathematics and science.

The *school questionnaire* sought information about the school's community, staff, students, curriculum and programs of study, and instructional resources and time. A school questionnaire was to be completed by the principal, headmaster, or other administrator of each school that participated in TIMSS.

Different from Populations 1 and 2, there was no teacher questionnaire administered to the teachers of students in their final year of secondary school.

## 2.3 Release Status for TIMSS Test Items and Background Questionnaires – Final-Year Assessment

Some TIMSS achievement test items have been released for unrestricted public use. At Population 3, all items in clusters A, D, G, H, K, and L are classified as public release and are available to secondary users (TIMSS, 1998). Information about how to obtain the released item sets is provided below. The student and school questionnaires are also classified as public release and are available to secondary users in Supplement 1 of this User Guide.

### TIMSS Released Item Set

The TIMSS released items for the assessment of students in their final year of secondary school are available in:

*TIMSS Released Item Set for the Final Year of Secondary School: Mathematics and Science Literacy, Advanced Mathematics, Physics*

To obtain copies of the TIMSS released item sets contact the International Study Center

TIMSS International Study Center  
Campion Hall Room 323 – CSTEPP  
Boston College  
Chestnut Hill, MA 02467  
United States  
Phone: +1-617-552-4521  
Fax: +1-617-552-8419  
e-mail: [timss@bc.edu](mailto:timss@bc.edu)

The TIMSS released items also are available on the World Wide Web at:  
[www.csteep.bc.edu/timss](http://www.csteep.bc.edu/timss)

---

# Chapter 3

## Sampling and Sampling Weights

This chapter describes the selection of school and student samples and the sampling weights included on the TIMSS data files. The TIMSS sample design is fully detailed in Foy, Rust, and Schleicher (1996) and Dumais (1998). The weighting procedures are described in Dumais and Foy (1998).

### 3.1 The Target Population

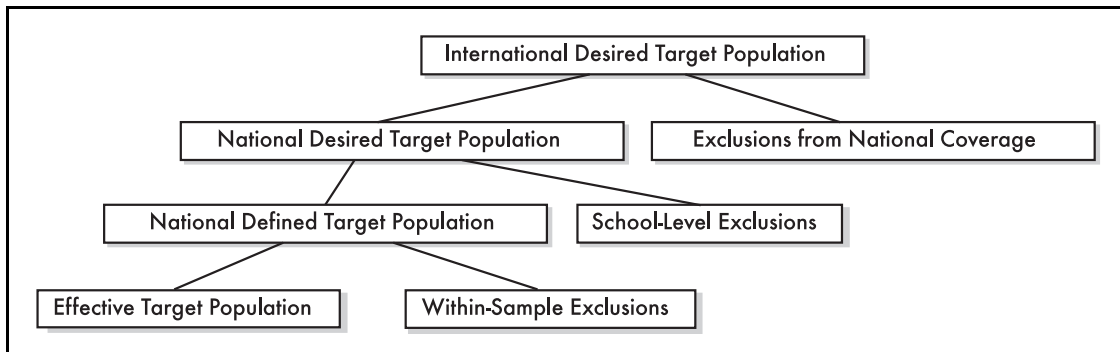
The selection of valid and efficient samples is crucial to the quality and success of an international comparative study such as TIMSS. The accuracy of the survey results depends on the quality of the available sampling information and of the sampling activities themselves. For TIMSS, NRCs worked on all phases of sampling with staff from Statistics Canada. NRCs were trained in how to select the school and student samples and in the use of the sampling software. In consultation with the TIMSS sampling referee (Keith Rust, Westat, Inc.), staff from Statistics Canada reviewed the national sampling plans, sampling data, sampling frames, and sample execution. This documentation was used by the International Study Center in consultation with Statistics Canada, the sampling referee, and the Technical Advisory Committee to evaluate the quality of the samples.

The intention of the assessment of final-year students was to measure what might be considered the “yield” of the elementary and secondary education systems of a country with regard to mathematics and science. This was accomplished by assessing the mathematics and science literacy of all students in the final year of secondary school, the advanced mathematics knowledge of students having taken advanced mathematics courses, and the physics knowledge of students having taken physics. The international desired population, then, was all students in the final year of secondary school, with students having taken advanced mathematics courses and students having taken physics courses as two overlapping sub-populations. Students repeating the final year were not part of the desired population. For each secondary education track in a country, the final grade of the track was identified as being part of the target population, allowing substantial coverage of students in their final year of schooling. For example, grade 10 could be the final year of a vocational program, and grade 12 the final year of an academic program. Both of these grade/track combinations are considered part of the population (but grade 10 in the academic track is not). Appendix A of Mullis et al. (1998) provides information about the structure of the upper secondary education systems and the students tested in each country.

### 3.2 Coverage of the TIMSS Population

The stated objective in TIMSS was that the effective population, the population actually sampled by TIMSS, be as close as possible to the International Desired Population. Figure 3.1 illustrates the relationship between the desired populations and the excluded populations at the country, school, and student levels.

**Figure 3.1**  
**Relationship Between the Desired Populations and Exclusions**



Using the International Desired Population as a basis, participating countries had to operationally define their population for sampling purposes. Some NRCs had to restrict coverage at the country level, for example, by excluding remote regions or a segment of the educational system. In these few situations, countries were permitted to define a national desired population that did not include part of the international desired population. Exclusions could be based on geographic areas or language groups. Table 3.1 shows differences in coverage between the international and national desired populations. Most participants achieved 100% coverage (20 out of 24). The countries with less than 100% coverage are footnoted in tables in the international report. Israel and Lithuania, as a matter of practicality, needed to define their tested populations according to the structure of their school systems. Latvia, which participated only in the physics assessment, also limited its testing to Latvian-speaking schools. Because coverage fell below 65%, the Latvian results have been labeled Latvia (LSS), for Latvian Speaking Schools, in the tables presenting results for the physics assessment. Italy was unable to include 4 of its 20 regions.

Within the national desired population, countries could define a population that excluded a small percentage (less than 10%) of certain kinds of schools or students that would be very difficult or resource-intensive to test (e.g., schools for students with special needs, or schools that were very small or located in extremely remote areas). Some countries also excluded students in particular tracks or school types. These exclusions are also shown in Table 3.1. The countries with particularly high exclusions are so footnoted in the achievement tables in the report.

**Table 3.1**  
**Coverage of TIMSS Target Population**

The International Desired Population is defined as follows:

Population 3 - All students in final year of secondary school

Country	International Desired Population		National Desired Population	
	Country Coverage	Notes on Coverage	Sample Exclusions	Notes on Exclusions
Australia	100%		5.5%	
Austria	100%		18.2%	Colleges and courses lasting less than 3 years excluded
Canada	100%		8.9%	
Cyprus	100%		22.0%	Private and vocational schools excluded
Czech Republic	100%		6.0%	
Denmark	100%		2.3%	
France	100%		1.0%	
Germany	100%		11.3%	
Greece	100%		85.0%	Only students having taken advanced mathematics and physics included
Hungary	100%		0.2%	
Iceland	100%		0.1%	
Israel	74%	Hebrew public education system	0.0%	
Italy	70%	Four regions did not participate	0.9%	
Latvia (LSS)	50%	Latvian speaking students	85.0%	Only students having taken physics included
Lithuania	84%	Lithuanian speaking students	0.0%	
Netherlands	100%		21.6%	Apprenticeship programs excluded
New Zealand	100%		0.0%	
Norway	100%		3.8%	
Russian Federation	100%		43.0%	Vocational schools and non-Russian speaking students excluded
Slovenia	100%		6.0%	
South Africa	100%		0.0%	
Sweden	100%		0.2%	
Switzerland	100%		2.5%	
United States	100%		3.7%	

Because population coverage falls below 65%, Latvia is annotated LSS for Latvian Speaking Schools only.

Participants could exclude schools from the sampling frame if they were in geographically remote regions, were extremely small, offered curriculum or structure different from the mainstream, or provided instruction only to students in the “within-school” exclusion categories. The general TIMSS rules for defining within-school exclusions follow.

- **Educable mentally disabled students.** These are students who are considered, in the professional opinion of the school principal or other qualified staff members, to be educable mentally disabled students, or who have been so diagnosed in psychological tests. This includes students who are emotionally or mentally unable to follow even the general instructions of the TIMSS test. It does not include students who merely exhibit poor academic performance or discipline problems.
- **Functionally disabled students.** These are students who are permanently physically disabled in such a way that they could not perform in the TIMSS tests.

Functionally disabled students who could perform in the TIMSS test were included in the testing.

- **Non-native-language speakers.** These are students who cannot read or speak the language of the test and so could not overcome the language barrier of testing. Typically, students who had received less than one year of instruction in the language of the test were excluded, but this definition was adapted in different countries. Some countries opted to test students in more than one language.

### 3.3 TIMSS Coverage Index

Historically, an important difference between education systems was the proportion of an age cohort that successfully completed upper secondary education. In order to avoid unwittingly comparing the elite students in one country with the more general population in another, therefore, it is important to be aware of the extent to which the upper secondary system in each country includes the total student population.

So as to learn how much of the school-leaving age cohort was still in school and represented by the TIMSS sample, a TIMSS Coverage Index (TCI) was computed for each country. The TCI is an estimate of the percentage of the school-leaving age cohort covered by the TIMSS final-year student sample. It reflects any omissions from the sample, such as students who were excluded because of handicap or who had dropped out of school, and, in some countries, tracks or educational programs that were not covered by the TIMSS sample. The TCI was computed by forming a ratio of the size of the student population covered by the TIMSS sample, as estimated from the sample itself, to the size of the school-leaving age cohort, which was derived from official population census figures supplied by each country. The TCI was defined as follows:

$$TCI = \frac{\text{Total Enrollment in TIMSS Grades 1995}}{(\text{Total National Population Aged 15-19 in 1995}) / 5}$$

The *numerator* in this expression is the total enrollment in the grades tested by TIMSS, estimated from the weighted sample data. This estimate corresponds to the size of the population to which the TIMSS results generalize, and makes appropriate provision for student non-response. It does not include students who are no longer attending school, or students who were excluded from the sample on grounds of physical or other disability. It also does not include students who were repeating the final grade.

The *denominator* in the expression is an estimate of the school-leaving age cohort size. Since the age at which students in upper secondary may leave school varies, TIMSS estimated the size of the school-leaving age cohort by taking the average of the size of the 1995 age cohorts for 15, 16, 17, 18, and 19-year-olds in each country. (Although the general procedure was to base the estimate on the 15-19 age group, there were exceptions in some countries. For example, in Germany, the estimate was based on the 17-19 age group.) This information was provided by National Research Coordinators from official population census figures in their countries. This approach reflects the fact that students in the final



year of secondary school are likely to be almost entirely a subset of the population of 15- to 19-year-olds in most countries.

Countries with high TCIs have most of their students still in school, and have covered this population with their TIMSS sample. Countries with low TCIs have fewer students still in school, or have excluded some components of their system from their sample (or both). Table 3.2 presents the TCI for each country, and also shows the two parts of the portion of the school-leaving age cohort not covered by the TIMSS sample: system components and students excluded by the country, and others – primarily young people who chose not to complete upper secondary education. The percentage of the age cohort covered by the TIMSS sample (the TCI), the percentage excluded from the sample, and the percentage of others not covered combine to form 100% of the school-leaving age cohort. For example, Australia has a TCI of 68.1%, which indicates that the TIMSS sample of final-year students covers just over two-thirds of the school-leaving age cohort. Of the remainder, 4% have been excluded from the sample, and the remaining 27.9% are presumably no longer attending school. The TCI for Cyprus is lower (47.9%), partly because Cyprus excluded students in private schools and in vocational programs (13.5%), and partly because a greater percentage of the age cohort is no longer attending school (38.6%).

**Table 3.2**  
**TIMSS Coverage Indices (TCIs)**

Country	TIMSS Coverage Index (TCI)*	Sample Exclusions <sup>†</sup>	Others Not Covered	Notes on Exclusions
Australia	68.1%	4.0%	27.9%	Colleges and courses lasting less than 3 years excluded
Austria	75.9%	16.8%	7.3%	
Canada	70.3%	6.8%	22.9%	
Cyprus	47.9%	13.5%	38.6%	
Czech Republic	77.6%	5.0%	17.4%	
Denmark	57.7%	1.3%	41.0%	Private and vocational schools excluded
France	83.9%	0.9%	15.3%	
Germany	75.3%	9.6%	-	
<sup>1</sup> Greece	10.0%	56.8%	33.2%	
Hungary	65.3%	0.1%	34.6%	
Iceland	54.5%	0.0%	45.4%	Only students having taken advanced mathematics and physics included
<sup>2</sup> Israel	-	-	-	
Italy	51.5%	0.5%	48.0%	
<sup>1</sup> Latvia	3.0%	16.8%	80.3%	
Lithuania	42.5%	0.0%	57.5%	
Netherlands	78.0%	21.5%	0.5%	Only students having taken physics included
New Zealand	70.5%	0.0%	29.5%	
Norway	84.0%	3.3%	12.7%	
Russian Federation	48.1%	36.3%	15.7%	
Slovenia	87.8%	5.6%	6.6%	
South Africa	48.9%	0.0%	51.1%	Apprenticeship programs excluded
Sweden	70.6%	0.2%	29.2%	
Switzerland	81.9%	2.1%	16.0%	
United States	63.1%	2.5%	34.5%	

\*TIMSS Coverage Index (TCI): Estimated percentage of school-leaving age cohort covered by TIMSS sample.

<sup>†</sup>Percentage different from that reported in Table 3.1 because this percentage is based on the entire school-leaving age cohort rather than the population of those students attending school.

<sup>1</sup>Results for Greece are reported only for advanced mathematics and physics; results for Latvia are available only for physics.

<sup>2</sup>The TCI could not be computed for Israel.

In order to quantify the coverage of the advanced mathematics and physics samples and help interpret the achievement results for these students, TIMSS computed a Mathematics TIMSS Coverage Index (MTCI) and a Physics TIMSS Coverage Index (PTCI), as shown in Table 3.3. The MTCI is the overall TCI multiplied by the percentage of the final-year sample having taken advanced mathematics. For example, in Australia 23.1% of the final-year sample had taken advanced mathematics. Multiplying this by the TCI (68.1%, from Table 3.2) gives a MTCI of 15.7%, as shown in the second column of Table 3.3. This implies that about 16% of the school-leaving age cohort in Australia had taken advanced mathematics in upper secondary school. Similarly, the PTCI for Australia is 12.6%, as shown in the fourth column of Table 3.3.

**Table 3.3**  
**TIMSS Coverage Indices (TCIs) for Advanced Mathematics and Physics**

Country	Percentage of Students in Sample Having Taken Advanced Mathematics	Mathematics TIMSS Coverage Index (MTCI)*	Percentage of Students in Sample Having Taken Physics	Physics TIMSS Coverage Index (PTCI) <sup>†</sup>
Australia	23.1%	15.7%	18.5%	12.6%
Austria	43.9%	33.3%	43.5%	33.1%
Canada	22.3%	15.6%	19.4%	13.7%
Cyprus	18.5%	8.8%	18.5%	8.8%
Czech Republic	14.1%	11.0%	14.1%	11.0%
Denmark	35.7%	20.6%	5.5%	3.2%
France	23.8%	19.9%	23.8%	19.9%
Germany	34.9%	26.3%	11.2%	8.4%
<sup>1</sup> Greece	-	10.0%	-	10.0%
<sup>2</sup> Israel	-	-	-	-
Italy	27.4%	14.1%	16.7%	8.6%
<sup>3</sup> Latvia	-	-	-	3.0%
Lithuania	6.1%	2.6%	-	-
Norway	-	-	10.0%	8.4%
Russian Federation	4.2%	2.0%	3.2%	1.5%
Slovenia	85.9%	75.4%	43.9%	38.6%
Sweden	23.0%	16.2%	23.1%	16.3%
Switzerland	17.4%	14.3%	17.3%	14.2%
United States	21.8%	13.7%	22.9%	14.5%

\*MTCI: Estimated percentage of school-leaving age cohort covered by TIMSS sample of advanced mathematics students.

<sup>†</sup>PTCI: Estimated percentage of school-leaving age cohort covered by TIMSS sample of physics students.

<sup>1</sup>Greece sampled only students having taken advanced mathematics and physics.

<sup>2</sup>The MTCI and the PTCI could not be computed for Israel.

<sup>3</sup>Latvia sampled only students having taken physics.

Note: Hungary, Iceland, the Netherlands, New Zealand, and South Africa did not participate in the advanced mathematics and physics testing.

Norway did not participate in the advanced mathematics testing and Lithuania did not participate in the physics testing.

### 3.4 Sample Design

In order to implement the TIMSS goal of assessing the mathematics and science literacy of all students while also assessing the advanced mathematics and physics knowledge of students with preparation in these subjects, it was necessary to develop a sampling design which ensured that students were stratified according to their level of preparation in mathematics and physics, so that appropriate test booklets could be assigned to them. According to the TIMSS design each student is characterized as having taken advanced mathematics (M) or not (O), and as having taken physics (P) or not (O). Combining these two-way classifications yields four mutually exclusive and exhaustive categories of students:

- OO Students having studied neither advanced mathematics nor physics
- OP Students having studied physics but not advanced mathematics
- MO Students having studied advanced mathematics but not physics
- MP Students having studied both advanced mathematics and physics

There were four kinds of student test booklets, and these were assigned to students on the basis of this classification scheme (OO, OP, MO, MP), so that each student completed one 90-minute test booklet. Students classified as OO received either booklet 1A or 1B, the two booklets containing items related to mathematics and science literacy. Students classified as OP received either booklet 1A or 1B, or one of the three booklets containing physics material (2A, 2B, or 2C). Students classified as MO received either booklet 1A or 1B, or one of the three booklets containing advanced mathematics material (3A, 3B, or 3C). Students classified as MP also received one booklet, although in this case it could have been any one of the booklets (1A, 1B, 2A, 2B, 2C, 3A, 3B, 3C or 4). Booklet 4 contained mathematics and science literacy, advanced mathematics, and physics items. The TIMSS test design is summarized in Chapter 2 of this User Guide and is described in detail in Adams and Gonzalez (1996).

**Table 3.4**  
**Assignment of Test Booklet According to Student Classification**

Student Type	Booklet Assigned
OO	1A or 1B
OP	1A, 1B, 2A, 2B or 2C
MO	1A, 1B, 3A, 3B or 3C
MP	1A, 1B, 2A, 2B, 2C, 3A, 3B, 3C or 4

The student samples used for estimating the parameters of the three populations of interest (all students in their final year of secondary school, final-year students having taken advanced mathematics, and final-year students having taken physics) were constructed by combining the students who had been assigned the appropriate booklets from each of the four groups of students (OO, OP, MO, and MP). For example, the student sample for estimating proficiency in mathematics and science literacy among all final-year students was made up of the students from each of the four student groups who were assigned one of the literacy booklets (booklets 1A or 1B) or the combined literacy, mathematics, and physics booklet (booklet 4). This ensured that each of the four types of students were properly represented in the final-year sample. The sample for estimating proficiency in advanced mathematics consisted of students from the MO (students having taken advanced mathematics but not physics) and MP (students having taken both advanced mathematics and physics) groups who were assigned one of the mathematics booklets (booklets 3A, 3B, or 3C) or the combined literacy, advanced mathematics, and physics booklet (booklet 4). Similarly, the sample for estimating proficiency in physics consisted of students from the OP (students having taken physics but not advanced mathematics) and MP (students having taken both advanced mathematics and physics) groups who were assigned one of the physics booklets (booklets 2A, 2B, or 2C) or the combined literacy, advanced mathematics, and physics booklet (booklet 4).

The basic sampling design was a two-stage process, with schools sampled with probability proportional to size (PPS) in the first stage, and a fixed number of students sampled in the second stage. To meet the TIMSS standard for sampling precision (see Dumais, 1998),

approximately 120 schools were required in each country. Within each sampled school, students were classified according to the four-group classification scheme (OO, OP, MO, MP), and the 40 students sampled at random, 10 from each of the four categories. Because the organization of school systems at upper secondary level varies considerably across countries, it was necessary for each country to work with the sampling consultants to customize the basic design for the local situation.

The basic design was well suited to comprehensive systems, where schools cater to all kinds of students, and students must be classified individually or on the basis of the courses they have taken. However, many of the TIMSS countries operate tracked systems, where students are assigned to particular types of schools on the basis of their academic interests and abilities. In such countries it was often possible to stratify whole schools in terms of whether or not they contained advanced mathematics or physics classes. In such systems it was sometimes possible to refine the basic design so as to achieve the required sampling precision with a smaller sample of schools. Although there was no analytic requirement to sample students in whole classes rather than individually, some countries found it more convenient to sample classes, even though the increased clustering effect sometimes necessitated larger sample sizes.

### 3.5 School Sampling

The sample-selection method used for first-stage sampling was based on a systematic probability-proportional-to-size technique. Countries were encouraged to stratify schools by important demographic variables (e.g., geographical region, public/private, etc.) as well as by school type. Small schools were handled either by assigning them to separate strata or by combining them with larger schools to form pseudo-schools for sampling purposes. In some very large countries, there was a preliminary sampling stage before schools were sampled, in which the country was divided into primary sampling units. Within each stratum, schools were listed in order of any implicit stratification variables, and then further sorted according to their measure of size (MOS). Schools were then sampled using a random-start fixed-interval procedure which ensured selection with probability proportional to size.

It was sometimes the case that a sampled school was unable to participate in the assessment. In such cases, this originally sampled school was replaced by a *replacement school*. The mechanism for selecting replacement schools, established a priori, identified the next school on the ordered school-sampling list as the replacement for each particular sampled school. The school after that was a second replacement, should it be necessary. Since schools were grouped by stratification variables and by size on the sampling frame, replacement schools should have characteristics similar to the originally selected school.

### 3.6 Student Sampling

Whereas schools were sampled with probability proportional to size, a fixed number of students was sampled within each school in the second stage of sampling. This gives selection probabilities for students which are inversely proportional to school size. The combined school and student selection probabilities result in an overall selection probability which is equal for all students in each explicit stratum. In untracked schools, students were classified according to their academic preparation into one of the four groups (OO, OP, MO, MP), and

a sample of 10 students was drawn from each group. If just three student types were present (for example if there were no OP students, as sometimes happened) three samples of 13 students were drawn. In schools with no advanced mathematics or physics students, all 40 students were sampled from the OO group. In some tracked systems, schools frequently consisted either of only OO students or of only MP students. In these situations all 40 students were sampled from whichever group was appropriate.

### 3.7 Response Rates

Weighted and unweighted response rates were computed for each participating country by grade, at the school level, and at the student level. Overall response rates (combined school and student response rates) also were computed.

#### 3.7.1 School-Level Response Rates

The minimum acceptable school-level response rate, before the use of replacement schools, was set at 85%. This criterion was applied to the unweighted school-level response rate. Both weighted and unweighted school-level response rates were reported, with and without replacement schools. The general formula for computing weighted school-level response rates is shown in the following equation:

$$R_{wgt}(sch) = \frac{\sum_{part} MOS_i / \pi_i}{\sum_{elig} MOS_i / \pi_i}$$

For each sampled school, the ratio of its measure of size (MOS) to its selection probability ( $\pi_i$ ) was computed. The weighted school-level response rate is the sum of the ratios for all participating schools divided by the sum of the ratios for all eligible schools. The unweighted school-level response rates were computed in a similar way, where all school ratios were set to one. This becomes simply the number of participating schools in the sample divided by the number of eligible schools in the sample. Since in most cases, in selecting the sample, the value of  $\pi_i$  was set proportional to  $MOS_i$  within each explicit stratum, it was generally the case that weighted and unweighted rates were similar.

#### 3.7.2 Student-Level Response Rates

Like the school-level response rate, the minimum acceptable student-level response rate was set at 85%. This criterion was applied to the unweighted student-level response rate. Both weighted and unweighted student-level response rates were reported. The general formula for computing student-level response rates is shown in the following equation:

$$R_{wgt}(stu) = \frac{\sum_{part} 1/p_j}{\sum_{elig} 1/p_j}$$

where  $p_j$  denotes the probability of selection of the student, incorporating all stages of selection. Thus the weighted student-level response rate is the sum of the inverse of the selection probabilities for all participating students divided by the sum of the inverse of the selection probabilities for all eligible students. The unweighted student response rates were computed in a similar way, but with each student contributing equal weight.

### 3.7.3 Overall Response Rates

The minimum acceptable overall response rate (combined school and student response rates) was set at 75%. This overall response rate for each grade was calculated as the product of the weighted school-level response rate at the grade without replacement schools and the weighted student-level response rate at the grade. School and student response rates are presented in Tables 3.5 to 3.7.

**Table 3.5**  
**Participation Rates – Mathematics and Science Literacy**

Country	School Participation		Student Participation (Weighted Percentage)	Overall Participation	
	School Participation Before Replacement (Weighted Percentage)	School Participation After Replacement (Weighted Percentage)		Overall Participation Before Replacement (Weighted Percentage)	Overall Participation After Replacement (Weighted Percentage)
Australia	48.8	66.2	78.1	38.1	51.8
Austria	35.9	90.9	79.7	28.6	72.5
Canada	82.2	82.6	82.7	68.0	68.3
Cyprus	100.0	100.0	98.2	98.2	98.2
Czech Republic	100.0	100.0	92.2	92.2	92.2
Denmark	54.9	54.9	88.9	48.8	48.8
France	80.3	80.3	85.6	68.7	68.7
Germany	88.7	100.0	80.1	71.0	80.1
Hungary	100.0	100.0	97.7	97.7	97.7
Iceland	100.0	100.0	73.6	73.6	73.6
Israel	48.8 **	48.8 **	98.3 **	48.0 **	48.0 **
Italy	59.9	65.0	94.8	56.8	61.6
Lithuania	97.1	97.1	87.9	85.4	85.4
Netherlands	35.8	56.3	87.6	31.3	49.3
New Zealand	87.0	100.0	80.6	70.1	80.6
Norway	74.1	80.0	88.9	65.9	71.1
Russian Federation	93.0	99.3	90.9	84.6	90.3
Slovenia	45.6	45.6	92.8	42.3	42.3
South Africa	65.0	65.0	99.4	64.6	64.6
Sweden	95.3	95.3	86.5	82.4	82.4
Switzerland	87.0	89.1	95.0	82.6	84.6
United States	77.1	85.1	74.6	57.6	63.5

\*\*Unweighted participation rates.

**Table 3.6**  
**Participation Rates – Advanced Mathematics**

Country	School Participation		Student Participation (Weighted Percentage)	Overall Participation	
	School Participation Before Replacement (Weighted Percentage)	School Participation After Replacement (Weighted Percentage)		Overall Participation Before Replacement (Weighted Percentage)	Overall Participation After Replacement (Weighted Percentage)
Australia	47.3	63.6	86.7	40.9	55.2
Austria	36.7	95.5	84.6	31.0	80.8
Canada	84.6	85.2	90.4	76.4	76.9
Cyprus	100.0	100.0	96.0	96.0	96.0
Czech Republic	100.0	100.0	92.1	92.1	92.1
Denmark	54.9	54.9	89.2	49.0	49.0
France	89.9	89.9	86.1	77.4	77.4
Germany	78.6	100.0	77.6	61.0	77.6
Greece	76.2	100.0	86.5	65.9	86.5
Israel	48.8 **	48.8 **	99.6 **	48.6 **	48.6 **
Italy	70.3	70.9	95.1	66.9	67.5
Lithuania	100.0	100.0	92.1	92.1	92.1
Russian Federation	97.6	99.4	96.5	94.2	95.9
Slovenia	45.6	45.6	93.0	42.4	42.4
Sweden	95.3	95.3	92.9	88.6	88.6
Switzerland	99.0	99.0	88.2	87.4	87.4
United States	75.7	84.7	79.6	60.2	67.4

\*\*Unweighted participation rates.



**Table 3.7**  
**Participation Rates – Physics**

Country	School Participation		Student Participation (Weighted Percentage)	Overall Participation	
	School Participation Before Replacement (Weighted Percentage)	School Participation After Replacement (Weighted Percentage)		Overall Participation Before Replacement (Weighted Percentage)	Overall Participation After Replacement (Weighted Percentage)
Australia	63.2	63.9	84.9	53.7	54.2
Austria	36.7	95.5	84.6	31.0	80.8
Canada	79.7	80.2	91.0	72.6	73.0
Cyprus	100.0	100.0	96.0	96.0	96.0
Czech Republic	100.0	100.0	92.1	92.1	92.1
Denmark	54.9	54.9	86.1	47.3	47.3
France	89.9	89.9	86.1	77.4	77.4
Germany	76.8	100.0	81.7	62.7	81.7
Greece	76.2	100.0	86.5	65.9	86.5
Israel	48.8 **	48.8 **	99.6 **	48.6 **	48.6 **
Italy	69.3	69.3	96.6	67.0	67.0
Latvia (LSS)	84.4	84.4	90.8	76.6	76.6
Norway	77.7	94.3	88.0	68.4	83.0
Russian Federation	97.6	98.8	96.2	93.9	95.1
Slovenia	45.6	45.6	94.2	43.0	43.0
Sweden	95.3	95.3	92.9	88.6	88.6
Switzerland	99.0	99.0	88.2	87.4	87.4
United States	77.0	84.3	80.3	61.8	67.7

\*\*Unweighted participation rates.

Because population coverage falls below 65%, Latvia is annotated LSS for Latvian Speaking Schools only.

### 3.8 Compliance with Sampling Guidelines

Figures 3.2, 3.3, and 3.4 show how countries have been grouped in tables reporting achievement results for literacy, advanced mathematics, and physics, respectively. Countries that complied with the TIMSS guidelines for school and student sampling, and that achieved acceptable participation rates – 85% of both the schools and students, or a combined rate (the product of school and student participation) of 75% – with or without replacement schools, are shown in the first panel. Countries that met the guidelines only after including replacement schools are so noted.

Countries that did not reach at least 50% school participation without the use of replacements schools, or that failed to reach the sampling participation standard even with their use, are shown in the second panel of Figures 3.2-3.4. Countries that did not meet the guidelines for student sampling are shown in the third panel, and countries that met neither these requirements nor participation rate requirements are shown in the bottom panel. Unweighted<sup>1</sup> results only are included for Israel because Israel had difficulties meeting several sampling guidelines.

<sup>1</sup> This is effectively implemented by assigning a weight of 1 to all students in the sample for Israel.

**Figure 3.2**  
**Countries Grouped for Reporting Achievement According to Their Compliance with Guidelines for Sample Implementation and Participation Rates Mathematics and Science Literacy**

Final Year of Secondary School	
<b>Countries satisfying guidelines for sample participation rates and sampling procedures</b>	
<sup>2</sup> Cyprus	<sup>†</sup> New Zealand
Czech Republic	<sup>2</sup> Russian Federation
Hungary	Sweden
<sup>1</sup> Lithuania	Switzerland
<b>Countries not satisfying guidelines for sample participation rates</b>	
Australia	Iceland
<sup>2</sup> Austria	<sup>1</sup> Italy
Canada	Norway
France	United States
<b>Countries with unapproved student sampling</b>	
<sup>†</sup> Germany	
<b>Countries with unapproved sampling procedures and low participation rates</b>	
Denmark	Slovenia
<sup>2</sup> Netherlands	South Africa

<sup>†</sup> Met guidelines for sample participation rates only after replacement schools were included.

<sup>1</sup> National Desired Population does not cover all of International Desired Population.

<sup>2</sup> National Defined Population covers less than 90 percent of National Desired Population.

**Figure 3.3**  
**Countries Grouped for Reporting Achievement According to Their Compliance with Guidelines for Sample Implementation and Participation Rates Advanced Mathematics**

Final Year of Secondary School	
<b>Countries satisfying guidelines for sample participation rates and sampling procedures</b>	
Canada	<sup>†</sup> Greece
<sup>2</sup> Cyprus	<sup>†</sup> Lithuania
Czech Republic	<sup>2</sup> Russian Federation
France	Sweden
<sup>†</sup> Germany	Switzerland
<b>Countries not satisfying guidelines for sample participation rates</b>	
Australia	<sup>†</sup> Italy
<sup>2</sup> Austria	United States
<b>Countries with unapproved sampling procedures and low participation rates</b>	
Denmark	Slovenia

<sup>†</sup> Met guidelines for sample participation rates only after replacement schools were included.

<sup>†</sup> National Desired Population does not cover all of International Desired Population.

<sup>2</sup> National Defined Population covers less than 90 percent of National Desired Population.

**Figure 3.4**  
**Countries Grouped for Reporting Achievement According to Their Compliance with Guidelines for Sample Implementation and Participation Rates**  
**Physics**

Final Year of Secondary School	
<b>Countries satisfying guidelines for sample participation rates and sampling procedures</b>	
Canada	<sup>1</sup> Latvia (LSS)
<sup>2</sup> Cyprus	<sup>†</sup> Norway
Czech Republic	<sup>2</sup> Russian Federation
France	Sweden
<sup>†</sup> Germany	Switzerland
<sup>†</sup> Greece	
<b>Countries not satisfying guidelines for sample participation rates</b>	
Australia	United States
<sup>2</sup> Austria	
<b>Countries with unapproved sampling procedures and low participation rates</b>	
Denmark	Slovenia

<sup>1</sup> Met guidelines for sample participation rates only after replacement schools were included.

<sup>†</sup> National Desired Population does not cover all of International Desired Population.

<sup>2</sup> National Defined Population covers less than 90 percent of National Desired Population.

### 3.9 Sampling Weights

Appropriate estimation of population characteristics based on the TIMSS samples requires that the TIMSS sample design be taken into account in all analyses. This is accomplished in part by assigning a weight to each respondent, where the sampling weight properly accounts for the sample design, takes into account any stratification or disproportional sampling of subgroups, and includes adjustments for non-response.<sup>2</sup>

The students within each country were selected using probability sampling. A consequence of this is that each student had a known probability of selection. The inverse of this selection probability is the sampling weight. In a properly selected and weighted sample, the sum of the weights for the sample approximates the size of the population. In TIMSS, the sum of the sampling weights for a country sample is an estimate of the size of the population of students within the country in the sampled grade(s). The sampling weights must be used whenever population estimates are required. The use of the appropriate sampling weights ensures that

<sup>2</sup> Sampling weights can only be computed when the probability of selection is known for all students.

the different subgroups that constitute the sample are properly and proportionally represented in the computation of population estimates.

Tables 3.8, 3.9, and 3.10 present the sample sizes and the estimate of the population size (sum of the weights) for the entire final-year sample, for advanced mathematics students, and for physics students, respectively, for each participating country.

**Table 3.8**  
**Sample Information for Students in their Final Year of Secondary School**

Country	Sample Size	Sum of Weights
Austria	1962	70,602.31
Australia	1941	170,846.70
Canada	5232	263,241.25
Cyprus	534	4,556.35
Czech Republic	2167	137,458.53
Denmark	2714	37,871.61
France	1590	637,935.42
Germany	2289	967,704.92
Hungary	5091	111,281.35
Iceland	1703	2,308.37
Israel	1357	1,357.00
Italy	1616	380,834.01
Lithuania	2887	22,160.53
Netherlands	1470	145,916.30
New Zealand	1763	37,548.58
Norway	2518	43,805.67
Russian Federation	2289	1,031,187.00
Slovenia	1622	26,643.83
South Africa	2757	374,618.41
Sweden	3068	71,243.28
Switzerland	3308	65,139.87
United States	5807	2,278,258.19

**Table 3.9**  
**Sample Information for Students in their Final Year of Secondary School Having Taken Advanced Mathematics**

Country	Sample Size	Sum of Weights
Australia	645	39,497.77
Austria	782	31,062.96
Canada	2781	58,605.69
Cyprus	391	837.10
Czech Republic	1101	19,446.23
Denmark	1388	13,526.75
France	1071	151,531.25
Germany	2296	262,789.31
Greece	456	14,620.21
Israel	953	953.00
Italy	398	104,476.74
Lithuania	734	1,360.12
Russian Federation	1638	42,858.20
Slovenia	1536	22,881.47
Sweden	1001	16,408.31
Switzerland	1404	11,343.08
United States	2785	496,852.18

**Table 3.10**  
**Sample Information for Students in their Final Year of Secondary School Having Taken Physics**

Country	Sample Size	Sum of Weights
Australia	661	31,619.49
Austria	777	30,794.92
Canada	2367	51,179.36
Cyprus	368	837.01
Czech Republic	1087	19,428.12
Denmark	654	2,073.34
France	1110	151,531.25
Germany	723	87,887.92
Greece	459	14,667.71
Israel	853	853.00
Italy	145	63,706.57
Latvia (LSS)	708	979.34
Norway	1048	4,368.70
Russian Federation	1233	32,975.43
Slovenia	747	11,706.46
Sweden	1012	16,458.54
Switzerland	1371	11,275.73
United States	3114	522,783.82

### 3.10 Sampling and Weight Variables Included in the Student Data Files<sup>3</sup>

There are several sampling weight variables included in the student data files. Some of these variables capture information from different stages of the sampling process, and others constitute the sampling weights themselves. The variables named in this section are included in the Student Background and Written Assessment datafiles. The weighting factors and their adjustment factors included in the student-level data files are as follows:

#### ***WGTFAC1 School Weighting Factor***

This variable corresponds to the inverse of the probability of selection for the school where the student is enrolled.

#### ***WGTADJ1 School Weighting Adjustment Factor***

This is an adjustment that is applied to WGTFAC1 to account for non-participating schools in the sample. If we were to multiply WGTFAC1 by WGTADJ1 we would obtain the sampling weight for the school, adjusted for non-participation.

#### ***WGTFAC2 Class Weighting Factor***

Although the sampling procedures called for the selection of students directly within schools, this was not always possible. Some countries sampled classrooms within schools, and then the students within these classrooms. This variable reflects the inverse of the probability of selection of the classroom within the school. In countries where students were selected directly within the schools, as the sampling procedures specified, or in schools where there was only one classroom from which to sample students, this variable takes the value of 1. In countries where a classroom was sampled prior to sampling the students, this variable takes values greater than 1.

#### ***WGTFAC3 Student Weighting Factor***

This is the inverse of the probability of selection for the individual student within a school, or classroom, where intact classes were selected. In cases where an intact classroom was selected, the value is set to 1 for all members of the classroom.

#### ***WGTADJ3 Student Weighting Adjustment Factor***

This is an adjustment applied to the variable WGTFAC3 to account for non-participating students in the selected school and/or classroom. If we were to multiply the variables WGTFAC2, WGTFAC3, and WGTADJ3 and sum them up within each school, we would obtain an estimate of the number of students within the sampled school.

The five variables listed above are all used in computing a student's basic sampling weight. The probability of selecting an individual student is the product of three independent events: selecting the school, the classroom, and the student. To obtain the probability of selection for an individual student into the overall sample it is necessary to multiply three selection

<sup>3</sup> For the user not familiar with the data files included in the TIMSS International Database we recommend reading Chapter 7 before proceeding with sections 3.10 and 3.11.

probabilities – school, classroom, and student – and their respective adjustment factors. The resulting product gives the individual probability of selection for the student. Inverting this probability gives the basic sampling weight for the student. However, students also were assigned a booklet at random, so that this probability also has to be incorporated into the student weight. The following variables reflect probability factors for students based on the assigned booklet type.

***WGTADJL4 Literacy Scale Adjustment Factor***

This is the inverse of the probability that a student in the sample was assigned booklet 1A, 1B, or 4. It is set to zero for students who were not assigned booklet 1A, 1B, or 4.

The full student weight for students in the sample for measuring performance in mathematics and science literacy is the product of six variables, as follows:

$$TOTWGTL4 = WGT FAC1 * WGTADJ1 * WGT FAC2 * WGT FAC3 * WGTADJ3 * WGTADJL4$$

***WGTADJM4 Advanced Mathematics Scale Adjustment Factor***

This is the inverse of the probability that a student in the sample was assigned booklet 3A, 3B, 3C, or 4. It is set to zero for students who were not assigned booklet 3A, 3B, 3C, or 4.

The full student weight for students in the sample for measuring performance in advanced mathematics is the product of six variables, as follows:

$$TOTWGTM4 = WGT FAC1 * WGTADJ1 * WGT FAC2 * WGT FAC3 * WGTADJ3 * WGTADJM4$$

***WGTADJP4 Physics Scale Adjustment Factor***

This is the inverse of the probability that a student in the sample was assigned booklet 2A, 2B, 2C, or 4. It is set to zero for students who were not assigned booklet 2A, 2B, 2C, or 4.

The full student weight for students in the sample for measuring performance in physics is the product of six variables, as follows:

$$TOTWGTP4 = WGT FAC1 * WGTADJ1 * WGT FAC2 * WGT FAC3 * WGTADJ3 * WGTADJP4$$

The full student weights for analyses involving scale scores in mathematics and science literacy, advanced mathematics, and physics, and for analyses of the student background data for all students in their final year of secondary school are *TOTWGTL4*, *TOTWGTM4*, and *TOTWGTP4*, respectively, as shown above. These weights may also be used for analyses involving individual item-level scores from the core clusters (cluster A for mathematics and science literacy, cluster I for advanced mathematics, and cluster E for physics). However, since the probability of items in non-core clusters being assigned to students was different from that for items in the core clusters, different weights are required for analyses involving item-level scores. These item-level weights are described in the next section.<sup>4</sup> Secondary users planning to analyze results for individual achievement items may find it convenient to use the item-level weights for analyses involving items from all item clusters, even though the scale weights (*TOTWGTL4*, *TOTWGTM4*, and *TOTWGTP4*) provide more accurate estimates for items in the core clusters (A, E, and I).

<sup>4</sup> The item-level weights were used in computing the percents correct for example items included in Chapters 3, 6, and 9 of the international report (Mullis, et. Al; 1998).



***WGTADJL Literacy Item Adjustment Factor***

This is the inverse of the probability that a student in the sample was assigned booklet 1A or 1B. It is set to zero for students who were not assigned booklet 1A, or 1B.

The full student weight to be used in analyzing performance on mathematics and science literacy items in clusters A, B, C, and D is the product of six variables, as follows:

$$TOTWGTL=WGTFAC1*WGTADJ1*WGTFAC2*WGTFAC3*WGTADJ3*WGTADJL$$

***WGTADJM Advanced Mathematics Item Adjustment Factor***

This is the inverse of the probability that a student in the sample was assigned booklet 3A, 3B, or 3C. It is set to zero for students who were not assigned booklet 3A, 3B, or 3C.

The full student weight to be used in analyzing performance on advanced mathematics items in clusters I, J, K, and L is the product of six variables, as follows:

$$TOTWGTM=WGTFAC1*WGTADJ1*WGTFAC2*WGTFAC3*WGTADJ3*WGTADJM$$

***WGTADJP Physics Item Adjustment Factor***

This is the inverse of the probability that a student in the sample was assigned booklet 2A, 2B, or 2C. It is set to zero for students who were not assigned booklet 2A, 2B, or 2C.

The full student weight to be used in analyzing performance on physics items in clusters E, F, G, and H is the product of six variables, as follows:

$$TOTWGTP=WGTFAC1*WGTADJ1*WGTFAC2*WGTFAC3*WGTADJ3*WGTADJP$$

There are three main types of sampling weights included in the data files: total student weights, “senate” weights, and “house” weights. These three types of sampling weights and their use are discussed in the following sections. Each one of the types of weights has 6 versions. The corresponding weight variable names are *TOTWGT<sub>x</sub>*, *SENWGT<sub>x</sub>*, and *HOUWGT<sub>x</sub>*, where *x* identifies the version of the weight. The names of the different versions and a brief description of how they are used are presented in Table 3.11.

**Table 3.11**  
**Versions of the Sampling Weights**

Version	Description	Use
L4	Literacy weights: includes students responding to booklets 1A, 1B, and 4	For estimates of scale scores in mathematics and science literacy and background data estimates for students in their final year of secondary school.
L	Literacy weights: includes students responding to booklets 1A and 1B	For estimates of item-level scores for the mathematics and science literacy items in clusters A, B, C, and D.
M4	Advanced Mathematics weights: includes students responding to booklets 3A, 3B, 3C, and 4	For estimates of scale scores in advanced mathematics and background data estimates for the advanced mathematics students in their final year of secondary school.
M	Advanced Mathematics weights: includes students responding to booklets 3A, 3B, and 3C	For estimates of item-level scores for the advanced mathematics items in clusters I, J, K, and L.
P4	Physics weights: includes students responding to booklets 2A, 2B, 2C, and 4	For estimates of scale scores in physics and background data estimates for the physics students in their final year of secondary school.
P	Physics weights: includes students responding to booklets 2A, 2B, and 2C	For estimates of item-level scores for the physics items in clusters E, F, G, and H.

## Total Student Weights

***TOTWGTL4*** Total Student Weight - Literacy (Booklets: 1A,1B,4)

***TOTWGTM4*** Total Student Weight - Advanced Mathematics (Booklets: 3A,3B,3C,4)

***TOTWGTP4*** Total Student Weight - Physics (Booklets: 2A,2B,2C,4)

***TOTWGTL*** Total Student Weight - Literacy (Booklets: 1A,1B)

***TOTWGTM*** Total Student Weight - Advanced Mathematics (Booklets: 3A,3B,3C)

***TOTWGTP*** Total Student Weight - Physics (Booklets: 2A,2B,2C)

These are the full student sampling weights. The sum of these weights within a sample provides an estimate of the size of the population. *TOTWGTL4* and *TOTWGTL* provide estimates of the number of students in their last year of secondary school within the country. *TOTWGTM4* and *TOTWGTM* provide estimates of the number of advanced mathematics students in the last year of secondary school. *TOTWGTP4* and *TOTWGTP* provide estimates of the number of students in their final year of secondary school who have taken physics.

Although the *TOTWGTLx* is a commonly used sampling weight, it sums to a very large number, and to a different number within each country. This is not always desirable. For example, if we want to compute a weighted estimate of the mean achievement in advanced mathematics for the students in their last year of secondary school across all countries, using the variable *TOTWGTM4* as our weight variable will lead each country to contribute proportionally to its population size, with the large countries contributing more to the estimate than small countries. Although this might be desirable in some circumstances (e.g., when computing the benchmarks for the population overall), this is not usually the case.

A key property of the sampling weights is that the same population estimates for means and proportions will be obtained as long as we use a weight variable proportional to the original weights (*TOTWGTLx*). For example, we could take the sampling weights for a large country and divide them by a constant to make them smaller. We could also take the weights of a smaller country and multiply them by a constant to make them bigger. Regardless of which constant is used within a country, the weighted estimates for means and proportions obtained from each of the transformations of the weights will be exactly the same. Capitalizing on this property, two other sets of weight variables (“senate” weights and “house” weights) were computed and included in the student data files. Each of these will yield exactly the same means and proportions within each country, but will give different results when estimates across countries are computed or significance tests performed.

## Senate Weights

- SENWGTL4*** *Senate Weight - Literacy (Booklets: 1A,1B,4)*  
***SENWGTM4*** *Senate Weight - Advanced Mathematics (Booklets: 3A,3B,3C,4)*  
***SENWGTP4*** *Senate Weight - Physics (Booklets: 2A,2B,2C,4)*  
***SENWGTL*** *Senate Weight - Literacy (Booklets: 1A,1B)*  
***SENWGTM*** *Senate Weight - Advanced Mathematics (Booklets: 3A,3B,3C)*  
***SENWGTP*** *Senate Weight - Physics (Booklets: 2A,2B,2C)*

These variables are computed as

$$SENWGT_x = \left( \frac{500}{\sum TOTWGT_x} \right) \times TOTWGT_x$$

within each country. The transformation of the weights will be different within each country, but in the end, the sum of the variable  $SENWGT_x$  within each country will add up to 500. The variable  $SENWGT_x$ , within each country, is proportional to  $TOTWGT_x$  by the ratio of 500 divided by the size of the population estimate. These sampling weights can be used when cross national comparisons are required and the user wants to have each country contribute the same amount to the cross national comparison. When this variable is used as the sampling weight for international estimates, the contribution of each country is the same (i.e., 500 students), regardless of the size of the population. The name “senate weight” is derived from an analogy with the U.S. Senate, where each state is represented equally (by two senators), regardless of the population of the state.

## House Weights

- HOUWGTL4*** *House Weight - Literacy (Booklets: 1A,1B,4)*  
***HOUWGTM4*** *House Weight - Advanced Mathematics (Booklets: 3A,3B,3C,4)*  
***HOUWGTP4*** *House Weight - Physics (Booklets: 2A,2B,2C,4)*  
***HOUWGTL*** *House Weight - Literacy (Booklets: 1A,1B)*  
***HOUWGTM*** *House Weight - Advanced Mathematics (Booklets: 3A,3B,3C)*  
***HOUWGTP*** *House Weight - Physics (Booklets: 2A,2B,2C)*

These variables are computed as

$$HOUWGT_x = \left( \frac{N}{\sum TOTWGT_x} \right) \times TOTWGT_x$$

within each country where  $N$  is the sample size for the country. The transformation of the weights will be different within each country, but in the end, the sum of the variables  $HOUWGT_x$  within each country will add up to the sample size for that country. The variable  $HOUWGT_x$  is proportional to  $TOTWGT_x$  by the ratio of the sample size divided by the size of the population. These sampling weights can be used when the user wants the actual sample size to be used in performing significance tests. Although some

statistical computer software allow the user to use the sample size as the divisor in the computation of standard errors, others will use the sum of the weights, and this results in severely deflated standard errors for the statistics if the *TOTWGT<sub>x</sub>* is used as the weighting variable. When performing analyses using such software, we recommend using the variable *HOUWGT<sub>x</sub>* as the weight variable. This ensures that the number of cases used in significance tests corresponds to the sample size. Because of the clustering effect in most TIMSS samples, it is usually also desirable to apply a correction factor such as a design effect to the *HOUWGT<sub>x</sub>* variable, or, preferably, to use the jackknife repeated replication method (see Chapter 8) to get more accurate estimates of sampling error. The name “house weight” is derived from an analogy with the U.S. House of Representatives, where the number of representatives for each state reflects the size of the population of the state.

### 3.11 Selecting Students for Analyses

To facilitate the selection of students for different types of analyses, a set of participation status variables have been created and included in the student background and written assessment data files. These variables can take a value of 1 or 0, depending on the participation status of each student in the mathematics and science literacy, advanced mathematics, or physics tests according to their test booklet assignment. The participation status variables are listed in Table 3.12, together with a brief description of how they are used in combination with weight variables in the analysis of achievement and student background data.<sup>5</sup>

<sup>5</sup> Some example analyses using the participation status and weight variables are presented in Chapter 9.

**Table 3.12**  
**Participation Status Variables for the Student Data Files**

Variable	Label	Use
INTMSL4	PARTICIPATION STATUS - LITERACY (Booklets:1A,1B,4)	Students with INTMSL4 set to one are those who responded to the literacy booklets 1A, 1B or 4, and who are also part of the TIMSS International sample. Otherwise this variable is set to zero. Use in combination with weights TOTWGTL4, SENWGTL4 or HOUWGTL4 for analyses involving scale scores in mathematics and science literacy or student background variables.
INTMSL	PARTICIPATION STATUS – LITERACY (Booklets:1A,1B)	Students with INTMSL set to one are those who responded to the literacy booklets 1A or 1B, and who are also part of the TIMSS International sample. Otherwise this variable is set to zero. Use in combination with weights TOTWGTL, SENWGTL or HOUWGTL for analyses involving item-level scores for mathematics and science literacy items in Clusters A, B, C, and D.
INTMSM4	PARTICIPATION STATUS – MATHEMATICS (Booklets:3A,3B,3C,4)	Students with INTMSM4 set to one are those who responded to the advanced mathematics booklets 3A, 3B, 3C or 4, and who are also part of the TIMSS International sample. Otherwise this variable is set to zero. Use in combination with weights TOTWGTM4, SENWGTM4 or HOUWGTM4 for analyses involving scale scores in advanced mathematics or student background variables.
INTMSM	PARTICIPATION STATUS – MATHEMATICS (Booklets:3A,3B,3C)	Students with INTMSM set to one are those who responded to the advanced mathematics booklets 3A, 3B or 3C, and who are also part of the TIMSS International sample. Otherwise this variable is set to zero. Use in combination with weights TOTWGTM, SENWGTM or HOUWGTM for analyses involving item-level scores for advanced mathematics items in Clusters I, J, K, and L.
INTMSP4	PARTICIPATION STATUS – PHYSICS (Booklets:2A,2B,2C,4)	Students with INTMSP4 set to one are those who responded to the physics booklets 2A, 2B, 2C or 4, and who are also part of the TIMSS International sample. Otherwise this variable is set to zero. Use in combination with weights TOTWGTP4, SENWGTP4 or HOUWGTP4 for analyses involving scale scores in physics or student background variables.
INTMSP	PARTICIPATION STATUS – PHYSICS (Booklets:2A,2B,2C)	Students with INTMSP set to one are those who responded to the physics booklets 2A, 2B or 2C, and who are also part of the TIMSS International sample. Otherwise this variable is set to zero. Use in combination with weights TOTWGTP, SENWGTP or HOUWGTP for analyses involving item-level scores for physics items in Clusters E, F, G, and H.

### 3.12 Weight Variables Included in the School Data Files

The TIMSS samples are designed to give representative samples of students within countries. Although students were selected by first sampling schools within the countries, the school samples were designed so that the sampling of students, rather than the sampling of schools, was optimized. In other words, the samples were selected to make statements about the students in the country rather than about the schools in the country.

Several weight variables included in the school files are listed below.

***WGTFAC1***     ***School Weighting Factor***

This variable corresponds to the inverse of the probability of selection for the school where the student is enrolled.

***WGTADJ1***     ***School Weighting Adjustment***

This is an adjustment that is applied to WGTFAC1 to account for non-participating schools in the sample. If we were to multiply WGTFAC1 by WGTADJ1 we would obtain the sampling weight for the school adjusted for non-participation.

***SCHWGT***     ***School-Level Weight***

The school sampling weight is the inverse of the probability of selection for the school, multiplied by its corresponding adjustment factor. It is computed as the product of WGTADJ1 and WGTFAC1. Although this weight variable can be used to estimate the number of schools with certain characteristics, it is important to keep in mind that the sample selected for TIMSS is a good sample of students, but not necessarily an optimal sample of schools. Schools are selected with probability proportional to their size, so it is expected that there is a greater number of large schools in the sample.

***STOWGTL4***     ***Sum of the Student Weights - Literacy (Booklets: 1A,1B,4)***

***STOWGTM4***     ***Sum of the Student Weights - Advanced Mathematics (Booklets: 3A,3B,3C,4)***

***STOWGTP4***     ***Sum of the Student Weights - Physics (Booklets: 2A,2B,2C,4)***

***STOWGTL***     ***Sum of the Student Weights - Literacy (Booklets: 1A,1B)***

***STOWGTM***     ***Sum of the Student Weights - Advanced Mathematics (Booklets: 3A,3B,3C)***

***STOWGTP***     ***Sum of the Student Weights - Physics (Booklets: 2A,2B,2C)***

These variables are the sum of the weights of the students within the school participating in the mathematics and science literacy, advanced mathematics, or physics tests.<sup>6</sup> If there are no students in the school participating in one or more of the TIMSS tests, then the

<sup>6</sup> The definition of the different versions of the school weight variables (STOWGTx) correspond to those shown for the student weights (TOTWGTx, SENWGTx, and HOUWGTx) given in Table 3.11.

corresponding variables are set to zero. These variables can be used to conduct analyses on questions like “How many students attend schools that have a certain characteristic?” Although weight variables have been included in the school files, there is no information in these files regarding the sampling zone in which the school is included (see Chapter 8 for discussion of sampling zones in the estimation of sampling variance). If the user is interested in computing proper standard errors of the population estimates, then the information in the school file should be merged with the student file and the analysis conducted using a student-level file.



---

# Chapter 4

## Data Collection, Materials Processing, Scoring, and Database Creation

Each country participating in TIMSS was responsible for collecting its national data and processing the materials in accordance with the international standards. In each country, a national research center and National Research Coordinator (NRC) were appointed to implement these activities. One of the main ways in which TIMSS sought to achieve uniform project implementation was by providing clear and explicit instructions on all operational procedures. Such instructions were provided primarily in the form of operations manuals, supported where possible by computer software systems that assisted NRCs in carrying out the specified field operations procedures. Forms accompanying some of the manuals served to document the implementation of the procedures in each country. Many of these forms were used to track schools, students, and teachers, and to ensure proper linkage of schools, students, and teachers in the database.

As part of the TIMSS quality assurance efforts, there also was a program of site visits by trained Quality Assurance Monitors representing the International Study Center. The data collected during these visits are presented in Martin and Mullis (1996) together with extensive documentation of the quality of the translation activities, population sampling, scoring, data checking, and database construction.

### 4.1 Data Collection and Field Administration

For the sake of comparability, all testing was conducted at the end of the school year. Most countries tested the mathematics and science achievement of their students at the end of the 1994-95 school year, most often in May and June of 1995. The three countries on a Southern Hemisphere school schedule (Australia, New Zealand, and South Africa) tested between August and December 1995, which was late in the school year in the Southern Hemisphere. Students in Australia were tested in September to October; students in New Zealand were tested in August; and students in South Africa were tested in August to December 1995. Three countries tested their final-year students (or a subset of them) at the end of the 1995-96 school year. Iceland tested its final-year students in 1996; Germany tested its gymnasium students in 1996; and Lithuania tested the students in vocational schools in 1996. In Germany and Lithuania, all other students included in the TIMSS assessment were tested in 1995.

In addition to selecting the sample of students to be tested, the NRCs were responsible for working with the School Coordinators, translating the test instruments, assembling and printing the test booklets, and packing and shipping the necessary materials to the sampled schools. They also were responsible for arranging for the return of the testing materials from the schools to the national center, preparing for and implementing the free-response scoring, entering the results into data files, conducting on-site quality assurance observations for a 10% sample of schools, and preparing a report on survey activities. The *Survey Operations Manual* was prepared by the IEA Data Processing Center for the NRCs and their colleagues who were responsible for implementing the TIMSS data collection and processing procedures. It describes the activities and responsibilities of the NRCs from the moment the international testing materials arrived at the national center to the moment the cleaned data sets and accompanying documentation were sent to the IEA Data

Processing Center. In addition to detailed within-school sampling instructions, the manual included:

- Procedures for translating and assembling the test instruments and questionnaires
- Instructions for obtaining cooperation from the selected schools
- Explicit procedures for packing and sending materials to the schools
- Preparations for test administration
- Instructions for data entry and verification

Included in this manual were a set of Survey Tracking Forms that were completed at various stages of the study to track schools and students, and ensure proper linkage between them in the database.

Each school was asked to appoint a coordinator to be the liaison between the national research center and the school. The *School Coordinator Manual* describes the steps the School Coordinator followed from the time the testing materials arrived at the school until they were returned to the NRC. Essentially, the School Coordinator was responsible for receiving the shipment of testing materials from the national center and providing for their security before and after the test date. The School Coordinator arranged the testing accommodations, trained Test Administrators, and arranged for make-up sessions when necessary.

The *Test Administrator Manual* covered the procedures from the beginning of testing to the return of the completed tests, questionnaires, and tracking forms to the School Coordinator. The manual includes a Test Administration Script to be read by the Test Administrator. The Test Administrators were responsible for activities preliminary to the testing session, including maintaining the security of the test booklets and ensuring adequacy of supplies and the testing environment. Activities during the testing session, included distribution of the test booklets to the appropriate students (using the Student Tracking Form), timing of the testing and breaks, and accurately reading the test administration script to the students.

## 4.2 Free-Response Scoring

Upon completion of each testing session, the School Coordinator shipped the booklets, questionnaires, and forms to the National Research Center. The NRCs then organized the instruments for scoring and data entry. These procedures were designed to maintain identification information that linked students to schools, minimize the time and effort spent handling the booklets, ensure reliability in the free-response scoring, and document the reliability of the scoring. Since approximately one-third of the written test time was devoted to free-response items, the free-response scoring was complex.

The free-response items were scored using item-specific rubrics. Scores were represented by two-digit codes. The first digit designates the correctness level of the 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. This coding approach was used with all free-response items, including both the short-answer and extended-response items.

The number of points specified in each rubric varies by item since each item is unique in terms of answer, approach, and types of misconceptions generated by students. Most items are worth one point. In these rubrics, correct student responses were coded as 10, 11, 12, and so on through 19

and earned one score point. The type of response in terms of the approach used or explanation provided is denoted by the second digit.

In all of the guides, incorrect student responses were coded as 70, 71, and so on through 79 and earned zero score points. However, as in the approach used for correct scores, the second digit in the code represents the type of misconception displayed, incorrect strategy used, or incomplete explanation given. An example guide for a short-answer science literacy item is shown below in Figure 4.1. In this guide, students received one point for a correct answer. There were several diagnostic codes for correct answers.

The rubrics for more complicated items were correspondingly more complex, having categories for full and partial credit. As shown in Figure 4.2, on this extended advanced mathematics task students received two points for a fully correct answer and one point for a partially correct answer. In some cases, the scoring guides include three points.

**Figure 4.1**  
**Example Coding Guide for Short-Answer Science Literacy Item**

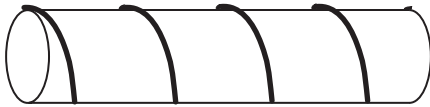
D3. José caught influenza. Write down one way he could have caught it.
Reproduced from TIMSS Population 3 Item Pool. Copyright © 1995 by IEA, The Hague

Code	Response
<b>Correct Response</b>	
<b>10</b>	Refers explicitly to transmission of germs.
<b>11</b>	Refers implicitly to transmission of germs by sneezing/coughing or close contact. <i>Examples: Drinking from the same cup.</i> <i>Shaking hands.</i> <i>Eating together or from same utensils.</i>
<b>12</b>	States only that he got it from someone who had the flu.
<b>19</b>	Other correct.
<b>Incorrect Response</b>	
<b>70</b>	Refers to being too cold. <i>Examples: He got it from being out in the cold.</i> <i>He got it from getting wet [or freezing].</i> <i>He got it because he did not wear enough clothes.</i>
<b>79</b>	Other incorrect.
<b>Nonresponse</b>	
<b>90</b>	Crossed out/erased, illegible, or impossible to interpret.
<b>99</b>	BLANK

**Figure 4.2**

### Example Coding Guide for Extended-Response Advanced Mathematics Item

K14. A string is wound symmetrically around a circular rod. The string goes exactly 4 times around the rod. The circumference of the rod is 4 cm and its length is 12 cm.



Find the length of the string. Show all your work.

Reproduced from TIMSS Population 3 Item Pool. Copyright © 1995 by IEA, The Hague

Code	Response
	<b>Correct Response</b>
20	Length of string = 20 cm. Method: <ul style="list-style-type: none"> <li>• The surface of the rod is represented as a rectangle 4 cm by 12 cm.</li> <li>• Four parallel congruent segments are drawn in the rectangle indicating the position of the string.</li> <li>• Length of one segment is calculated using Pythagorean theorem <math>\sqrt{3^2 + 4^2} = 5</math>. Total length of string = <math>4 \times 5</math> cm = 20 cm.</li> </ul>
21	Length of string = 20 cm. Method: <ul style="list-style-type: none"> <li>• Half of surface of rod represented as rectangle 2 cm by 12 cm.</li> <li>• Eight congruent segments drawn in the rectangle indicating position of string.</li> <li>• Length of one segment calculated using Pythagorean theorem <math>\sqrt{2^2 + 1.5^2} = 2.5</math>. Total length of string = <math>8 \times 2.5</math> cm = 20 cm.</li> </ul>
22	Length of string = 20 cm. Method used: <ul style="list-style-type: none"> <li>• Situation represented either by rectangle 16 x 12 with string as its diagonal OR by right triangle with sides 16 and 12 and string as its hypotenuse.</li> <li>Pythagorean theorem used to calculate length of string <math>\sqrt{16^2 + 12^2} = 20</math> cm.</li> </ul>
29	All other fully correct solutions.
	<b>Partial Response</b>
10	Length of string = 20 cm. No work shown.
11	Surface of rod represented by rectangle with correct dimensions and position of string correctly indicated, but numerical error in the calculation of the length of string.
19	All other partially correct solutions with <u>correct method</u> and minor error.
	<b>Incorrect Response</b>
70	Incorrect answer. No work shown.
71	Length of string = 16 cm. Argument: It is the same as 4 circles.
72	Length of string = 28 cm. Argument: "If the string were wound 4 times around the same place, its length would be $4 \times 4$ . But since it "moves" along the rod which is 12 cm long, we must add these 12 cm to the length of the string."

<b>73</b>	<p>Estimation methods:  Length of 1 revolution estimated or stated but not calculated; then it is multiplied by 4.</p> <p><i>Examples: 1 revolution is approx. 6 cm long, length of string is <math>4 \times 6 = 24</math> cm.</i></p> <p><i>1 revolution is <math>(4 + 1.5)</math> cm long, length of string is <math>4 \times 5.5 = 22</math> cm.</i></p> <p><i>Length of string must be greater than 16 cm (it would be 16 cm if it were 4 circles) and/or</i></p> <p><i>Length of string <u>must be less than</u> <math>16 + 12 = 28</math> cm.</i></p> <p><i><math>16 \text{ cm} &lt; L &lt; 28 \text{ cm}</math></i></p>
<b>74</b>	String is represented by a curve, e.g. parts of a circle or an ellipse.
<b>79</b>	All other incorrect attempts with some work shown.
<b>Nonresponse</b>	
<b>90</b>	Crossed-out, illegible, or impossible to interpret.
<b>99</b>	BLANK

To meet the goal of reliable scoring, TIMSS had an ambitious training program. The training sessions were designed to familiarize representatives of the national centers with the scoring rubrics. These national representatives were then responsible for training personnel in their own countries to apply the two-digit codes reliably. TIMSS recommended that scorers be organized into teams of about six, headed by a team leader who monitored the progress and reliable use of the codes. The team leaders were to continually check and reread the responses coded by their team, systematically covering the daily work of each scorer. To document information about the within-country agreement among scorers, TIMSS developed a procedure whereby approximately 10% of the student responses were to be scored independently by two readers.

As shown in Table 4.1, the within-country percentages of exact agreement for both the correctness score and the full two-digit diagnostic code revealed a high degree of agreement for the countries that documented the reliability of their coding. Unfortunately, lack of resources precluded several countries from providing this information. At Population 3, a very high percentage of exact agreement was observed on the mathematics and science literacy items for all countries, with averages across items for the correctness score ranging from 91% to 98% and an overall average of 95% across the 13 countries. The corresponding average across all countries for diagnostic code agreement was 87%, with a range of 81% to 95%. For advanced mathematics, the average agreement for correctness was 96% overall for the 10 countries, ranging from 93% to 99%. For diagnostic codes, the overall average was 88%, ranging from 81% to 97%. For physics, the average agreement for correctness was 95% overall for the 11 countries, with average agreement ranging from 89% to 100%. For diagnostic codes, the overall average was 86% and ranged from 77% to 97%.

**Table 4.1**  
**TIMSS Within-Country Free-Response Coding Reliability Data**

	Correctness Score Agreement			Diagnostic Code Agreement		
	Overall Average Percent of Exact Agreement Across All Items	Range Across Countries of Average Percent of Exact Agreement		Overall Average Percent of Exact Agreement Across Items	Range Across Countries of Average Percent of Exact Agreement	
		Min	Max		Min	Max
<b>Mathematics and Science Literacy</b>	95	91	98	87	81	95
<b>Advanced Mathematics</b>	96	93	99	88	81	97
<b>Physics</b>	95	89	100	86	77	97

Note: Overall averages are based on all countries providing coding reliability: mathematics and science literacy based on 13 countries; advanced mathematics based on 10 countries; physics based on 11 countries.

For more details about the TIMSS free-response scoring, refer to:

- *Scoring Techniques and Criteria* (Lie, Taylor, and Harmon, 1996)
- *Training Sessions for Free-Response Scoring and Administration of the Performance Assessment* (Mullis, Jones, and Garden, 1996)
- *Guide to Checking, Coding, and Entering the TIMSS Data* (TIMSS, 1995)
- *Quality Control Steps for Free-Response Scoring* (Mullis and Smith, 1996)

### 4.3 Data Entry

To maintain equality among countries, very little optical scanning and no image processing of item responses was permitted. Essentially all of the student test information was recorded in the student booklets or on separate coding sheets, and similar procedures were used for the questionnaires. Entry of the achievement and background data was facilitated by the *International Codebooks*, and the DATAENTRYMANAGER software program.

The background questionnaires were stored with the various tracking forms so that the data entry staff could control the number of records to enter and transcribe the necessary information during data entry. NRCs were asked to arrange for double-entry of a random sample of at least 5% of the test instruments and questionnaires. An error rate of 1% was considered acceptable.

After entering data files in accordance with the international procedures, countries submitted their data files to the IEA Data Processing Center.

## 4.4 Database Creation

Even though extreme care was taken in developing manuals and software for use by the 24 countries participating in the assessment of final-year students, the national centers inadvertently introduced various types of inconsistencies in the data, which needed to be thoroughly investigated by the IEA Data Processing Center and the International Study Center at Boston College.

The TIMSS data underwent an exhaustive cleaning process designed to identify, document, and correct deviations from the international instruments, file structures, and coding schemes. The process also emphasized consistency of information with national data sets and appropriate linking among the many data files. The national centers were contacted regularly throughout the cleaning process and were given multiple opportunities to review the data for their countries.

## 4.5 Instrument Deviations and National Adaptations

Ensuring the international comparability of both the cognitive and contextual variables was an important aspect of TIMSS. A number of data management steps were focused on evaluating the international comparability of the TIMSS items, and any deviations for specific items were handled on an individual basis. An overview of the procedures and policies applied to ensuring international comparability is provided in the following sections relating to the test items and the background questionnaire items.

### 4.5.1 Cognitive Items

All TIMSS written assessment test items were originally developed in English and then translated by the participating TIMSS countries into the other languages for testing. In addition to the translation verification steps used for all TIMSS test items (Maxwell, 1996), a thorough item review process was also used to further evaluate any items that were functioning differently in different countries according to the international item statistics (Mullis and Martin, 1998). As a result of this review process, a few items were identified as not being internationally comparable in certain countries and were deleted from the international data files and from the analyses for the international reports. A list of all these deleted items as well as any items omitted in the test instruments for specific countries is given in Table 4.2. Table 4.3 shows any recodes made to the cognitive items as a result of the item review.

**Table 4.2**  
**List of Deleted Cognitive Items—Population 3**

	Country	Item	Variable Name
Mathematics and Science Literacy	All	A09, Part A C10	CSEGA09A CSMGC10
	Cyprus	C05 D12	CSMGC05 CSMGD12
	Greece	C05 D12 A11, Part C	CSMGC05 CSMGD12 CSEGA11C
	France	B04 B06	CSMGB04 CSMGB06
	Hungary	B08 B21 B26 C20 D15, Part B D16, Part A D16, Part B	CSMGB08 CSMGB21 CSSGB26 CSSGC20 CSSGD15B CSSGD16A CSSGD16B
	Switzerland	B06	CSMGB06
	Slovenia	A11, Part C	CSEGA11C
	Advanced Mathematics	Cyprus	J02
France		J18	CSEMJ18
Greece		J02	CSMMJ02
Israel		J14 J16, Part B L08	CSMMJ14 CSSMJ16B CSMML08
Lithuania		K09	CSMMK09
Switzerland		J02 J17	CSMMJ02 CSSMJ17
United States		J08	CSMMJ08
Physics	All	H11	CSMPH11
	Australia	H19, Part A	CSEPH19A
	Czech Republic	F06	CSMPF06
	Denmark	F07 H14	CSMPF07 CSSPH14
	France	F15	CSEPF15
	Germany	G16 H14	CSEPG16 CSSPH14



**Table 4.3**  
**List of Recodes Made to Free-Response Item Codes**

	Item	Variable	Recodes	Comment
Mathematics and Science Literacy	B25	CSSGB25	20 → 10 21 → 11 22 → 12 10 → 13 29 → 19	Category 10 was only 1 point category and generally had less than 1 percent of the students, which made distinction between 1 and 2 points unclear.
	B26	CSSGB26	10 → 23 19 → 29	Categories 10 and 19 contain correct answer.
	D02	CSSGD02	20 → 12 21 → 13	Discrimination between 20s and 10s not clear.
	D04	CSEGD04	20 → 10 21 → 11 22 → 12 29 → 19 10 → 73 11 → 74 19 → 75	Is a link item with Y01 at Population 2 and similar to Population 2 only 20s had positive point-biserials in many countries.
	D17	CSSGD17	13 → 22	In some countries 10s had almost or even higher point-biserials than 20s.

## 4.5.2 Background Questionnaire Items

As was the case with the test instruments, the international versions of the student and school background questionnaires were also developed in English and then translated into other languages by the TIMSS countries. While the intent of TIMSS was to provide internationally comparable data for all variables, there were many contextual differences among countries so that the international version of the questions were not always appropriate in all countries. Therefore, the international versions of the questionnaires were designed to provide an opportunity for individual countries to modify some questions or response options in order to include the appropriate wording or options most consistent with their own national systems. In the international versions of the questionnaires (Supplement 1), such questions contain instructions to NRCs to substitute the appropriate wording for their country or to modify or delete any inappropriate questions or options. These instructions were indicated in two ways in the questionnaires:

- 1) NRC NOTE:
- 2) <International Option> (indicating that the NRC was to substitute, if necessary, an appropriate national option that would retain the same basic interpretation as the international version)

In addition to some of the questions designed to contain some nationally modified portions, some countries also included revised versions of other questions as well as some inadvertent deviations from the international version (e.g. option reversals). Although no formal translation verification step was used for the background questionnaires, countries were to provide a National Deviations Report and a Data Management Form that included information about how any revised or deleted items were to be handled during data entry and data processing. The information contained in these reports was used, together with the international univariate summary statistics for background questionnaire items and direct comparisons of the international and national versions of the

questionnaires, in order to evaluate the international comparability of any modified items. This review process was instituted for the majority of the background questionnaire items and, in particular, for any of the variables that were reported in the international report. Each item deviation or national adaptation was individually reviewed in order to determine which of the following actions should be taken:

- National data deleted as not being internationally comparable
- National data recoded to match the international version
- National data retained with some documentation describing modifications

Whenever possible, national data were retained by recoding to match as closely as possible the international version of the items and/or by documenting minor deviations. NRCs were contacted to resolve questions regarding the international comparability of revised items, and no changes were made to the international data files without first informing the NRC and receiving confirmation whenever possible. A summary of all available documentation for the deleted or modified background questionnaire items in the international data files is provided in Supplement 2.<sup>1</sup>

<sup>1</sup> This documentation was sent to all NRCs for review and verification prior to release of the international data files.

---

# Chapter 5

## TIMSS Scaling Procedures

The principal method by which student achievement is reported in TIMSS is through scale scores derived using Item Response Theory (IRT) scaling. With this approach, the performance of a sample of students in a subject area can be summarized on a common scale or series of scales even when different students have been administered different items. The common scale makes it possible to report on relationships between students' characteristics (based on their responses to the background questionnaires) and their overall performance in mathematics and science.

Because of the need to achieve broad coverage of both subjects within a limited amount of student testing time, each student was administered relatively few items within each of the content areas of each subject. In order to achieve reliable indices of student proficiency in this situation it was necessary to make use of multiple imputation or "plausible values" methodology. Further information on plausible value methods may be found in Mislevy (1991), and in Mislevy, Johnson, and Muraki (1992). The proficiency scale scores or plausible values assigned to each student are actually random draws from the estimated ability distribution of students with similar item response patterns and background characteristics. The plausible values are intermediate values that may be used in statistical analyses to provide good estimates of parameters of student populations. Although intended for use in place of student scores in analyses, plausible values are designed primarily to estimate population parameters, and are not optimal estimates of individual student proficiency.

This chapter provides details of the IRT model used in TIMSS to scale the achievement data. For those interested in the technical background of the scaling, the chapter describes the model itself, the method of estimating the parameters of the model, and the construction of the international scale.

### 5.1 The TIMSS Scaling Model

The scaling model used in TIMSS was the multidimensional random coefficients logit model as described by Adams, Wilson, and Wang (1997), with the addition of a multivariate linear model imposed on the population distribution. The scaling was done with the *ConQuest* software (Wu, Adams, and Wilson, 1997) that was developed in part to meet the needs of the TIMSS study.

The multidimensional random coefficients model is a generalization of the more basic unidimensional model.

## 5.2 The Unidimensional Random Coefficients Model

Assume that  $I$  items are indexed  $i=1,\dots,I$  with each item admitting  $K_i + 1$  response alternatives  $k = 0, 1, \dots, K_i$ . Use the vector valued random variable,  $\mathbf{X}_i = (X_{i1}, X_{i2}, \dots, X_{iK_i})$ ,

$$\text{where } X_{ij} = \begin{cases} 1 & \text{if response to item } i \text{ is in category } j \\ 0 & \text{otherwise} \end{cases} \quad (1)$$

to indicate the  $K_i + 1$  possible responses to item  $i$ .

A response in category zero is denoted by a vector of zeroes. This effectively makes the zero category a reference category and is necessary for model identification. The choice of this as the reference category is arbitrary and does not affect the generality of the model. We can also collect the  $\mathbf{X}_i$  together into the single vector  $\mathbf{X}' = (\mathbf{X}'_1, \mathbf{X}'_2, \dots, \mathbf{X}'_I)$ , which we call the response vector (or pattern). Particular instances of each of these random variables are indicated by their lower case equivalents:  $\mathbf{x}$ ,  $\mathbf{x}_i$  and  $x_{ik}$ .

The items are described through a vector  $\xi^T = (\xi_1, \xi_2, \dots, \xi_p)$  of  $p$  parameters. Linear combinations of these are used in the response probability model to describe the empirical characteristics of the response categories of each item. These linear combinations are defined by design vectors  $\mathbf{a}_k$ , ( $j = 1, \dots, I; k = 1, \dots, K_i$ ), each of length  $p$ , which can be collected to form a design matrix  $\mathbf{A}' = (\mathbf{a}_{11}, \mathbf{a}_{12}, \dots, \mathbf{a}_{1K_1}, \mathbf{a}_{21}, \dots, \mathbf{a}_{2K_2}, \dots, \mathbf{a}_{IK_I})$ . Adopting a very general approach to the definition of items, in conjunction with the imposition of a linear model on the item parameters, allows us to write a general model that includes the wide class of existing Rasch models.

An additional feature of the model is the introduction of a scoring function, which allows the specification of the score or "performance level" that is assigned to each possible response to each item. To do this we introduce the notion of a response score  $b_{ij}$ , which gives the performance level of an observed response in category  $j$  of item  $i$ . The  $b_{ij}$  can be collected in a vector as

$\mathbf{b}^T = (b_{11}, b_{12}, \dots, b_{1K_1}, b_{21}, b_{22}, \dots, b_{2K_2}, \dots, b_{IK_I})$ . (By definition, the score for a response in the zero category is zero, but other responses may also be scored zero.)

In the majority of Rasch model formulations there has been a one-to-one match between the category to which a response belongs and the score that is allocated to the response. In the simple logistic model, for example, it has been standard practice to use the labels 0 and 1 to indicate both the categories of performance and the scores. A similar practice has been followed with the rating scale and partial credit models, where each different possible response is seen as indicating a different level of performance, so that the category indicators 0, 1, 2, that are used serve as both scores and labels. The use of  $\mathbf{b}$  as a scoring function allows a more flexible relationship between the qualitative aspects of a response and the level of performance that it reflects. Examples of where this is applicable are given in Kelderman and Rijkes (1994) and Wilson (1992). A primary reason for implementing this feature in the model was to facilitate the analysis of the two-digit coding scheme that was used in the TIMSS short-answer and extended-response items. In the final

analyses, however, only the first digit of the coding was used in the scaling, so this facility in the model and scaling software was not used in TIMSS.

Letting  $\theta$  be the latent variable the item response probability model is written as:

$$\Pr(\mathbf{X}_{ij} = 1; \mathbf{A}, \mathbf{b}, \xi | \theta) = \frac{\exp(b_{ij}\theta + \mathbf{a}_{ij}^T \xi)}{\sum_{k=1}^{K_i} \exp(b_{ik}\theta + \mathbf{a}_{ij}^T \xi)}, \quad (2)$$

and a response vector probability model as

$$f(\mathbf{x}; \xi | \theta) = \Psi(\theta, \xi) \exp[\mathbf{x}^T (\mathbf{b}\theta + \mathbf{A}\xi)], \quad (3)$$

with

$$\Psi(\theta, \xi) = \left\{ \sum_{z \in \Omega} \exp[\mathbf{z}^T (\mathbf{b}\theta + \mathbf{A}\xi)] \right\}^{-1}, \quad (4)$$

where  $\Omega$  is the set of all possible response vectors.

### 5.3 The Multidimensional Random Coefficients Multinomial Logit Model

The multidimensional form of the model is a straightforward extension of the model that assumes that a set of  $D$  traits underlie the individuals' responses. The  $D$  latent traits define a  $D$ -dimensional latent space, and the individuals' positions in the  $D$ -dimensional latent space are represented by the vector  $\theta = (\theta_1, \theta_2, \dots, \theta_D)$ . The scoring function of response category  $k$  in item  $i$  now corresponds to a  $D$  by 1 column vector rather than a scalar as in the unidimensional model. A response in category  $k$  in dimension  $d$  of item  $i$  is scored  $b_{ikd}$ . The scores across  $D$  dimensions can be collected into a column vector  $\mathbf{b}_{ik} = (b_{ik1}, b_{ik2}, \dots, b_{ikD})^T$ , again be collected into the scoring sub-matrix for item  $i$ ,  $\mathbf{B}_i = (\mathbf{b}_{i1}, \mathbf{b}_{i2}, \dots, \mathbf{b}_{iK_i})^T$  and then collected into a scoring matrix  $\mathbf{B} = (\mathbf{B}_1^T, \mathbf{B}_2^T, \dots, \mathbf{B}_I^T)^T$  for the whole test. If the item parameter vector,  $\xi$ , and the design matrix,  $\mathbf{A}$ , are defined as they were in the unidimensional model, the probability of a response in category  $k$  of item  $i$  is modeled as

$$\Pr(\mathbf{X}_{ij} = 1; \mathbf{A}, \mathbf{B}, \xi | \theta) = \frac{\exp(\mathbf{b}_{ij}\theta + \mathbf{a}_{ij}^T \xi)}{\sum_{k=1}^{K_i} \exp(\mathbf{b}_{ik}\theta + \mathbf{a}_{ik}^T \xi)} \quad (5)$$

And for a response vector we have:

$$f(\mathbf{x}; \xi | \theta) = \Psi(\theta, \xi) \exp[\mathbf{x}'(\mathbf{B}\theta + \mathbf{A}\xi)], \quad (6)$$

with

$$\Psi(\theta, \xi) = \left\{ \sum_{z \in \Omega} \exp[\mathbf{z}^T (\mathbf{B}\theta + \mathbf{A}\xi)] \right\}^{-1} \quad (7)$$

The difference between the unidimensional model and the multidimensional model is that the ability parameter is a scalar,  $\theta$ , in the former, and a  $D$  by 1 column vector,  $\boldsymbol{\theta}$ , in the latter. Likewise, the scoring function of response  $k$  to item  $i$  is a scalar,  $b_{ik}$ , in the former, whereas it is a  $D$  by 1 column vector,  $\mathbf{b}_{ik}$ , in the latter.

## 5.4 The Population Model

The item response model is a conditional model in the sense that it describes the process of generating item responses conditional on the latent variable,  $\boldsymbol{\theta}$ . The complete definition of the TIMSS model, therefore, requires the specification of a density,  $f_{\boldsymbol{\theta}}(\boldsymbol{\theta}; \boldsymbol{\alpha})$ , for the latent variable,  $\boldsymbol{\theta}$ . We use  $\boldsymbol{\alpha}$  to symbolize a set of parameters that characterize the distribution of  $\boldsymbol{\theta}$ . The most common practice when specifying unidimensional marginal item response models is to assume that the students have been sampled from a normal population with mean  $\mu$  and variance  $\sigma^2$ . That is:

$$f_{\boldsymbol{\theta}}(\boldsymbol{\theta}; \boldsymbol{\alpha}) \equiv f_{\boldsymbol{\theta}}(\boldsymbol{\theta}; \mu, \sigma^2) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-\frac{(\boldsymbol{\theta} - \mu)^2}{2\sigma^2}\right] \quad (8)$$

or equivalently

$$\boldsymbol{\theta} = \mu + E \quad (9)$$

where  $E \sim N(0, \sigma^2)$ .

Adams, Wilson, and Wu (1997) discuss how a natural extension of (8) is to replace the mean,  $\mu$ , with the regression model,  $\mathbf{Y}_n^T \boldsymbol{\beta}$ , where  $\mathbf{Y}_n$  is a vector of  $u$  fixed and known values for student  $n$ , and  $\boldsymbol{\beta}$  is the corresponding vector of regression coefficients. For example,  $\mathbf{Y}_n$  could be constituted of student variables such as gender, socio-economic status, or major. Then the population model for student  $n$ , becomes

$$\theta_n = \mathbf{Y}_n^T \boldsymbol{\beta} + E_n, \quad (10)$$

where we assume that the  $E_n$  are independently and identically normally distributed with mean zero and variance  $\sigma^2$  so that (10) is equivalent to

$$f_{\boldsymbol{\theta}}(\boldsymbol{\theta}_n; \mathbf{Y}_n, b, \sigma^2) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left[-\frac{1}{2\sigma^2} (\boldsymbol{\theta}_n - \mathbf{Y}_n^T \boldsymbol{\beta})^T (\boldsymbol{\theta}_n - \mathbf{Y}_n^T \boldsymbol{\beta})\right], \quad (11)$$

a normal distribution with mean  $\mathbf{Y}_n^T \boldsymbol{\beta}$  and variance  $\sigma^2$ . If (11) is used as the population model then the parameters to be estimated are  $\boldsymbol{\beta}$ ,  $\sigma^2$ , and  $\xi$ .

The TIMSS scaling model takes the generalization one step further by applying it to the vector valued  $\boldsymbol{\theta}$  rather than the scalar valued  $\theta$  resulting in the multivariate population model

$$f_{\boldsymbol{\theta}}(\boldsymbol{\theta}_n; \mathbf{W}_n, \boldsymbol{\gamma}, \boldsymbol{\Sigma}) = (2\pi)^{-\frac{D}{2}} |\boldsymbol{\Sigma}|^{-\frac{1}{2}} \exp\left[-\frac{1}{2} (\boldsymbol{\theta}_n - \boldsymbol{\gamma} \mathbf{W}_n)^T \boldsymbol{\Sigma}^{-1} (\boldsymbol{\theta}_n - \boldsymbol{\gamma} \mathbf{W}_n)\right], \quad (12)$$

where  $\gamma$  is a  $u \times D$  matrix of regression coefficients,  $\Sigma$  is a  $D \times D$  variance-covariance matrix and  $\mathbf{W}_n$  is a  $u \times 1$  vector of fixed variables. If (12) is used as the population model then the parameters to be estimated are  $\gamma$ ,  $\Sigma$ , and  $\xi$ . In TIMSS we refer to the  $\mathbf{W}_n$  variables as conditioning variables.

## 5.5 Estimation

The ConQuest software uses maximum likelihood methods to provide estimates of  $\gamma$ ,  $\Sigma$ , and  $\mathbf{x}$ . Combining the conditional item response model (6) and the population model (12) we obtain the unconditional or marginal response model,

$$f_{\mathbf{x}}(\mathbf{x}; \xi, \gamma, \Sigma) = \int_{\theta} f_{\mathbf{x}}(\mathbf{x}; \xi | \theta) f_{\theta}(\theta; \gamma, \Sigma) d\theta, \quad (13)$$

and it follows that the likelihood is,

$$\Lambda = \prod_{n=1}^N f_{\mathbf{x}}(\mathbf{x}_n; \xi, \gamma, \Sigma), \quad (14)$$

where  $N$  is the total number of sampled students.

Differentiating with respect to each of the parameters and defining the marginal posterior as

$$h_{\theta}(\theta_n; \mathbf{W}_n, \xi, \gamma, \Sigma | \mathbf{x}_n) = \frac{f_{\mathbf{x}}(\mathbf{x}_n; \xi | \theta_n) f_{\theta}(\theta_n; \mathbf{W}_n, \gamma, \Sigma)}{f_{\mathbf{x}}(\mathbf{x}_n; \mathbf{W}_n, \xi, \gamma, \Sigma)} \quad (15)$$

provides the following system of likelihood equations:

$$\mathbf{A}' \sum_{n=1}^N \left[ \mathbf{x}_n - \int_{\theta_n} E_{\mathbf{z}}(\mathbf{z} | \theta_n) h_{\theta}(\theta_n; \mathbf{Y}_n, \xi, \gamma, \Sigma | \mathbf{x}_n) d\theta_n \right] = \mathbf{0}, \quad (16)$$

$$\hat{\gamma} = \left( \sum_{n=1}^N \bar{\theta}_n \mathbf{W}_n^T \right) \left( \sum_{n=1}^N \mathbf{W}_n \mathbf{W}_n^T \right)^{-1}, \quad (17)$$

$$\text{and} \quad \hat{\Sigma} = \frac{1}{N} \sum_{n=1}^N \int_{\theta_n} (\theta_n - \gamma \mathbf{W}_n) (\theta_n - \gamma \mathbf{W}_n)^T h_{\theta}(\theta_n; \mathbf{Y}_n, \xi, \gamma, \Sigma | \mathbf{x}_n) d\theta_n, \quad (18)$$

$$\text{where} \quad E_{\mathbf{z}}(\mathbf{z} | \theta_n) = \Psi(\theta_n, \xi) \sum_{\mathbf{z} \in \Omega} \mathbf{z} \exp[\mathbf{z}'(\mathbf{b}\theta_n + \mathbf{A}\xi)] \quad (19)$$

$$\text{and} \quad \bar{\theta}_n = \int_{\theta_n} \theta_n h_{\theta}(\theta_n; \mathbf{Y}_n, \xi, \gamma, \Sigma | \mathbf{x}_n) d\theta_n. \quad (20)$$

The system of equations defined by (16), (17), and (18) is solved using an EM algorithm (Dempster, Laird, and Rubin, 1977) following the approach of Bock and Aitken (1981).

## 5.6 Latent Estimation and Prediction

The marginal item response (13) does not include parameters for the latent values  $\theta_n$  and hence the estimation algorithm does not result in estimates of the latent values. For TIMSS, expected a posteriori (EAP) estimates of each student's latent achievement were produced. The EAP prediction of the latent achievement for case  $n$  is

$$\theta_n^{EAP} = \sum_{r=1}^P \Theta_r h_{\Theta}(\Theta_r; \mathbf{W}_n, \hat{\xi}, \hat{\gamma}, \hat{\Sigma} | \mathbf{x}_n). \quad (21)$$

Variance estimates for these predictions were estimated using

$$\text{var}(\theta_n^{EAP}) = \sum_{r=1}^P (\Theta_r - \theta_n^{EAP})(\Theta_r - \theta_n^{EAP})^T h_{\Theta}(\Theta_r; \mathbf{W}_n, \hat{\xi}, \hat{\gamma}, \hat{\Sigma} | \mathbf{x}_n). \quad (22)$$

## 5.7 Drawing Plausible Values

Plausible values are random draws from the marginal posterior of the latent distribution, (15), for each student. For details on the uses of plausible values the reader is referred to Mislevy (1991) and Mislevy, Beaton, Kaplan, and Sheehan (1992).

Unlike previously described methods for drawing plausible values (Beaton, 1987; Mislevy et al., 1992) ConQuest does not assume normality of the marginal posterior distributions. Recall from (15) that the marginal posterior is given by

$$h_{\theta}(\theta_n; \mathbf{W}_n, \xi, \gamma, \Sigma | \mathbf{x}_n) = \frac{f_{\mathbf{x}}(\mathbf{x}_n; \xi | \theta_n) f_{\theta}(\theta_n; \mathbf{W}_n, \gamma, \Sigma)}{\int_{\theta} f_{\mathbf{x}}(\mathbf{x}; \xi | \theta) f_{\theta}(\theta, \gamma, \Sigma) d\theta}. \quad (23)$$

The ConQuest procedure begins drawing  $M$  vector valued random deviates,  $\{\phi_{nm}\}_{m=1}^M$  from the multivariate normal distribution  $f_{\theta}(\theta_n, \mathbf{W}_n, \gamma, \Sigma)$  for each case  $n$ . These vectors are used to approximate the integral in the denominator of (23) using the Monte Carlo integration

$$\int_{\theta} f_{\mathbf{x}}(\mathbf{x}; \xi | \theta) f_{\theta}(\theta, \gamma, \Sigma) d\theta \approx \frac{1}{M} \sum_{m=1}^M f_{\mathbf{x}}(\mathbf{x}; \xi | \phi_{nm}) \equiv \mathfrak{S}. \quad (24)$$

At the same time the values

$$p_{mn} = f_{\mathbf{x}}(\mathbf{x}_n; \xi | \phi_{mn}) f_{\theta}(\phi_{mn}; \mathbf{W}_n, \gamma, \Sigma) \quad (25)$$

are calculated, so that we obtain the set of pairs  $\left(\phi_{nm}, \frac{p_{mn}}{\mathfrak{S}}\right)_{m=1}^M$  that can be used as an approximation to the posterior density (23); and the probability that  $\phi_{nj}$  could be drawn from this density is given by

$$q_{nj} = \frac{p_{mn}}{\sum_{m=1}^M p_{mn}}. \quad (26)$$



At this point  $L$  uniformly distributed random numbers,  $\{\eta_i\}_{i=1}^L$ , are generated, and for each random draw the vector,  $\Phi_{ni_0}$ , that satisfies the condition

$$\sum_{s=1}^{i_0-1} q_{sn} < \eta_i \leq \sum_{s=1}^{i_0} q_{sn} \quad (27)$$

is selected as a plausible vector.

## 5.8 Scaling Steps

The item response model described above was fit to the data in two steps. In the first step a calibration of the items on each scale was undertaken using all available students from the samples of the participating countries. During this step the items coded as not reached were treated as not administered or missing. In a second step the model was fitted separately to the data for each country with the item parameters fixed at values estimated in the first step. During the second step the model was fitted to all students having taken items in the corresponding scale. Items coded as not reached were treated as incorrect in this second step.

The international calibration samples from each country were weighted so that each country contributed proportionally the same to the estimates.

Because of the nature of the tests and the populations of interests, there were a total of 5 separate scalings that were undertaken. For the mathematics and science literacy tests, achievement was summarized on two separate scales computed by means of a multidimensional item response model—science literacy and mathematics literacy. The composite results for mathematics and science literacy represents an average of the results on the mathematics and science literacy scales. For the advanced mathematics test, achievement was summarized on three separate scales computed using a multidimensional item response model—numbers and equations, calculus, and geometry. The overall results for advanced mathematics were derived from all of the mathematics items scaled together using a unidimensional item response model. They include those items in the three content areas and 10 items that were not reported in those content areas. For the physics test, achievement was summarized on five separate scales computed using a multidimensional item response model—mechanics; electricity and magnetism; heat; wave phenomena; and modern physics. The overall results for physics were derived from all of the physics items scaled together using a unidimensional item response model.

## 5.9 Standardizing the International Scale Scores

The plausible values produced by the scaling procedure were in the form of logit scores that were on a scale that ranged generally between -3 and +3. For reporting purposes the scores for each of the scalings were mapped by a linear transformation onto an international scale with a mean of 500 and a standard deviation of 100. Each country was weighted to contribute the same when the international mean and standard deviation were set. The transformation applied to the logit scores was

$$S_{ijk} = 500 + 100 * \left( \frac{q_{ijk} - \bar{q}_j}{SD_{q_j}} \right)$$

where  $S_{ijk}$  is the standardized scale score for student  $i$ , in plausible value  $j$ , in country  $k$ ;  $\theta_{ijk}$  is the logit score for the same student,  $\bar{\theta}_j$  is the weighted average across all countries on plausible value  $j$ , and  $SD_{\theta_j}$  is the standard deviation across all countries on plausible value  $j$ . Note that while the scale for the standardized scale scores has a mean of 500 and a standard deviation of 100, the international student mean is not always exactly 500. This is because all the data was not available at the time of standardizing the scales.

Since the scale scores were actually plausible values drawn randomly for individual students from the posterior distribution of achievement, it was possible to obtain scores that were unusually high or low. If a transformed score was below 10, the score was recoded to 10, therefore making 10 the lowest score on the transformed scale. This happened in very few cases across the sample. If a transformed score was above 990, the score was recoded to 990, therefore making 990 the lowest score on the transformed scale.

Five plausible values were drawn for each student on each scale included in the test administered to the student. Each of the plausible values was transformed onto the international scale. The variation between results computed using each separate value reflects the error due to imputation. All five values are provided in the International Database for each scale and these should be used to obtain estimates of achievement for the students.

---

# Chapter 6

## Student Achievement Scores

The TIMSS international database for the final year of secondary school contains several student-level achievement scores. These scores were computed at different stages of the project to serve specific purposes. This chapter presents a description of these achievement scores, how they were derived, how they were used by TIMSS, and how users of the database can use them. For identification purposes, the first letter for the variable name identifies the population for which the score was computed. The scores computed for students in their final year of secondary school have the letter C as the first character in their name (The letters A and B were used to identify the scores for the students in the primary school years and in the middle school year, respectively). This convention was also followed with other background and derived variables and with the files included in the database.

### 6.1 Achievement Scores in the Student Files<sup>1</sup>

Four types of achievement scores are included in the student data files: raw scores, national Rasch scores, international EAP scores, and plausible values. Each type is described below.

#### 6.1.1 Raw Scores

CSLSCPT	Number of raw score points on the mathematics and science literacy items
CSMSCPT	Number of raw score points on the advanced mathematics items
CSPSCPT	Number of raw score points on the physics items

After the items were scored as right or wrong, or to their level of correctness in the case of the open-ended items, raw scores were computed by adding the number of points obtained across the items for the subject. Multiple-choice items received a score of either 1 or 0. Open-ended response items received score points from 0 to 3 depending on their coding guide. Open-ended items with a first digit of 7 or 9, indicating an incorrect/incomplete answer, were given zero points. The value of the first digit of the code for the open-ended items determines the number of score points assigned to it. A description of the algorithm used to score the items can be found in Chapter 9 in the section “Scoring the Items.”

Since students were assigned different booklets, booklets contain different numbers of items for each subject, and the specific items contained in each booklet had varying difficulty, raw scores should not be used to compare students’ performance. It is recommended that these scores be used only to verify whether the items have been recoded correctly when a user decides to recompute test scores. Raw scores can be found in the Student Background data files and in the Written Assessment data files.

<sup>1</sup> The organization of the data files included in the International Database is described in Chapter 7.

### 6.1.2 National Rasch Scores

CSLNRSC	National Rasch Mathematics and Science Literacy Score (ML)
CSMNRSC	National Rasch Advanced Mathematics Score (ML)
CSPNRSC	National Rasch Physics Score (ML)

The national Rasch scores were designed for preliminary analyses. They were computed by standardizing logit scores on each of the scales to have a weighted mean of 150 and standard deviation of 10 within each country. The national Rasch scores provided a basic Rasch score for conducting preliminary analyses within countries, but could not be used for international comparisons since each country has been assigned the same mean score. The logit scores were computed using the Quest Rasch analysis software. Quest provides maximum likelihood (ML) estimates, based on the Rasch model, for the performance of the students on a set of items. The computation of these scores took into account the varying difficulty of the items across test booklets, and the performance and ability of the students responding to each set of items. These logit scores were obtained using item difficulties that were computed separately for each country. When computing the item difficulties, responses marked as "not reached" were treated as not administered. This avoided giving inflated item difficulties to the items located at the end of the test in cases where students systematically did not reach the end of the test. The country specific item difficulties were then used to compute logit scores for each student.

Although "not reached" items were treated as not administered in the item calibration stage, when computing the student logit scores the responses marked as "not reached" were treated as incorrect. This avoided unfairly favoring students who started answering the test and stopped as soon as they did not know the answer to a question. Logit scores for the students generally ranged between -4 and +4. Given that when using maximum likelihood estimation procedures it is not possible to obtain finite logit scores for those students who correctly answered all or none of the items, scores for these students were set to +5 and -5 logits prior to standardizing the scales to the weighted mean of 150 and standard deviation of 10.

These national Rasch scores should not be used for international comparisons for two reasons: they were computed with a different set of item difficulties for each country, and the weighted mean score within each country is always equal to 150. National Rasch scores can be found in the Student Background data files and in the Written Assessment data files.

### 6.1.3 International Expected a Posteriori (EAP) Scores

CSLEAP0	Mathematics and Science Literacy EAP Score
CSLEAP1	Mathematics Literacy EAP Score
CSLEAP2	Science Literacy EAP Score
CSMEAP0	Advanced Mathematics EAP Score
CSMEAP1	Numbers, Equations, and Functions EAP Score
CSMEAP2	Calculus EAP Score
CSMEAP3	Geometry EAP Score

CSPEAP0	Overall Physics EAP Score
CSPEAP1	Mechanics EAP Score
CSPEAP2	Electricity and Magnetism EAP Score
CSPEAP3	Heat EAP Score
CSPEAP4	Wave Phenomena EAP Score
CSPEAP5	Modern Physics EAP Score

With IRT scales composed of large numbers of items, say 30 or more, the expected a posteriori (EAP) score (see Chapter 5) is usually a sufficiently reliable indicator of a student's proficiency on the latent attribute, or ability. However, when, as in TIMSS, an individual student responds to just a few items on each scale, EAP scores are unreliable estimates of ability, and do not provide consistent estimates of population parameters or of their error variance. International expected a posteriori (EAP) scores are the average of the distribution from which the corresponding plausible values are drawn. The EAP scores can be used if scores for individual students are necessary, but should not be used to obtain population estimates of proficiency. These scores are included in the Student Background and the student Written Assessment files.

#### 6.1.4 Plausible Values

CSLPV01 – CSLPV05	Mathematics and Science Literacy Plausible Value 1 to 5
CSLPV11 – CSLPV15	Mathematics Literacy Plausible Value 1 to 5
CSLPV21 – CSLPV25	Science Literacy Plausible Value 1 to 5
CSMPV01 – CSMPV05	Advanced Mathematics Plausible Value 1 to 5
CSMPV11 – CSMPV15	Numbers, Equations, and Functions Plausible Value 1 to 5
CSMPV21 – CSMPV25	Calculus Plausible Value 1 to 5
CSMPV31 – CSMPV35	Geometry Plausible Value 1 to 5
CSPPV01 – CSPPV05	Overall Physics Plausible Value 1 to 5
CSPPV11 – CSPPV15	Mechanics Plausible Value 1 to 5
CSPPV21 – CSPPV25	Electricity and Magnetism Plausible Value 1 to 5
CSPPV31 – CSPPV35	Heat Plausible Value 1 to 5
CSPPV41 – CSPPV45	Wave Phenomena Plausible Value 1 to 5
CSPPV51 – CSPPV55	Modern Physics Plausible Value 1 to 5

These are the international mathematics and science achievement scores used to report achievement at the international level. It is recommended that these scores be used for both international and within-country comparisons. Not only do they allow for comparisons across countries, but they also take into account the specific difficulty of the items attempted by each student and their relative difficulty internationally, and reflect the measurement error component.

As described in Chapter 5, TIMSS made use of multiple imputation or “plausible values” methodology to provide estimates of student proficiency in the different achievement scales. The need for plausible values arises from the fact that any student was administered only a fraction of the items in the assessment, as described in Chapter 2. Time constraints during the test administration did not allow for all the items to be administered to each student. Because of the error involved in the imputation process,

TIMSS produced not one but five imputed values for each student in the different achievement scales. A plausible value is an estimate of how the individual student would have performed on a test that included all items in the assessment (see Chapter 5). Since no student responded to all items, this estimate is based on the responses to the items that were included in the test booklet that the student actually took and the performance of students with similar characteristics.

Plausible values have been shown to provide better estimation of population parameters than EAP scores. They were developed during the analysis of the 1983-84 NAEP data. The general theory of the NAEP plausible values can be attributed to Mislevy (Mislevy and Sheehan, 1987; 1989) based on Rubin's work (Rubin, 1987) on multiple imputations.

Within a subject area and across the sample, one set of plausible values can be considered as good as another. Each of these sets is equally well designed to estimate population parameters, although the estimates will differ somewhat when using different plausible values. This difference is attributable to imputation error. Five sets of plausible values are provided so that analyses may be replicated as many as five times. Results which vary greatly from replication to replication may be influenced by unreliability in the achievement measures, and considered to be suspect. In the international analyses for the TIMSS primary and middle school populations, the reliability of the achievement measures was sufficiently high that the imputation error could be ignored, and analyses were therefore based on just the first plausible value. However, in reporting achievement at the final year of secondary school the imputation error could not be ignored, and consequently, all five plausible values were used for each scale that was reported, and the error of the achievement estimate reflects the sampling variance as well as the imputation variance. The plausible values are included in the Student Background data files and in the Written Assessment files. Descriptive statistics for each of the plausible values, for each of the scales are presented in Tables 6.1 through Table 6.13.

## 6.2 Additional Scores Included in the Database

Twelve items in the mathematics and science literacy test were originally classified as measuring reasoning and social utility. In the analysis and reporting of the TIMSS results, six of the items were included in the mathematics literacy scale and six of the items were included in the science literacy scale. The raw scores, national Rasch scores, EAP scores, and plausible values for mathematics literacy and science literacy in the International Database include these items. For research purposes, however, the International Database also includes scores for a separate reasoning and social utility scale, comprising the twelve items. These scores are: raw scores (CSRSCPT), national Rasch scores (CSRNRSC), EAP Scores (CSREAP0), and plausible values (CSRPV01, CSRPV02, CSRPV03, CSRPV04, CSRPV05). Descriptive statistics for the plausible values for the reasoning and social utility scale are presented in Table 6.14.

**Table 6.1**  
**Descriptive Statistics for Mathematics and Science Literacy Plausible Values**

Country	Plausible Value 1			Plausible Value 2			Plausible Value 3			Plausible Value 4			Plausible Value 5		
	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.
Australia	524.7	95.1	839.1	524.5	93.7	805.3	525.1	96.4	874.3	524.0	94.4	815.8	524.8	94.7	873.2
Austria	519.4	80.3	792.0	517.9	79.2	789.9	519.0	80.8	787.7	519.7	81.8	814.4	519.1	80.0	786.5
Canada	525.8	83.1	872.1	526.7	83.0	844.6	525.1	82.5	855.4	525.6	83.1	825.0	525.0	84.3	869.5
Cyprus	446.3	73.0	702.7	447.1	72.6	759.8	447.6	73.7	760.9	446.8	73.3	708.5	445.7	72.8	697.0
Czech Republic	476.4	92.9	880.3	476.2	91.2	928.6	475.7	92.9	926.8	476.9	91.0	964.6	476.2	91.0	910.1
Denmark	528.2	80.7	782.0	528.5	81.7	784.6	527.6	81.5	788.4	528.4	81.1	775.8	527.9	80.8	754.2
France	504.3	73.6	757.3	504.7	73.8	707.4	505.9	74.1	728.5	505.3	74.7	713.5	505.9	74.0	721.2
Germany	495.8	89.6	792.5	494.7	89.1	822.8	497.8	88.7	855.7	494.1	88.9	790.3	496.3	88.4	828.3
Hungary	477.2	84.4	827.5	476.3	84.6	821.3	477.1	84.2	804.3	476.8	84.1	797.0	477.4	84.3	827.0
Iceland	540.7	77.5	809.9	541.1	77.5	794.6	541.2	77.5	813.5	542.1	76.8	778.5	541.3	77.9	798.2
Israel	484.5	123.4	843.6	482.6	122.4	795.0	483.3	123.6	836.1	483.9	124.4	947.0	484.3	123.2	850.4
Italy	475.3	81.7	795.9	474.6	83.5	812.9	475.1	82.6	791.3	475.6	83.6	747.3	476.4	84.9	827.2
Lithuania	464.9	81.0	716.5	465.4	80.5	747.7	465.4	80.6	703.4	464.4	80.1	712.7	464.9	79.8	747.6
Netherlands	558.3	84.4	801.8	559.7	84.9	831.3	558.7	83.8	807.0	559.7	83.8	822.3	557.7	83.9	802.5
New Zealand	524.6	91.5	888.5	524.4	92.6	818.7	525.4	91.2	793.4	525.3	92.0	832.0	525.8	92.7	838.6
Norway	536.6	87.9	845.9	535.8	88.0	917.4	536.3	88.1	859.6	535.3	87.7	860.6	535.8	87.2	908.9
Russian Federation	475.1	82.6	791.5	476.3	83.9	796.0	475.6	83.5	777.5	476.9	83.2	778.5	475.3	82.6	790.4
Slovenia	515.2	83.3	774.7	514.7	81.5	754.3	514.2	81.6	767.7	514.1	82.3	768.4	514.0	83.0	745.5
South Africa	353.0	87.1	761.1	352.7	87.2	798.5	351.9	87.3	760.0	352.4	88.7	770.0	352.1	87.7	765.9
Sweden	555.5	91.2	849.5	555.2	91.7	864.9	554.7	90.4	835.7	555.5	91.3	852.2	555.9	92.0	875.0
Switzerland	532.0	88.6	839.0	532.6	87.7	861.1	530.1	87.8	840.6	531.2	87.2	850.8	531.4	86.5	876.7
United States	470.9	88.7	923.5	471.0	89.5	852.4	470.7	89.4	822.4	469.9	89.1	821.7	470.8	88.8	817.0

**Table 6.2**  
**Descriptive Statistics for Mathematics Literacy Plausible Values**

Country	Plausible Value 1			Plausible Value 2			Plausible Value 3			Plausible Value 4			Plausible Value 5							
	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.					
Australia	521.7	96.2	192.7	847.5	96.3	232.5	820.8	521.9	99.3	219.4	858.1	522.2	97.1	244.1	818.8	522.4	97.1	202.8	866.9	
Austria	519.0	80.0	274.4	744.4	79.2	286.4	791.1	518.1	80.0	291.5	816.6	517.8	80.4	273.8	765.7	517.6	79.6	294.9	756.0	
Canada	519.3	89.7	243.7	835.5	90.0	237.2	816.2	518.9	89.8	233.6	871.2	519.1	90.0	248.6	832.1	518.4	91.1	251.9	881.8	
Cyprus	445.2	72.9	251.5	745.4	46.2	71.9	264.2	802.3	446.6	74.0	267.3	810.3	446.0	73.6	280.9	718.4	445.2	72.0	271.1	729.9
Czech Republic	465.5	100.7	225.7	962.8	98.9	222.8	990.0	465.6	100.7	232.7	990.0	467.3	98.1	230.7	990.0	465.8	98.6	250.3	990.0	
Denmark	547.5	86.9	238.2	845.6	87.0	269.3	812.1	546.6	87.2	252.0	820.3	547.7	85.8	252.1	811.4	546.9	86.8	284.6	814.8	
France	522.6	78.0	246.1	812.9	79.2	199.8	786.2	524.3	79.5	247.2	766.6	522.8	80.2	183.6	757.5	523.8	79.3	211.3	766.9	
Germany	495.2	95.1	149.9	827.8	93.4	184.4	858.5	497.6	93.3	175.7	873.8	493.3	93.5	146.3	889.4	495.4	93.3	156.0	844.8	
Hungary	483.6	92.1	210.1	873.7	92.4	230.9	823.5	483.3	92.2	199.5	816.5	483.2	92.2	220.6	808.0	483.9	92.4	216.7	839.7	
Iceland	533.1	88.1	276.6	851.2	88.3	290.5	831.5	533.5	87.2	320.4	820.5	534.3	86.9	302.2	789.4	533.7	88.7	270.5	839.5	
Israel	480.5	119.4	10.0	828.7	118.7	10.0	780.1	480.0	119.4	31.1	854.7	480.7	121.3	10.9	905.8	480.3	119.0	42.4	808.3	
Italy	475.4	86.1	205.7	804.4	87.9	198.7	801.5	475.0	86.5	192.4	781.4	476.5	87.7	182.1	745.0	477.4	89.1	200.1	820.0	
Lithuania	469.3	85.4	168.4	736.4	84.5	127.7	780.1	470.5	84.8	151.0	733.9	468.8	84.6	157.8	747.5	469.0	83.8	168.1	739.8	
Netherlands	559.6	89.7	302.8	807.5	90.7	300.9	857.7	560.0	89.8	307.0	813.0	560.9	89.6	301.4	882.9	559.1	90.0	296.4	829.3	
New Zealand	521.7	97.8	241.1	900.0	98.5	246.7	840.5	522.3	96.9	220.2	828.1	521.1	99.4	218.3	845.6	522.2	98.2	176.2	837.2	
Norway	528.9	94.3	236.1	831.4	94.8	228.0	894.4	528.5	93.8	234.2	831.8	527.8	93.9	218.9	876.3	527.8	92.9	239.2	867.3	
Russian Federation	470.2	85.6	180.0	765.5	85.8	242.0	821.7	469.6	85.8	184.0	820.0	472.0	85.5	238.1	772.4	470.3	84.7	210.6	784.0	
Slovenia	356.2	81.1	205.5	749.3	80.2	203.2	801.4	355.2	80.9	177.8	776.4	356.7	82.7	203.3	783.6	356.1	81.7	200.0	756.6	
South Africa	551.8	97.8	258.9	860.8	99.7	235.4	874.7	551.0	97.7	255.9	834.7	551.9	98.4	286.8	868.1	552.1	99.7	269.6	884.8	
Sweden	540.1	89.3	242.8	864.9	88.6	258.6	852.4	539.1	88.7	231.8	863.5	539.8	88.6	268.5	894.2	540.4	87.3	254.9	911.4	
Switzerland	461.8	90.3	184.0	884.0	90.9	209.0	817.9	460.9	91.5	177.3	819.2	460.1	91.6	204.6	797.0	461.2	91.1	185.5	808.1	
United States	512.4	87.6	251.5	757.5	86.3	254.3	790.4	511.6	85.3	205.4	769.5	510.6	86.7	239.7	755.5	511.5	88.0	230.4	739.8	



**Table 6.3**

**Descriptive Statistics for Science Literacy Plausible Values**

Country	Plausible Value 1			Plausible Value 2			Plausible Value 3			Plausible Value 4			Plausible Value 5		
	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.
Australia	527.8	100.7	884.6	526.4	98.5	848.6	528.2	100.5	946.1	525.8	98.9	844.9	527.2	99.4	893.6
Austria	519.8	87.0	873.5	518.4	85.4	877.8	520.0	87.8	940.9	521.6	88.7	898.5	520.7	86.2	817.0
Canada	532.3	85.1	908.8	533.5	84.4	902.2	531.3	83.7	871.5	532.1	84.4	838.8	531.7	86.2	887.8
Cyprus	447.4	82.9	726.3	448.0	82.7	734.0	448.5	82.9	758.7	447.5	82.3	776.1	446.2	83.7	764.4
Czech Republic	487.3	92.1	943.1	486.2	90.4	975.1	485.8	92.3	905.7	486.6	90.8	939.2	486.7	90.5	881.7
Denmark	508.9	85.9	795.3	509.3	87.7	787.5	508.5	86.7	820.0	509.2	87.8	786.8	508.8	86.4	795.6
France	486.0	78.8	738.5	485.7	78.8	699.6	487.5	78.5	763.9	487.8	78.9	720.1	487.9	78.9	714.5
Germany	496.3	91.0	800.8	496.1	91.5	818.4	498.0	90.7	865.5	494.9	90.7	803.1	497.3	89.8	811.9
Hungary	470.7	86.4	842.2	470.1	86.5	846.6	470.8	86.0	812.5	470.3	86.0	821.8	470.9	86.0	817.1
Iceland	548.4	75.5	781.8	548.6	74.8	776.9	549.0	76.4	806.6	549.9	74.7	812.6	549.0	75.5	791.0
Israel	488.4	134.4	892.5	485.9	133.4	846.2	486.7	135.3	855.0	487.1	134.6	988.2	488.3	134.7	927.3
Italy	475.2	85.0	787.5	474.1	86.9	824.2	475.1	86.4	802.8	474.7	87.5	765.4	475.4	87.8	834.3
Lithuania	460.4	84.1	757.8	461.0	84.5	755.0	460.3	84.3	750.2	460.0	83.5	738.7	460.8	83.5	755.3
Netherlands	556.9	86.1	867.8	559.5	86.4	849.6	557.3	85.1	810.9	558.5	85.0	820.1	556.3	85.1	825.5
New Zealand	527.5	94.0	877.0	528.2	95.0	809.3	528.5	94.5	819.1	529.5	92.6	838.7	529.5	95.9	880.5
Norway	544.2	91.2	916.2	543.6	90.9	940.5	544.2	91.9	900.8	542.8	90.8	866.6	543.9	91.4	950.5
Russian Federation	480.1	90.1	817.4	481.5	91.9	850.8	481.5	91.0	835.9	481.7	91.2	842.2	480.3	90.6	809.9
Slovenia	518.0	84.9	798.8	517.7	82.6	753.5	516.9	84.1	792.1	517.6	83.8	788.6	516.6	83.8	784.3
South Africa	349.9	99.1	772.8	349.3	99.9	795.7	348.6	99.3	770.0	348.2	100.7	756.3	348.2	99.2	782.0
Sweden	559.2	91.4	873.8	558.9	90.8	886.0	558.5	90.2	901.5	559.1	91.2	857.7	559.6	91.6	878.4
Switzerland	523.8	95.0	895.2	523.5	93.7	902.6	521.1	93.5	878.2	522.5	93.1	870.3	522.5	92.8	842.1
United States	480.0	93.9	963.0	480.4	94.9	886.8	480.5	94.0	832.5	479.7	93.6	892.0	480.5	93.1	862.8

**Table 6.4**  
**Descriptive Statistics for Advanced Mathematics Plausible Values**

Country	Plausible Value 1				Plausible Value 2				Plausible Value 3				Plausible Value 4				Plausible Value 5			
	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.
Australia	523.8	108.2	231.1	990.0	525.1	109.2	191.7	990.0	526.0	109.3	217.3	990.0	523.4	111.1	190.8	990.0	525.3	108.0	206.9	990.0
Austria	435.2	89.7	140.2	673.4	436.0	91.0	142.3	704.4	436.8	92.5	149.9	702.7	435.9	93.0	132.4	667.5	436.3	90.0	145.2	696.9
Canada	509.1	98.6	133.3	967.9	509.8	98.0	155.5	889.6	508.4	98.8	191.1	940.7	510.3	96.7	144.7	959.1	509.8	99.6	178.3	916.9
Cyprus	518.1	84.0	218.8	709.1	516.0	86.3	197.5	726.8	519.6	86.8	231.3	740.5	518.3	83.5	222.6	749.6	519.1	86.3	182.8	729.3
Czech Republic	470.4	106.8	101.3	962.8	469.8	105.7	105.3	963.3	468.0	105.9	149.4	919.2	467.5	106.8	156.1	976.0	468.7	106.1	142.3	933.2
Denmark	522.7	73.9	283.4	749.0	522.6	73.1	293.6	751.3	522.2	72.0	279.5	752.1	523.4	72.6	274.5	781.1	521.4	73.1	282.9	765.2
France	557.5	70.2	276.5	782.7	557.6	70.7	313.0	785.4	556.2	68.9	291.1	774.6	556.0	70.4	244.8	778.3	557.4	70.5	311.7	807.1
Germany	465.9	85.1	10.0	722.0	463.9	84.9	10.0	732.7	464.1	83.8	104.3	723.2	465.3	85.2	66.0	700.3	464.9	86.0	82.4	733.1
Greece	511.3	105.7	10.0	887.4	512.2	103.8	10.0	858.3	512.2	104.4	10.0	846.5	514.3	103.4	10.0	913.7	512.8	107.2	10.0	809.7
Israel	557.6	75.5	104.5	882.2	557.4	76.3	155.0	859.8	555.9	76.2	88.8	872.6	558.0	77.7	70.8	903.3	557.5	75.5	149.0	860.3
Italy	474.9	94.5	93.9	745.4	474.5	94.6	71.7	738.0	473.5	96.0	84.0	752.0	473.7	95.1	67.9	758.1	474.5	95.9	72.1	747.1
Lithuania	516.2	86.3	201.8	792.4	517.0	85.4	183.4	769.0	517.3	85.4	209.1	836.7	515.1	84.2	177.2	778.8	516.1	84.4	189.2	767.9
Russian Federation	542.3	112.0	219.9	928.4	542.1	112.3	189.3	972.2	540.8	110.0	258.0	921.5	543.7	111.7	232.3	949.5	542.3	112.0	205.9	979.8
Slovenia	474.6	93.5	146.7	857.9	473.7	93.4	139.2	823.6	475.9	93.4	138.8	877.8	475.1	94.8	136.0	860.5	474.2	93.8	175.3	849.7
Sweden	512.2	86.6	183.3	788.0	512.9	86.7	209.6	839.7	511.6	86.8	215.3	791.8	510.9	85.4	172.0	789.9	512.3	83.4	234.9	775.2
Switzerland	533.5	90.2	238.6	990.0	532.0	89.2	167.4	990.0	534.0	92.6	247.4	990.0	534.5	90.7	198.0	990.0	531.7	89.6	250.5	990.0
United States	441.2	97.0	63.2	831.3	443.5	99.6	10.6	861.3	441.9	97.8	87.8	852.3	441.1	97.9	105.6	838.4	442.4	97.6	72.9	843.5

**Table 6.5**  
**Descriptive Statistics for Numbers, Equations, and Functions Plausible Values**

Country	Plausible Value 1				Plausible Value 2				Plausible Value 3				Plausible Value 4				Plausible Value 5			
	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.
Australia	517.8	96.6	134.3	990.0	517.0	99.8	112.9	990.0	517.1	96.6	153.6	990.0	519.5	99.2	177.6	990.0	516.0	98.0	185.9	990.0
Austria	411.8	91.1	44.2	689.2	413.1	91.6	88.1	656.2	410.8	91.0	94.1	659.7	411.9	88.7	105.0	675.2	412.5	94.6	18.0	644.8
Canada	513.5	86.0	147.5	845.2	512.5	83.6	172.6	888.9	511.6	84.5	168.5	806.3	511.3	83.7	154.1	834.7	511.5	83.9	163.5	829.9
Cyprus	508.1	91.9	148.9	776.7	508.5	92.9	236.7	776.3	510.1	92.9	144.6	750.7	510.6	93.2	200.1	811.0	511.7	92.9	208.3	747.6
Czech Republic	460.0	103.9	112.1	816.2	458.9	104.7	96.6	823.6	460.7	102.8	82.3	839.0	460.2	103.3	144.4	849.2	462.2	103.2	116.8	853.1
Denmark	504.3	61.4	233.3	674.3	504.1	61.1	186.5	708.4	503.9	62.5	251.1	701.5	503.8	62.5	245.9	701.7	502.7	61.2	266.5	688.2
France	548.0	55.5	202.9	752.2	548.0	56.6	238.4	763.0	548.6	55.3	222.7	749.7	548.4	58.0	198.0	745.7	548.6	55.8	192.4	782.7
Germany	457.0	79.6	10.0	711.1	455.8	78.1	27.5	665.6	457.3	80.5	46.4	702.2	457.9	81.7	17.2	674.9	457.0	80.8	10.0	691.6
Greece	540.4	113.9	10.0	990.0	540.3	112.7	10.0	990.0	538.6	113.5	10.0	990.0	535.3	111.3	17.1	975.6	540.4	112.2	33.4	990.0
Israel	548.4	78.2	24.7	899.1	545.5	76.8	45.8	871.2	546.8	74.8	121.0	874.2	548.5	76.5	100.6	861.9	545.9	77.6	42.2	940.6
Italy	460.2	103.6	10.0	773.5	459.8	103.1	10.0	750.6	459.9	102.7	10.0	750.3	460.0	104.0	10.0	758.6	459.4	102.6	10.0	762.6
Lithuania	546.1	85.1	164.1	813.8	548.8	86.0	130.7	829.7	549.2	83.0	117.1	821.6	546.1	84.2	196.7	781.9	546.7	84.6	148.9	782.5
Russian Federation	553.6	108.3	215.6	990.0	556.4	105.2	220.5	990.0	554.6	107.6	217.0	990.0	556.6	106.3	256.0	990.0	554.8	106.3	251.8	990.0
Slovenia	492.0	106.5	124.8	877.3	490.3	104.9	49.0	859.1	491.1	105.7	74.9	896.0	491.2	106.4	92.5	876.5	490.7	105.7	125.1	846.4
Sweden	523.1	88.6	209.5	863.3	523.0	87.4	176.4	793.6	524.2	87.3	240.8	816.5	522.5	90.1	268.7	805.0	523.3	88.9	207.5	818.6
Switzerland	513.1	86.9	152.7	990.0	513.1	87.4	157.7	990.0	514.3	88.3	122.8	990.0	515.5	88.0	61.6	990.0	514.0	87.0	120.1	990.0
United States	459.9	86.2	69.6	829.0	459.7	86.8	148.6	863.7	457.8	86.8	124.6	813.7	458.8	86.6	84.5	884.0	457.8	86.5	139.9	836.4

**Table 6.6**  
**Descriptive Statistics for Calculus Plausible Values**

Country	Plausible Value 1				Plausible Value 2				Plausible Value 3				Plausible Value 4				Plausible Value 5			
	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.	Mean	S.D.	Min.	Max.
Australia	530.3	100.2	176.8	990.0	529.3	100.8	178.6	990.0	530.0	99.6	146.2	990.0	531.1	100.2	212.8	990.0	528.5	100.8	186.3	990.0
Austria	439.2	88.9	118.1	824.5	438.5	87.9	108.2	794.1	440.2	89.9	112.9	779.5	437.9	87.7	116.8	721.6	440.7	90.2	55.4	761.5
Canada	505.2	94.0	153.9	984.2	503.3	92.0	139.6	947.2	502.8	93.0	128.9	919.5	503.3	91.8	180.8	924.0	502.4	93.2	157.7	861.6
Cyprus	558.8	97.4	222.3	867.9	559.7	99.8	194.4	867.2	561.4	100.7	236.2	836.3	561.7	101.4	204.9	897.6	561.5	100.0	222.6	845.0
Czech Republic	445.2	97.4	184.5	990.0	445.3	97.1	156.7	990.0	445.6	97.4	177.8	990.0	446.2	97.2	213.1	990.0	446.7	97.7	207.0	990.0
Denmark	508.5	84.5	206.4	768.9	508.6	86.3	222.2	797.7	508.1	85.5	171.8	812.8	508.3	85.6	187.1	752.4	507.9	86.1	182.3	860.1
France	560.4	63.8	354.0	766.0	559.1	65.7	288.0	755.9	559.9	62.6	307.8	748.6	558.9	64.8	343.6	771.1	559.4	63.2	364.1	785.1
Germany	454.8	84.7	132.2	740.7	453.7	83.8	57.4	672.5	453.3	84.5	142.6	728.1	454.6	86.3	96.7	695.9	453.9	85.8	76.6	702.1
Greece	539.4	101.3	85.4	800.4	540.4	98.9	118.1	831.7	538.1	98.1	78.8	820.3	534.2	96.8	134.4	769.1	540.1	96.2	72.6	823.5
Israel	538.8	69.5	240.3	839.3	536.4	68.1	255.4	872.4	537.9	66.3	321.7	814.5	539.2	69.2	299.6	856.9	536.3	68.8	233.6	812.0
Italy	520.1	107.2	183.9	849.8	520.1	107.8	191.3	883.9	520.5	107.4	166.3	895.8	520.5	107.0	179.2	872.9	519.5	107.1	164.6	881.0
Lithuania	497.0	78.6	173.3	720.6	500.5	79.7	238.1	796.7	498.7	76.6	216.6	748.6	498.5	77.8	209.5	781.8	497.2	77.6	212.8	763.2
Russian Federation	535.9	108.2	196.1	990.0	538.8	104.2	229.7	990.0	536.1	106.8	192.2	990.0	537.9	106.4	197.5	990.0	537.0	106.5	200.9	990.0
Slovenia	472.1	71.1	201.0	736.3	470.4	70.5	206.7	754.0	470.7	71.0	182.8	747.7	471.7	72.2	227.2	748.3	471.2	71.0	214.5	754.0
Sweden	480.2	87.3	171.0	819.0	480.3	88.2	122.2	818.4	481.1	88.5	208.2	802.0	479.4	88.2	183.1	853.9	480.9	87.7	175.1	786.8
Switzerland	511.5	97.2	10.0	990.0	511.6	95.1	10.0	990.0	513.6	97.6	10.0	990.0	513.6	97.4	10.0	990.0	512.2	95.6	10.0	990.0
United States	449.7	96.1	133.1	868.4	448.8	97.1	96.7	822.5	448.9	96.2	185.7	841.8	451.0	97.6	94.5	828.9	449.3	99.7	141.4	793.6

**Table 6.7**  
**Descriptive Statistics for Geometry Plausible Values**

Country	Plausible Value 1			Plausible Value 2			Plausible Value 3			Plausible Value 4			Plausible Value 5		
	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.
Australia	497.2	120.4	184.0 990.0	495.0	123.2	96.1 990.0	497.1	123.6	83.7 990.0	498.1	123.7	132.0 990.0	493.9	121.9	70.6 990.0
Austria	461.8	98.7	114.0 785.8	462.2	96.3	116.6 765.7	462.3	96.3	120.0 813.9	461.9	97.9	75.5 775.6	461.0	99.2	126.7 781.2
Canada	500.4	96.2	66.7 844.9	499.9	94.7	127.7 943.6	498.9	94.6	105.6 833.4	498.4	93.7	138.0 852.6	499.3	95.7	90.1 820.0
Cyprus	515.0	98.0	206.4 841.4	516.4	101.2	204.8 866.8	516.3	99.3	200.2 825.2	518.3	98.1	205.2 830.0	518.5	99.1	209.9 805.4
Czech Republic	494.2	102.4	227.4 990.0	493.3	102.8	238.3 990.0	494.3	101.4	220.4 990.0	494.5	102.0	218.3 990.0	495.5	102.2	228.4 990.0
Denmark	526.1	69.7	260.4 782.7	527.8	69.6	203.4 751.9	526.8	71.3	193.9 750.1	528.7	71.8	196.5 744.5	526.6	69.9	281.8 755.3
France	543.9	75.5	152.8 749.2	543.9	78.4	183.7 768.4	544.8	76.1	159.5 803.0	542.9	77.5	146.5 793.5	544.9	76.5	160.8 789.3
Germany	487.2	75.1	64.9 698.0	486.1	74.2	110.6 711.9	486.9	75.0	121.3 701.5	487.2	76.9	87.8 706.1	486.9	74.4	101.0 692.5
Greece	500.1	119.9	47.9 804.3	499.3	115.5	10.0 797.9	495.9	116.2	17.8 850.7	493.2	114.5	25.8 801.0	499.7	115.6	35.3 788.2
Israel	562.4	73.9	158.0 874.5	560.8	74.0	147.1 862.5	562.0	71.2	233.9 864.8	563.8	73.1	201.4 867.5	560.6	74.0	117.0 838.9
Italy	480.0	103.7	47.9 784.4	480.2	103.7	43.6 790.1	480.3	104.5	25.2 784.7	479.9	103.4	31.6 786.0	479.1	103.5	43.3 773.7
Lithuania	514.7	81.4	136.0 780.9	515.2	81.8	124.2 820.6	517.8	82.0	129.6 774.2	513.9	84.4	104.6 765.2	514.8	82.0	135.4 779.6
Russian Federation	546.5	105.9	217.0 974.4	549.8	102.3	223.7 990.0	546.9	104.3	237.9 990.0	549.5	102.2	269.1 980.5	548.3	103.3	231.6 980.6
Slovenia	477.7	83.5	100.1 820.9	475.6	83.3	144.7 819.5	476.3	83.1	74.8 873.0	475.9	83.9	108.3 817.5	476.4	83.5	107.6 849.1
Sweden	492.4	82.8	187.9 861.7	492.3	83.7	185.0 860.0	493.1	82.9	193.6 837.0	491.7	82.5	211.7 806.7	492.5	83.0	192.0 849.2
Switzerland	546.2	87.5	223.1 990.0	546.8	85.8	247.0 990.0	547.1	87.2	202.1 990.0	548.8	86.7	168.3 990.0	547.2	88.0	190.8 990.0
United States	423.9	95.4	44.4 852.7	423.6	96.9	52.7 901.3	422.9	96.8	10.0 904.8	424.6	96.6	10.0 887.4	423.4	96.8	20.2 856.4

**Table 6.8**  
**Descriptive Statistics for Physics Plausible Values**

Country	Plausible Value 1			Plausible Value 2			Plausible Value 3			Plausible Value 4			Plausible Value 5					
	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.			
Australia	514.8	81.8	790.5	517.8	81.7	787.8	518.7	84.5	297.7	806.0	518.2	80.9	277.9	815.8	520.0	82.2	300.0	816.1
Austria	434.9	83.7	740.7	435.6	83.7	754.5	433.7	83.2	232.4	797.5	434.8	84.0	176.7	764.0	434.2	81.7	192.7	767.9
Canada	483.5	84.4	919.6	487.2	89.0	990.0	486.9	87.6	191.9	945.0	484.3	86.9	234.0	990.0	485.0	85.7	226.2	990.0
Cyprus	495.0	105.1	919.4	493.9	104.8	895.6	493.3	104.7	137.6	917.5	494.4	105.8	104.5	908.3	495.1	106.2	148.5	932.3
Czech Republic	450.8	82.8	856.1	450.8	81.7	851.8	451.4	81.5	213.5	846.5	451.7	82.0	220.9	867.8	452.3	83.0	166.8	853.0
Denmark	534.0	84.4	786.8	534.9	82.7	790.3	535.2	85.8	298.7	783.7	535.4	84.1	306.8	815.8	532.4	85.9	274.2	837.3
France	465.1	64.4	673.3	466.3	68.5	707.9	466.5	65.5	215.3	674.0	468.5	64.4	209.1	678.8	465.2	66.6	187.6	682.3
Germany	520.6	96.9	800.0	522.5	93.8	777.9	522.8	94.2	270.3	791.3	522.9	92.4	249.8	779.3	520.8	93.5	297.9	783.1
Greece	487.3	86.3	762.8	485.4	87.5	830.5	487.9	87.1	125.9	780.0	486.4	87.8	120.9	791.4	485.4	87.9	149.7	748.5
Israel	505.4	86.5	755.5	507.2	89.1	776.1	504.4	86.3	29.6	776.5	506.7	86.2	43.6	796.7	504.6	88.3	44.8	782.6
Italy	436.0	79.7	661.7	435.1	80.8	670.6	435.1	80.6	229.3	662.4	437.1	80.0	235.3	685.7	436.2	79.7	249.2	673.1
Latvia (LSS)	489.0	101.6	875.9	487.6	99.9	938.4	488.2	100.8	302.6	884.3	487.9	100.5	261.6	929.0	488.2	98.8	244.7	880.3
Norway	581.0	90.7	875.2	578.0	90.9	926.3	580.2	90.5	251.1	953.4	581.8	90.9	335.2	894.8	583.7	89.8	354.8	866.1
Russian Federation	548.5	109.8	990.0	546.0	110.2	972.4	543.1	109.2	248.5	923.4	544.0	111.3	192.6	948.8	544.4	110.9	224.6	990.0
Slovenia	524.4	108.2	884.5	523.4	107.6	875.2	523.8	108.2	135.6	841.0	518.8	111.1	115.2	868.1	522.8	109.6	126.8	871.3
Sweden	572.8	92.3	865.7	573.8	92.6	887.9	572.0	91.8	264.7	818.3	573.5	92.2	272.3	857.0	572.6	91.4	238.0	909.1
Switzerland	489.0	88.6	797.4	488.3	89.8	808.3	487.6	87.7	249.9	778.2	488.8	87.9	233.8	816.1	487.4	87.6	219.4	812.9
United States	422.3	60.2	819.3	422.2	59.9	801.5	422.2	60.6	136.4	845.5	422.2	59.1	160.1	794.2	423.7	59.7	151.1	810.5

**Table 6.9**  
**Descriptive Statistics for Mechanics Plausible Values**

Country	Plausible Value 1			Plausible Value 2			Plausible Value 3			Plausible Value 4			Plausible Value 5							
	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.	Mean	S.D.	Max.					
Australia	505.1	89.5	207.9	896.8	507.8	86.5	191.2	904.8	506.9	86.3	219.2	902.3	507.8	89.0	215.7	895.1	506.7	86.8	205.8	909.7
Austria	420.6	77.9	209.2	742.5	420.6	78.0	213.4	722.0	420.5	77.6	178.3	745.3	420.4	78.9	193.6	733.6	420.2	77.5	202.2	742.1
Canada	473.4	90.0	211.8	990.0	473.5	90.9	200.4	990.0	474.0	88.5	149.7	990.0	473.7	87.7	163.0	990.0	472.1	88.0	198.9	990.0
Cyprus	530.2	117.3	48.1	990.0	529.8	117.0	70.9	990.0	530.3	117.1	55.3	990.0	530.0	117.0	46.5	990.0	530.8	116.1	61.4	990.0
Czech Republic	468.7	80.8	76.9	990.0	468.9	80.7	168.7	990.0	467.8	80.2	83.6	990.0	469.5	79.9	108.7	990.0	467.9	80.9	92.1	990.0
Denmark	528.9	86.3	198.0	883.2	529.6	88.5	214.0	904.3	529.0	87.0	203.2	927.4	528.4	87.9	169.5	904.4	531.2	86.8	183.6	879.0
France	459.2	73.4	159.5	682.2	456.8	72.8	170.6	736.0	456.6	75.2	134.7	719.4	454.6	76.1	130.3	689.3	456.9	76.5	167.1	715.0
Germany	495.9	91.3	103.6	759.4	495.6	90.1	120.1	747.4	496.0	89.6	119.1	779.1	493.9	87.9	165.9	736.6	492.5	92.6	54.7	765.6
Greece	512.6	91.0	94.0	833.9	512.9	91.1	87.9	782.8	515.3	91.1	135.8	792.4	512.1	90.6	72.5	788.4	515.6	89.9	129.2	790.1
Israel	547.6	86.9	83.9	990.0	548.1	85.4	82.1	990.0	548.0	85.8	112.4	990.0	547.7	85.4	103.4	990.0	547.3	86.3	72.0	990.0
Italy	421.1	101.7	10.0	734.0	421.6	95.6	61.1	718.8	418.6	106.1	51.7	743.4	416.0	100.4	10.0	723.7	424.1	103.9	53.1	779.8
Latvia (LSS)	489.2	90.5	271.4	990.0	488.3	93.2	264.6	990.0	488.7	92.3	210.9	990.0	488.2	91.9	264.7	990.0	489.5	89.9	231.7	990.0
Norway	568.9	90.0	214.4	885.2	573.9	87.7	253.5	898.4	569.9	89.3	281.0	938.1	572.6	90.7	249.0	976.8	572.9	89.7	275.5	924.0
Russian Federation	538.1	91.3	288.2	943.9	537.0	91.9	278.6	914.2	537.6	90.9	299.7	956.7	537.9	91.0	307.9	929.6	536.9	91.6	311.9	982.7
Slovenia	554.2	122.3	10.0	922.9	550.4	119.6	10.0	897.9	551.1	121.6	10.0	895.1	553.9	118.3	10.0	885.9	551.8	119.3	61.7	899.9
Sweden	563.2	78.9	215.2	833.2	562.7	80.1	242.8	811.5	563.5	81.4	210.8	815.1	563.6	81.1	208.9	811.2	563.0	80.5	218.2	854.4
Switzerland	481.9	85.3	151.7	897.8	482.6	86.6	168.8	919.5	482.1	86.7	196.8	895.3	482.6	85.9	169.7	922.2	482.2	86.7	195.3	958.6
United States	420.4	58.3	123.9	703.5	419.8	58.1	141.3	686.2	420.8	60.2	125.3	696.7	421.0	59.1	151.4	776.3	420.2	59.1	166.3	691.1

**Table 6.10**  
**Descriptive Statistics for Electricity and Magnetism Plausible Values**

Country	Plausible Value 1			Plausible Value 2			Plausible Value 3			Plausible Value 4			Plausible Value 5		
	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.
Australia	510.7	91.5	161.4 795.8	512.4	89.9	172.0 818.7	512.4	93.2	189.7 842.5	513.6	88.7	202.3 807.0	512.6	91.8	136.3 818.0
Austria	431.5	93.1	63.0 862.5	433.4	93.7	88.0 835.0	432.6	93.0	77.5 818.7	431.2	95.2	38.4 857.8	430.8	92.9	77.2 851.0
Canada	484.9	82.1	239.4 990.0	485.9	86.6	203.9 990.0	483.3	80.3	220.6 990.0	485.4	83.8	223.8 990.0	483.0	83.3	219.7 990.0
Cyprus	502.2	114.8	10.0 990.0	503.2	114.2	10.0 990.0	502.2	114.5	10.0 990.0	500.6	115.4	10.0 990.0	502.5	113.8	10.0 990.0
Czech Republic	465.2	74.8	177.5 837.4	465.6	75.3	233.0 837.8	464.0	75.2	173.8 850.2	465.3	75.3	202.3 810.1	464.8	75.3	191.9 807.3
Denmark	512.1	79.3	196.2 842.5	512.8	81.4	263.9 814.8	512.7	80.2	217.8 836.5	510.9	77.6	230.5 795.7	514.6	78.4	210.8 760.4
France	494.1	60.0	227.2 710.6	494.0	59.4	243.5 715.8	493.8	59.5	223.4 696.6	493.1	60.7	279.6 700.8	493.6	59.8	289.4 703.5
Germany	515.1	91.6	78.4 804.7	511.3	93.0	110.4 772.7	513.3	92.5	118.8 789.8	509.7	88.6	99.8 787.9	511.2	94.2	82.3 801.2
Greece	519.5	104.7	117.0 830.0	517.7	106.3	106.5 821.9	519.2	103.7	125.0 859.3	520.0	105.0	124.3 801.5	522.0	105.7	117.3 881.0
Israel	556.3	106.5	10.0 885.7	557.5	104.6	10.0 893.4	556.5	104.8	10.0 876.8	556.5	104.8	10.0 871.9	556.2	105.6	10.0 885.0
Italy	468.4	96.2	131.3 758.9	467.1	108.0	182.9 762.3	470.9	103.7	96.6 806.4	477.8	103.0	102.6 759.1	478.7	98.3	157.4 708.9
Latvia (LSS)	484.5	94.7	189.4 893.3	483.5	94.8	249.8 883.1	486.2	93.4	225.6 886.7	485.0	95.1	215.7 877.3	485.3	93.4	250.0 908.8
Norway	563.6	93.3	272.4 841.0	567.3	91.5	308.9 853.8	563.1	93.0	289.9 898.1	564.4	94.6	245.5 906.2	565.2	92.4	224.7 898.3
Russian Federation	550.3	107.3	210.7 990.0	546.8	108.0	175.2 990.0	550.3	106.9	180.5 990.0	550.3	106.6	175.3 990.0	548.2	107.0	223.0 990.0
Slovenia	509.6	111.5	49.1 990.0	507.3	108.3	10.0 955.5	508.5	111.7	35.7 936.9	510.9	109.7	86.9 973.8	508.3	108.4	103.2 961.8
Sweden	569.7	87.4	182.0 839.0	569.1	87.1	176.1 863.3	570.3	87.4	206.9 865.0	570.8	90.1	148.9 892.6	569.7	88.4	186.4 878.9
Switzerland	479.0	92.9	54.8 937.9	480.2	93.0	98.3 938.4	480.8	94.5	46.4 951.9	479.4	94.2	58.4 969.1	479.9	94.1	53.8 934.2
United States	419.1	58.0	161.3 746.2	420.5	58.7	178.2 746.4	419.0	59.1	151.9 743.1	420.2	58.4	150.5 767.3	419.3	59.0	122.7 721.3



**Table 6.11**  
**Descriptive Statistics for Heat Plausible Values**

Country	Plausible Value 1			Plausible Value 2			Plausible Value 3			Plausible Value 4			Plausible Value 5		
	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.
Australia	515.0	83.7	135.8 829.9	517.4	82.5	133.7 809.8	517.3	80.8	156.3 803.1	517.8	83.3	161.1 810.4	517.7	82.3	152.3 810.9
Austria	443.7	94.3	10.0 847.5	445.6	92.8	10.0 856.0	445.2	91.9	10.0 839.6	444.5	93.5	10.0 876.6	443.7	92.4	10.0 899.0
Canada	508.3	84.3	210.7 900.0	506.8	86.8	222.8 980.3	510.5	83.9	186.7 926.1	507.3	86.0	159.2 875.9	505.6	85.4	220.7 924.4
Cyprus	475.8	146.5	10.0 958.8	476.3	145.2	10.0 942.7	475.4	145.2	10.0 959.2	475.3	145.3	10.0 938.3	475.8	145.9	10.0 945.0
Czech Republic	487.8	80.6	10.0 990.0	488.5	80.2	71.4 990.0	487.0	79.1	10.0 990.0	488.6	79.8	10.0 990.0	488.4	80.9	10.0 990.0
Denmark	512.7	97.0	240.4 890.0	513.0	101.3	182.8 956.1	511.1	99.4	201.2 876.7	510.8	97.3	204.5 881.1	514.3	96.2	157.6 852.9
France	491.2	68.3	157.9 682.1	490.3	69.3	164.2 711.4	489.5	69.4	157.4 709.4	490.9	69.1	134.5 711.3	491.9	67.5	185.9 714.7
Germany	497.6	97.9	133.9 909.5	497.3	96.4	175.3 806.8	497.9	97.6	135.8 795.5	493.4	97.3	109.2 848.0	493.0	99.0	165.5 920.3
Greece	480.4	116.2	10.0 848.8	479.9	116.8	10.0 839.2	480.1	118.4	16.6 833.8	479.6	118.7	38.6 797.7	482.7	118.5	10.0 841.5
Israel	478.6	83.9	10.0 847.7	479.2	82.3	10.0 788.1	478.3	83.6	10.0 844.0	477.9	82.9	10.0 874.0	477.7	84.6	10.0 829.2
Italy	489.8	85.3	233.3 721.4	490.7	82.9	244.4 673.6	487.9	90.2	210.1 725.2	485.6	86.7	231.8 702.7	494.6	85.2	226.5 697.9
Latvia (LSS)	503.0	113.0	192.5 961.6	501.8	112.0	183.9 990.0	505.0	114.3	178.9 990.0	504.2	112.5	175.0 933.7	503.8	112.2	222.4 990.0
Norway	534.0	69.4	237.9 768.8	538.8	70.6	262.4 781.5	535.5	70.1	302.2 746.8	535.1	71.4	269.3 780.6	537.5	69.4	296.1 778.4
Russian Federation	531.5	105.1	119.8 990.0	528.6	106.6	157.2 990.0	531.4	104.9	140.8 990.0	531.6	104.5	137.5 990.0	529.2	105.7	140.0 990.0
Slovenia	521.9	118.6	10.0 990.0	519.0	116.6	10.0 990.0	518.8	118.8	10.0 990.0	522.9	116.3	10.0 990.0	520.9	118.6	10.0 990.0
Sweden	522.4	81.3	168.8 785.4	520.9	80.2	159.1 825.6	521.2	83.4	188.5 761.6	523.0	82.9	157.9 756.5	522.0	81.6	168.1 754.5
Switzerland	508.6	88.9	179.3 990.0	510.5	90.2	135.8 990.0	508.9	89.4	190.9 990.0	508.9	89.2	114.4 990.0	508.3	89.9	182.7 990.0
United States	477.2	59.2	89.2 791.6	475.8	58.3	136.0 778.5	477.0	58.6	171.6 806.9	477.4	58.1	162.7 807.0	476.7	57.9	146.4 773.5

**Table 6.12**  
**Descriptive Statistics for Wave Phenomena Plausible Values**

Country	Plausible Value 1					Plausible Value 2					Plausible Value 3					Plausible Value 4					Plausible Value 5				
	Mean	S.D.	Min.	Max.		Mean	S.D.	Min.	Max.		Mean	S.D.	Min.	Max.		Mean	S.D.	Min.	Max.		Mean	S.D.	Min.	Max.	
Australia	516.7	100.2	131.9	990.0		519.5	98.0	163.4	990.0		517.4	95.9	210.9	990.0		520.4	98.5	186.4	990.0		520.1	97.2	175.2	990.0	
Austria	468.4	87.4	169.1	785.9		467.8	87.4	165.8	775.6		468.2	84.5	182.3	791.1		467.5	88.3	158.6	767.9		467.3	86.4	171.3	798.4	
Canada	488.4	80.1	182.7	949.6		488.8	81.3	173.2	958.6		489.6	79.9	214.4	990.0		487.1	81.4	190.7	963.7		486.3	77.3	207.3	942.2	
Cyprus	507.3	119.2	108.5	990.0		506.9	118.7	146.4	986.5		506.9	119.0	127.4	990.0		506.5	119.3	146.7	990.0		507.4	119.1	137.4	990.0	
Czech Republic	446.9	76.2	208.9	732.4		447.3	75.8	199.2	726.7		446.3	74.7	183.9	730.4		447.1	75.2	222.9	732.2		446.8	77.6	168.2	713.3	
Denmark	536.7	97.6	10.0	862.0		537.8	99.0	104.9	919.6		535.2	95.8	96.5	907.4		534.9	97.8	69.4	889.4		538.6	95.6	10.0	920.8	
France	463.4	73.3	10.0	747.2		463.1	73.6	16.3	664.6		463.5	72.7	10.0	671.0		463.5	73.7	65.5	693.9		463.7	72.8	115.0	713.5	
Germany	531.6	98.9	227.3	888.9		531.5	97.3	220.4	855.2		531.8	96.9	180.3	794.0		527.9	94.3	211.6	797.7		527.8	101.0	249.8	795.9	
Greece	453.2	92.5	104.7	776.5		452.5	91.8	141.9	779.7		453.4	93.9	104.2	834.9		452.9	93.4	123.1	786.4		454.6	94.8	89.3	829.3	
Israel	444.1	92.4	10.0	990.0		444.8	90.8	10.0	990.0		444.2	92.3	14.5	990.0		443.6	91.0	10.0	990.0		443.7	92.7	10.0	990.0	
Italy	446.6	103.2	10.0	807.6		448.1	97.9	51.8	726.9		439.6	107.4	10.0	826.3		437.7	101.1	10.0	800.4		450.8	108.8	10.0	909.3	
Latvia (LSS)	498.7	91.1	271.2	950.3		497.4	89.6	276.1	948.2		498.2	91.9	212.2	947.9		497.0	91.2	262.7	988.2		499.1	91.9	243.5	954.7	
Norway	558.5	90.7	281.3	857.6		562.1	89.1	246.9	908.1		558.9	92.1	208.0	908.6		560.2	91.1	175.3	844.3		561.3	89.2	194.6	905.4	
Russian Federation	515.4	105.7	173.7	952.9		512.8	105.5	171.2	990.0		517.4	106.2	223.8	950.5		516.0	105.8	174.6	990.0		514.0	106.3	236.8	990.0	
Slovenia	514.2	115.1	10.0	906.5		512.4	115.9	10.0	930.9		512.3	117.6	10.0	897.1		515.3	113.4	10.0	858.9		513.5	113.8	10.0	888.4	
Sweden	559.8	106.8	69.6	954.5		559.1	106.3	160.2	939.2		560.2	108.7	155.5	893.5		561.1	108.6	177.4	944.2		560.1	107.2	126.2	932.4	
Switzerland	497.0	87.4	200.9	937.7		498.0	89.7	161.4	879.2		497.7	90.0	166.7	897.9		498.6	89.0	152.4	895.6		497.1	89.3	175.1	929.8	
United States	451.4	53.1	216.4	768.5		450.5	53.5	241.1	743.8		451.4	53.5	242.4	795.1		451.6	53.3	214.4	780.1		450.5	52.9	231.3	765.8	

**Table 6.13**  
**Descriptive Statistics for Modern Physics Plausible Values**

Country	Plausible Value 1			Plausible Value 2			Plausible Value 3			Plausible Value 4			Plausible Value 5		
	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.	Mean	S.D.	Min. Max.
Australia	518.8	89.0	165.6 867.0	521.2	86.6	218.1 858.7	521.1	88.9	209.7 885.1	521.8	88.0	178.2 854.9	521.5	87.7	196.2 838.4
Austria	479.6	84.1	144.2 890.7	481.2	82.2	135.0 915.1	479.7	80.7	159.3 913.9	478.4	83.4	114.1 942.8	480.9	82.3	119.9 897.2
Canada	495.1	77.1	204.0 785.2	494.5	80.7	138.7 806.2	494.4	81.1	185.0 791.9	493.9	81.7	184.8 839.2	494.0	80.4	205.8 812.6
Cyprus	434.8	131.7	10.0 818.5	434.3	130.7	10.0 787.5	434.2	131.3	10.0 801.1	433.9	131.5	10.0 791.0	434.4	131.3	10.0 808.6
Czech Republic	453.5	86.9	10.0 829.2	454.1	87.2	108.8 933.9	452.1	86.9	10.0 855.5	454.3	87.1	26.3 888.7	453.0	88.3	10.0 896.9
Denmark	542.7	81.4	106.1 795.7	543.6	83.8	86.9 793.5	543.5	80.7	105.8 773.6	542.9	80.0	114.4 794.5	545.4	81.1	87.6 801.2
France	473.1	61.7	217.4 700.8	473.8	59.6	188.2 684.9	474.8	61.3	213.0 671.0	474.1	60.7	209.2 671.8	474.4	61.2	223.3 663.7
Germany	546.8	107.1	220.8 876.5	546.2	107.3	155.2 931.9	546.7	109.0	167.2 897.9	543.5	106.5	200.8 870.3	541.8	108.2	172.3 886.3
Greece	445.8	92.8	10.0 753.3	446.2	93.4	10.0 765.0	447.3	93.6	10.0 756.2	446.5	92.1	10.0 752.8	446.9	92.3	10.0 784.7
Israel	475.8	80.4	16.7 714.6	476.4	79.6	42.5 749.5	476.4	79.7	61.5 728.6	475.6	79.6	30.4 760.9	475.4	78.2	50.6 738.6
Italy	422.8	93.8	10.0 632.0	418.5	94.7	10.0 636.1	418.4	99.9	10.0 661.5	422.0	96.6	10.0 628.0	423.1	97.5	10.0 626.6
Latvia (LSS)	488.6	96.3	255.0 940.6	486.6	95.7	218.2 976.2	488.4	93.6	252.4 926.6	486.1	95.3	262.6 990.0	488.4	95.5	227.6 975.6
Norway	574.7	84.2	179.6 947.8	577.4	85.7	99.2 946.8	574.4	84.3	127.4 980.6	576.3	85.0	153.0 945.5	576.8	81.0	140.9 990.0
Russian Federation	543.2	98.4	109.3 990.0	540.5	97.1	147.9 980.7	543.3	96.8	210.9 985.5	543.5	97.9	189.6 975.8	541.6	99.8	188.4 990.0
Slovenia	513.0	114.1	130.3 801.9	509.2	111.6	85.9 785.1	509.6	115.0	114.1 841.8	512.8	112.5	115.6 790.0	511.2	111.0	142.6 807.7
Sweden	559.7	75.0	274.3 766.0	558.7	77.1	281.6 759.7	559.8	77.2	299.4 780.1	560.6	78.1	280.3 782.8	560.5	78.0	305.7 771.5
Switzerland	486.9	82.4	69.8 990.0	488.5	84.4	88.1 990.0	488.0	83.0	116.9 990.0	488.6	82.8	118.5 990.0	487.3	84.1	157.5 990.0
United States	456.7	48.2	301.5 740.6	457.0	50.1	298.1 724.2	456.2	49.9	303.7 777.8	456.2	48.9	272.2 757.4	455.5	49.5	300.9 717.1

**Table 6.14**  
**Descriptive Statistics for the Reasoning and Social Utility Plausible Values**

Country	Plausible Value 1					Plausible Value 2					Plausible Value 3					Plausible Value 4					Plausible Value 5				
	Mean	S.D.	Min.	Max.		Mean	S.D.	Min.	Max.		Mean	S.D.	Min.	Max.		Mean	S.D.	Min.	Max.		Mean	S.D.	Min.	Max.	
Australia	523.6	91.6	176.9	806.7		523.1	91.6	238.1	898.4		523.9	92.3	199.3	832.6		525.6	91.4	190.3	852.6		526.1	92.2	247.3	877.5	
Austria	514.7	83.5	193.5	827.4		516.7	84.3	236.5	822.1		515.9	85.2	195.9	922.4		514.9	82.5	210.1	801.7		515.7	86.6	251.0	805.9	
Canada	531.8	80.2	189.6	832.1		531.1	79.4	224.2	861.0		530.5	80.6	195.5	811.7		529.9	80.4	249.7	858.6		530.2	79.7	237.6	809.0	
Cyprus	450.6	83.1	238.4	722.5		455.1	82.1	202.2	723.9		453.2	81.2	197.9	724.8		453.6	79.4	176.7	722.3		452.6	83.8	152.0	699.0	
Czech Republic	484.4	92.5	192.2	860.6		482.1	92.3	213.2	939.9		483.6	92.1	211.8	923.9		483.7	92.7	248.0	929.0		484.3	90.5	253.6	959.6	
Denmark	536.6	72.3	226.4	736.6		536.4	72.8	281.0	749.6		535.6	73.5	270.8	761.6		536.6	72.4	278.1	780.2		535.2	73.1	266.0	748.3	
France	511.8	67.0	215.1	710.4		512.3	68.1	190.4	728.5		509.6	67.2	154.5	744.3		511.9	68.4	151.2	735.7		510.5	67.3	180.9	715.8	
Germany	499.6	93.9	114.2	854.6		495.3	93.4	90.9	843.6		497.4	93.2	121.8	798.6		497.1	95.7	135.5	832.0		498.5	93.4	173.8	783.4	
Hungary	487.2	84.9	188.5	827.1		486.4	86.1	188.8	791.4		486.2	84.4	196.9	785.8		485.9	84.6	231.3	770.7		485.7	84.2	185.9	776.9	
Iceland	551.0	74.7	342.8	823.5		550.5	74.6	342.6	880.7		551.7	73.9	311.4	813.6		549.6	73.8	291.3	808.5		549.8	72.5	325.3	788.9	
Israel	484.8	126.8	21.2	856.8		485.0	126.1	10.0	839.2		485.9	126.7	10.0	813.5		485.2	129.8	10.0	876.6		484.1	126.2	10.0	808.0	
Italy	466.5	89.2	142.5	752.5		467.7	88.3	147.8	751.2		468.2	87.5	148.3	735.3		467.2	88.4	141.1	711.0		468.2	88.6	63.7	725.0	
Lithuania	459.6	91.3	91.0	730.4		460.5	93.2	81.2	820.4		461.8	92.2	103.9	816.0		461.4	92.0	33.7	764.3		460.2	92.4	76.5	768.3	
Netherlands	568.9	79.0	295.2	786.3		567.1	79.6	286.7	802.3		568.7	81.5	298.7	790.7		568.0	80.6	308.3	781.8		567.6	80.6	318.3	797.6	
New Zealand	525.4	89.2	220.1	810.1		527.4	88.6	257.4	753.8		526.3	89.0	237.0	754.9		526.6	89.0	235.6	847.0		526.8	90.4	159.4	791.5	
Norway	540.6	85.9	105.0	819.9		541.7	85.8	70.6	874.1		541.8	86.8	95.6	864.3		542.9	87.1	46.7	831.3		541.6	86.9	57.5	879.2	
Russian Federation	470.8	85.5	187.8	803.4		472.4	86.8	173.1	868.5		469.7	84.2	157.6	857.5		468.7	85.0	155.3	820.4		471.8	87.3	180.3	866.8	
Slovenia	495.5	80.3	209.0	749.4		496.1	81.1	195.7	778.8		495.9	79.5	194.8	752.5		496.2	80.9	169.8	776.1		495.9	80.5	161.6	778.3	
South Africa	320.1	107.9	98.7	691.5		319.3	108.3	75.6	773.8		319.6	109.9	60.0	705.1		320.7	107.2	93.8	752.1		319.3	106.8	66.8	792.7	
Sweden	557.7	82.0	314.4	817.3		557.4	81.1	318.5	843.3		557.0	82.2	334.4	882.3		557.6	83.2	299.5	830.7		556.2	82.5	232.6	842.0	
Switzerland	533.5	81.2	243.0	903.0		532.4	80.8	245.7	890.8		534.0	80.5	242.9	890.9		533.2	82.1	240.6	869.6		534.0	80.1	247.9	897.4	
United States	470.1	88.0	165.9	849.7		469.1	87.7	142.6	809.1		469.6	87.3	158.4	837.9		468.6	89.6	183.4	775.5		469.8	86.9	147.8	820.3	

---

# Chapter 7

## Content and Format of Database Files

### 7.1 Introduction

This chapter describes the content and format of the TIMSS International Database for the Final Year of Secondary School (Population 3). This chapter is organized in five major sections corresponding to the types of files included in the database. Within each section, the contents of the files are described. These file types are:

- Data Files
- Codebook Files
- Program Files
- Data Almanacs
- Test-Curriculum Matching Analysis Files

The TIMSS international *Data Files* reflect the result of an extensive series of data management and quality control steps taken to insure the international comparability, quality, accuracy, and general utility of the database in order to provide a strong foundation for secondary analyses.<sup>1</sup> As part of the international data files, all variables derived for reporting in the international reports are included. In addition, analysis notes are provided for all reporting variables, allowing users to replicate these computations. These analysis notes are included in Supplement 3 of this User Guide.

Also included in the database are *Codebook Files*. These specifically document the format of all variables in each of the data files. Several *Program Files* are provided for use in secondary analyses, including Data Access Control files for converting the raw data files provided in ASCII format into SAS or SPSS files, macro programs for computing statistics using the jackknife repeated replication method discussed in Chapter 8, and macro programs for converting cognitive item response codes to score values used in the computation of international scores. *Data Almanac Files* contain unweighted summary statistics for each participating country, on each variable in the student and school background questionnaires, as well as derived variables. To investigate the match between the TIMSS achievement tests (advanced mathematics and physics) and the curriculum in participating countries, country representatives indicated whether or not each test item addressed a topic in their curriculum. The *Test-Curriculum Matching Analysis Files* contain this information for each country for the items in the advanced mathematics and physics tests for Population 3. The following sections describe each of the file types and how they can be used to access and analyze the TIMSS international data for students and schools.

<sup>1</sup> For more detailed information about data entry, data processing, data cleaning, data management, database structure, as well as analysis and reporting of the TIMSS data, see the *TIMSS Technical Reports, Volumes I, II and III* (Martin and Kelly, 1996; 1997; 1998) and *TIMSS: Quality Assurance in Data Collection* (Martin and Mullis, 1996).

## 7.2 Data Files

There are three basic types of data files in the TIMSS International Database:

- Background Files
- Written Assessment Files
- Coding Reliability Files

These files and the variables contained in each are described in Sections 7.2.1 through 7.2.3.

Data files are provided for each country that participated in TIMSS at Population 3 and for which internationally comparable data are available. At Population 3 there are four different data file types, reflecting achievement data collected from students, background data collected from students and schools, and coding reliability data. For each file type, a separate data file is provided for each participating country that collected the information. Table 7.1 shows the Population 3 data files provided.

**Table 7.1**  
**TIMSS Population 3 Data Files**

File Type	File Name
Written Assessment File	CSA<Country>1.DAT
Student Background File	CSG<Country>1.DAT
School Background File	CCG<Country>1.DAT
Coding Reliability File	CSR<Country>1.DAT

The data files for each country may be identified by a three-digit alphabetic country abbreviation designated as <country> in the general file names shown in Table 7.1. The three-digit abbreviations used for each TIMSS country are listed in Table 7.2 along with the numeric code values used in the country identification variable contained in the background data files (see the following section discussing identification variables). Table 7.2 also shows the inclusion status of each country in the Population 3 data files and in the variables containing achievement score and cognitive item data<sup>2</sup> for the three Population 3 tests: mathematics and science literacy, advanced mathematics, and physics.

<sup>2</sup> See Sections 7.2.1.4 and 7.2.2 for descriptions of achievement score and cognitive item variables.

**Table 7.2**  
**Country Identification and Inclusion Status in Population 3 Data Files**

Country	Abbreviation	Numeric Code	Inclusion in Population 3 Data Files				Inclusion in Population 3 Tests		
			C S A	C S G	C C G	C S R	Mathematics and Science Literacy	Advanced Mathematics	Physics
Australia	AUS	036	✓	✓	✓	✓	✓	✓	✓
Austria	AUT	040	✓	✓	✓		✓	✓	✓
Canada	CAN	124	✓	✓	✓	✓	✓	✓	✓
Cyprus	CYP	196	✓	✓	✓		✓	✓	✓
Czech Republic	CSK	200	✓	✓	✓	✓	✓	✓	✓
Denmark	DNK	208	✓	✓	✓	✓	✓	✓	✓
France	FRA	250	✓	✓	✓	✓	✓	✓	✓
Germany	DEU	280	✓	✓	✓	✓	✓	✓	✓
Greece	GRC	300	✓	✓	✓			✓	✓
Hungary	HUN	348	✓	✓	✓		✓		
Iceland	ISL	352	✓	✓	✓		✓		
Israel	ISR	376	✓	✓	✓		✓	✓	✓
Italy	ITA	380	✓	✓	✓	✓	✓	✓	✓
Latvia	LVA	428	✓	✓	✓				✓
Lithuania	LTU	440	✓	✓	✓		✓	✓	
Netherlands	NLD	528	✓	✓	✓	✓	✓		
New Zealand	NZL	554	✓	✓	✓	✓	✓		
Norway	NOR	578	✓	✓	✓	✓	✓		✓
Russian Federation	RUS	643	✓	✓	✓	✓	✓	✓	✓
Slovenia	SVN	890	✓	✓	✓		✓	✓	✓
South Africa	ZAF	717	✓	✓	✓		✓		
Sweden	SWE	752	✓	✓	✓	✓	✓	✓	✓
Switzerland	CHE	756	✓	✓	✓	✓	✓	✓	✓
United States	USA	840	✓	✓	✓	✓	✓	✓	✓

All TIMSS international data files are provided in ASCII format, enabling users to access the data directly without having to go through intermediate steps. All details of the file structure are provided in codebook files related to each of the data files listed in Table 7.1. The use of these codebooks is described in Section 7.3.

## 7.2.1 Background Files

There are two different types of TIMSS background files – student and school.

### 7.2.1.1 Student Background File

Students who participated in TIMSS were administered a background questionnaire with questions related to home background and school experiences. The Student Background file contains students' responses to these questions as well as a series of identification variables, tracking variables, sampling and weighting variables, achievement variables, and the derived variables that were used in creating tables for the international report. In addition to the questions included the background questionnaire, students were also asked four questions located at the end

of their test booklets regarding calculator usage during the TIMSS test. The student responses to these questions are contained in four calculator variables (CSBGCAL1 - CSBGCAL4) also included in the student background file.

### 7.2.1.2 School Background File

The principals or administrators of the schools in the TIMSS sample were administered a school background questionnaire with questions about school policy and school environment. The School Background data file contains the responses given to the questions in this questionnaire. This file also contains a series of identification variables, tracking variables, sampling, and weighting variables. The school data files can be merged with the student data files by using the country and school identification variables.

### 7.2.1.3 Identification Variables

In all background files, several identification variables are included that provide information used to identify students or schools, and to link cases between the different data files. The identification variables have prefixes of ID and are listed below.

#### Variables Included in Student and School Background Files

IDCNTRY	Three-digit country identification code (see Table 7.2 for a list of country codes). This variable should always be used as one of the link variables whenever files will be linked across countries.
IDPOP	Identifies the population; takes the value of 3 for all records.
IDSTRAT	This variable identifies the sampling stratum for each student and school. This variable was used in some but not all countries to identify specific strata within the population for that country. Definitions of the IDSTRAT options included for each country are shown in Table 7.3.
IDSCHOOL	Identification number that uniquely identifies the school within each country. The codes for the school are not unique across countries. Schools can be uniquely identified only by the IDCNTRY and IDSCHOOL combination.

#### Additional Variables Included in the Student Background Files

IDSTUD	Identification number that uniquely identifies each student in the country sampled. Students can be uniquely identified in the database by the combination of IDSTUD and IDCNTRY.
IDBOOK	Identifies the specific test booklet that was administered to the student (1A, 1B, 2A, 2B, 2C, 3A, 3B, 3C, 4).
IDSUBPOP	Identifies the subpopulation to which each student belongs based on their mathematics and physics preparation as described in Chapter 3 (1 = OO; 2 = OP; 3 = MO; 4 = MP).



In the Student Background file, the IDSTUD variable provides a unique identification number to identify each student within each country. The variable IDSCHOOL, contained in both the student and school background files, is a unique identification number for each school within a country that may be used to link school background data to corresponding students.

**Table 7.3**  
**Definitions of Stratification Options (IDSTRAT) for Population 3 Countries**

Country	Status	Options
Australia	Stratification data not available.	
Austria	Stratification by school-type	1 = Academic school (AHS) 2 = Higher technical vocational school (BHS) 3 = Intermediate technical/vocational school (BMS) 4 = Apprenticeship (BS)
Canada	Stratification by province	125 = Alberta 126 = British Columbia 127 = New Brunswick 128 = Ontario - English 129 = Ontario - French 130 = Newfoundland 131 = Other
Cyprus	Stratification by districts and community type	11 = Nicosia (capital); urban 12 = Nicosia (capital); rural 21 = Limassol; urban 22 = Limassol; rural 31 = Lamaka-Famagusta; urban 32 = Lamaka-Famagusta; rural 41 = Paphos, urban 42 = Paphos, rural
Czech Republic	Stratification by school type and community type	11 = Vocational/technical school; city 21 = Vocational/technical school; town 31 = Vocational/technical school; small town 14 = Gymnasium; city 24 = Gymnasium; town 34 = Gymnasium; small town [City = >1 million inhabitants; Town = 90,000 to 1 million inhabitants; Small Town = <90,000 inhabitants.]
Denmark	Stratification data not available.	
France	Stratification data not available.	
Germany	Stratification data not available.	
Greece	Stratification data not available.	
Hungary	Stratification by school/program type and community type	1 = Academic school; capital region 2 = Academic school; city/town region 3 = Academic school; other settlement 4 = Academic school (religious/alternative) 5 = Vocational school (industrial) 6 = Vocational school (agricultural) 7 = Vocational school (merchant) 8 = Vocational school (humanities) 11 = Trade school (trade) 33 = Trade school (merchant) 34 = Trade school (agricultural)
Iceland	Stratification data not available.	
Israel	Stratification by school type and size of settlement	Strata defined by three digits: First digit: 1 = State school 2 = State religious school Second Digit: 1 = Academic school 2 = Technical single-track school (vocational or agricultural) 3 = Comprehensive school (academic and/or technical/vocational and/or agricultural tracks) Third digit: 1 = Population > 100,000 2 = Population > 25,000 3 = Population > 5,000 4 = Village (usually regional schools) 5 = Boarding schools/youth villages
Italy	Stratification by school/program type	11 = Classical schools (classics, languages, teaching programs) 12 = Classical schools (science program) 20 = Art schools 30 = Vocational schools 40 = Technical schools
Latvia	Stratification data not available	
Lithuania	Stratification data not available	

**Table 7.3 (Continued)**  
**Definitions of Stratification Options (IDSTRAT) for Population 3 Countries**

Country	Status	Options
Netherlands	Stratification by school/program type	1 = MBO - Technology 2 = MBO - Economics 3 = MBO - Agriculture 4 = MBO - Home Economics 5 = HAVO/VWO 6 = Short MBO - Technology 7 = Short MBO - Economics 8 = Short MBO - Home Economics [School types: VWO (senior academic); HAVO (senior general); MBO (senior vocational).]
New Zealand	Stratification not used.	
Norway	Stratification by program/track	31 = Academic track (AF) - math/science specialization 32 = Academic track (AF) - general 33 = Vocational track (HK) - Trade/office 34 = Vocational track (HI) - Craft/Industry 35 = Other (HS, HH, and other non-specialists)
Russian Federation	Stratification data not available.	
Slovenia	Stratification not used.	
South Africa	Stratification by province	1 = Eastern Cape Province 2 = Free State Province 3 = Gauteng Province 4 = KwaZulu-Natal Province 5 = Mpumalanga Province 6 = Northern Cape Province 7 = Northern Province 8 = North-West Province 9 = Western Cape Province
Sweden	Stratification not used.	
Switzerland	Stratification by program/track within school	1 = Academic - Math/Science Programs (Type C) 2 = Academic - Language or Economic Programs (Type A/B, D, or E) 3 = Academic - Cantonal Types and Primary Teacher Education 4 = Other Academic Programs 5 = Selected Technical Vocational Programs 6 = Commercial Vocational Programs 7 = Other Vocational Programs
United States	Stratification by geographic regions	1 = Northeast 2 = Southeast 3 = Central 4 = West [States assigned to geographic regions according to the NAEP sampling framework.] <sup>1</sup>

<sup>1</sup> United States stratification by geographic region is described in the NAEP 1992 Technical Report (Johnson and Carlson, 1994).

## 7.2.1.4 Achievement Scores

### Student-Level Achievement Score Variables

The variables used in the reporting of international achievement results, as well as other achievement variables representing students' proficiency in mathematics and science described in Chapter 6, are provided in the Population 3 Student Background files. The achievement variables can be found under the following variable names organized into four major categories.

#### Raw Scores

CSLSCPT	Number of raw score points on mathematics and science literacy items
CSMSCPT	Number of raw score points on advanced mathematics items
CSPSCPT	Number of raw score points on physics items
CSRSCPT	Number of raw score points on reasoning and social utility items

---

**National Rasch Scores**

CSLNRSC	National Rasch Mathematics and Science Literacy Score
CSMNRSC	National Rasch Advanced Mathematics Score
CSPNRSC	National Rasch Physics Score
CSRNRSC	National Rasch Reasoning and Social Utility Score

**Expected-a-Posteriori Scores (EAP): one for each overall scale and content area subscales**

CSLEAP0	Mathematics and Science Scale Literacy EAP Score
CSLEAP1	Mathematics Literacy Scale EAP Score
CSLEAP2	Science Literacy Scale EAP Score
CSMEAP0	Overall Advanced Mathematics Scale EAP Score
CSMEAP1	Numbers, Equations and Functions Subscale EAP Score
CSMEAP2	Calculus Subscale EAP Score
CSMEAP3	Geometry Subscale EAP Score
CSPEAP0	Overall Physics Scale EAP Score
CSPEAP1	Mechanics Subscale EAP Score
CSPEAP2	Electricity and Magnetism Subscale EAP Score
CSPEAP3	Heat Subscale EAP Score
CSPEAP4	Wave Phenomena Subscale EAP Score
CSPEAP5	Modern Physics Subscale EAP Score
CSREAP0	Reasoning and Social Utility Subscale EAP Score

Plausible Values: Five for each overall scale and content area subscale

CSLPV01-CSLPV05	Mathematics and Science Literacy Scale Plausible Values (1-5)
CSLPV11-CSLPV55	Mathematics Literacy Scale Plausible Values (1-5)
CSLPV21-CSLPV25	Science Literacy Scale Plausible Values (1-5)
CSMPV01-CSMPV05	Overall Advanced Mathematics Scale Plausible Values (1-5)
CSMPV11-CSMPV15	Numbers, Equations and Functions Subscale Plausible Values (1-5)
CSMPV21-CSMPV25	Calculus Subscale Plausible Values (1-5)
CSMPV31-CSMPV35	Geometry Subscale Plausible Values (1-5)
CSPPV01-CSPPV05	Overall Physics Scale Plausible Values (1-5)
CSPPV11-CSPPV15	Mechanics Subscale Plausible Values (1-5)
CSPPV21-CSPPV25	Electricity and Magnetism Subscale Plausible Values (1-5)
CSPPV31-CSPPV35	Heat Subscale Plausible Values (1-5)
CSPPV41-CSPPV45	Wave Phenomena Subscale Plausible Values (1-5)
CSPPV51-CSPPV55	Modern Physics Subscale Plausible Values (1-5)
CSRPV01-CSRPV05	Reasoning and Social Utility Subscale Plausible Values (1-5)

### 7.2.1.5 Tracking Variables

Information about students and schools provided on the survey tracking form<sup>3</sup> is included in tracking variables. These variables have prefixes of IL or IT and are listed below.

#### Tracking Variables Included in the Student Background Files

ITSEX	Gender of each student
ITBIRTHM	Month of birth of each student
ITBIRTHY	Year of birth of each student
ITDATEM	Month of testing for each student
ITDATEY	Year of testing for each student
ITLANG	Language of testing for each student. Data are included for four countries that administered the TIMSS test in more than one language: Canada: 1 = English; 2 = French Norway: 1 = Bokmål; 2 = Nynorsk South Africa: 1 = English; 2 = Afrikaans Switzerland: 1 = German; 2 = French; 3 = Italian

<sup>3</sup> Survey Tracking Forms are listings of students or schools used for sampling and administration purposes.

ITPART	Participation status variable indicating whether each student participated in a TIMSS achievement test session; only those students with ITPART equal to 3 (participated in a TIMSS session) are included in the Student Background files
ITPART2	Participation status variable indicating whether each student participated in a TIMSS background questionnaire session.
ILRELIAB	Linking variable indicating the inclusion status of each student in the Coding Reliability File containing double-coded free-response items (see Section 7.2.3)

### Tracking Variables in the School Background Files

ITREPLAC	Replacement status variable indicating whether each school is an original sampled school or a replacement school
ILREPLAC	School identification number of original school replaced by each replacement school
ITPART	Participation status variable indicating whether each school participated and returned a questionnaire

## 7.2.1.6 International Background Variables

### General Information

International background variables obtained from the student and school questionnaires are provided in the corresponding background data files. In general, the background variables are provided for all countries where the data are considered internationally comparable. The assessment of international comparability for background variables was based on information provided by NRCs regarding any national adaptations of the background questionnaire items. In a few cases, some slightly modified specific country options were retained in the international variables. Additional national variables not included in the international version of the questionnaire are not contained in the international background files. For a description of the information obtained from the international student and school background questionnaire items, see Chapter 2. Copies of the international versions of the questionnaires are provided in Supplement 1. More information regarding how national adaptations of background questionnaire items were handled in creating the TIMSS background data files is provided in Supplement 2.

### International Background Variable Values

The values for the background variables are either categorical options or open-ended numerical values, in accordance with the corresponding background questionnaire item formats. The codebook files for students and schools contain the international background variable names, descriptive labels, response code definitions, formats, and field locations corresponding to each questionnaire item (codebook files are described in Section 7.3).

## Identifying Background Variables by Questionnaire Numbers

The international background variables are listed in the codebook files in order of the corresponding questions in the international version of the background questionnaires. For each background variable, the corresponding international questionnaire location is given. A set of tables mapping each background variable to its location in the questionnaire is included in Supplement 1. The questionnaire item numbers associated with each variable contained in Supplement 1 are indicated by field locations according to the formats given in Table 7.4.

**Table 7.4**

### Background Questionnaire Item Field Location Format Conventions for Population 3

Questionnaire	Field Location Format	Example Questions
Student Questionnaire	SQ3-xxxx	SQ3-2 Are you a female or a male?
		SQ3-11/1 How far in school did your mother go?
School Questionnaire	SCQ3-xxxx	SCQ3-1 In what type of community is school located?
		SCQ3-19A1 What is the total school enrollment of boys?

## International Background Variable Names

The naming system for the background variables permits the identification of the population and questionnaire from which it was obtained based on 7- or 8-digit codes according to the general definitions given in Table 7.5.

**Table 7.5**  
**International Background Variable Naming Conventions for Population 3**

Character Position	Definition	Values	
		Student Variables	School Variables
1	Population	C = Population 3	C = Population 3
2	Questionnaire Type	S	C
3	Background Variable	B	B
4	Subject Area	G = General M = Mathematics S = General Science B = Biology C = Chemistry E = Earth Science P = Physics	G = General M = Mathematics S = Science P = Physics
5-8	Abbreviated Question Reference	****	****

For example:

CSBGEDUM = Population 3 students’ reports of mother’s highest level of education (SQ3-11/1)

CCBMNTEA = Population 3 schools’ reports of how many teachers in the school teach the most advanced mathematics (SCQ3-6)

### 7.2.1.7 Variables Derived from Student Background Data

#### General Information

In addition to the background variables contained in the student and school questionnaires, a number of derived variables based on the student background variables were computed for use in the international report. These derived variables, many of which use combined response options or data from more than one item, are also included in the International Database.

#### Use of Derived Variables in the International Reports

The derived variables can be used to reproduce the values shown in the international report tables by applying the appropriate student filters and weights. A description of each derived variable is included in Supplement 3. The index to Supplement 3 lists all derived variables along with their corresponding international table reference location (international table number in which the derived variable data are reported). The order of variables in this index and in the Student Background file codebooks is based on the referenced table in the international report. The

nomenclature used to indicate the international table location reference for each derived variable is based on 7- to 9-digit codes according to the general definitions given in Table 7.5.

## Table 7.6

### International Report Table Location Reference Definition for Derived Variables in Population 3

Character Position	Definition	Value(s)
1	Population	C = Population 3
2-3	International Report Section	LR = Mathematics and Science Literacy Section MR = Advanced Mathematics Section SR = Physics Section
4-9	Table/Figure Reference	Tx.yy (a,b,..) = Table Number [x.yy = chapter.number; (a,b,...) = columns]

For example:

CLRT4.17e = Population 3, Literacy Section, Table 4.17, Column 5

CMRT7.12 = Population 3, Advanced Mathematics Section, Table 7.12

CSRT10.6 = Population 3, Physics Section, Table 10.6

### Derived Variable Names

The derived variable names are based on 7- or 8-digit codes according to the same general definitions as are used for the international background variables (see Table 7.5). In the case of the derived variables, however, the third character is a D (derived variable) instead of B (background variable). For example, CSDGEDUP is a derived variable including Population 3 students' reports of parents' education.

### Information About How the Derived Variables Were Computed

Descriptions of each of the derived variables based on student background data, their associated international table/figure references, and analysis notes indicating how they were computed using the associated student source variables are provided in Supplement 3. Supplement 3 is presented in alphabetical order by derived variable name. The documentation for derived variables reflects the rules for construction from the internationally defined background questions. Due to national adaptations of some items in the background questionnaires, some countries have been omitted or handled somewhat differently for certain report variables. Documentation describing how specific national adaptations were handled in the international derived variables is also summarized in Supplement 3 for each of the report variables. The response option definitions for all derived variables as well as their associated international report table and figure references also are included in the student codebook files, which are described in Section 7.3.



### 7.2.1.8 Sampling and Weight Variables

Variables related to the samples selected in each country are included in the student and school background files. These sampling and weighting variables can be found under the following variable names organized into three major categories: Jackknife Variables, Participation Status Variables, and Weighting Variables.

More detailed descriptions of the sampling and weight variables and their use in analyses involving scale scores, item scores, and background variables to compute the population estimates and standard errors presented in the international reports are described in Chapter 3 and Chapter 8. In addition, Chapter 9 includes some sample analyses using these variables.

#### Jackknife Variables (Student Background File)

JKZONE	Jackknife zone
JKINDIC	Replicate code

#### Participation Status Variables (Student Background File)

INTMSL4	Participation Status - Literacy (Set to 1 for students assigned Booklet 1A, 1B or 4.)
INTMSM4	Participation Status - Advanced Mathematics (Set to 1 for students assigned Booklet 3A, 3B, 3C or 4.)
INTMSP4	Participation Status - Physics (Set to 1 for students assigned Booklet 2A, 2B, 2C or 4.)
INTMSL	Participation Status –Literacy (Set to 1 for students assigned Booklet 1A or 1B.)
INTMSM	Participation Status - Advanced Mathematics (Set to 1 for students assigned Booklet 3A, 3B, or 3C.)
INTMSP	Participation Status - Physics (Set to 1 for students assigned Booklet 2A, 2B, or 2C.)

#### Weighting Variables (Student Background File)

TOTWGTL4	Total Student weight - Literacy (Booklets: 1A, 1B, 4)
TOTWGTM4	Total Student Weight - Advanced Mathematics (Booklets: 3A, 3B, 3C, 4)
TOTWGTP4	Total Student Weight - Physics (Booklets: 2A, 2B, 2C, 4)
TOTWGTL	Total Student Weight - Literacy (Booklets: 1A, 1B)
TOTWGTM	Total Student Weight - Advanced Mathematics (Booklets: 3A, 3B, 3C)
TOTWGTP	Total Student Weight - Physics (Booklets: 2A, 2B, 2C)
HOUWGTL4	House Weight - Literacy (Booklets: 1A, 1B, 4)
HOUWGTM4	House Weight - Advanced Mathematics (Booklets: 3A, 3B, 3C, 4)

HOUWGTP4	House Weight - Physics (Booklets: 2A, 2B, 2C, 4)
HOUWGTL	House Weight - Literacy (Booklets: 1A, 1B)
HOUWGTM	House Weight - Advanced Mathematics (Booklets: 3A, 3B, 3C)
HOUWGTP	House Weight - Physics (Booklets: 2A, 2B, 2C)
SENWGTL4	Senate Weight - Literacy (Booklets: 1A, 1B, 4)
SENWGTM4	Senate Weight - Advanced Mathematics (Booklets: 3A, 3B, 3C, 4)
SENWGTP4	Senate Weight - Physics (Booklets: 2A, 2B, 2C, 4)
SENWGTL	Senate Weight - Literacy (Booklets: 1A, 1B)
SENWGTM	Senate Weight - Advanced Mathematics (Booklets: 3A, 3B, 3C)
SENWGTP	Senate Weight - Physics (Booklets: 2A, 2B, 2C)
WGTFAC1	School Weighting Factor
WGTFAC2	Class Weighting Factor
WGTFAC3	Student Weighting Factor
WGTADJ1	School Weighting Adjustment Factor
WGTADJ3	Student Weighting Adjustment Factor
WGTADJL4	Literacy Scale Adjustment Factor (Booklets: 1A, 1B, 4)
WGTADJM4	Advanced Mathematics Scale Adjustment Factor (Booklets: 3A, 3B, 3C, 4)
WGTADJP4	Physics Scale Adjustment Factor (Booklets: 2A, 2B, 2C, 4)
WGTADJL	Literacy Item Adjustment Factor (Booklets: 1A, 1B)
WGTADJM	Advanced Mathematics Item Adjustment Factor (Booklets: 3A, 3B, 3C)
WGTADJP	Physics Item Adjustment Factor (Booklets: 2A, 2B, 2C)

### Weighting Variables (School Background File)

WGTFAC1	School Weighting Factor
WGTADJ1	School Weighting Adjustment Factor
SCHWGT	School-Level Weight
STOWGTL4	Sum of Student Weights – Literacy (Booklets: 1A, 1B, 4)
STOWGTM4	Sum of Student Weights – Advanced Mathematics (Booklets: 3A, 3B, 3C, 4)
STOWGTP4	Sum of Student Weights – Physics (Booklets: 2A, 2B, 2C, 4)
STOWGTL	Sum of Student Weights – Literacy (Booklets: 1A, 1B)
STOWGTM	Sum of Student Weights – Advanced Mathematics (Booklets: 3A, 3B, 3C)
STOWGTP	Sum of Student Weights – Physics (Booklets: 2A, 2B, 2C)

## 7.2.2 Written Assessment Files

Student Written Assessment files containing the student response data for the individual cognitive items in the TIMSS Written Assessment are provided for each country. Students who participated in TIMSS were administered one of nine test booklets with questions in mathematics and science literacy, advanced mathematics, or physics. The specific booklet type administered to each student depended on their academic preparation as identified by their subpopulation assignment (see Chapters 2 and 3).

Some of the questions administered in the TIMSS tests were multiple-choice questions and some were open-ended. The responses to the open-ended questions were coded using a two-digit diagnostic code scoring system described in Chapter 4. The student Written Assessment data files contain the responses selected by students to the multiple-choice questions and the two-digit scoring codes assigned by the scorers to the student responses to the free-response items. The specific test booklet that was administered to the student is coded in the variable IDBOOK. The Written Assessment data files also contain a series of identification variables, sampling and weighting variables, and achievement variables. The data contained in this file can be linked to the Student Background data files by using the variables IDCNTRY and IDSTUD.

### 7.2.2.1 Cognitive Item Variable Names

The cognitive item variable names are based on 7-digit alphanumeric codes according to the general definitions given in Table 7.7.

**Table 7.7**  
**Variable Name Definitions for the Written Assessment Items in Population 3**

Character Position	Definition	Value(s)
1	Population	C = Population 3
2	Student Variable	S
3	Item Format	M = Multiple Choice S = Short Answer E = Extended Response
4	Subject Area	G = Mathematics and Science Literacy M = Advanced Mathematics P = Physics
5	Cluster Location	A, B, ..., L
6-7	Item Number	01-99
8	Item Part (where applicable)	A, B, C, ...

For example:

CSMGA03 = Mathematics and Science Literacy test, multiple-choice format,  
item number A03

CSSMJ16A = Advanced Mathematics test , short-answer format, item number J16, Part A

CSEPH18 = Physics test, extended-response format, item number H18

### 7.2.2.2 Cognitive Item Response Code Values

The values assigned to each of the cognitive item variables depend on the item format. For the multiple-choice items, one-digit numerical values of 1-5 are used to correspond to the response options a - e. For these items, the correct response is included as part of the item-variable label in the codebook files. For the free-response items two-digit numerical codes are used that correspond to the diagnostic scoring rubrics used to determine fully-correct, partially-correct, and incorrect responses for each item. As described in Chapter 4, the correctness score level may be determined by the first digit of these codes (3\* = 3 points; 2\* = 2 points; 1\* = 1 point; 7\* or 9\* = 0 points). In addition to the correctness score information, specific missing codes are also defined that are described in the section discussing missing codes (Section 7.2.5). Since all cognitive item variables are included for all students in the assessment files regardless of which test booklet they completed, a 'Not Administered' code is given to all items that were not included in the test booklet assigned to each student.<sup>4</sup>

### 7.2.2.3 Analysis By Mathematics and Science Literacy, Physics, and Advanced Mathematics Content Area Reporting Categories

The TIMSS cognitive items measured student achievement in different content areas within mathematics and science literacy, advanced mathematics, and physics. In the international report, results were reported separately for 10 different content area subscales: mathematics literacy, science literacy, three advanced mathematics content areas, and five physics content areas. In order to permit secondary analyses involving the items in each of these content areas, the classification of TIMSS written assessment items into each of the Population 3 content area reporting categories is presented in Tables 7.8 through 7.10.

<sup>4</sup> See Chapter 2 for a discussion of the design of the TIMSS written assessment.

**Table 7.8**  
**Classification of Population 3 Items by Mathematics and Science Literacy Reporting Categories**

Items in Subject Area Reporting Category			
Mathematics Literacy (44 items)		Science Literacy (32 items)	
A03	C05	A01	B10
A04	C06	A02	B11
A05	C07	A06A	B12
A08	C08	A06B	B13
A10	C09	A07	C14
A12	C10	A09A	C15
B14	C11	A09B	C16
B15	C12	A11A	C17
B16	C13	A11B	C18
B17	D06	A11C	C19
B18	D07	B01	C20
B19	D08	B02	C21
B20	D09	B03	D01
B21	D10	B04	D02
B22	D11	B05	D03
B23	D12	B06	D04
B24	D13	B07	D05A
B25	D14	B08	D05B
B26	D15A	B09	
C01	D15B		
C02	D16A		
C03	D16B		
C04	D17		

\*Items in cluster A are classified as Reasoning and Social Utility (RSU). These items were included in the Mathematics and Science Literacy scales. Achievement score variables for a separate RSU subscale are included in the Student Background and Written Assessment files, but these subscale results were not included in the international report.

**Table 7.9**  
**Classification of Population 3 Items by Advanced Mathematics Content Area Reporting Categories**

Items in Content Area Reporting Category				
Number, Equations, and Functions  (17 items)	Calculus  (15 items)	Geometry  (23 items)	Probability and Statistics*  (7 items)	Validation and Structure*  (3 items)
I01	I04	I07	I05	I10
I02	I06	I08	J12	J18
I03	J05	I09	J13	L11
J01	J06	J07	K11	
J02	J14	J08	L10	
J03	J15A	J09	L14	
J04	J15B	J10	L15A	
K01	J17	J11	L15B	
K02	K03	J16A		
K13	K04	J16B		
K15	K05	J19		
K16	K06	K07		
L01	K17	K08		
L02	L05	K09		
L03	L06	K10		
L04	L07	K12		
L16		K14		
		K18		
		L08		
		L09		
		L12		
		L13		
		L17		
		L18		

\*Items in the Probability and Statistics and Validation and Structure content areas were included in the overall Advanced Mathematics scale. Too few items were included in these content areas to report as separate subscales.

**Table 7.10**  
**Classification of Population 3 Items by Physics Content Area Reporting Categories**

Items in Content Area Reporting Category				
Mechanics  (16 Items)	Electricity and Magnetism  (16 items)	Heat  (9 items)	Wave Phenomena  (10 items)	Modern Physics: Particle, Quantum and Astrophysics, and Relativity  (14 items)
E03	E04	E08	E01	E02
E05	E06	F05	E10	E07
F02	E09	F12	F01	F03
F04	F06	G02	F11	F07
F10	F08	G03	F13	F09
F17A	F14	G11	G05	F15
F17B	F16	H02	G13	G06
G07	G01	H07	H09	G10
G08	G04	H14	H12	G14
G09	G17		H19A	G18
G12	G19		H19B	H03
G15	H06			H05
G16	H08			H15
H01	H10			H18
H04	H16			
H11	H17			
H13				

### 7.2.2.4 Release Status of TIMSS Test Items

To aid in the interpretation of the student cognitive item response data, a large number of TIMSS items have been released to the public. Written assessment items in Clusters A, D, G, H, K, and L are available, along with the scoring guides. See Chapter 2 for information on how to obtain the released item sets.

### 7.2.2.5 Other Variables in the Written Assessment Files

In addition to the written assessment item variables, a number of other variables described in previous sections are included for each student, to aid in case identification and in linking to the data in other files:

- Identification variables
- Tracking variables
- Sampling and weighting variables
- Score variables

The codebooks for student assessment files include a complete list of all of these variables as well as item and code value labels to assist in the interpretation of the student assessment data.

## 7.2.3 Coding Reliability Files

### General Information

The coding reliability files contain data that can be used to investigate the reliability of the TIMSS free-response item scoring. These files contain one record for each student included in the reliability sample for the written assessment. For each free-response item in the written assessment, the following three variables are included:

- Original Code Variable (cognitive item response codes obtained from the first coder)
- Second Code Variable (cognitive item response codes obtained from second coder)
- Response Code Agreement Variable (degree of agreement between the two codes)

It should be noted that the Second Code Variable data were used only to evaluate the within-country coding reliability. Only the first codes contained in the student assessment files were used in computing the achievement scores reflected in the Student Background files and the international report.



## Reliability Variable Names

The variable names for the Original Code, Second Code, and Agreement Code variables are based on the same general naming system as that for the written assessment item variables as shown in Table 7.7. The three reliability variables may be differentiated by the second character in the variable name:

- Original Code Variable: Second Character = S (e.g. CSSGA07)
- Second Code Variable: Second Character = R (e.g. CRSGA07)
- Agreement Code Variable: Second Character = I (e.g. CISGA07)

## Reliability Variable Code Values

The values contained in both the Original Code and Second Code variables are the two-digit diagnostic codes obtained using the TIMSS scoring rubrics. The Agreement Code variable has three different values depending on the degree of agreement between the two coders:

- 0 = Identical codes (both digits in the diagnostic codes are identical)
- 1 = Identical score but different diagnostic code (first digits are the same; second digits are different)
- 2 = Different score (both the first and second digits are different)

In general, the response code data contained in the Original Code variables in the Coding Reliability files are identical to those contained in the Student Written Assessment files. In a few cases, however, the response codes for specific items were recoded after a review of the international item statistics revealed inconsistencies in the original coding guides or showed that the original codes were not functioning as desired. The recodings made to the Population 3 written assessment items are shown in Table 4.3.

For the items indicated in Table 4.3, the response codes in the Written Assessment files reflect the recoded diagnostic code values which were used in computing the achievement scores reported in the international reports. In contrast, the Original Code Variables in the Coding Reliability files contain the original unrecoded response codes. This was done so that the coding reliability measure indicated in the Agreement Code Variables was based on the original coding guides used during the free-response coding sessions conducted in each country. One exception to this is that any nationally defined diagnostic codes employed in individual countries (\*7 or \*8) were recoded to the "other" category (\*9) within the same correctness level prior to the computation of the Code Agreement Variables (see Chapter 4).

## Other Variables in the Reliability Files

In addition to the coding reliability variables, the Coding Reliability files also include identification variables to aid in case identification. Some tracking variables are also included that were used in conducting the coding reliability study within each country, including the reliability booklet set to which each student was assigned (ITBSET) and the identification of the first and second coders (ITCODE, ITCODE2).

## 7.2.4 Missing Codes in the International Data Files

All values assigned to variables in the TIMSS international data files are numeric, and a subset of the numeric values for each of the variable types is reserved for specific codes related to different categories of missing data.<sup>5</sup> The missing categories defined below are assigned different values depending on the field width of the variable and the variable type.

### Omitted Response Codes (9, 99, 999, ...)

Omitted response codes are used for questions/items that a student or school principal should have answered but did not answer. These are indicated as "missing" in the codebooks. For questionnaire data, no differentiation has been made between no answer and invalid answers, such as checking two or more response options in a categorical question or unreadable or uninterpretable responses to open-ended questions. In a few cases, data received from a country in an invalid or inconsistent way were also recoded to "missing". For cognitive items, an Omitted Response Code was given only in cases in which the item was left blank; a special code was used for invalid answers as described below. The specific Omitted Response Code value given depends on the number of valid codes available for each item.

*For Identification, Tracking, or Background Questionnaire Items:* An Omitted Response Code value of 9 is used for categorical items with 7 or less valid response options; for categorical items with more than 7 categories, a code value of 99 is used. For open-ended background questionnaire items or other items containing non-categorical values, the omit code is the next available 9\* code greater than the highest response in the defined valid range. Background questionnaire item values outside the valid ranges were recoded to missing.

*For Cognitive Items:* An Omitted Response Code value of 9 is used for multiple-choice cognitive items. For free-response written assessment items, the two-digit 99 code is used for omitted/blank responses.

### Uninterpretable Response Codes (7, 90)

For the cognitive items, separate codes were established to distinguish between totally blank responses (omitted/missing) and uninterpretable or invalid responses. For multiple-choice items, cases where more than one response option was checked were classified as uninterpretable and given a code 7. For the free-response items in the written assessment, uninterpretable student responses were given a code 90, which is distinguished from the code 99 given to items that were left blank.

### Not Administered Codes (8, 98, 998, ...)

Special codes were given for items that were "not administered" to distinguish these cases from data that are missing due to non-response. The specific Not Administered Code value given depends on the number of valid codes available for each item, as described above for the Omitted Response Codes.

<sup>5</sup> The SAS and SPSS control statement files will recode these missing categories to special numeric missing codes in SAS and explicit missing codes in SPSS (see Section 7.4).

There are two general cases when the Not Administered Codes are used.

1) Data were not collected for a variable for specific individuals. Reasons for this include:

- *Booklet not assigned to the student* - Only one of the nine rotated test booklets was assigned to each student. All variables corresponding to items which were not given to a student have been coded to "Not Administered."
- *Booklet not received / booklet lost* - If a respondent did not receive the instruments assigned to him/her, or the instruments were lost after administration, all items have been coded to "Not Administered."
- *Student absent from background questionnaire session* - If a student was not present for the administration of the background questionnaire, they were coded to "Not Administered" for all background questionnaire items. However, if a student participated in the background questionnaire session but did not answer the questions, the items were coded to "Omit."
- *Item left out or misprinted* - If a particular question or item ( or a whole page) was misprinted or not available to the student or school, the corresponding variables have been coded to "Not administered".

2) An item was omitted for all cases in a country. All cases are coded to "Not Administered."

- *Cognitive items omitted or mistranslated in student test booklets* - Any items identified during the translation verification or item analysis processes that were mistranslated such that the nature of the question was altered were coded to "Not Administered."
- *Background questionnaire items were omitted* - Questions in the student or school background questionnaires that were considered not applicable in some countries were not included in their questionnaires and were coded to "Not Administered."
- *Background questionnaire items were mistranslated or not internationally comparable* - In some cases, questions in the international version of the questionnaires were mistranslated or modified to fit the national situation. Whenever possible, modified background questionnaire items were recoded to match as closely as possible the international version of the items. This could not be done in all cases, however, and some national data were recoded to "not administered" in order to include only the internationally comparable data. Background questionnaire items that have been omitted or where internationally-comparable data are not available for some countries are documented in Supplement 2 of this user guide.

**Not Applicable Response Codes (6, 96, 996,...)**

The Not Applicable Response Codes are used only for the background questionnaire items in which responses are dependent on a filter question. If a dependent question was “not applicable” to a respondent because he/she answered a filter question negatively, the dependent question(s) have been coded to "not applicable." Also, if a respondent was not meant to answer a variable because of its logical relationship to other variables in the questionnaire design, these variables also have been recoded to "not applicable". The specific Not Applicable Code value given depends on the number of valid codes available for each item in the same fashion as was described above for the Omitted Response Codes.

**Not Reached Item Codes (6, 96)**

The Not Reached Item Codes are used only for cognitive items. Test items at the end of each test booklet in each testing session which were left blank were considered "not reached" due to the fact that the student did not complete the test. These responses are distinguished from the "missing" responses, as they are handled differently during the item calibration process (see Chapter 5). They are treated as incorrect responses, however, in computing achievement scores. For the multiple-choice items, a Not Reached Item Code value of 6 is used. For the free-response written assessment items, a Not Reached Item Code value of 96 is used.

## 7.3 Codebook Files

All information related to the structure of the data files as well as the source, format, descriptive labels, and response option codes for all variables discussed in Section 7.2 is contained in codebook files. Two versions the codebook files are provided for each of the data file types: one ASCII Text Format file for obtaining codebook printouts and one Machine-Readable ASCII Format file for use in programming. The codebook file extension (CDT or CDF) is used to identify the respective file format. Table 7.11 lists all codebook files provided.

**Table 7.11**  
**Population 3 Codebook Files**

<b>Data File</b>	<b>ASCII Text Format File Name</b>	<b>Machine-Readable ASCII Format File Name</b>
Written Assessment File	CSACODE.CDT	CSACODE.CDF
Student Background File	CSGCODE.CDT	CSGCODE.CDF
School Background File	CCGCODE.CDT	CCGCODE.CDF
Coding Reliability File	CSRCODE.CDT	CSRCODE.CDF

### 7.3.1 Accessing the Codebook Files

Both codebook file types are included in the database CD in ASCII format. They can be read and edited with any text editor or word processing software that can read files of their size. Each is designed with a specific purpose in mind.

#### Printout Format (\*.CDT)

The printout format is a text file containing the information from the codebook in a printout format. This format can be read with any word processing software and printed after some minor formatting. We suggest using a mono-spaced font, and a font size and page layout combination that will accommodate 132 characters per line. The information for each variable is presented in several lines of text. The lines for each variable are properly labeled.

#### Machine Readable Format (\*.CDF)

A second formatted version of the codebooks is also included in the database. In this version each variable occupies one line and the following fields are included: variable name, question location, starting column, ending column, number of digits, number of decimals, variable label, and value labels. These files can be used by those who want to use programming languages other than SAS or SPSS to conduct analysis with the TIMSS data. The value labels in these files are separated by semicolons. Table 7.12 describes the structure of the machine-readable codebook files.

**Table 7.12**  
**File Structure of Machine-Readable Codebook Files**

Field Name	Location
Variable name	1-8
Question location	10-20
Starting at	22-24
Ending at	27-29
Number of digits	31-32
Number of decimals	34
Variable label	36-74
Value labels	77-200

### 7.3.2 Using the Codebooks

The variables in the codebooks appear in order by variable name within sections for each codebook type. The major sections of each codebook type are as follows:

#### Student and School Background File Codebooks

- Identification Variables
- Tracking Variables
- International Background Variables (in order of questionnaire item location)
- Derived Variables ( in order of international table/figure reference)
- Sampling and Weighting Variables
- Score Variables (Student Files only)

#### Student Written Assessment File Codebooks

- Identification Variables
- Tracking Variables
- Cognitive Item Variables (in order by item within clusters)
- Sampling and Weighting Variables
- Score Variables

### Coding Reliability File Codebooks

- Identification Variables
- Tracking Variables
- Reliability Variables (organized into sets of three variables described previously in order by item within cluster)

An example printout of a page from the codebook for the student background data (CSGCODE) is shown in Figure 7.1.

**Figure 7.1**  
**Example Printout of the Student Background Codebook**

```
Codebook, Date: 28.05.98 File: csgcode.SDB
```

Var. Question No.	Variable Name	Variable Label	Code	Option	Location/Format
1	VERSION	*FILE VERSION*	VALUE	Version number	1-2 / C 2.0
2	COUNTRY	*COUNTRY ID*	VALUE	Country ID	3-5 / N 3.0
3	POP	*POPULATION ID*		1 population 1 2 population 2 3 population 3	6 / C 1.0
4	STRAT	*STRATUM ID*	VALUE	Stratum ID 999 missing 998 not admin.	7-9 / N 3.0
5	SCHOOL	*SCHOOL ID*	VALUE	School ID	10-15 / N 6.0
6	STUD	*STUDENT ID*	VALUE	Student ID	16-25 / N 10.0
7	RELIABLE	*RELIABILITY CODING STATUS*		0 student did not participate in reliability coding 1 student participated in reliability coding	26 / C 1.0
	COMMENT: new variable, indicates whether a student participated in the reliability coding or not				
8	SUBPOP	*SUB-POPULATION*		1 non specialists (OO) 2 physics specialists (OP) 3 mathematics specialists (MO) 4 mathematics & physics specialists (MP)	27 / C 1.0
9	BOOK	*BOOKLET ID*		1A booklet 1A 1B booklet 1B 2A booklet 2A 2B booklet 2B 2C booklet 2C 3A booklet 3A 3B booklet 3B 3C booklet 3C 4 booklet 4	28-29 / C 2.0

The fields contained in the codebook files and shown in Figure 7.1 are identified as follows:

**Variable Number:** The first column contains a sequential number for each variable in each codebook file.

**Question:** The second column contains an abbreviated variable identifier providing descriptive information needed to identify the content of the question and/or the source for each type of variable, e.g., questionnaire item location for background questionnaire item variables, table/figure reference for derived variables, or booklet location for cognitive item variables.

**Variable Name:** The third column contains the variable name associated with each variable included in the international data files. The naming system used for each variable type is described in the previous sections on the contents of data files.

**Variable Label:** The fourth column contains an extended textual variable label of up to 40 characters associated with each variable, providing more descriptive information about the content of each variable. For multiple-choice cognitive items, the variable label includes the correct response option enclosed in brackets. During data analysis, the variable labels can be printed out to enhance understanding of the results.

**Code:** The fifth column contains the codes used for variable responses. For variables containing numeric data such as that supplied for score variables, sampling and weighting variables, some identification variables, or background questionnaire variables containing responses to open-ended questions, the keyword VALUE is entered in the Code column. For categorical variables, all possible response options are listed. Any missing codes described in Section 7.2.4 are also included for either numerical or categorical variables. For example, for multiple-choice cognitive items, the code options are 1,2,3,4,5 while for the free-response cognitive items, the code options are the two-digit numerical codes described in Chapter 4.

**Option:** The sixth column includes a textual description of each type of response option. For variables containing numeric data, it contains an explanation of the values contained in the variable.

**Location/Format:** The seventh column presents the location and format of each variable in the raw data files. The location/format indicates the pattern used to write each value of a numeric or categorical variable, with a general structure of

$$XX-YY / <N \text{ or } C> ZZ.Z$$

The numbers preceding the slash (/) indicate the location of the variable and refer to its position in the raw data file (starting (XX) - ending (YY) column positions). The <N or C> after the slash identifies the variable as numerical (N) or categorical (C). The numeric code after the slash (ZZ.Z) indicates the total number of digits (including the decimal point) and the number of decimal places associated with each variable (e.g. 2.0 = 2 integer digits, 0 decimal places; 6.2 = six total digits: 3 integer digits, decimal point, and two decimal digits).



## 7.4 Program Files

Three different types of program files are provided for use in analyses of the TIMSS data files:

- Data Access Control Files
- Jackknife Statistics Program Files
- Scoring Program Files

The Data Access Control files are provided to convert the ASCII-format raw data files into SAS data sets or SPSS system files. A different control file is required for each data file, and the control files are named so that the first three characters match the first three characters of the respective data file. The Jackknife Statistics Program files are used to compute the percentage of students within defined subgroups and the mean value for each subgroup on specified continuous variables, as well as the standard errors associated with these statistics, using the jackknife repeated replication (JRR) method discussed in Chapter 8. The Scoring Program files are required to score written assessment cognitive item response codes in order to compute the raw scores. For all program files, two versions are provided: one for SAS programs and one for SPSS programs. The file extension (SAS or SPS) is used to identify the respective SAS and SPSS program files. Table 7.13 lists all program files provided.

**Table 7.13**  
**Population 3 Program Files**

Data Files	SAS File Name	SPSS File Name
<b>I. Data Access Control Files</b>		
Control Files for the Written Assessment File	CSACONTR.SAS	CSACONTR.SPS
Control Files for Student Background File	CSGCONTR.SAS	CSGCONTR.SPS
Control Files for School Background File	CCGCONTR.SAS	CCGCONTR.SPS
Control Files for Coding Reliability File	CSRCONTR.SAS	CSRCONTR.SPS
<b>II. Jackknife Statistics Program Files</b>		
	JACKGEN.SAS	JACKGEN.SPS
	JACKPV.SAS	JACKPV.SPS
<b>III. Scoring Program Files for Written Assessment Cognitive Items</b>		
	CSASCORE.SAS	CSASCORE.SPS

Chapter 9 further describes the SAS and SPSS program files and how they are applied through the use of specific example analyses using the TIMSS student and school data files.

## 7.5 Data Almanacs

Data almanacs are included for all student and school background variables. The data almanacs are text files that display weighted summary statistics for each participating country on each variable included in the background questionnaires administered to students and school administrators or principals. There are a total of six Data Almanac files containing data for student or school background variables for each of three different samples:

- All students in their final year of upper-secondary school (assigned booklet 1A, 1B, or 4 and included in the mathematics and science literacy results)
- Students in their final year of upper-secondary school having taken advanced mathematics (assigned booklet 3A, 3B, 3C, or 4 and included in the advanced mathematics results)
- Students in their final year of upper-secondary school having taken physics (assigned booklet 2A, 2B, 2C, or 4 and included in the physics results)

The data almanac files corresponding to each questionnaire type and student subpopulation are listed in Table 7.14.

**Table 7.14**  
**Data Almanac Files for Population 3**

Almanac File	Contents
CSGALML.ALM	Student background data - Mathematics and Science Literacy
CSGALMM.ALM	Student background data - Advanced Mathematics
CSGALMP.ALM	Student background data - Physics
CCGALML.ALM	School background data - Mathematics and Science Literacy
CCGALMM.ALM	School background data - Advanced Mathematics
CCGALMP.ALM	School background data - Physics

Since Booklet 4 includes mathematics and science literacy, advanced mathematics, and physics items, students who took Booklet 4 are included in all three student Data Almanac files. The inclusion of schools in the three different school Data Almanac files is based on the students within the schools. If a school contains students in all three samples, then that school will be represented in all Data Almanac files. The data reflected in the almanacs use the student as the unit of analysis, even if the information is based on school questionnaire items. Therefore, the weighted percentages in the school data almanacs reflect the percentage of students in each sample to whom the data apply.

There are two types of displays in the data almanacs, depending on whether data are obtained from a categorical or open-ended question. The display for categorical variables includes:

- Student population estimated from the sum of weights for all students in the sample
- Sample size (number of students or schools included in the sample)
- Weighted percentage of students for whom the item was coded as “Not administered”
- Weighted percentage of students corresponding to each valid response option (percentages based only on the students with valid data)
- Weighted percentage of students for whom none of the valid options were selected (“Not administered” or “Omitted”)

In cases where the question was Not Applicable based on the response to an explicit filter question, the percentage of students with valid data for whom the question was Not Applicable is also presented. An example of the almanac display for a categorical variable in the student background file is shown in Figure 7.2.

**Figure 7.2**  
**Example Data Almanac Display for a Categorical Student Background Variable**

1Third International Mathematics and Science Study - 1995 Assessment								May 29, 1998 61	
Student Background Variables - Students in the Final Year of Secondary School (INTMSL4=1)									
Question: Do you intend to attend a four-year college or university? (CSBGFE4)									
Location: SQ3-14A4									
								3.POPULATION 3	
								GEN\CONT EDUC\ATTEND <4 YR COLLEGE OR U>	
Country	Population	Sample	%NA	1.YES %	2.NO %	NOT APP. %	Other %		
Australia	170847	1941	2.0	65.7	25.5	8.8	2.9		
Austria	70602	1962	3.7	32.1	41.1	26.7	4.4		
Canada	263241	5232	0.7	62.5	33.7	3.7	5.3		
Cyprus	4556	534	0.4	61.4	21.3	17.3	7.0		
Czech Republic	137459	2167	0.4	24.6	35.8	39.6	0.8		
Denmark	37872	2714	9.5	38.5	35.8	25.7	12.8		
France	637935	1590	0.5	26.8	62.7	10.5	2.5		
Germany	967705	2289	73.2	62.1	34.5	3.4	74.2		
Hungary	111281	5091	3.2	9.3	74.0	16.6	12.0		
Iceland	2308	1703	1.9	65.4	28.1	6.5	5.9		
Israel	1357	1357	10.1	64.4	28.4	7.2	19.4		
Italy	380834	1616	0.8	44.2	23.2	32.6	7.0		
Lithuania	22161	2887	2.1	50.5	36.2	13.3	8.3		
Netherlands	145916	1470	2.2	17.1	60.6	22.3	3.6		
New Zealand	37549	1763	0.7	53.0	37.7	9.3	2.8		
Norway	43806	2518	1.5	43.0	46.1	10.9	6.1		
Russian Federation	1031187	2289	0.3	60.6	38.0	1.4	6.0		
Slovenia	26644	1622	2.7	62.5	25.9	11.6	5.4		
South Africa	374618	2757	2.1	75.2	19.1	5.7	8.8		
Sweden	71243	3068	2.0	43.8	40.7	15.5	5.9		
Switzerland	65140	3308	0.6	19.4	50.4	30.1	1.9		
United States	2278258	5807	2.8	69.3	26.8	3.9	7.0		

A second type of data display is used for numerical variables obtained from open-ended items. This type of display includes:

- Student population estimated from the sum of weights for all students in the sample
- Sample size (students or schools)
- Weighted percentage of students for whom the items was coded as “Not Administered”
- Weighted percentage of students for whom the item is omitted
- Weighted percentage of students for whom the question was “Not Applicable”
- The mean, mode, minimum, maximum, and the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles across students

An example of the almanac display for a numerical variable in the school background file is shown in Figure 7.3.

Figure 7.3

Example Data Almanac Display for a Numerical School Background Variable

Third International Mathematics and Science Study - 1995 Assessment  
 School Background Variables - Students in the Final Year of Secondary School (INTMSL4=1) May 29, 1998 20  
 Question: How many classroom teachers are on the staff of your school? (FTE) (CCBGFTB4)  
 Location: SCQ3-3D  
 \*POPULATION ID\*=3,POPULATION 3

Country	Population	Cases	%Not Ad.	%Omit	%Not Ap.	Mean	Mode	Min	P5	P10	Q1	Median	Q3	P90	P95	Max
Australia	170847	86	4.9	2.3	0.0	53.6	43.0	19.0	27.4	27.4	36.0	50.0	60.0	80.0	85.3	180.0
Austria	70602	169	8.5	3.0	0.0	45.0	45.0	4.0	13.3	23.0	35.0	59.4	86.0	110.0	120.0	385.0
Canada	263228	336	17.9	4.7	0.0	59.6	52.0	2.0	16.0	23.7	38.5	56.2	76.9	99.3	120.0	153.0
Cyprus	4556	28	12.9	0.0	0.0	52.6	29.0	20.0	20.0	29.0	41.7	51.4	61.0	75.0	76.0	113.0
Czech Republic	137459	150	0.2	1.5	0.0	29.3	50.0	1.0	9.0	10.5	19.0	28.1	37.3	49.0	53.0	91.0
Denmark	29095	91	0.8	1.8	0.0	60.4	46.0	6.5	28.0	36.0	52.0	62.0	73.0	80.0	85.0	110.0
France	637935	56	4.5	8.5	0.0	116.5	35.0	10.0	18.0	38.0	79.0	113.0	134.0	190.0	260.0	260.0
Germany	967705	152	35.3	3.7	0.0	56.5	48.0	2.0	15.0	25.0	34.0	53.0	75.0	103.0	106.0	112.0
Hungary	111281	204	15.0	5.8	0.5	47.4	35.0	1.0	15.5	21.0	34.3	46.0	60.5	76.5	82.5	94.0
Iceland	2308	30	5.7	10.2	0.0	43.1	8.5	8.0	8.5	10.0	33.2	44.6	55.0	59.0	77.5	77.5
Israel	1357	52	8.8	15.9	0.0	103.2	70.0	8.0	34.0	59.0	90.0	120.0	180.0	300.0	300.0	300.0
Italy	380834	101	1.4	8.9	0.0	78.3	69.0	9.0	15.0	30.0	48.0	73.0	107.0	130.0	154.0	190.0
Lithuania	22161	142	24.4	4.6	0.0	62.7	30.0	15.0	30.0	30.0	47.0	63.0	75.0	91.0	108.0	145.0
Netherlands	145916	79	100.0	0.0	0.0	117.3	999.0	9.6	24.0	27.8	36.0	48.0	64.0	89.7	999.0	999.0
New Zealand	37549	79	4.5	0.0	0.0	58.2	40.0	16.0	23.0	28.0	40.0	51.8	70.0	97.0	120.0	160.0
Norway	43806	131	9.1	3.1	0.0	54.1	65.0	7.0	19.0	22.0	35.0	52.0	71.0	87.0	96.0	132.0
Russian Federation	1031187	163	0.0	8.4	0.0	54.1	33.0	18.0	31.0	32.0	41.0	55.0	74.0	89.0	91.0	152.0
Slovenia	26644	79	66.9	2.6	0.0	58.2	33.0	18.0	31.0	32.0	41.0	55.0	74.0	89.0	91.0	152.0
South Africa	374618	90	69.9	0.7	0.0	23.3	16.0	1.0	1.0	12.0	16.0	23.0	28.0	40.0	51.0	51.0
Sweden	62244	126	4.5	13.4	0.0	84.6	70.0	3.0	13.0	37.0	61.0	80.0	110.0	134.0	148.0	190.0
Switzerland	65140	383	17.5	9.6	0.0	53.7	40.0	0.3	7.0	8.0	13.0	39.0	70.0	102.0	200.0	320.0
United States	2278258	211	17.1	5.7	0.0	71.2	42.0	1.0	13.0	19.0	35.5	63.0	93.0	126.0	156.0	263.0

## 7.6 Test-Curriculum Matching Analysis Files

To investigate the match of the TIMSS advanced mathematics and physics tests to the curricula of each of the participating countries, TIMSS carried out a special analysis called the Test-Curriculum Matching Analysis (TCMA). Each country was to identify, for each advanced mathematics and physics item, whether the topic of the item was intended in the curriculum for the majority of the students. Results based on items considered appropriate are presented in Appendix C of Mullis, et al. (1998). The selection of the items by each country is included as part of the International Database.

There are two files that contain data indicating the items selected by each country. These files are located in the subdirectory called TCMA in the corresponding CD. The two files are:

TCMAMP3.CSV	Test-Curriculum Matching Analysis for Advanced Mathematics
TCMASP3.CSV	Test-Curriculum Matching Analysis for Physics

These files are in text format, with their fields separated by commas. The first record for each file contains the labels for each field in the file. Each row in the file contains a country's selection status for all items in the appropriate subject area:

- 1 = Item Included
- 0 = Item Excluded

---

# Chapter 8

## Estimating Sampling and Imputation Variance

When analyzing data from complex designs such as TIMSS, it is important to compute correct error variance estimates for the statistics of interest. In TIMSS this error variance comes from two sources: the sampling process and the imputation process. This chapter describes the methods used to estimate each of these error variance components.

### 8.1 Estimating Sampling Variance

With complex sampling designs that involve more than simple random sampling, as in the case of TIMSS where a multi-stage cluster design was used, there are several methods for estimating the sampling error of a statistic that avoid the assumption of simple random sampling. One such method is the jackknife repeated replication (JRR) technique (Wolter, 1985). The particular application of the JRR technique used in TIMSS is termed a paired selection model because it assumes that the sampled population can be partitioned into strata, with the sampling in each stratum consisting of two primary sampling units (PSU), selected independently. Following this first-stage sampling, there may be any number of subsequent stages of selection that may involve equal or unequal probability selection of the corresponding elements. The TIMSS design called for a total of 150 schools for the target population. These schools constituted the PSUs in most countries, and were paired sequentially after sorting by a set of implicit stratification variables. This resulted in the implicit creation of 75 strata, with two schools selected per stratum.

The jackknife repeated replication (JRR) method is suitable for estimating sampling errors in the TIMSS design because it provides approximately unbiased estimates of the sampling error arising from the complex sample selection procedure for estimates such as percentages, totals and means. In addition, this method can also be readily adapted to the estimation of sampling errors for parameters estimated using other statistical modeling procedures, such as percent-correct technology. The general use of the JRR entails systematically assigning pairs of schools to sampling zones, and the random selection of one of these schools to have its contribution doubled, and the other zeroed, so as to construct a number of “pseudo-replicates” of the original sample. The statistic of interest is computed once for all of the original sample, and once more for each of the pseudo-replicate samples. The variation between the original sample estimate and the estimates from each of the replicate samples is the jackknife estimate of the sampling error of the statistic.

#### 8.1.1 Computing Sampling Variance Using the JRR Method

When implementing the JRR method in TIMSS, for each country we assume that there were up to 75 strata or zones ( $H$ ) within each country, each one containing two sampled schools selected independently. When computing a statistic " $t$ " from the sample for a country, the formula for the JRR variance estimate of the statistic  $t$  is then given by the following equation:

$$Var_{jrr}(t) = \frac{1}{H} \sum_{h=1}^H [t(J_h) - t(S)]^2$$

where  $H$  is the number of pairs in the entire sample for the country. The term  $t(S)$  corresponds to

the statistic computed for the whole sample (computed with any specific weights that may have been used to compensate for the unequal probability of selection of the different elements in the sample or any other post-stratification weight). The element  $t(J_h)$  denotes the same statistic using the  $h$ th jackknife replicate, computed for all cases except those in the  $h$ th stratum of the sample, removing all cases associated with one of the randomly selected units of the pair within the  $h$ th stratum, and including, twice, the elements associated with the other unit in the  $h$ th stratum. In practice, this is effectively accomplished by recoding to zero the weights for the cases of the element of the pair to be excluded from the replication, and multiplying by two the weights of the remaining element within the  $h$ th pair. This results in a set of  $H$  replicate weights which may be used in computing the JRR variance estimate.

As we can see from the above formula, the computation of the JRR variance estimate for any statistic from the TIMSS database requires the computation of any statistic up to 76 times for any given country: once to obtain the statistic for the full sample, and up to 75 times to obtain the statistics for each of the jackknife replicates ( $J_h$ ). The number of times a statistic needs to be computed for a given country will depend on the number of implicit strata or sampling zones defined for the sample.

Note that when using the JRR technique for the estimation of sampling variability, the approach will appropriately reflect the combined effect of the between- and within-sampling zone contributions to the variance.

Doubling and zeroing the weights of the selected units within the strata or “zones” is accomplished effectively with the creation of replicate weights which are then used in the calculations. The next chapter shows how this approach allows standard statistical software such as SAS or SPSS to be used to compute JRR estimates of sampling variability. The replicate weight approach requires the user to temporarily create a new set of weights for each pseudo-replicate sample. Each replicate weight is equal to  $k$  times the overall sampling weight, where  $k$  can take values of zero, one, or two, depending on whether or not the case is to be removed from the computation, left as it is, or have its weight doubled. The value of  $k$  for an individual student record for a given replicate depends on the assignment of the record to the specific PSU and zone.

## 8.1.2 Construction of Sampling Zones for Sampling Variance Estimation

An important step in applying the JRR technique to the estimation of sampling variability consists of assigning the schools to implicit strata known as sampling zones. Most TIMSS sample designs in the participating countries called for a total of 150 sampled schools per target population. Each of these 150 school were assigned to one of 75 sampling zones. These zones were constructed by pairing the sequentially sampled schools and assigning these pairs to a sampling zone. Since schools were generally sorted by a set of implicit stratification variables, the resulting assignment of sequentially sampled schools to sampling zones takes advantage of any benefit due to this implicit stratification. In cases when more than 75 pairs of schools were sampled within a country, schools were then assigned to sampling zones in a way such that no more than 75 sampling zones were defined. In some cases this meant assigning more than two schools to the same sampling zone.

Sampling zones were constructed within explicit strata. In cases when there was an odd number



of schools in an explicit stratum, either by design or because of school-level non-response, the students in the remaining school were randomly divided into two "quasi" schools for the purposes of calculating the jackknife standard error. Each zone then consisted of a pair of schools or "quasi" schools. When computing replicate weights for the estimation of JRR sampling error, one member of each pair of schools is randomly selected to have its weights doubled, while the weights of the other member are set to zero.

The variable JKZONE indicates the sampling zone to which the student's school is assigned. The sampling zones can have values from 1 to 75 in the Student Background and Written Assessment data files. This variable is included in the Student Background and the student Written Assessment data files. For each individual student, this variable is identical in these files.

The variable JKINDIC indicates how the student is to be used in the computation of the replicate weights. This variable can have values of either 1 or 0. Those student records with a value of 0 should be excluded from the corresponding replicate weight, and those with a value of 1 should have their weights doubled. This variable is included in the Student Background and the student Written Assessment data files. For each individual student, this variable is identical in these two files.

### 8.1.3 Computing the JRR Replicate Weights

Having assigned the schools to zones, if it is desired to use standard statistical software such as SAS or SPSS for sampling variance estimation, a convenient way to proceed is to construct a set of replicate weights for each pseudo-replicate sample. In TIMSS, the schools in the sample were assigned in pairs to one of the 75 zones indicated by the variable JKZONE, and within each zone the pair members were randomly assigned an indicator ( $u_i$ ) represented by the variable JKINDIC, coded randomly to 1 or 0 so that one of the members of each pair had values of 1 on the variable  $u_i$ , and the remaining one a value of 0. This indicator determined whether the weights for the elements in the school in this zone was to be doubled or zeroed. The replicate weight ( $W_h^{g,i,j}$ ) for the elements in a school assigned to zone  $h$  is computed as the product of  $k_h$  times their overall sampling weight, where  $k_h$  could take values of zero, one, or two depending on if the case was not to contribute, be left unchanged, or have it count twice for the computation of the statistic of interest. In TIMSS, the replicate weights are not permanent variables, but are created temporarily by the sampling variance estimation program as a useful computing device. An example program which makes use of replicate weights in computing JRR estimates is provided in the next chapter.

When creating the replicate weights the following procedure is followed:

1. Each sampled student is assigned a vector of 75 weights or  $W_h^{g,i,j}$ , where  $h$  takes values from 1 to 75.
2. The value of  $W_0^{g,i,j}$  is the overall sampling weight which is simply the product of the final school weight, the appropriate final classroom weight, and the appropriate final student weight as defined in Chapter 3.
3. The replicate weights for a single case are then computed as:

$$W_h^{g,i,j} = W_0^{g,i,j} * k_{hi},$$

where the variable  $k_h$  for an individual  $i$  takes the value  $k_{hi}=2*u_i$  if the record belongs to zone  $h$ ,

and  $k_h=1$  otherwise.

In TIMSS, a total of 75 replicate weights were computed regardless of the number of actual zones within a country. If a country had fewer than 75 zones, then the replicate weights  $W_h$ , where  $h$  was greater than the number of zones within the country, were each the same as the overall sampling weight. Although this involves some redundant computation, having 75 replicate weights for each country has no effect on the size of the error variance computed using the jackknife formula, but facilitates the computation of standard errors for a number of countries at one time.

## 8.2 Estimating Imputation Variance

The general procedure for estimating the imputation variance using plausible values is as follows: first estimate the statistic ( $t$ ), each time using a different set of the plausible values ( $M$ ). The statistics  $t_m$  can be anything estimable from the data, such as a mean, the difference between means, percentiles, etc. If all of the ( $M=5$ ) plausible values in the TIMSS database are used, the parameter will be estimated five times, once using each set of plausible values. Each of these estimates will be called  $t_m$ , where  $m=1,2,\dots,5$ .

Once the statistics are computed the imputation variance is then computed as:

$$Var_{imp} = \left(1 + \frac{1}{M}\right) \times Var(t_m)$$

where  $M$  is the number of plausible values used in the calculation, and  $Var(t_m)$  is the variance of the estimates computed using each plausible value.

## 8.3 Combining Sampling and Imputation Variance

When reporting standard errors for proficiency estimates using plausible values the user will need to combine the sampling and imputation components of the error variance for the estimate. Under ideal circumstances, and with unlimited computing resources, the user would compute the imputation variance for the plausible values, and the JRR sampling variance for each of the plausible values. This would be equivalent to computing the same statistic up to 380 times (once for each of the five plausible values, and 75 times for each set of replicate weights for each plausible value). However, an acceptable shortcut is to compute the JRR variance component using one plausible value, and then the imputation variance using the five plausible values. Using this approach, the same statistic needs to be computed only 80 times. Under this procedure the error variance component for a statistic computed using plausible values is computed using the following formula:

$$Var(t_{pv}) = Var_{jrr}(t_1) + Var_{imp}$$

where  $Var_{jrr}(t_1)$  is the sampling variance for the first plausible value. This User Guide contains programs in SAS and SPSS that compute each of these variance components for the TIMSS data. Tables 8.1 through 8.14 show for each set of plausible values, the mean of the first plausible value, the JRR error for the first plausible value, the mean of the five plausible values, the standard error of this mean, the JRR sampling variance for the first plausible value, and the imputation variance.

**Table 8.1**  
**Sampling and Imputation Variance for the Mathematics and Science Literacy Scale**  
**Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	1941	524.71	9.5	89.88	0.21	524.60	9.5
Austria	1962	519.41	5.3	28.20	0.60	519.03	5.4
Canada	5232	525.80	2.5	6.24	0.51	525.64	2.6
Cyprus	534	446.27	2.4	5.62	0.64	446.68	2.5
Czech Republic	2167	476.39	10.5	110.17	0.23	476.28	10.5
Denmark	2714	528.20	3.1	9.84	0.19	528.11	3.2
France	1590	504.28	4.8	23.17	0.60	505.20	4.9
Germany	2289	495.78	5.2	26.48	2.45	495.74	5.4
Hungary	5091	477.17	3.0	8.93	0.21	476.94	3.0
Iceland	1703	540.72	1.5	2.24	0.30	541.30	1.6
Israel	1357	484.45	12.1	145.66	0.69	483.71	12.1
Italy	1616	475.28	5.2	27.32	0.56	475.40	5.3
Lithuania	2887	464.88	5.7	32.98	0.21	464.98	5.8
Netherlands	1470	558.29	4.8	22.81	0.95	558.83	4.9
New Zealand	1763	524.58	4.6	21.28	0.44	525.10	4.7
Norway	2518	536.55	4.0	16.10	0.28	535.97	4.1
Russian Federation	2289	475.11	5.8	33.11	0.64	475.83	5.8
Slovenia	1622	515.20	8.1	66.16	0.30	514.45	8.2
South Africa	2757	353.04	9.3	86.92	0.24	352.43	9.3
Sweden	3068	555.48	4.3	18.40	0.21	555.36	4.3
Switzerland	3308	531.95	5.3	28.39	1.10	531.44	5.4
United States	5807	470.91	3.1	9.59	0.23	470.67	3.1

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance  
Some results may appear inconsistent because of rounding.

**Table 8.2**  
**Sampling and Imputation Variance for the Mathematics Literacy Scale Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	1941	521.67	9.3	86.75	0.14	522.13	9.3
Austria	1962	518.98	5.3	27.85	0.49	517.96	5.3
Canada	5232	519.32	2.8	7.71	0.36	519.11	2.8
Cyprus	534	445.16	2.4	5.95	0.49	445.84	2.5
Czech Republic	2167	465.52	12.3	150.80	0.61	466.05	12.3
Denmark	2714	547.48	3.3	10.69	0.26	547.27	3.3
France	1590	522.60	5.0	25.22	0.61	523.42	5.1
Germany	2289	495.22	5.6	31.34	3.72	494.96	5.9
Hungary	5091	483.59	3.2	9.90	0.31	483.31	3.2
Iceland	1703	533.10	2.0	3.95	0.24	533.65	2.1
Israel	1357	480.51	12.1	147.30	0.37	480.15	12.2
Italy	1616	475.35	5.3	28.46	1.31	475.90	5.5
Lithuania	2887	469.33	6.0	36.12	0.54	469.46	6.1
Netherlands	1470	559.64	4.7	21.71	0.51	559.93	4.7
New Zealand	1763	521.70	4.4	19.75	0.64	521.56	4.5
Norway	2518	528.92	4.1	16.91	0.28	528.22	4.2
Russian Federation	2289	470.15	6.1	37.09	1.03	470.64	6.2
Slovenia	1622	512.40	8.3	68.54	0.51	511.56	8.3
South Africa	2757	356.19	8.3	68.95	0.36	356.03	8.3
Sweden	3068	551.82	4.3	18.43	0.24	551.67	4.3
Switzerland	3308	540.10	5.7	32.03	1.18	540.22	5.8
United States	5807	461.79	3.1	9.63	0.52	461.10	3.2

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance

Some results may appear inconsistent because of rounding.

**Table 8.3**  
**Sampling and Imputation Variance for the Science Literacy Scale Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	1941	527.75	9.7	94.96	1.18	527.07	9.8
Austria	1962	519.83	5.5	29.82	1.67	520.10	5.6
Canada	5232	532.29	2.5	6.02	0.78	532.18	2.6
Cyprus	534	447.39	2.8	8.03	0.93	447.52	3.0
Czech Republic	2167	487.26	8.8	77.43	0.37	486.51	8.8
Denmark	2714	508.91	3.6	12.92	0.14	508.95	3.6
France	1590	485.97	5.0	24.67	1.38	486.98	5.1
Germany	2289	496.33	4.9	24.15	1.64	496.53	5.1
Hungary	5091	470.74	3.0	8.79	0.15	470.57	3.0
Iceland	1703	548.35	1.4	1.96	0.41	548.96	1.5
Israel	1357	488.40	12.3	151.14	1.36	487.27	12.4
Italy	1616	475.21	5.3	27.95	0.35	474.89	5.3
Lithuania	2887	460.44	5.7	31.99	0.20	460.51	5.7
Netherlands	1470	556.94	5.1	25.90	1.98	557.73	5.3
New Zealand	1763	527.47	5.1	25.98	0.91	528.63	5.2
Norway	2518	544.19	4.1	16.76	0.37	543.73	4.1
Russian Federation	2289	480.07	5.6	31.80	0.71	481.02	5.7
Slovenia	1622	518.00	8.1	66.12	0.41	517.35	8.2
South Africa	2757	349.89	10.4	108.56	0.66	348.84	10.5
Sweden	3068	559.15	4.4	19.15	0.20	559.04	4.4
Switzerland	3308	523.80	5.2	26.98	1.40	522.67	5.3
United States	5807	480.03	3.2	10.45	0.14	480.23	3.3

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance  
 Some results may appear inconsistent because of rounding.

**Table 8.4**  
**Sampling and Imputation Variance for the Advanced Mathematics Scale Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	645	523.77	11.6	134.10	1.40	524.71	11.6
Austria	782	435.24	7.2	51.76	0.38	436.06	7.2
Canada	2781	509.13	4.2	17.53	0.68	509.49	4.3
Cyprus	391	518.05	4.1	16.56	2.25	518.21	4.3
Czech Republic	1101	470.42	11.1	123.39	1.73	468.89	11.2
Denmark	1388	522.68	3.3	10.76	0.64	522.45	3.4
France	1071	557.53	3.8	14.46	0.75	556.94	3.9
Germany	2296	465.90	5.6	30.76	0.81	464.81	5.6
Greece	456	511.31	5.9	35.14	1.45	512.56	6.1
Israel	953	557.62	5.3	27.56	0.82	557.27	5.3
Italy	398	474.85	9.6	91.11	0.39	474.23	9.6
Lithuania	734	516.24	2.4	5.82	0.85	516.35	2.6
Russian Federation	1638	542.33	9.2	84.05	1.30	542.25	9.2
Slovenia	1536	474.55	9.1	83.63	0.81	474.69	9.2
Sweden	1001	512.24	4.3	18.53	0.71	511.99	4.4
Switzerland	1404	533.50	4.8	23.04	1.83	533.13	5.0
United States	2785	441.15	5.8	33.40	1.19	441.99	5.9

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance  
 Some results may appear inconsistent because of rounding.

**Table 8.5**  
**Sampling and Imputation Variance for the Numbers, Equations, and Functions Scale Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	645	517.76	9.2	85.45	2.09	517.46	9.4
Austria	782	411.76	7.4	54.29	0.88	412.00	7.4
Canada	2781	513.46	3.8	14.12	0.99	512.06	3.9
Cyprus	391	508.05	5.4	29.29	2.69	509.77	5.7
Czech Republic	1101	460.04	11.6	134.58	1.79	460.41	11.7
Denmark	1388	504.28	2.6	6.92	0.45	503.76	2.7
France	1071	547.95	4.1	16.65	0.13	548.31	4.1
Germany	2296	457.01	5.0	24.53	0.68	457.00	5.0
Greece	456	540.39	6.7	45.39	5.81	538.99	7.2
Israel	953	548.38	4.3	18.19	2.33	547.03	4.5
Italy	398	460.24	9.2	84.82	0.11	459.87	9.2
Lithuania	734	546.08	2.3	5.28	2.83	547.37	2.9
Russian Federation	1638	553.62	8.7	75.27	1.91	555.19	8.8
Slovenia	1536	492.01	9.8	96.62	0.52	491.04	9.9
Sweden	1001	523.12	4.7	22.02	0.45	523.23	4.7
Switzerland	1404	513.14	5.1	25.70	1.15	514.01	5.2
United States	2785	459.88	5.2	26.66	1.22	458.78	5.3

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance  
 Some results may appear inconsistent because of rounding.

**Table 8.6**  
**Sampling and Imputation Variance for the Calculus Scale Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	645	530.29	11.6	134.86	1.11	529.84	11.7
Austria	782	439.23	6.4	41.09	1.68	439.30	6.5
Canada	2781	505.19	3.4	11.40	1.39	503.39	3.6
Cyprus	391	558.82	5.0	24.82	1.99	560.62	5.2
Czech Republic	1101	445.20	9.7	94.02	0.49	445.79	9.7
Denmark	1388	508.54	3.3	10.91	0.10	508.27	3.3
France	1071	560.40	2.9	8.29	0.47	559.53	3.0
Germany	2296	454.78	4.3	18.43	0.51	454.03	4.4
Greece	456	539.42	6.8	45.96	7.76	538.44	7.3
Israel	953	538.79	4.2	17.46	2.10	537.69	4.4
Italy	398	520.06	10.4	107.58	0.19	520.11	10.4
Lithuania	734	496.99	2.0	4.02	2.33	498.36	2.5
Russian Federation	1638	535.88	9.0	81.02	1.76	537.13	9.1
Slovenia	1536	472.13	6.6	42.95	0.59	471.22	6.6
Sweden	1001	480.17	4.3	18.48	0.55	480.36	4.4
Switzerland	1404	511.46	5.6	30.87	1.31	512.50	5.7
United States	2785	449.65	4.0	15.71	0.96	449.54	4.1

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance  
Some results may appear inconsistent because of rounding.

**Table 8.7**  
**Sampling and Imputation Variance for the Geometry Scale Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	645	497.21	12.4	153.40	3.74	496.26	12.5
Austria	782	461.79	7.8	61.40	0.29	461.85	7.9
Canada	2781	500.40	3.7	13.63	0.75	499.38	3.8
Cyprus	391	515.01	4.7	21.63	2.55	516.89	4.9
Czech Republic	1101	494.24	9.8	95.26	0.74	494.36	9.8
Denmark	1388	526.07	2.9	8.21	1.35	527.20	3.1
France	1071	543.89	3.7	13.39	0.80	544.08	3.8
Germany	2296	487.21	5.4	29.64	0.24	486.88	5.5
Greece	456	500.14	8.0	64.17	10.81	497.66	8.7
Israel	953	562.39	5.3	27.82	1.97	561.90	5.5
Italy	398	479.95	9.5	89.61	0.30	479.89	9.5
Lithuania	734	514.74	2.3	5.39	2.64	515.30	2.8
Russian Federation	1638	546.52	9.0	81.24	2.62	548.22	9.2
Slovenia	1536	477.66	7.5	56.50	0.75	476.36	7.6
Sweden	1001	492.35	4.4	19.24	0.32	492.37	4.4
Switzerland	1404	546.15	4.1	16.67	1.11	547.19	4.2
United States	2785	423.93	5.1	25.61	0.49	423.68	5.1

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance  
Some results may appear inconsistent because of rounding.

**Table 8.8**  
**Sampling and Imputation Variance for the Physics Scale Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	661	514.77	5.8	34.05	4.51	517.88	6.2
Austria	777	434.89	6.4	40.42	0.63	434.64	6.4
Canada	2367	483.51	2.8	7.66	3.14	485.38	3.3
Cyprus	368	494.99	5.7	32.77	0.67	494.35	5.8
Czech Republic	1087	450.84	6.2	38.06	0.49	451.40	6.2
Denmark	654	534.01	4.0	15.56	1.76	534.36	4.2
France	1110	465.10	3.5	11.90	2.18	466.32	3.8
Germany	723	520.60	11.8	140.28	1.49	521.91	11.9
Greece	459	487.33	5.5	29.68	1.51	486.49	5.6
Israel	853	505.36	6.3	39.64	1.82	505.65	6.4
Italy	145	435.97	10.3	105.91	0.89	435.89	10.3
Latvia (LSS)	708	488.97	21.5	462.57	0.32	488.18	21.5
Norway	1048	580.96	6.1	37.12	5.24	580.92	6.5
Russian Federation	1233	548.52	11.4	129.82	5.54	545.20	11.6
Slovenia	747	524.42	15.3	233.13	6.03	522.62	15.5
Sweden	1012	572.81	3.8	14.63	0.60	572.94	3.9
Switzerland	1371	489.01	3.4	11.63	0.59	488.23	3.5
United States	3114	422.29	3.2	10.21	0.50	422.53	3.3

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance  
Some results may appear inconsistent because of rounding.

**Table 8.9**  
**Sampling and Imputation Variance for the Mechanics Scale Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	661	505.12	5.9	35.22	1.46	506.87	6.1
Austria	777	420.58	4.9	23.99	0.04	420.44	4.9
Canada	2367	473.41	3.5	12.36	0.66	473.35	3.6
Cyprus	368	530.21	6.6	43.69	0.17	530.23	6.6
Czech Republic	1087	468.70	5.9	35.02	0.57	468.55	6.0
Denmark	654	528.87	4.7	22.15	1.40	529.42	4.9
France	1110	459.19	3.9	14.99	3.18	456.83	4.3
Germany	723	495.87	9.2	84.81	2.81	494.79	9.4
Greece	459	512.59	6.3	39.39	3.11	513.69	6.5
Israel	853	547.64	5.5	29.97	0.12	547.77	5.5
Italy	145	421.05	14.0	194.85	11.46	420.26	14.4
Latvia (LSS)	708	489.20	18.1	327.74	0.39	488.78	18.1
Norway	1048	568.91	6.0	35.40	5.32	571.64	6.4
Russian Federation	1233	538.09	9.3	86.44	0.34	537.49	9.3
Slovenia	747	554.16	17.2	297.15	3.43	552.25	17.3
Sweden	1012	563.16	4.0	15.94	0.17	563.19	4.0
Switzerland	1371	481.88	3.5	12.10	0.12	482.28	3.5
United States	3114	420.43	2.7	7.47	0.28	420.46	2.8

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance  
Some results may appear inconsistent because of rounding.



**Table 8.10**  
**Sampling and Imputation Variance for the Electricity and Magnetism Scale**  
**Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	661	510.66	4.2	17.65	1.30	512.30	4.4
Austria	777	431.48	6.2	38.81	1.34	431.90	6.3
Canada	2367	484.91	3.4	11.38	2.01	484.50	3.7
Cyprus	368	502.18	6.2	38.87	1.08	502.11	6.3
Czech Republic	1087	465.22	5.4	29.65	0.49	464.96	5.5
Denmark	654	512.14	3.5	12.46	2.15	512.63	3.8
France	1110	494.13	4.0	16.29	0.20	493.71	4.1
Germany	723	515.08	9.6	92.35	5.19	512.12	9.9
Greece	459	519.51	6.4	40.85	2.86	519.68	6.6
Israel	853	556.31	6.7	44.62	0.33	556.59	6.7
Italy	145	468.40	8.7	76.28	34.60	472.58	10.5
Latvia (LSS)	708	484.52	17.3	300.75	1.19	484.89	17.4
Norway	1048	563.55	5.9	35.15	3.28	564.72	6.2
Russian Federation	1233	550.25	9.0	80.80	3.04	549.16	9.2
Slovenia	747	509.57	14.6	211.74	2.23	508.91	14.6
Sweden	1012	569.67	3.2	10.44	0.53	569.93	3.3
Switzerland	1371	479.01	4.5	19.98	0.60	479.87	4.5
United States	3114	419.09	2.9	8.62	0.55	419.60	3.0

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance  
Some results may appear inconsistent because of rounding.

**Table 8.11**  
**Sampling and Imputation Variance for the Heat Scale Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	661	515.03	4.1	16.96	1.54	517.02	4.3
Austria	777	443.72	5.6	30.75	0.87	444.54	5.6
Canada	2367	508.31	3.7	13.61	4.22	507.70	4.2
Cyprus	368	475.76	6.7	44.47	0.19	475.68	6.7
Czech Republic	1087	487.75	4.6	21.51	0.60	488.06	4.7
Denmark	654	512.71	4.0	15.65	2.50	512.39	4.3
France	1110	491.17	3.3	10.55	0.98	490.74	3.4
Germany	723	497.56	5.8	34.04	7.01	495.85	6.4
Greece	459	480.36	7.1	50.10	1.90	480.51	7.2
Israel	853	478.57	4.0	16.22	0.40	478.31	4.1
Italy	145	489.79	7.6	57.16	13.73	489.73	8.4
Latvia (LSS)	708	503.00	21.3	454.99	1.81	503.57	21.4
Norway	1048	533.98	3.7	14.02	4.45	536.17	4.3
Russian Federation	1233	531.45	10.3	106.66	2.49	530.45	10.5
Slovenia	747	521.85	10.2	103.63	3.71	520.69	10.4
Sweden	1012	522.37	4.2	17.79	0.82	521.89	4.3
Switzerland	1371	508.55	3.4	11.85	0.88	509.04	3.6
United States	3114	477.16	2.9	8.55	0.47	476.80	3.0

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance  
 Some results may appear inconsistent because of rounding.

**Table 8.12**  
**Sampling and Imputation Variance for the Wave Phenomena Scale Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	661	516.68	6.6	43.92	3.36	518.81	6.9
Austria	777	468.39	7.2	52.37	0.25	467.83	7.3
Canada	2367	488.40	2.8	8.05	2.18	488.05	3.2
Cyprus	368	507.31	6.5	42.20	0.15	507.01	6.5
Czech Republic	1087	446.93	5.4	29.30	0.18	446.88	5.4
Denmark	654	536.72	5.2	27.12	3.06	536.65	5.5
France	1110	463.43	3.6	13.27	0.04	463.44	3.7
Germany	723	531.55	10.0	100.54	5.08	530.09	10.3
Greece	459	453.22	5.2	27.47	0.74	453.30	5.3
Israel	853	444.11	6.1	36.79	0.29	444.08	6.1
Italy	145	446.62	14.2	201.58	38.34	444.56	15.5
Latvia (LSS)	708	498.73	17.6	307.95	0.92	498.10	17.6
Norway	1048	558.45	5.1	25.93	2.87	560.19	5.4
Russian Federation	1233	515.43	9.2	84.05	3.84	515.13	9.4
Slovenia	747	514.18	11.4	130.89	1.96	513.54	11.5
Sweden	1012	559.78	4.6	21.42	0.65	560.05	4.7
Switzerland	1371	496.99	3.0	8.95	0.51	497.68	3.1
United States	3114	451.36	2.2	4.62	0.34	451.08	2.2

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance  
 Some results may appear inconsistent because of rounding.

**Table 8.13**  
**Sampling and Imputation Variance for the Modern Physics Scale Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	661	518.83	5.6	31.42	1.67	520.88	5.8
Austria	777	479.57	5.9	34.96	1.47	479.95	6.0
Canada	2367	495.13	2.7	7.09	0.28	494.40	2.7
Cyprus	368	434.78	5.2	26.69	0.13	434.31	5.2
Czech Republic	1087	453.49	4.9	23.55	0.89	453.39	4.9
Denmark	654	542.72	4.7	22.36	1.34	543.60	4.9
France	1110	473.12	3.3	10.84	0.49	474.05	3.4
Germany	723	546.83	12.8	164.57	6.17	545.00	13.1
Greece	459	445.84	4.9	24.08	0.41	446.54	5.0
Israel	853	475.82	7.2	51.48	0.27	475.94	7.2
Italy	145	422.79	8.9	79.08	6.54	420.95	9.3
Latvia (LSS)	708	488.62	19.0	360.22	1.71	487.62	19.0
Norway	1048	574.73	5.1	25.64	2.09	575.92	5.3
Russian Federation	1233	543.21	9.8	95.73	2.07	542.45	9.9
Slovenia	747	512.98	15.0	225.24	3.67	511.17	15.1
Sweden	1012	559.73	3.4	11.83	0.69	559.84	3.5
Switzerland	1371	486.87	3.7	13.97	0.68	487.83	3.8
United States	3114	456.72	2.4	5.84	0.36	456.31	2.5

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance  
 Some results may appear inconsistent because of rounding.

**Table 8.14**  
**Sampling and Imputation Variance for the Reasoning and Social Utility Scale**  
**Plausible Values**

Country	Number of Cases	Mean of 1st Plausible Value	JRR Error of 1st Plausible Value	JRR Sampling Variance	Imputation Variance	Mean of 5 Plausible Values	Standard Error*
Australia	1941	523.60	8.4	70.71	2.06	524.46	8.5
Austria	1962	514.73	5.7	32.57	0.77	515.59	5.8
Canada	5232	531.77	2.5	6.30	0.66	530.68	2.6
Cyprus	534	450.56	2.5	6.48	3.22	453.00	3.1
Czech Republic	2167	484.41	9.9	97.51	1.04	483.61	9.9
Denmark	2714	536.59	2.3	5.30	0.51	536.09	2.4
France	1590	511.83	3.9	15.19	1.49	511.24	4.1
Germany	2289	499.60	5.5	30.31	3.14	497.58	5.8
Hungary	5091	487.22	3.1	9.39	0.43	486.29	3.1
Iceland	1703	550.97	1.3	1.63	0.87	550.51	1.6
Israel	1357	484.76	12.2	148.40	0.52	485.01	12.2
Italy	1616	466.48	5.6	31.30	0.64	467.56	5.7
Lithuania	2887	459.61	5.8	34.02	0.92	460.71	5.9
Netherlands	1470	568.87	4.1	16.90	0.64	568.04	4.2
New Zealand	1763	525.41	4.3	18.24	0.64	526.50	4.3
Norway	2518	540.63	4.2	17.85	0.75	541.71	4.3
Russian Federation	2289	470.83	6.0	36.07	2.77	470.67	6.2
Slovenia	1622	495.47	7.6	57.34	0.09	495.91	7.6
South Africa	2757	320.06	11.0	120.75	0.43	319.81	11.0
Sweden	3068	557.71	3.9	14.98	0.48	557.18	3.9
Switzerland	3308	533.54	4.3	18.80	0.55	533.44	4.4
United States	5807	470.11	3.1	9.56	0.43	469.43	3.2

\*Standard error is computed as the square root of the sum of the JRR Sampling Variance and the Imputation Variance

Some results may appear inconsistent because of rounding.



---

# Chapter 9

## Performing Analyses with the TIMSS Data: Some Examples

This chapter presents some basic examples of analyses that can be performed with the TIMSS International Database using the sampling weights, plausible values, and selection variables discussed in previous chapters. It also provides details on some SPSS and SAS code that can be used to conduct such analyses, and how to interpret the output from these programs. Although the analyses presented here are simple in nature, they are designed to familiarize the user with the different files and their structure, as well as the relevant variables that need to be included in most analyses. All the analyses presented here compute the percent of students in specified subgroups, the mean achievement in mathematics and science literacy or on another continuous variable for each subgroup, and the corresponding standard errors for the percent and mean statistics. Most of the example analyses presented in this chapter, based on student data, replicate analyses that are included in the TIMSS international report (Mullis et al., 1998). Copies of the tables from the international report that are replicated in the examples that follow also are included in this chapter. The user is invited to compare the results from these analyses to the tables in the reports, and is encouraged to practice analyzing the TIMSS data by trying to replicate some of the tables that are presented in the international report.<sup>1</sup>

In our examples we use macros written for SAS and SPSS that can be used to perform any of the analyses that are described below. These are general subroutines that can be used for many purposes, provided the user has some basic knowledge of the SAS or SPSS macro language. The user with some programming experience in either one of these statistical packages will be able to make the necessary modifications to the macros to obtain the desired results. When using these macros, the code assumes that the user has already created a system file in SPSS, or a data set if using SAS, that contains the variables necessary for the analysis. As part of this chapter we also describe the control files included in the CD that can be used to create SAS and SPSS data set system files.

### 9.1 Contents of the CD

One CD accompanies this User Guide. The CD has the following internal file structure:

- A main directory identifying the population (POP3).
- Within each main directory, there are five sub-directories.
  - DATA : Contains ASCII data files
  - PROGRAMS : Contains SPSS and SAS programs
  - CODEBOOK : Contains codebooks
  - ALMANACS : Contains data almanacs
  - TCMA : Contains Test–Curriculum Matching Analysis Data

<sup>1</sup> Documentation regarding the computational methods used to obtain any derived variables included in the international report is presented in Supplement 3 of this User Guide.

The directory names within the CD and the file names follow the DOS naming convention: file names with up to eight characters, followed by a three-character extension (as in FILENAME.EXT). Files with the same names are related to each other, and the extension identifies their function. The extensions used in the files contained in the CDs are indicated in Table 9.1 below.

**Table 9.1**  
**Three-letter Extension Used to Identify the Files Contained in the CD**

Extension	Description
.SAS	SAS Control File or program
.SPS	SPSS Control File or program
.DAT	ASCII data file
.ALM	Almanac
.CDT	Codebook in text format
.CDF	Codebook in machine readable format

The DATA sub-directory contains the TIMSS data files in ASCII format. The different data file types that are in this directory are described in Chapter 7. Each of these files has two corresponding control files in the PROGRAMS sub-directory, as shown in Table 7.13. One of these two files reads the ASCII data and creates a SAS data set, the second one reads the ASCII data and creates an SPSS system file. There are several other programs in this directory. The other programs that can be found in this directory are the following:

#### ***JACKGEN.SAS and JACKGEN.SPS***

Both of these are macro programs, one in SAS and one in SPSS, that can be used to compute the weighted percent of students within defined groups, and their mean on a specified continuous variable (other than the plausible value scales). These macros also generate replicate weights and compute JRR sampling variance for the percent of students within the group defined by the classification variable(s) specified, and JRR sampling variance for the mean estimates on the continuous variable. These macros should not be used with plausible values, since they make no provision for estimating imputation error. Otherwise, the continuous variable can be any continuous variable in the file. Instructions on how to use these macros are included later in this chapter.

#### ***JACKPV.SAS and JACKPV.SPS***

Both of these are macro programs, one in SAS and one in SPSS, that can be used to compute the weighted percent of students within defined groups, and their mean achievement score on an achievement scale using the available plausible values. These macros make use of the plausible values provided for each subject in computing the mean achievement score. These macros also generate replicate weights and compute the JRR sampling variance for the percent of students within the group, and the JRR and imputation variance for the mean achievement



score. These macros should only be used when multiple plausible values are used in the analysis.

***EXAMPLE1.SAS, EXAMPLE2.SAS and EXAMPLE3.SAS***  
***EXAMPLE1.SPS, EXAMPLE2.SPS and EXAMPLE3.SPS***

These are the programs used in the examples presented later in this chapter. These programs are included in the CD.

***CSASCORE.SAS***  
***CSASCORE.SPS***

These files contain control code in SAS and SPSS that can be used to convert the response codes to the cognitive items discussed in Chapter 7 to their corresponding correctness score levels. The use of these programs is described later in this chapter.

## **9.2 Creating SAS Data Sets and SPSS System Files**

The CD contains SAS and SPSS control code to read each one of the ASCII data files and create a SAS data set or an SPSS system file. An extract of the main sections of the control code for reading the Student Background data file using SAS is presented in Figure 9.1, and for creating the file in SPSS in Figure 9.2. Each of these files contain information on the location for each variable in the file, its format, a descriptive label for each variable and their categories (in the case of categorical variables), and code for handling missing data. The control and data files have been created to facilitate access of the data on a country by country basis. The command lines in the control files need to be edited to produce programs that will create SAS or SPSS files for any specified country. While most of the program code is functional, users will need to edit input and output file names. Performing analyses that require the data from more than one country will necessitate merging the respective data files into a larger one. Alternatively, the user can access the data and compute the necessary statistics on a country by country basis by reading one file at a time, computing the necessary statistics, and then moving on to the next country's data. The method chosen by the user will depend greatly on the storage and processing capacity of the computer system that is used. For the examples that we present in this User Guide we have combined the data files of individual countries into one larger data file that contains the data for all participating countries. The three character identifier for this file is "ALL."

When creating a dataset using SAS you will need to do the following:

1. Open the corresponding control file. For example, CSGCONTR.SPS, as in our example shown in Figure 9.1, to create the dataset for the student background data for a particular country.
2. In the line where it reads LIBNAME LIBDATA substitute <path> with the directory location where you want the SAS dataset saved. In the example in Figure 9.1 we'll be storing the dataset in a directory called "C:\TIMSS\POP3." The directory location must be included in quotes.

3. In the line where it reads `LIBNAME LIBRARY` substitute `<path>` with the directory location where you want the SAS formats to be saved. In our example in Figure 9.1 we'll be storing the format in a directory called "C:\TIMSS\POP3\FORMATS". The directory location must be included in quotes.
4. In the line where it reads `FILENAME RAWINP` you will need to substitute `<path><filename>` with the name and location of the raw data file that you will be reading. Keep in mind that the first three characters of the raw data file should match the first three characters of the control file provided in the CD. In our example in Figure 9.1 we will be reading the file "C:\TIMSS\POP3\DATA\CSGAUS1.DAT". The name of the data file should be included in quotes.
5. In the line where it reads `%LET FNAME` you will need to enter the name you want to use for the data set that will be created with the data that will be read with the program. Although it is not strictly necessary, we recommend that you use the same naming convention to name this file. In our example we will be calling the dataset created `CSGAUS1`. This dataset will be stored in the directory specified in the `LIBNAME LIBDATA` command.
6. If you do not want to store the formats with the variables in your file, you will need to either delete or comment out all the lines that start with the word `FORMAT`.
7. Submit the code for processing. After processing is complete you will find the dataset in the directory you specified in numeral 2 above, under the name you specified in numeral 5 above.

When creating a system file using SPSS you will need to do the following:

1. Open the corresponding control file. For example, `CSGCONTR.SPS`, as in our example shown in Figure 9.2, to create the dataset for the student background data for a particular country.
2. In the line where it reads `FILE HANDLE RAWINP` you will need to substitute `<path><filename>` with the name and location of the raw data file that you will be reading. Keep in mind that the first three characters of the raw data file should match the first three characters of the control file provided in the CD. In our example in Figure 9.2 we will be reading the file "C:\TIMSS\POP3\DATA\CSGAUS1.DAT". The name of the data file should be included in quotes.
3. In the line where it reads `FILE HANDLE SYSOUT` you will need to substitute `<path><filename>` with the name and location where you want your system file stored. Although it is not strictly necessary, we recommend that you use the same naming convention to name this file. In our example we will be calling the dataset created `CSGAUS1.SAV`.
4. Submit the code for processing. After processing is complete you will find the system file in the location and under the name you specified in numeral 3 above.

**Figure 9.1****Extract of the SAS Control Code for Creating a Student Background SAS Dataset**

```

Options errorabend nosource ;
LIBNAME libdat "c:\timss\pop3" ;
LIBNAME library "c:\timss\pop3\formats" ;
*NOTE: Please change in the following lines the country code ;
*      according to your country ;
FILENAME rawinp "c:\timss\pop3\DATA\CSGAUS1.DAT" ;

%LET FNAME = CSGAUS ;

PROC FORMAT LIBRARY=library ;
.
. { Formats are included here}
.
DATA LIBDAT.&FNAME ;

*NOTE: LRECL may need to be modified ;
INFILE rawinp LRECL=1400 END=eof LENGTH=lnline LINE=ln MISSOEVER ;
*LENGTH STATEMENTS FOR ALL VARIABLES ;
LENGTH DEFAULT=5 ;
.
. {Length statements are included here}
.

*MISSING CATEGORIES ;
MISSING A B R N ;
      INPUT
.
. {Variable names, location and format are include here}
.

%PUT INFO: assign value labels to variables ;
.
. {Assignment of formats to variables are included here}
.

%PUT INFO: assign labels to variables ;
      LABEL
.
. {Labels are assigned here}
.
.
. {Missing values are recoded to character values}
.

```

## Figure 9.2

### Extract of the SPSS Control Code for Creating a Student Background SPSS System File

```

/* Program cration, Date: 28.05.98 */
/* File: CSGCODE.SDB */
/* Please read the NOTES below carefully */

* NOTE: Assignment of variable names.
* NOTE: Please change <path> and <filename> according to your system.
*       Change this in two of the following lines of this program.
file handle rawinp / name ="c:\timss\pop3\DATA\CSGAUS1.DAT"
                / lrecl=1400.
file handle sysout / name ="CSGAUS1.SAV".

data list file = rawinp
.
. {Variable names, location and format are include here}
.
*NOTE: Assignment of variable labels.
variable labels      VERSION      "*FILE VERSION*"
.
. {Variable labels are assigned here}
.

VALUE LABELS
.
. {Value Labels are assigned here}
.

*RECODE of sysmis-values.
.
. {Missing values are recoded here to SYSMIS}
.

*RECODE of other missing-values.
.
. {Other missing value are recoded here}
.

* Save the data in SYSOUT
save outfile = sysout.

```

## 9.3 Computing JRR Replicate Weights, Sampling, and Imputation Variance for the Achievement Scores Using SPSS and SAS

In this section, example SAS and SPSS code that may be used to compute the JRR standard errors for mean plausible values and percentages is described. This code is provided in the form of SAS and SPSS macros. These macros compute the percentage of students within a subgroup defined by a set of classification variables, the JRR standard error of this percentage, the mean for the group on one of the achievement scales using plausible values, and the standard error of that mean including the sampling and imputation variance components.

Each of these macros operate as follows:

1. Computes a set of replicate weights specified by the parameter NJKR for each record using the procedure described in the Chapter 8.
2. Aggregates or summarizes the file computing the sum of the weights for each category, the sum of the weights overall, and the sum of a weighted set of plausible values.

3. Computes the percent of people within each group, their mean on the plausible values, and their corresponding standard errors. In SPSS, the resulting working file contains the corresponding statistics. In SAS, the data set FINAL in the work library contains the corresponding statistics.

When using each of these macros, the user needs to specify a set of classification variables, the name of the plausible values and how many there are, the number of replicate weights to be generated, the variable that contains the sampling information such as JKZONE and JKINDIC, and the sampling weight that is to be used for the analysis. When using the SAS macro the user will also need to specify the data file that contains the data that is to be processed. In SPSS the macro uses the current working file, which is the file that was last read.

### **9.3.1 SAS Macro for Computing Mean Plausible Values and Percents with Corresponding Standard Errors (JACKPV.SAS)**

The CD containing the TIMSS International Database contains a SAS macro program called JACKPV.SAS. This macro can be used to compute weighted percents and means within categories defined by variables specified by the user, and the JRR standard error estimate for these statistics. The control code for the macro JACKPV.SAS is presented in Figure 9.3.

**Figure 9.3**  
**SAS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKPV.SAS)**

```

%macro jackpv (wgt,jkz,jki,njkr,cvar,rootpv,npv,infile);
%let i = 1;
%do %while(%length(%scan(&cvar,&i)));
  %let i = %eval(&i + 1);
%end;
%let niv = %eval(&i-1);
%let cvar2 = %substr(&cvar,1,%eval(%length(&cvar)-%length(%scan(&cvar,&niv))));

%let allpv = ;

%do i = 1 %to &npv;
  %let pv&i = &rootpv.&i;
  %let allpv = &allpv &&pv&i ;
%end;

data a;
  set &infile (keep= &wgt &jkz &jki &cvar &allpv) ;
  array rwt rwt1 - rwt&njkr ; * Replicate Weight ;
  array wtx wtx1 - wtx&njkr ; * Weighted x ;
  k = 1;
  if &pv1 > .z;
do i=1 to &njkr;
  if &jkz <>i then rwt(i) = &wgt * 1;
  if (&jkz = i & &jki = 1) then rwt(i) = &wgt * 2;
  if (&jkz = i & &jki = 0) then rwt(i) = &wgt * 0;
  wtx(i) = rwt(i) * &pv1 ;
end ;
  wtx0 = &wgt * &pv1;
  N = 1;

%do i = 1 %to &npv;
  wtpv&i = &wgt * &&pv&i ;
%end;

proc summary data=a missing;
  class k &cvar;
  var N &wgt rwt1-rwt&njkr wtx0 wtx1-wtx&njkr
  %do i = 1 %to &npv;
    wtpv&i
  %end;
  ;
  output out=b sum=;
run;

proc summary data=b;
  var k _type_;
  output out=maxtype max=k maxtype;

data b;
  merge b maxtype (drop= _type_ _freq_);
  by k;
  maxtype2 = maxtype-1;

data type2 (keep = k &cvar2 t2wt t2rwt1-t2rwt&njkr);
  set b ;
  where _type_ = maxtype2 ;
  array rwt rwt1 - rwt&njkr ;
  array t2rwt t2rwt1 - t2rwt&njkr ;
  t2wt = &wgt ;
  do i = 1 to &njkr;
    t2rwt(i) = rwt(i);
  end;

```

**Figure 9.3 (Continued)****SAS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKPV.SAS)**

```

proc sort data = b out = d;
  by k &cvar2;
  where _type_=maxtype;

data e (drop = t2wt t2rwt1-t2rwt&njkr);
  merge d type2 ;
  by k &cvar2 ;
  twt = t2wt ;
  array rwt trwt1-trwt&njkr ;
  array t2rwt t2rwt1-t2rwt&njkr ;
  do i = 1 to &njkr;
    rwt(i) = t2rwt(i);
  end;

data final (keep = N &cvar &wgt mnx mnx_se
              mnpv mnpv_se mnx_jvr mnpv_ivr
              pct pct_se );
  set e ;
  mnx = wtx0 / &wgt ;
  pct = 100 * (&wgt / twt) ;
  array wtx wtx1 - wtx&njkr ;
  array rwt rwt1 - rwt&njkr ;
  array rmnx rmnx1 - rmnx&njkr ;
  array trwt trwt1 - trwt&njkr ;
  array rp rpl - rp&njkr ;
  mnx_jvr = 0 ;
  pct_var = 0 ;
  do i = 1 to &njkr ;
    if rwt(i) > 0 then
      rmnx(i) =wtx(i)/rwt(i) ; * Compute the mean for each replicate;
    else rmnx(i) = . ;
    rp(i) =100*(rwt(i)/trwt(i)) ; * Compute the proportion of each
replicate;
    if rmnx(i) <> . then
      mnx_jvr = mnx_jvr + (rmnx(i) - mnx )**2 ; * Compute JK_Var for Mean X;
      pct_var = pct_var + (rp(i) - pct)**2 ; * Compute JK_VAR for P ;
  end;
  %do i = 1 %to &npv;
    mnpv&i = wtpv&i / &wgt ;
  %end;
  mnpv_ivr = 0;
  %if %eval(&npv) > 1 %then %do;
  mnpv_ivr = (1+(1 / %eval(&npv))) * var(of %do i = 1 %to &npv; mnpv&i %end;);
  %end;
  mnx_se = sqrt(mnx_jvr);
  mnpv_se = sqrt(mnx_jvr + mnpv_ivr);
  mnpv = mean(of %do i = 1 %to &npv; mnpv&i %end;);
  pct_se = sqrt(pct_var);
run;

%mend jackpv;

```

The user needs to know some basic SAS macro language in order to use the macro JACKPV.SAS. The macro needs to be first included in the program file where it will be used. If the user is operating in batch mode, the macro needs to be called once in every batch. If the user is using SAS interactively, the macro needs to be called once at the beginning of the session and it will remain active throughout the session. If the session is terminated and restarted at a later time, the macro needs to be called once again.

The macro is included in the program file or session where it will be used by issuing the following command as part of the SAS syntax:

```
%include "<path>jackpv.sas";
```

where <path> points to the specific drive and directory where the macro JACKPV.SAS can be found. The macro requires that several parameter arguments be submitted when it is invoked. These parameters are:

### ***WGT***

The sampling weight to be used in the analysis. Depending on the Population of interest the user will generally choose between TOTWGTL4, TOTWGTM4 or TOTWGTP4.

### ***JKZ***

The variable that captures the assignment of the student to a particular sampling zone. The name of this variable in all TIMSS files is JKZONE.

### ***JKI***

The variable that captures whether the case is to be dropped or have its weight doubled for the corresponding replicate weight. The name of this variable in all TIMSS files is JKINDIC.

### ***NJKR***

This indicates the number of replicate weights to be generated when computing the JRR error estimates. When conducting analyses using the data from all countries the value of NJKR should be set to 75 for the student and school background data. The user working with the data for only one country should set the NJKR argument to as many replicates as are needed in the country. The minimum number of replicates needed by country is shown in Table 9.2. If the data from two or more countries is being used for an analysis, then the larger number of jackknife zones should be used. When in doubt about what number to set the NJKR parameter, it should be set to 75. The error variance will always be estimated correctly if more replicate weights than necessary are computed, but will be underestimated if the user specifies less replicate weights than necessary.



***CVAR***

This lists the variables that are to be used to classify the students in the data file. This can be a single variable, or a list of variables. The maximum number of variables will depend mostly on the computer resources available to the user at the time. It is recommended always to include the variable that identifies the country. At least one variable has to be specified, usually `IDCOUNTRY`.

***ROOTPV***

This is the prefix used to identify the plausible values for the achievement scale of interest. A list of the prefixes used to identify the plausible values is shown in Table 9.3.

***NPV***

This is the number of plausible values that will be used for the analysis. Generally the user will want to use all five plausible values for the analysis although under special circumstances less can be used. The user should always use at least two plausible values for any analysis involving plausible values.

***FNAME***

The name of the data set that contains the variables necessary for the analysis. The name of this file has to be in either the single or multilevel SAS format. The multilevel SAS file format contains the library as well as the data set. It is important to emphasize that this data set must include only those cases that are of interest in the analysis. If the user wants to have specific cases excluded from the analysis, as for example, students with missing variables or select students from a specific group, the selection should be done before the file is used in the analysis, prior to invoking the macro. The macro does not provide an option to select students.

**Table 9.2**

**Number of Replicate Weights Needed for Computing the JRR Error Variance Estimate for Each Sample**

Country	Final Year of Secondary School: All Students	Final Year of Secondary School: Advanced Mathematics Students	Final Year of Secondary School: Physics Students
Australia	49	49	49
Austria	75	58	58
Canada	75	75	75
Cyprus	28	21	21
Czech Republic	75	45	45
Denmark	65	65	65
France	70	31	31
Germany	75	75	75
Greece	-	30	30
Hungary	75	-	-
Iceland	30	-	-
Israel	30	30	30
Italy	52	51	38
Latvia	-	-	19
Lithuania	73	29	-
Netherlands	40	-	-
New Zealand	39	-	-
Norway	66	-	33
Russian Federation	41	41	41
Slovenia	39	39	-
South Africa	52	-	-
Sweden	75	75	75
Switzerland	75	75	75
United States	33	33	33

**Table 9.3**  
**Values for the Parameter ROOTPV**

ROOTPV	Achievement Scale
CSLPV0	Mathematics and Science Literacy
CSLPV1	Mathematics Literacy
CSLPV2	Science Literacy
CSMPV0	Advanced Mathematics
CSMPV1	Numbers, Equations and Functions
CSMPV2	Calculus
CSMPV3	Geometry
CSPPV0	Physics
CSPPV1	Mechanics
CSPPV2	Electricity and Magnetism
CSPPV3	Heat
CSPPV4	Wave Phenomena
CSPPV5	Particle, Quantum Physics, Astrophysics and & Relativity
CSRPV0	Reasoning and Social Utility

The simplest and most straightforward way to use this macro is to invoke it using the conventional SAS notation for invoking macros followed by the list of arguments for the analysis enclosed in parenthesis, separated by commas, and in the same order as they are listed above. For example, if the macro is invoked as:

```
%jackpv (TOTWGTL4 , JKZONE, JKINDIC, 75, IDCNTY ITSEX, CSLPV0, 5, CSGALL1);
```

it will compute the mean mathematics and science literacy achievement score using 5 plausible values, and its corresponding standard error, for boys and girls within each country, using the variable TOTWGTL4 as the sampling weight. It will also compute the percent of boys and girls within the country, and its corresponding standard error. The data will be read from the data set CSGALL1, and the standard error of the statistics will be computed based on 75 replicate weights.

The file that contains these results is called FINAL and can be found in the WORK library. The variables that are contained in this file are:

### ***Classification Variables***

Each of the classification variables are kept in the resulting file. In the above example there are two classification variables in the WORK.FINAL data set, IDCNTY and ITSEX. There is one unique occurrence for each combination of the categories for these variables.

***Weight Variable***

Contains the estimate in the population that belongs to the group defined by the specific combination of the classification variable categories. In the example this variable is called TOTWGTL4 because we wanted an estimate for the students in the final year of secondary school that took the mathematics and science literacy test.

***N***

Contains the number of cases in the group defined by the specific combination of categories for the classification variables. In the example it is the number of boys and girls in the sample for each country.

***PCT***

Contains the weighted percent of people in the group for the classification variable listed last, within the specific combination of the categories defined by the classification variables. In the example, it is the percent of boys and girls, within each country.

***PCT\_SE***

Contains the standard error of PCT computed using the JRR method for computing the standard error.

***MNX***

Contains the weighted mean for the first plausible value for the group defined by the corresponding combination of classification variable categories.

***MNX\_SE***

Contains the JRR standard error of the mean for first plausible value computed using the JRR method for computing the standard error. This does not include the imputation error that should be computed when using plausible values.

***MNPV***

Contains the mean of the plausible values used in the analysis.

***MNPV\_SE***

Contains the standard error for the mean of the plausible values. This standard error contains the sampling and the imputation component of the error of the estimate.

The file resulting from using this macro can be printed using a SAS procedure of choice. An example call to this macro, and a subset of the resulting file is presented in Figure 9.4. In this example the macro computes the percent of boys and girls by country, and their mean achievement in the mathematics and science literacy scale. The listing presented in Figure 9.4 is interpreted in the following way. The first line shows that there were 1,166 students in the sample for IDCNTY=36 (Australia) and who had ITSEX=1 (Girls). It is estimated that there are 99,777 girls in the final year of secondary school in Australia, their mean mathematics and science literacy score is 511.212, with a standard error of 9.2578. The mean and standard error using only the first

plausible value are 511.030 and 9.2551 respectively. We can also tell from this line of data that it is estimated that 58 (58.4013) percent of the students in the final year of secondary school in Australia are girls, and the standard error of this estimate is 2.9. The second line shows the same information, but for the boys (ITSEX=2) in Australia. It is estimated that there are 71,070 boys in their final year of secondary school in Australia, their mean achievement on mathematics and science literacy is 543.396 with a standard error or 10.7393. It is estimated that 42 (41.5987) percent of students in the final year of secondary school are boys, and the standard error of this estimate is 2.9. These results are presented in Table 1.3 of the TIMSS international report “Mathematics and Science Achievement in the Final Year of Secondary School” and are shown in Figure 9.5 of this User Guide. Notice that for all cases the values under the column MNPV\_SE are always greater than the values under the column MNX\_SE. This is because the MNPV\_SE also include the error variance due to the uncertainty in the imputation.

**Figure 9.4**  
**SAS Control Code and Output File for Using the Macro JACKPV.SAS**

```

libname pop3 "c:\timss\data\pop3";

options nocenter mprint;
* Read the variables from the student file;
data student;
  set pop3.ugfile
    (keep=idcntry idpop itsex jkindic jkzone intmsl4 totwgtl4
     csbgday1 csplpv01--csplpv05);
  if intmsl4 = 1 ; * Select only Literacy Sample;
  if itsex in(1,2); * Select only if not missing ITSEX;

* Now use the macro JACKPV to get the results;
%include jackpv.sas";

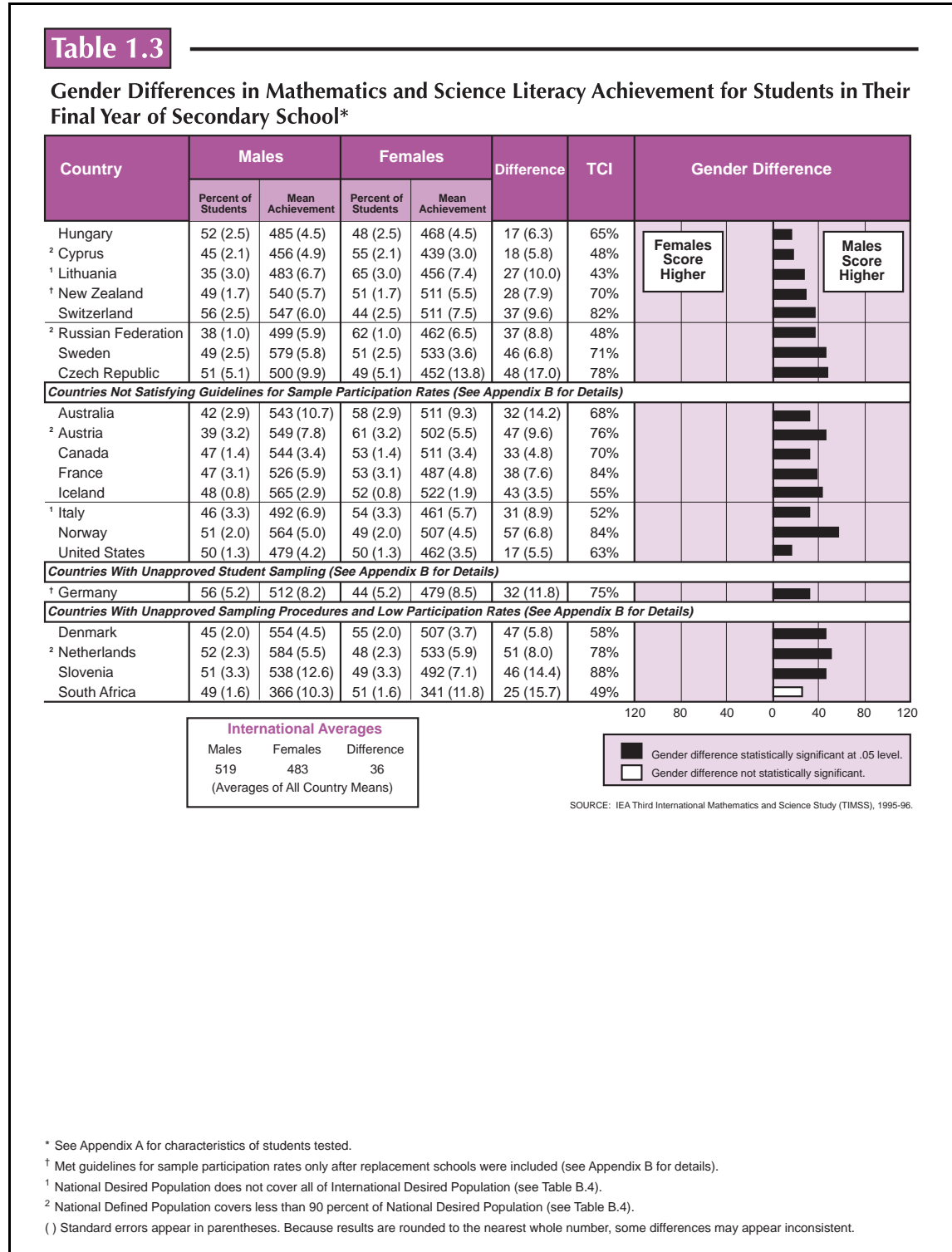
%jackpv(totwgtl4,jkzone,jkindic,75,idcntry itsex,csplpv0,5,student);

proc print data=final noobs;
  var idcntry itsex N totwgtl4 pct pct_se mnx mnx_se mnpv mnpv_se;

```

IDCNTRY	ITSEX	N	TOTWGTL4	PCT	PCT_SE	MNX	MNX_SE	MNPV	MNPV_SE
36	1	1166	99776.74	58.4013	2.86776	511.030	9.2551	511.212	9.2578
36	2	775	71069.97	41.5987	2.86776	543.910	10.6619	543.396	10.7339
40	1	1041	42427.56	61.4688	3.21786	503.435	5.3782	502.115	5.4735
40	2	878	26595.37	38.5312	3.21786	548.006	7.7213	549.063	7.8465
124	1	2533	137971.73	52.9164	1.42824	510.934	3.2329	510.751	3.3963
124	2	2672	122763.41	47.0836	1.42824	543.798	3.3055	543.657	3.3965
196	1	280	2491.02	54.9106	2.05533	438.334	2.9283	438.908	3.0247
196	2	251	2045.49	45.0894	2.05533	456.437	4.5596	456.455	4.9128
200	1	1052	67023.77	48.7593	5.13805	450.725	13.7310	451.565	13.7695
200	2	1115	70434.77	51.2407	5.13805	500.809	9.9132	499.796	9.9496
208	1	1446	20328.95	54.7867	2.00175	506.467	3.6312	506.759	3.7122
208	2	1206	16776.65	45.2133	2.00175	554.045	4.3407	553.502	4.4600
250	1	759	331891.27	52.6527	3.06406	486.757	4.6813	487.231	4.8014
250	2	813	298448.96	47.3473	3.06406	524.470	5.7714	525.644	5.9032
280	1	1108	416372.35	44.2963	5.23380	477.974	8.3847	479.234	8.5159
280	2	1071	523598.85	55.7037	5.23380	512.423	7.8948	511.518	8.1761
348	1	2542	51285.02	47.7515	2.45457	467.625	4.4431	467.948	4.4885
348	2	2370	56114.74	52.2485	2.45457	485.636	4.4044	484.902	4.4746
352	1	887	1199.15	52.4969	0.79161	521.348	1.8238	521.871	1.9403
352	2	800	1085.08	47.5031	0.79161	563.936	2.7168	564.595	2.8633
380	1	840	206434.31	54.2059	3.30550	460.903	5.6604	460.994	5.6923
380	2	776	174399.69	45.7941	3.30550	492.302	6.8036	492.446	6.8817
440	1	1938	14488.91	65.4235	3.00418	455.569	7.4104	455.641	7.4244
440	2	948	7657.44	34.5765	3.00418	482.437	6.6843	482.600	6.7039
528	1	725	70693.81	48.4482	2.31716	531.602	5.7129	532.522	5.8670
528	2	745	75222.48	51.5518	2.31716	583.362	5.4261	583.549	5.4786
554	1	911	19114.58	50.9063	1.67912	510.234	5.3856	511.111	5.4902
554	2	852	18434.00	49.0937	1.67912	539.461	5.4960	539.604	5.7331
578	1	1328	21663.70	49.4541	1.98219	507.440	4.5232	507.016	4.5374
578	2	1190	22141.97	50.5459	1.98219	565.038	4.9612	564.308	5.0153
643	1	1448	638710.12	61.9393	1.02371	460.780	6.4533	461.591	6.4942
643	2	841	392476.88	38.0607	1.02371	498.432	5.6956	499.012	5.8924
717	1	1370	187961.73	51.3089	1.56874	341.526	11.7877	340.552	11.8259
717	2	1315	178371.67	48.6911	1.56874	366.183	10.2998	366.019	10.3202
752	1	1606	36456.65	51.2554	2.53266	532.314	3.5641	532.654	3.5764
752	2	1461	34670.77	48.7446	2.53266	579.381	5.6824	578.794	5.7991
756	1	1623	28239.84	43.8510	2.45855	510.569	7.3600	510.786	7.4977
756	2	1660	36159.75	56.1490	2.45855	548.454	5.9099	547.320	6.0355
840	1	2968	1145052.11	50.2600	1.29751	462.372	3.5199	462.224	3.5279
840	2	2839	1133206.08	49.7400	1.29751	479.536	4.1081	479.198	4.1979
890	1	735	12601.33	48.5842	3.32623	491.661	7.0602	491.880	7.0763
890	2	828	13335.76	51.4158	3.32623	539.686	12.5400	538.145	12.5865

**Figure 9.5**  
**Sample Table for Analysis Using Plausible Values Taken From the TIMSS**  
**International Report “Mathematics and Science Achievement in the Final Year of**  
**Secondary School”**



### 9.3.2 SPSS Macro for Computing Mean and Percents with Corresponding Standard Errors (JACKPV.SPS)

The CD containing the TIMSS International Database also contains an SPSS macro program called JACKPV.SPS. This macro can be used to compute weighted percents and means within categories defined by variables specified by the user, and the JRR standard error estimate for these statistics. The control code for the macro JACKPV.SPS is presented in Figure 9.6 of this User Guide.

**Figure 9.6**

#### SPSS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKPV.SPS)

```

set mprint=on.
* This macro computes the mean DVAR within the groups specified by the variable
  list in CVAR.
* The parameters needed to run this macro are:
  CVAR   :The list of classification variables;
  PVS    :The plausible values in the form "PV1 TO PVx";
  NPV    :The number of Plausible values (generally 5);
  NJKR   :The Number of Jackknife replicates;
  JKZ    :The variable with the assignment to JK zones;
  JKI    :The variable with the JK Indicator (0 or 1);
  WGT    :The sampling weight to use.

define jackpv (cvar   = !charend('/') /
              pvs    = !charend('/') /
              NPV    = !charend('/') !default(5) /
              njkr   = !charend('/') !default(75) /
              jkz    = !charend('/') !default(jkzone) /
              jki    = !charend('/') !default(jkindic) /
              wgt    = !charend('/') !default(fullwgt) ).

weight off.

compute k = 1.

sort cases by !cvar.

vector pv = !pvs.
vector wgtpv(!NPV).
loop #i = 1 to !NPV.
compute wgtpv(#i) = !wgt * pv(#i).
end loop.

compute wgtx = !wgt * !head(!pvs).

vector rwgt(!njkr).
vector rwgtx(!njkr).
loop #i = 1 to !njkr.
if (!jkz = #i and !jki = 0) rwgt(#i) = !wgt * 0.
if (!jkz = #i and !jki = 1) rwgt(#i) = !wgt * 2.
if (!jkz <> #i                ) rwgt(#i) = !wgt * 1.

compute rwgtx(#i) = rwgt(#i) * !head(!pvs).

end loop.

```



**Figure 9.6 (Continued 1)****SPSS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKPV.SPS)**

```

!let !nway = tmpjck1.
aggregate outfile = !nway
  / break = !cvar k
  /      !wgt rwgt1 to !concat(rwgt,!njkr) wgtx rwgtx1 to !concat(rwgtx,!njkr)
        wgtpv1 to !concat(wgtpv,!NPV)
  = sum(!wgt rwgt1 to !concat(rwgt,!njkr) wgtx rwgtx1 to !concat(rwgtx,!njkr)
        wgtpv1 to !concat(wgtpv,!NPV))
  / n = n(!wgt).

!let !cvar1 = !null
!do !i !in(!cvar)
!let !cvar1 = !concat(!i,!blank(1),!cvar1)
!doend

!let !cvar2a = !concat(!cvar1,!blank(1),k).
!let !cvar2b = !concat(!tail(!cvar1),!blank(1),k).

!do !i !in(!cvar2a)

!if (!cvar2b = !null) !then !break.
!ifend.

get file = !nway.
compute k = 1.

!let !seq = !index(!cvar1,!i).

aggregate outfile = tmpjckA
  / break = !cvar2a
  /      !wgt rwgt1 to !concat(rwgt,!njkr) wgtx rwgtx1 to !concat(rwgtx,!njkr)
        wgtpv1 to !concat(wgtpv,!NPV)
  = sum(!wgt rwgt1 to !concat(rwgt,!njkr) wgtx rwgtx1 to !concat(rwgtx,!njkr)
        wgtpv1 to !concat(wgtpv,!NPV))
  / n = sum(n).

aggregate outfile = tmpjckB
  / break = !cvar2b
  /      twgt twgt1 to !concat(twgt,!njkr)
  = sum(!wgt rwgt1 to !concat(rwgt,!njkr)).

get file = tmpjckA.
sort cases by !cvar2B.
save outfile = tmpjckA.

match files
  / file = tmpjckA
  / table = tmpjckB
  / by !cvar2b.

compute type = !seq.

```

**Figure 9.6 (Continued 2)****SPSS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKPV.SPS)**

```

!if (!head(!cvar1)=!head(!cvar2a)) !then
save outfile = jackfnl.
!else
add files
  / file = jackfnl
  / file = *.
save outfile = jackfnl.
!ifend.
!let !cvar2a = !tail(!cvar2a).
!let !cvar2b = !tail(!cvar2b).

!doend

do if (!wgt > 0).
compute mnx = wgtx / !wgt.
compute pct = 100 * (!wgt / twgt).
end if.

vector wgtpv = wgtpv1 to !concat(wgtpv,!NPV).
vector mnpv(!NPV).
loop #i = 1 to !NPV.
compute mnpv(#i) = wgtpv(#i) / !wgt.
end loop.

compute mnpv = mean(mnpv1 to !concat(mnpv,!NPV)).

compute mnpv_ivr = ((1+!NPV)/!NPV)*(variance(mnpv1 to !concat(mnpv,!NPV))).

compute mnx_jvr = 0.
compute pct_var = 0.

vector rwgt = rwgt1 to !concat(rwgt,!njkr).
vector twgt = twgt1 to !concat(twgt,!njkr).
vector wgtx = wgtx1 to !concat(wgtx,!njkr).
loop #i = 1 to !njkr.
do if (rwgt(#i)) > 0.
compute mnx_jvr = mnx_jvr + (mnx - (wgtx(#i) / rwgt(#i)))*2.
compute pct_var = pct_var + (pct - (100 * (rwgt(#i) / twgt(#i))))*2.
end if.
end loop.

compute mnx_se = sqrt(mnx_jvr).
compute mnpv_se = sqrt(mnx_jvr + mnpv_ivr).
compute pct_se = sqrt(pct_var).
autorecode type
  / into level.

execute.

select if level=1.

save outfile = final.

!enddefine.

```

The user needs to know some basic SPSS macro language in order to use the macro JACKPV.SPS. The macro should be first included in the program file where it is going to be used. If the user is operating in batch mode, then the macro should be called in every batch. If the user is using SPSS interactively then the macro should be called once at the beginning of the session and it will remain active throughout the session. If the session is terminated and restarted at

a later time the macro should be called once again. Once the macro is included in a specific session the word “JACKPV” should not be used within that program because doing so will invoke the macro.

The macro is included in the program file where it will be used by issuing the following command under SPSS:

```
include "<path>jackpv.sps".
```

where <path> points to the specific drive and directory where the macro JACKPV.SPS can be found. The macro requires that several parameters be submitted when it is invoked. These parameters are:

### ***WGT***

The sampling weight to be used in the analysis. Depending on the population of interest the user will generally choose between TOTWGTL4, TOTWGTM4 or TOTWGTP4.

### ***JKZ***

The variable that captures the assignment of the student to a particular sampling zone. The name of this variable in all TIMSS files is JKZONE.

### ***JKI***

The variable that captures whether the case is to be dropped or have its weight doubled for the corresponding replicate weight. The name of this variable in all TIMSS files is JKINDIC.

### ***NJKR***

This indicates the number of replicate weights to be generated when computing the JRR error estimates. When conducting analyses using the data from all countries, the value of NJKR should be set to 75 for the student and school background data. The user working with the data for only one country should set the NJKR argument to as many replicates as are needed in the country. The minimum number of replicates needed by country is shown in Table 9.2. If the data from two or more countries is being used for an analysis, then the larger number of jackknife zones should be used. When in doubt about what number to set the NJKR parameter, it should be set to 75. The error variance will always be estimated correctly if more replicate weights than necessary are computed, but will be underestimated if the user specifies less replicate weights than necessary.

### ***CVAR***

This lists the variables that are to be used to classify the students in the data file. This can be a single variable, or a list of variables. The maximum number of variables will depend mostly on the computer resources available to the user at the time. It is recommended to always include the variable that identifies the country. At least one variable has to be specified, usually IDCNTY.

**PVS**

These are the plausible values to be used in the analysis. The plausible values need to be specified in the form “Plausible Value 1 to Plausible Value 5” as in “CSLPV21 to CSLPV25” when analyzing the science literacy scale. Although in most cases the user will want to use all five available plausible values, the programs will also work when less are specified. The user should always use at least two plausible values.

**NPV**

This is the number of plausible values that will be used for the analysis. Generally the user will want to use all five plausible values for the analysis although under some circumstances less can be used. The user should always use at least two plausible values for any analysis.

Unlike the SAS macro (JACKPV.SAS), the JACKPV macro in SPSS does not require that the data file that contains the data of interest be specified when calling the macro. By default, SPSS uses the current working file. This needs to be read with the GET FILE command prior to invoking the macro.

The simplest and most straightforward way is to invoke the macro using the conventional SPSS notation for invoking macros. This involves listing the macro name followed by the corresponding list of arguments for the analysis, each separated by a slash. For example, if the macro is invoked as:

```
get file   = CSGALL1.
jackpv cvar = IDCNTRY ITSEX /
  pvs     = CSLPV01 TO CSLPV05 /
  npv    = 5           /
  jkz    = JKZONE     /
  jki    = JKINDIC   /
  njkr   = 75         /
  WGT    = TOTWGTL4 / .
```

it will compute the mean mathematics and science literacy achievement score using five plausible values, and its corresponding standard error, for boys and girls within each country, using the variable TOTWGTL4 as the sampling weight. It will also compute the percent of boys and girls within the country, and its corresponding standard error. The data will be read from the data set CSGALL1, and the standard error of the statistics will be computed based on 75 replicate weights.

The file that contains these results is then called FINAL and is saved to the default directory being used by SPSS. The variables that are contained in this file are:

**Classification Variables**

Each of the classification variables are kept in the resulting file. In the above example there are two classification variable in the WORK.FINAL data set, IDCNTRY and ITSEX. There is one unique occurrence for each combination of the categories for these variables.

***Weight Variable***

Contains the estimate in the population that belongs to the group defined by the specific combination of the classification variable categories. In the example this variable is called TOTWGTL4 because we wanted an estimate for the students in the final year of secondary school that took the mathematics and science literacy test.

***N***

Contains the number of cases in the group defined by the specific combination of categories for the classification variables. In the example this is the number of boys and girls in the sample for each country.

***PCT***

Contains the weighted percent of people in the group for the classification variable listed last, within the specific combination of the categories defined by the classification variables. In the example, it is the percent of boys and girls, within each country.

***PCT\_SE***

Contains the standard error of PCT computed using the JRR method for computing the standard error.

***MNX***

Contains the weighted mean for the first plausible value for the group defined by the corresponding combination of classification variable categories.

***MNX\_SE***

Contains the JRR standard error of the mean for first plausible value computed using the JRR method for computing the standard error. This does not include the imputation error that should be computed when using plausible values.

***MNPV***

Contains the mean of the plausible values used in the analysis.

***MNPV\_SE***

Contains the standard error for the mean of the plausible values. This standard error contains the sampling and the imputation component of the error of the estimate.

The file resulting from using this macro can be printed using a SPSS procedure of choice. An example call to this macro, and a printout of the resulting file is presented in Figure 9.7. In this example the macro computes the percent of boys and girls by country, and their mean achievement in the mathematics and science literacy scale. The listing presented in Figure 9.7 is interpreted in the following way. The first line shows that there were 1,166 students in the sample for IDCNTY=36 (Australia) and who had ITSEX=1 (Girls). It is estimated that there are 99,777 girls in the final year of secondary school in Australia, and their mean mathematics and science

literacy score is 511.212, with a standard error of 9.2578. The mean and standard error using only the first plausible value are 511.030 and 9.2551, respectively. We can also tell from this line of data that it is estimated that 58 (58.4013) percent of the students in the final year of secondary school in Australia are girls, and the standard error of this estimate is 2.9. The second line shows the same information, but for the boys (ITSEX=2) in Australia. It is estimated that there are 71070 boys in their final year of secondary school in Australia, their mean achievement on mathematics and science literacy is 543.396 with a standard error or 10.7393. It is estimated that 42 (41.5987) percent of students in the final year of secondary school are boys, and the standard error of this estimate is 2.9. These results are the ones presented in Table 1.3 of the report titled “Mathematics and Science Achievement in the Final Year of Secondary School” that and that are shown in Figure 9.5 of this User Guide. Notice that for all cases the values under the column MNPV\_SE are always greater than the values under the column MNX\_SE. This is because the MNPV\_SE also include the error variance due to the uncertainty in the imputation.

**Figure 9.7****SPSS Control Code and Extract of Output File for Using the Macro JACKPV.SPS**

```

get file = "csgall1.sys"
  / keep=idcntry idstud itsex jkindic jkzone totwgtl4
  cslpv01 to cslpv05 intmsl4.

select if intmsl4 = 1.
select if not(missing(itsex)).

* Now use the macro JACKPV to get the results.
include "jackpv.sps".

jackpv  cvar = idcntry itsex /
        pvs  = cslpv01 to cslpv05 /
        npv  = 5 /
        njkr = 75 /
        jkz  = jkzone /
        jki  = jkindic /
        wgt  = totwgtl4 / .

formats idcntry (f3.0) itsex (f1.0) n (f5.0) totwgtl4 (f10.2).
report format=list automatic
  / var = idcntry itsex n totwgtl4 pct pct_se mnx mnx_se mnpv mnpv_se

```

IDCNTRY	ITSEX	N	TOTWGTL4	PCT	PCT_SE	MNX	MNX_SE	MNPV	MNPV_SE
36	1	1166	99776.74	58.40	2.87	511.03	9.26	511.21	9.26
36	2	775	71069.97	41.60	2.87	543.91	10.66	543.40	10.73
40	1	1041	42427.56	61.47	3.22	503.44	5.38	502.11	5.47
40	2	878	26595.37	38.53	3.22	548.01	7.72	549.06	7.85
124	1	2533	137971.73	52.92	1.43	510.93	3.23	510.75	3.40
124	2	2672	122763.41	47.08	1.43	543.80	3.31	543.66	3.40
196	1	280	2491.02	54.91	2.06	438.33	2.93	438.91	3.02
196	2	251	2045.49	45.09	2.06	456.44	4.56	456.46	4.91
200	1	1052	67023.77	48.76	5.14	450.72	13.73	451.57	13.77
200	2	1115	70434.77	51.24	5.14	500.81	9.91	499.80	9.95
208	1	1446	20328.95	54.79	2.00	506.47	3.63	506.76	3.71
208	2	1206	16776.65	45.21	2.00	554.04	4.34	553.50	4.46
250	1	759	331891.27	52.65	3.06	486.76	4.68	487.23	4.80
250	2	813	298448.96	47.35	3.06	524.47	5.77	525.64	5.90
280	1	1108	416372.35	44.30	5.23	477.97	8.38	479.23	8.52
280	2	1071	523598.85	55.70	5.23	512.42	7.89	511.52	8.18
348	1	2542	51285.02	47.75	2.45	467.62	4.44	467.95	4.49
348	2	2370	56114.74	52.25	2.45	485.64	4.40	484.90	4.47
352	1	887	1199.15	52.50	.79	521.35	1.82	521.87	1.94
352	2	800	1085.08	47.50	.79	563.94	2.72	564.60	2.86
380	1	840	206434.31	54.21	3.31	460.90	5.66	460.99	5.69
380	2	776	174399.69	45.79	3.31	492.30	6.80	492.45	6.88
440	1	1938	14488.91	65.42	3.00	455.57	7.41	455.64	7.42
440	2	948	7657.44	34.58	3.00	482.44	6.68	482.60	6.70
528	1	725	70693.81	48.45	2.32	531.60	5.71	532.52	5.87
528	2	745	75222.48	51.55	2.32	583.36	5.43	583.55	5.48
554	1	911	19114.58	50.91	1.68	510.23	5.39	511.11	5.49
554	2	852	18434.00	49.09	1.68	539.46	5.50	539.60	5.73
578	1	1328	21663.70	49.45	1.98	507.44	4.52	507.02	4.54
578	2	1190	22141.97	50.55	1.98	565.04	4.96	564.31	5.02
643	1	1448	638710.12	61.94	1.02	460.78	6.45	461.59	6.49
643	2	841	392476.88	38.06	1.02	498.43	5.70	499.01	5.89
717	1	1370	187961.73	51.31	1.57	341.53	11.79	340.55	11.83
717	2	1315	178371.67	48.69	1.57	366.18	10.30	366.02	10.32
752	1	1606	36456.65	51.26	2.53	532.31	3.56	532.65	3.58
752	2	1461	34670.77	48.74	2.53	579.38	5.68	578.79	5.80
756	1	1623	28239.84	43.85	2.46	510.57	7.36	510.79	7.50
756	2	1660	36159.75	56.15	2.46	548.45	5.91	547.32	6.04
840	1	2968	1145052.11	50.26	1.30	462.37	3.52	462.22	3.53
840	2	2839	1133206.08	49.74	1.30	479.54	4.11	479.20	4.20
890	1	735	12601.33	48.58	3.33	491.66	7.06	491.88	7.08
890	2	828	13335.76	51.42	3.33	539.69	12.54	538.15	12.59

## 9.4 Computing JRR Replicate Weights and Sampling Variance for Variables other than the Plausible Value Scores Using SPSS and SAS

In this section example SAS and SPSS code that may be used to compute the JRR standard errors for means and percentages is described. This code is provided in the form of SAS and SPSS macros. These macros compute the percentage of students within a subgroup defined by a set of classification variables, the JRR standard error of this percentage, the mean for the group on a variable of choice (not the plausible value scale scores), and the JRR standard error of that mean.

Each of these macros operate as follows:

1. Computes a set of 75 replicate weights for the record using the procedure described in the previous chapter.
2. Aggregates or summarizes the file computing the sum of the weights for each category, the sum of the weights overall, and the sum of a weighted analysis variable.
3. Computes the percent of people within each group, their mean on the analysis variable, and their corresponding standard errors. In SPSS the resulting working file contains the corresponding statistics. In SAS the file FINAL contains the corresponding statistics.

When using each of these macros, the user needs to specify a set of classification variables, one analysis variable, the number of replicate weights to be generated, the variable that contains the sampling information such as JKZONE and JKINDIC, and the sampling weight that is to be used for the analysis. When using the SAS macro the user will also need to specify the data file that contains the data that is to be processed. In SPSS the macro uses the current working file, which is the file that was last read.

### 9.4.1 SAS Macro for Computing Mean and Percents with Corresponding Standard Errors (JACKGEN.SAS)

The CD containing the TIMSS International Database contains a SAS macro program called JACKGEN.SAS. This macro can be used to compute weighted percents and means within categories defined by variables specified by the user, and the JRR standard error estimate for these statistics. The control code for the macro JACKGEN.SAS is presented in Figure 9.5. This macro should not be used to compute standard errors for the plausible values. When looking to estimate the standard errors of plausible values the user should use the macro JACKPV described in section 9.3 of this User Guide.



**Figure 9.8**  
**SAS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKGEN.SAS)**

```

* Computes Jackknife Standard Errors for DV;

%macro jackgen (wgt,jkz,jki,njkr,cvar,dvar,fname);

%let i = 1;                                * Start counter at 1;
%do %while(%length(%scan(&cvar,&i)));
  %let i = %eval(&i + 1);
%end;
%let niv = %eval(&i-1);                    * Diminish by 1 at the end;
%let cvar2 = %substr(&cvar,1,%eval(%length(&cvar) - %length(%scan(&cvar,&niv))));

data a;
  set &fname (keep= &wgt &jkz &jki &cvar &dvar);
  array rwt rwt1 - rwt&njkr ;              * Replicate Weight      ;
  array wtx wtx1 - wtx&njkr ;             * Weighted X              ;
  k = 1 ;                                  * computes a constant     ;
  if &dvar >.Z;                            * Keep only non-missing;
  do i=1 to &njkr;
    if &jkz <>i then rwt(i) = &wgt * 1;
    if (&jkz = i & &jki = 1) then rwt(i) = &wgt * 2;
    if (&jkz = i & &jki = 0) then rwt(i) = &wgt * 0;
    wtx(i) = rwt(i) * &dvar ;
  end;

  N = 1;
  wtx0 = &wgt * &dvar;

proc summary data=a missing;
  class k &cvar;
  var N &wgt rwt1-rwt&njkr wtx0 wtx1-wtx&njkr ;
  output out=b sum=;
run;

proc summary data=b;
  var k _type_;
  output out=maxtype max=k maxtype;

data b;
  merge b maxtype (drop= _type_ _freq_);
  by k;
  maxtype2 = maxtype-1;

data type2 (keep = k &cvar2 t2wt t2rwt1-t2rwt&njkr);
  set b ;
  where _type_ = maxtype2 ;
  array rwt rwt1-rwt&njkr ;
  array t2rwt t2rwt1 - t2rwt&njkr;
  t2wt = &wgt;
  do i = 1 to &njkr;
    t2rwt(i) = rwt(i);
  end;

proc sort data = b out=d;
  by k &cvar2;
  where _type_ = maxtype;

```

**Figure 9.8 (Continued)****SAS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKGEN.SAS)**

```

data e (drop = t2wt t2rwt1-t2rwt&njkr);
  merge d type2 ;
  by k &cvar2 ;
  twt = t2wt;
  array rwt trwt1-trwt&njkr ;
  array t2rwt t2rwt1-t2rwt&njkr ;
  do i= 1 to &njkr;
    rwt(i) = t2rwt(i);
  end;

data final (keep = N &cvar mnx mnx_se pct pct_se &wgt);
  set e;
  mnx = wtx0 / &wgt ;
  pct = 100* (&wgt/ twt);
  array wtx wtx1 - wtx&njkr ;
  array rwt rwt1 - rwt&njkr ;
  array rmn rmn1 - rmn&njkr ;
  array trwt trwt1 - trwt&njkr ;
  array rp rp1 - rp&njkr ;
  mn_jvar = 0;
  p_jvar = 0;
  n_jvar = 0;
  do i = 1 to &njkr ;
    if rwt(i) > 0 then
      rmn(i) = wtx(i) / rwt(i) ;
    else rmn(i) = . ;
    rp(i) = 100 * (rwt(i) / trwt(i)) ;
    if rmn(i) <> . then
      mn_jvar = mn_jvar + (rmn(i) - mnx )**2;
      p_jvar = p_jvar + (rp(i) - pct )**2;
    if rwt(i) > 0 then
      n_jvar = n_jvar + (&wgt - rwt(i))**2;
  end;
  mnx_se = sqrt(mn_jvar ) ;
  pct_se = sqrt(p_jvar ) ;
  n_se = sqrt(n_jvar ) ;
  run;

%mend jackgen;

```

The user needs to know some basic SAS macro language in order to use JACKGEN.SAS. The macro needs to be first included in the program file where it will be used. If the user is operating in batch mode, the macro needs to be called in every batch. If the user is using SAS interactively, the macro needs to be called once at the beginning of the session and it will remain active throughout the session. If the session is terminated and restarted at a later time, the macro needs to be called once again.

The macro is included in the program file or session where it will be used by issuing the following command as part of the SAS syntax:

```
%include "<path>jackgen.sas";
```

where <path> points to the specific drive and directory where the macro JACKGEN.SAS can be found. The macro requires that several parameter arguments be submitted when it is invoked. These parameters are:

**WGT**

The sampling weight to be used in the analysis. Depending on the population of interest the user will generally choose between TOTWGTL4, TOTWGTM4 or TOTWGTP4.

***JKZ***

The variable that captures the assignment of the student to a particular sampling zone. The name of this variable in all TIMSS files is JKZONE.

***JKI***

The variable that captures whether the case is to be dropped or have its weight doubled for the corresponding replicate weight. The name of this variable in all TIMSS files is JKINDIC.

***NJKR***

This indicates the number of replicate weights to be generated when computing the JRR error estimates. When conducting analyses using the data from all countries the value of NJKR should be set to 75 for the student and school. The user working with the data for only one country should set the NJKR argument to as many replicates as are needed for one country. The maximum number of replicates by country is shown in Table 9.2. If the data from two or more countries is being used for an analysis, then the larger number of jackknife zones should be used. When in doubt on what number to set the NJKR parameter, it should be set to 75. The error variance will always be estimated correctly if more replicate weights than necessary are computed, but will be underestimated if the user specifies less replicate weights than necessary.

***CVAR***

This lists the variables that are to be used to classify the students in the data file. This can be a single variable, or a list of variables. The maximum number of variables will depend mostly on the computer resources available to the user at the time. It is recommended to always include the variable that identifies the country. At least one variable has to be specified, usually IDCNTY.

***DVAR***

This is the variable for which means are to be computed. Only one variable can be listed here. This variable should not be any of the plausible values. This variable should be considered as a continuous variable.

***FNAME***

The name of the data set that contains the variables necessary for the analysis. The name of this file has to be in either the single or multilevel SAS format. The multilevel SAS file format contains the library as well as the data set. It is important to emphasize that this data set must include only those cases that are of interest in the analysis. If the user wants to have specific cases excluded from the analysis, as for example, students with missing variables or select students from a specific grade, this should be done prior to invoking the macro.

The simplest and most straightforward way is to invoke the macro using the conventional SAS notation for invoking macros followed by the list of arguments for the analysis enclosed in parenthesis, separated by commas, and in the same order as they are listed above. For example, if the macro is invoked as:

```
%jackgen(TOTWGTL4 ,JKZONE,JKINDIC,75,IDCNTRY, CDSGSTUD, CSGALL1);
```

it will compute the mean number of hours the students within each country spend studying or doing homework, and its corresponding standard error, using the variable TOTWGTL4 as the sampling weight. In this case it will also compute the percent of students within the country compared to the overall student population (an irrelevant figure for this analysis), and its corresponding standard error. The data will be read from the data set CSGALL1, and the standard error of the statistics will be computed based on 75 replicate weights. The results from this analysis are those presented in the last column of Table 4.15 from the report “Mathematics and Science Achievement in the Final Year of Secondary School” and which is shown in Figure 9.10.

The file that contains these results is called FINAL and can be found in the WORK library. The variables that are contained in this file are:

### ***Classification Variables***

Each of the classification variables are kept in the resulting file. In the above example there was only one classification variable in the WORK.FINAL data set. This is IDCNTRY. There is one unique occurrence for each category for this variable. If there were more variables included in the CVAR parameter then there would be one unique occurrence for each combination of the categories of the classification variables.

### ***Weight Variable***

Contains the estimate in the population that belongs to the group defined by the specific combination of the classification variable categories. In the example this variable is called TOTWGTL4.

### ***N***

Contains the number of cases in the group defined by the specific combination of categories for the classification variables.

### ***MNX***

Contains the weighted mean of the variable DVAR for the group defined by the corresponding combination of classification variable categories.

### ***MNX\_SE***

Contains the standard error of the mean for variable specified in DVAR, computed using the JRR method for computing the standard error.

***PCT***

Contains the weighted percent of people in the group for the classification variable listed last, within the specific combination of the categories defined by the groups defined initially. In the example, it is the percent of boys and girls, within each combination of country and grade.

***PCT\_SE***

Contains the standard error of PCT computed using the JRR method for computing the standard error.

The file resulting from using this macro can be printed using a SAS procedure of choice. An example call to this macro, and a subset of the resulting file is presented in Figure 9.9. In this example the macro computes the percent of students within each country (an irrelevant figure for our analysis) and the mean number of hours spent studying or doing homework. The listing presented in Figure 9.9 is interpreted in the following way. The first line shows that there were 1,831 students in the sample for IDCNTRY=36 (Australia) for whom there was data available in the variables for this analysis. It is estimated that on average Australian students spend about 3.3288 hours each day studying or doing homework, with a standard error of 0.129. The columns titled PCT and PCT\_SE in this case are irrelevant as they indicate the percent of students in the overall population that are Australian, and the standard error of the percent. The second line shows the same information but for IDCNTRY=40 (Austria) where students report that on average they spent 1.993 hours per day studying or doing homework. Again, the columns titled PCT and PCT\_SE in this case are irrelevant as they indicate the percent of students in the overall population that are Austrian and the standard error for the percent. The results from this analysis are the ones presented in Table 4.15 of the report titled “Mathematics and Science Achievement in the Final Year of Secondary School” that and that are shown in Figure 9.10 of this User Guide.

**Figure 9.9**  
**SAS Control Code and Extract of Output File for Using the Macro JACKGEN.SAS**

```

libname pop3 "C:\TIMSS\POP3\DATA";

options nocenter;
* Read the variables from the student file;
data student;
  set pop3.csgall1
    (keep=idcntry idpop jkindic jkzone intmsl4 totwgtl4 csdgstud);
  if intmsl4 = 1 ; * Select only Literacy Sample;

* Now use the macro jackgen to get the results;
%include "jackgen.sas";

%jackgen(totwgtl4, jkzone, jkindic, 75, idcntry, csdgstud, student);

proc print data=final noobs;
  var idcntry N totwgtl4 pct pct_se mnx mnx_se;

```

IDCNTRY	N	TOTWGT4	PCT	PCT_SE	MNX	MNX_SE
36	1831	159626.80	3.0232	0.25966	3.32880	0.12946
40	1792	65333.56	1.2373	0.07283	1.99342	0.07288
124	4825	241477.00	4.5733	0.29424	2.74447	0.11121
196	511	4342.42	0.0822	0.00291	3.15567	0.10511
200	2071	129408.30	2.4508	0.30102	1.39505	0.06790
208	1998	28029.64	0.5308	0.03578	2.39070	0.06478
250	950	387655.05	7.3417	0.77456	3.42605	0.11214
348	4745	103208.90	1.9547	0.07385	2.94807	0.06831
352	1554	2101.36	0.0398	0.00142	2.08553	0.04549
380	1542	363367.49	6.8818	0.31252	4.04203	0.13652
440	2701	20808.68	0.3941	0.02709	3.17745	0.11068
528	1406	139934.13	2.6502	0.09715	1.71150	0.05911
554	1677	35665.95	0.6755	0.04405	2.23366	0.05504
578	2245	39265.14	0.7436	0.03777	1.91681	0.05053
643	2167	978597.41	18.5335	1.12430	3.51437	0.10417
717	2080	290355.85	5.4990	0.46651	4.76653	0.17059
752	2910	67563.13	1.2796	0.06388	1.89564	0.06706
756	3073	57399.66	1.0871	0.05263	1.97698	0.06477
840	5533	2141931.58	40.5658	1.56629	1.73739	0.06442
890	1383	24075.40	0.4560	0.04106	2.15454	0.12391

**Figure 9.10**

**Sample Table for Analysis of Continuous Variables Taken from the TIMSS International Report “Mathematics and Science Achievement in the Final Year of Secondary School”**

**Table 4.15**

**Students’ Reports on the Hours Per Day Spent Studying or Doing Homework<sup>†</sup>  
Mathematics and Science Literacy  
Final Year of Secondary School\***

Country	Less Than One Hour		At Least 1 Hour But Less Than 2 Hours		2-3 Hours		More Than 3 Hours		Average Hours <sup>1</sup>
	Percent of Students	Mean Achievement	Percent of Students	Mean Achievement	Percent of Students	Mean Achievement	Percent of Students	Mean Achievement	
<i>Australia</i>	13 (1.6)	489 (15.8)	20 (1.5)	527 (12.6)	22 (1.5)	525 (6.5)	44 (2.7)	545 (9.5)	3.3 (0.13)
<i>Austria</i>	31 (2.0)	508 (7.7)	28 (1.4)	533 (6.8)	20 (1.8)	524 (7.9)	20 (1.4)	529 (8.7)	2.0 (0.07)
<i>Canada</i>	18 (1.2)	505 (3.3)	31 (1.7)	527 (5.9)	21 (1.4)	540 (4.7)	30 (2.0)	537 (5.8)	2.7 (0.11)
<i>Cyprus</i>	12 (1.6)	414 (6.4)	20 (1.6)	441 (6.8)	28 (2.1)	444 (5.4)	40 (2.1)	465 (4.0)	3.2 (0.11)
<i>Czech Republic</i>	39 (2.5)	468 (7.7)	31 (1.3)	482 (13.2)	20 (3.0)	498 (9.8)	10 (1.3)	481 (34.2)	1.4 (0.07)
<i>Denmark</i>	r 13 (1.1)	496 (7.1)	31 (1.6)	546 (4.9)	35 (1.6)	543 (4.1)	21 (1.4)	543 (5.1)	r 2.4 (0.06)
<i>France</i>	s 7 (1.1)	498 (11.1)	23 (1.7)	536 (7.9)	19 (1.5)	521 (8.4)	52 (1.9)	523 (6.1)	s 3.4 (0.11)
<i>Germany</i>	- -	- -	- -	- -	- -	- -	- -	- -	- -
<i>Hungary</i>	13 (0.9)	454 (4.4)	26 (0.9)	469 (4.1)	26 (0.9)	482 (4.6)	36 (1.2)	492 (4.2)	2.9 (0.07)
<i>Iceland</i>	21 (1.0)	547 (4.3)	37 (1.2)	552 (3.0)	23 (1.1)	546 (4.3)	19 (1.2)	532 (4.7)	2.1 (0.05)
<i>Italy</i>	8 (1.1)	461 (13.9)	15 (1.6)	472 (7.5)	21 (1.6)	483 (7.2)	56 (2.4)	481 (6.2)	4.0 (0.14)
<i>Lithuania</i>	16 (1.7)	451 (13.7)	21 (1.1)	469 (6.3)	21 (0.9)	473 (6.4)	41 (2.0)	471 (5.5)	3.2 (0.11)
<i>Netherlands</i>	25 (2.2)	527 (8.0)	46 (1.4)	569 (5.2)	16 (1.6)	574 (9.9)	13 (1.3)	572 (12.6)	1.7 (0.06)
<i>New Zealand</i>	20 (2.0)	489 (10.1)	34 (1.9)	539 (7.1)	25 (1.4)	538 (5.2)	22 (1.1)	536 (5.8)	2.2 (0.06)
<i>Norway</i>	27 (1.2)	522 (5.7)	37 (1.0)	547 (5.0)	17 (1.3)	556 (5.8)	19 (1.1)	535 (6.7)	1.9 (0.05)
<i>Russian Federation</i>	10 (0.8)	463 (7.6)	24 (1.4)	467 (6.8)	21 (1.2)	479 (7.0)	45 (1.9)	486 (6.0)	3.5 (0.10)
<i>Slovenia</i>	19 (2.3)	493 (11.7)	36 (3.0)	541 (9.8)	23 (2.1)	508 (7.8)	22 (2.2)	527 (8.7)	2.2 (0.12)
<i>South Africa</i>	r 8 (0.8)	353 (11.9)	13 (1.5)	389 (18.3)	20 (1.5)	370 (15.9)	59 (2.2)	360 (9.2)	r 4.8 (0.17)
<i>Sweden</i>	28 (1.6)	533 (5.5)	34 (1.2)	575 (5.7)	20 (1.2)	565 (5.9)	19 (1.4)	560 (6.7)	1.9 (0.07)
<i>Switzerland</i>	28 (1.9)	520 (6.1)	34 (1.2)	544 (8.0)	21 (1.4)	535 (7.5)	17 (1.1)	545 (7.6)	2.0 (0.06)
<i>United States</i>	34 (1.7)	452 (3.7)	34 (1.1)	481 (4.9)	18 (0.9)	479 (5.9)	15 (1.1)	501 (7.9)	1.7 (0.06)

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1995-96.

<sup>†</sup> Study time is defined as the sum of time reported spent studying or doing homework in mathematics, science, and other subjects.

\* See Appendix A for characteristics of the students sampled.

<sup>1</sup> Based on sum of responses to three questions about time spent studying or doing homework for mathematics, science, and other subjects. Categorization and average hours based on: No time = 0; Less than 1 hour = .5; 1-2 hours = 1.5; 3-5 hours = 4; More than 5 hours = 7.

Countries shown in italics did not satisfy one or more guidelines for sample participation rates or student sampling (see Figure B.4).

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

An "r" indicates a 70-84% student response rate. An "s" indicates a 50-69% student response rate.

A dash (-) indicates data are not available.

## 9.4.2 SPSS Macro for Computing Mean and Percents with Corresponding Standard Errors (JACKGEN.SPS)

The CD containing the TIMSS International Database also contains an SPSS macro program called JACKGEN.SPS. This macro can be used to compute weighted percents and means within categories. Although the user can compute weighted percent and mean estimates using other basic SPSS commands, the macro JACKGEN.SPS also computes the JRR error estimate for these means and percents. The control code for the macro JACKGEN.SPS is presented in Figure 9.7. This macro should not be used to compute standard errors for the plausible values. Standard errors of plausible values should be computed with the macro JACKPV described in section 9.3 of this User Guide.

**Figure 9.11**  
**SPSS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKGEN.SPS)**

```

set mprint=on.

define jackgen (cvar = !charend('/') /
               dvar = !charend('/') /
               njkr = !charend('/') /
               jkz = !charend('/') /
               jki = !charend('/') /
               wgt = !charend('/') ) .

weight off.
select if not(missing(!dvar)).

sort cases by !cvar.

compute k = 1.

compute wgtx = !wgt * !dvar.

vector rwgt(!njkr).
vector rwgtx(!njkr).
loop #i = 1 to !njkr.
if (!jkz = #i and !jki = 0) rwgt(#i) = !wgt * 0.
if (!jkz = #i and !jki = 1) rwgt(#i) = !wgt * 2.
if (!jkz <>#i
    ) rwgt(#i) = !wgt * 1.

compute rwgtx(#i) = rwgt(#i) * !dvar.

end loop.

!let !nway = tmpjck1

aggregate outfile = !nway
/ presorted
/ break = !cvar k
/ !wgt rwgt1 to !concat(rwgt,!njkr) wgtx rwgtx1 to !concat(rwgtx,!njkr)
= sum(!wgt rwgt1 to !concat(rwgt,!njkr) wgtx rwgtx1 to !concat(rwgtx,!njkr))
/ n = n(!wgt).

!let !cvar1 = !null
!do !i !in(!cvar)
!let !cvar1 = !concat(!i,!blank(1),!cvar1)
!doend

```



**Figure 9.11 (Continued 1)****SPSS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKGEN.SPS)**

```

!let !cvar2a = !concat(!cvar1,!blank(1),k)
!let !cvar2b = !concat(!tail(!cvar1),!blank(1),k)

!do !i !in(!cvar2a)

!if (!cvar2b = !null) !then !break
!ifend

get file = !nway.
compute k = 1.

!let !seq = !index(!cvar1,!i)

aggregate outfile = tmpjckA
  / break = !cvar2a
  /      = sum(!wgt rwgt1 to !concat(rwgt,!njkr) wgtx rwgtx1 to !concat(rwgtx,!njkr)
  / n = sum(n).

aggregate outfile = tmpjckB
  / break = !cvar2b
  /      = sum(!wgt rwgt1 to !concat(rwgt,!njkr)
  /      = sum(!wgt rwgt1 to !concat(rwgt,!njkr)).

get file = tmpjckA.
sort cases by !cvar2B.
save outfile = tmpjckA.

match files
  / file = tmpjckA
  / table = tmpjckB
  / by !cvar2b.

compute type = !seq.

!if (!head(!cvar1)=!head(!cvar2a)) !then
save outfile = jackfnl.
!else
add files
  / file = jackfnl
  / file = *.
save outfile = jackfnl.
!ifend.

!let !cvar2a = !tail(!cvar2a).
!let !cvar2b = !tail(!cvar2b).

!doend

do if (!wgt > 0).
compute mnx = wgtx / !wgt.
compute pct = 100 * (!wgt / twgt).
end if.

compute mnx_var = 0.
compute pct_var = 0.

```

## Figure 9.11 (Continued 2)

### SPSS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACKGEN.SPS)

```

vector rwgt = rwgt1 to !concat(rwgt,!njkr).
vector twgt = twgt1 to !concat(twgt,!njkr).
vector wgtx = rwgtx1 to !concat(rwgtx,!njkr).
loop #i = 1 to !njkr.
do if (rwgt(#i)) > 0.
compute mnx_var = mnx_var + (mnx - (wgtx(#i) / rwgt(#i)))*2.
compute pct_var = pct_var + (pct - (100 * (rwgt(#i) / twgt(#i))))*2.
end if.
end loop.

compute mnx_se = sqrt(mnx_var).
compute pct_se = sqrt(pct_var).

autorecode type
  / into level.
execute.

select if level=1.

save outfile = final.

!enddefine.

```

The user needs to know some basic SPSS macro language in order to use the macro. It should be first included in the program file where it is going to be used. If the user is operating in batch mode, then the macro should be called in every batch. If the user is using SPSS interactively then the macro should be called once at the beginning of the session and it will remain active throughout the session. If the session is terminated and restarted at a later time the macro should be called once again. Once the macro is included in a specific session the word “JACKGEN” should not be used within that program because doing so will invoke the macro.

The macro is included in the program file where it will be used by issuing the following command under SPSS:

```
include "<path>jackgen.sps".
```

where <path> points to the specific drive and directory where the macro JACKGEN.SPS can be found. The macro requires that several parameters be submitted when it is invoked. These parameters are:

#### **WGT**

The sampling weight to be used in the analysis. Depending on the population of interest the user will generally choose between TOTWGTL4, TOTWGTM4 or TOTWGTP4.

#### **JKZ**

The variable that captures the assignment of the student to a particular sampling zone. The name of this variable in all TIMSS files is JKZONE.

***JKI***

The variable that captures whether the case is to be dropped or have its weight doubled for the corresponding replicate weight. The name of this variable in all TIMSS files is JKINDIC.

***NJKR***

This indicates the number of replicate weights to be generated when computing the JRR error estimates. When conducting analyses using the data from all countries the value of NJKR should be set to 75 for the student and school. The user working with the data for only one country should set the NJKR argument to as many replicates as are needed for one country. The maximum number of replicates by country is shown in Table 9.2. If the data from two or more countries is being used for an analysis, then the larger number of jackknife zones should be used. When in doubt on what number to set the NJKR parameter, it should be set to 75. The error variance will always be estimated correctly if more replicate weights than necessary are computed, but will be underestimated if the user specifies less replicate weights than necessary.

***CVAR***

This lists the variables that are to be used to classify the students in the data file. This can be a single variable, or a list of variables. The maximum number of variables will depend mostly on the computer resources available to the user at the time. It is recommended to always include the variable that identifies the country. At least one variable has to be specified, usually IDCNTY.

***DVAR***

This is the variable for which means are to be computed. Only one variable can be listed here. This variable should not be any of the plausible values. This variable should be considered as a continuous variables.

Unlike the SAS macro (JACKGEN.SAS), the JACKGEN macro in SPSS does not require that the data file that contains the data of interest be specified when calling the macro. By default, SPSS uses the current working file. This needs to be read with the GET FILE command prior to invoking the macro.

The simplest and most straightforward way is to invoke the macro using the conventional SPSS notation for invoking macros. This involves listing the macro name followed by the corresponding list of arguments for the analysis, each separated by a slash. For example, if the macro is invoked as:

```
get file = "CSGALL1.SAV".
jackgen cvar = IDCNTY /
  dvar = CSDGSTUD /
  jkz = JKZONE /
  jki = JKINDIC /
  njkr = 75 /
  WGT = TOTWGTL4 / .
```

it will compute the mean number of hours the students within each country spend studying or doing homework, and its corresponding standard error, using the variable TOTWGTL4 as the

sampling weight. In this case it will also compute the percent of students within the country compared to the overall student population (an irrelevant figure for this analysis), and its corresponding standard error. The data to be used is that contained in the file `CSGALL1.SAV`, and the standard error of the statistics will be computed based on 75 replicate weights. The results from this analysis are those presented in the last column of Table 4.15 from the TIMSS international report “Mathematics and Science in the Final Year of Secondary School,” shown in Figure 9.10 of this User Guide.

The file that contains these results is then called `FINAL` and is saved to the default directory being used by SPSS. The variables that are contained in this file are:

### ***Classification Variables***

Each of the classification variables are kept in the resulting file. In the above example there was only one classification variable in the `WORK.FINAL` data set. This is `IDCOUNTRY`. There is one unique occurrence for each category for this variable. If there were more variables included in the `CVAR` parameter then there would be one unique occurrence for each combination of the categories of the classification variables.

### ***Weight Variable***

Contains the estimate in the population that belongs to the group defined by the specific combination of the classification variable categories. In the example this variable is called `TOTWGTL4`.

### ***N***

Contains the number of cases in the group defined by the specific combination of categories for the classification variables.

### ***MNX***

Contains the weighted mean of the variable `DVAR` for the group defined by the corresponding combination of classification variable categories.

### ***MNX\_SE***

Contains the standard error of the mean for variable specified in `DVAR`, computed using the `JRR` method for computing the standard error.

### ***PCT***

Contains the weighted percent of people in the group for the classification variable listed last, within the specific combination of the categories defined by the groups defined initially. In the example, we would obtain the percent of boys and girls, within each country.

### ***PCT\_SE***

Contains the standard error of `PCT` computed using the `JRR` method for computing the standard error.

The file resulting from using this macro can be printed using an SPSS procedure of choice. An example call to this macro, and a subset of the resulting file is presented in Figure 9.12. In this example the macro computes the percent of students within each country (an irrelevant figure for our analysis) and the mean number of hours spent studying or doing homework. The listing presented in Figure 9.12 is interpreted in the following way. The first line shows that there were 1831 students in the sample for IDCNTRY=36 (Australia) for whom there was data available in the variables for this analysis. It is estimated that on average Australian students spend about 3.3288 hours each day studying or doing homework, with a standard error of 0.129. The columns titled PCT and PCT\_SE in this case are irrelevant as they indicate the percent of students in the overall population that are Australian and the standard error of this estimate. The second line shows the same information but for IDCNTRY=40 (Austria) where students report that on average they spent 1.993 hours per day studying or doing homework. Again, the columns titled PCT and PCT\_SE in this case are irrelevant as they indicate the percent of students in the overall population that are Austrian and the standard error of this estimate.

**Figure 9.12****SPSS Control Code and Extract of Output File for Using the Macro JACK.SPS**

```

set noverify
spss

get file = "csgall1.sys"
  / keep= idcntry idstud jkindic jkzone totwgtl4
        csdgstud intmsl4.

select if intmsl4 = 1.

* Now use the macro JACKGEN to get the results.
include "jackgen.sps".

jackgen cvar = idcntry   /
        dvar = csdgstud /
        njkr = 75       /
        jkz  = jkzone   /
        jki  = jkindic  /
        wgt  = totwgtl4 /.

formats idcntry (f3.0) n (f5.0) totwgtl4 (f10.2).

report format = list automatic
  / var = idcntry n totwgtl4 pct pct_se mnx mnx_se

```

IDCNTRY	N	TOTWGT4	PCT	PCT_SE	MNX	MNX_SE
36	1831	159626.80	3.02	.26	3.33	.13
40	1792	65333.56	1.24	.07	1.99	.07
124	4825	241477.00	4.57	.29	2.74	.11
196	511	4342.42	.08	.00	3.16	.11
200	2071	129408.30	2.45	.30	1.40	.07
208	1998	28029.64	.53	.04	2.39	.06
250	950	387655.05	7.34	.77	3.43	.11
348	4745	103208.90	1.95	.07	2.95	.07
352	1554	2101.36	.04	.00	2.09	.05
380	1542	363367.49	6.88	.31	4.04	.14
440	2701	20808.68	.39	.03	3.18	.11
528	1406	139934.13	2.65	.10	1.71	.06
554	1677	35665.95	.68	.04	2.23	.06
578	2245	39265.14	.74	.04	1.92	.05
643	2167	978597.41	18.53	1.12	3.51	.10
717	2080	290355.85	5.50	.47	4.77	.17
752	2910	67563.13	1.28	.06	1.90	.07
756	3073	57399.66	1.09	.05	1.98	.06
840	5533	2141931.58	40.57	1.57	1.74	.06
890	1383	24075.40	.46	.04	2.15	.12

## 9.5 Performing Analyses with Student-Level Variables Using Plausible Values

Many analyses of the TIMSS data can be undertaken using student-level data and plausible values. We have already presented two examples in the previous section when explaining how to use the two macros provided with the data files. We will now proceed to work out another example where all the steps are undertaken, including the invocation of the corresponding SAS and SPSS macro. For example, suppose we want to replicate one of the results presented in the international report. We are interested in analyzing students' reports of the hours spent each day watching television and videos, and their achievement in mathematics and science literacy. These are the results that are presented in Figure 9.13 below and that appear in Table 4.20 of the TIMSS international report

“Mathematics and Science Achievement in the Final Year of Secondary School.” Since this table includes results using the plausible values we need to use the macro JACKPV.

To replicate the results from this table we need to undertake several steps. After reviewing the codebooks and the questionnaire information we identify the question asking about the number of hours students watch television (see Supplement 1). Students’ responses to this question are captured in the variable CSBGDAY1, which is found in the Student Background data file. Our next step is to review the documentation of national adaptations to the questionnaires to ensure that there were no deviations listed for this variable (see Supplement 2). If no changes were made we can continue with our analysis without any modifications.

We then proceed to read from the SPSS system file or SAS data set that contains this variable, the plausible values for mathematics and science literacy (CSLPV01, CSLPV02, CSLPV03, CSLPV04 and CSLPV05), the sampling weight for the student (TOTWGTL4), the variables that contain the jackknife replication information (JKZONE and JKINDIC), and the variable containing the country identification code (IDCOUNTRY). In this analysis we will use the data for all countries in the database, although the same steps should be taken to analyze these variables within a single country or for a select group of countries.

The SAS code used to perform this analysis is presented in Figure 9.14, and the SPSS code is presented in Figure 9.15. A selection of the results obtained from these programs are displayed in Figure 9.16 and Figure 9.17. We have included as part of the programs the corresponding value labels and format statements so that the different categories or groups are labeled appropriately.

Notice that one of the steps in each of these programs was to select only those students in the sample who were assigned INTMSL4 = 1. These are those students who were administered the mathematics and science literacy test, our group of interest for this analysis. We also selected those who answered the question about the number of hours they watch television each day, and we recoded the variable so that there would be one category for those who reported watching television or video less than one hour per day, as was done for the report table shown in Figure 9.13.

Note also in this analysis that although we have used the data from all the countries, we selected from the system file or SAS data set only those variables that are relevant to our analysis. Although, in general, this analysis is quite feasible with a powerful desktop computer, the user needs to keep in mind that computing and storage requirements for this type of analysis are quite demanding so it is always prudent to select out only the variables that are necessary for an analysis.

In general, to perform analyses using the Student Background data files and involving the use of plausible values, the user needs to do the following:

- Identify the variable or variables of interest in the student file and find out about any specific national adaptations to the variable.
- Retrieve the relevant variables from the data files, including the plausible values, sampling weights, JRR replication information, and any other variables used in the selection of cases.
- Use the macro JACKPV with the corresponding arguments and parameters.
- Print out the result file.



## Figure 9.13

Sample Table for Analysis of Student Level Variables and Plausible Values Taken from the TIMSS International Report "Mathematics and Science Achievement in the Final Year of Secondary School"

Table 4.20

Students' Reports on the Hours Per Day Spent Watching Television or Videos  
Mathematics and Science Literacy  
Final Year of Secondary School\*

Country	Less Than One Hour		1-2 Hours		3-5 Hours		More Than 5 Hours	
	Percent of Students	Mean Achievement	Percent of Students	Mean Achievement	Percent of Students	Mean Achievement	Percent of Students	Mean Achievement
<i>Australia</i>	34 (2.1)	532 (11.0)	44 (2.2)	530 (8.9)	17 (1.5)	527 (7.3)	5 (0.6)	476 (10.5)
<i>Austria</i>	37 (2.2)	532 (8.2)	47 (1.8)	518 (5.0)	14 (1.0)	507 (6.8)	2 (0.4)	--
<i>Canada</i>	38 (1.1)	531 (3.3)	44 (1.4)	528 (3.7)	15 (1.3)	512 (5.7)	3 (0.5)	502 (12.2)
<b>Cyprus</b>	38 (1.9)	451 (5.5)	43 (2.3)	447 (4.2)	16 (1.6)	434 (6.7)	2 (0.7)	--
<b>Czech Republic</b>	21 (2.3)	512 (11.1)	51 (2.7)	479 (10.6)	22 (1.6)	450 (11.6)	5 (1.0)	429 (5.9)
<i>Denmark</i>	32 (1.4)	541 (4.4)	51 (1.3)	526 (3.5)	15 (1.1)	519 (5.9)	2 (0.3)	--
<i>France</i>	49 (2.1)	512 (5.0)	41 (1.7)	503 (6.6)	8 (0.8)	491 (9.4)	2 (0.4)	--
<i>Germany</i>	x x	x x	x x	x x	x x	x x	x x	x x
<b>Hungary</b>	31 (0.9)	505 (4.4)	44 (0.9)	475 (3.4)	19 (0.7)	451 (4.3)	6 (0.5)	426 (5.0)
<i>Iceland</i>	38 (1.1)	555 (4.0)	47 (1.4)	536 (2.9)	14 (1.0)	528 (4.3)	2 (0.3)	--
<i>Italy</i>	36 (1.4)	477 (6.8)	50 (1.6)	477 (6.3)	12 (0.9)	473 (6.9)	2 (0.4)	--
<b>Lithuania</b>	21 (1.0)	473 (6.5)	50 (1.3)	469 (5.3)	24 (1.3)	459 (7.5)	6 (0.7)	439 (12.5)
<i>Netherlands</i>	19 (1.4)	569 (10.4)	49 (1.6)	567 (5.0)	27 (1.7)	547 (5.9)	4 (0.7)	503 (14.0)
<b>New Zealand</b>	28 (1.4)	537 (6.1)	45 (1.6)	527 (4.9)	20 (1.4)	511 (9.7)	7 (1.4)	506 (12.5)
<i>Norway</i>	32 (1.5)	549 (5.0)	50 (1.2)	540 (4.5)	16 (1.0)	505 (5.3)	2 (0.4)	--
<b>Russian Federation</b>	15 (1.2)	490 (8.9)	46 (1.4)	482 (6.4)	31 (1.4)	469 (6.5)	7 (0.8)	451 (8.9)
<i>Slovenia</i>	44 (2.6)	520 (9.2)	44 (2.3)	520 (8.2)	10 (1.1)	486 (14.4)	3 (0.6)	484 (19.0)
<i>South Africa</i>	57 (2.3)	345 (8.9)	29 (2.0)	377 (12.4)	11 (1.1)	389 (17.3)	3 (0.5)	342 (15.5)
<b>Sweden</b>	33 (1.1)	576 (5.3)	51 (1.0)	554 (4.5)	15 (0.7)	526 (7.2)	1 (0.2)	--
<b>Switzerland</b>	55 (2.1)	545 (4.8)	36 (1.7)	521 (7.8)	8 (0.9)	505 (9.3)	2 (0.4)	--
<i>United States</i>	40 (1.5)	483 (4.2)	39 (1.1)	473 (4.1)	15 (0.9)	458 (4.5)	5 (0.4)	424 (7.9)

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1995-96.

\* See Appendix A for characteristics of the students sampled.

Countries shown in italics did not satisfy one or more guidelines for sample participation rates or student sampling (see Figure B.4).

( ) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

An "x" indicates data available for <50% students.

A tilde (-) indicates insufficient data to report achievement.

**Figure 9.14****SAS Control Statements for Performing Analyses with Student-Level Variables Using Plausible Values (EXAMPLE1.SAS)**

```

libname pop3 "{directory location}";

options nocenter;
* Read the variables from the student file;
data student;
  set pop3.csgall1
    (keep=idcntry idpop jkindic jkzone intmsl4 totwgtl4
      csbgday1 cs1pv01--cs1pv05);
  if intmsl4 = 1 ;          * Select only Literacy Sample;
  if nmiss(csbgday1)=0;
  if csbgday1=1 then csbgday1=2;

* Define the format for the variables used;
proc format library=work;
  value idpop 1 = "Population 1"
             2 = "Population 2"
             3 = "Population 3";
  value seetv 2 = "Less than 1 Hr."   3 = "1 to 2 hrs.      "
             4 = "3 to 5 hrs.      "  5 = "More than 5 hrs.";

  value country
    036 = "Australia"      040 = "Austria"      056 = "Belgium (Fl)"
    057 = "Belgium (Fr)"  100 = "Bulgaria"     124 = "Canada"
    170 = "Colombia"      196 = "Cyprus"       200 = "Czech Republic"
    208 = "Denmark"      826 = "England"     250 = "France"
    280 = "Germany"     300 = "Greece"      344 = "Hong Kong"
    348 = "Hungary"     352 = "Iceland"     364 = "Iran, Islamic Rep."
    372 = "Ireland"     376 = "Israel"      380 = "Italy"
    392 = "Japan"       410 = "Korea"       414 = "Kuwait"
    428 = "Latvia (LSS)" 440 = "Lithuania"   528 = "Netherlands"
    554 = "New Zealand"  578 = "Norway"      608 = "Philippines"
    620 = "Portugal"    642 = "Romania"     643 = "Russian Federation"
    827 = "Scotland"    702 = "Singapore"  201 = "Slovak Republic"
    890 = "Slovenia"    717 = "South Africa" 724 = "Spain"
    752 = "Sweden"     756 = "Switzerland" 764 = "Thailand"
    840 = "United States";

* Now use the macro JACKPV to get the results;
%include "jackpv.sas";

%jackpv(totwgtl4,jkzone,jkindic,75,idcntry idpop csbgday1,cs1pv0,5,student);

proc print data=final noobs;
  by idcntry idpop;
  where idpop = 3;
  var csbgday1 N totwgtl4 pct pct_se mnpv mnpv_se;
  format idcntry country. idpop idpop. csbgday1 seetv.;

```

**Figure 9.15****SPSS Control Statements for Performing Analyses with Student-Level Variables Using Plausible Values (EXAMPLE1.SPS)**

```

get file = "csgall1.sys"
  / keep=idcntry idstud idpop jkindic jkzone totwgtl4
    csbgday1 cs1pv01 to cs1pv05 intmsl4.

recode csbgday1 (1,2=2) (8,9=sysmis) (else=copy).

select if not(missing(csbgday1)) and intmsl4 = 1.

value labels
  idpop 1 "Population 1"
        2 "Population 2"
        3 "Population 3" /
  csbgday1 2 "Less than 1 Hr." 3 "1 to 2 hrs."
           4 "3 to 5 hrs." 5 "More than 5 hrs." /
  idcntry
036 "Australia"      040 "Austria"      056 "Belgium (Fl)"
057 "Belgium (Fr)"  100 "Bulgaria"     124 "Canada"
170 "Colombia"      196 "Cyprus"       200 "Czech Republic"
208 "Denmark"       826 "England"     250 "France"
280 "Germany"       300 "Greece"      344 "Hong Kong"
348 "Hungary"       352 "Iceland"     364 "Iran, Islamic Rep."
372 "Ireland"       376 "Israel"      380 "Italy"
392 "Japan"         410 "Korea"       414 "Kuwait"
428 "Latvia (LSS)"  440 "Lithuania"   528 "Netherlands"
554 "New Zealand"  578 "Norway"      608 "Philippines"
620 "Portugal"     642 "Romania"     643 "Russian Federation"
827 "Scotland"     702 "Singapore"  201 "Slovak Republic"
890 "Slovenia"     717 "South Africa" 724 "Spain"
752 "Sweden"       756 "Switzerland" 764 "Thailand"
840 "United States".

* Now use the macro JACKPV to get the results.
include "jackpv.sps".

jackpv  cvar = idcntry idpop csbgday1 /
        pvs  = cs1pv01 to cs1pv05 /
        npv  = 5 /
        njkr = 75 /
        jkz  = jkzone /
        jki  = jkindic /
        wgt  = totwgtl4 / .

sort cases by idcntry idpop.

formats n (f5.0) totwgtl4 (f10.2).

report format=list automatic
  / var = csbgday1 (label) n totwgtl4 pct pct_se mnpv mnpv_se
  / break = idcntry idpop.

```

**Figure 9.16****Extract of SAS Computer Output for Performing Analyses with Student-Level Variables Using Plausible Values**

```

*COUNTRY ID*=Australia *POPULATION ID*=Population 3

```

CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Less than 1 Hr.	677	56672.35	34.0968	2.09173	531.878	10.9588
1 to 2 hrs.	823	73786.77	44.3936	2.18748	530.078	8.8530
3 to 5 hrs.	319	28046.65	16.8742	1.45446	526.983	7.2784
More than 5 hrs.	82	7704.60	4.6355	0.61063	476.364	10.4701

```

*COUNTRY ID*=Austria *POPULATION ID*=Population 3

```

CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Less than 1 Hr.	681	25172.86	37.0117	2.16233	531.681	8.2482
1 to 2 hrs.	888	32046.58	47.1181	1.75478	517.650	4.9906
3 to 5 hrs.	287	9613.64	14.1349	1.03042	506.785	6.7653
More than 5 hrs.	40	1180.25	1.7353	0.42252	480.065	15.4169

```

*COUNTRY ID*=Canada *POPULATION ID*=Population 3

```

CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Less than 1 Hr.	2023	99128.02	38.2982	1.07370	530.957	3.3267
1 to 2 hrs.	2240	114990.45	44.4266	1.38480	528.318	3.6657
3 to 5 hrs.	752	38069.53	14.7082	1.27147	512.252	5.6964
More than 5 hrs.	133	6644.14	2.5670	0.49118	501.749	12.2315

```

.
.
.
.
.

```

```

*COUNTRY ID*=Switzerland *POPULATION ID*=Population 3

```

CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Less than 1 Hr.	1936	34865.69	54.9477	2.05323	545.258	4.7887
1 to 2 hrs.	1069	22687.18	35.7546	1.66312	520.715	7.7783
3 to 5 hrs.	212	4918.94	7.7522	0.89440	505.269	9.3335
More than 5 hrs.	32	980.72	1.5456	0.38643	446.025	19.1202

```

*COUNTRY ID*=United States *POPULATION ID*=Population 3

```

CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Less than 1 Hr.	2162	883587.80	40.1030	1.52005	483.354	4.17041
1 to 2 hrs.	2197	868250.45	39.4069	1.14639	473.083	4.11831
3 to 5 hrs.	966	341157.85	15.4840	0.88328	458.392	4.54326
More than 5 hrs.	363	110301.67	5.0062	0.44188	424.331	7.85913

```

*COUNTRY ID*=Slovenia *POPULATION ID*=Population 3

```

CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Less than 1 Hr.	648	11282.19	43.5595	2.55658	520.081	9.2402
1 to 2 hrs.	660	11369.69	43.8973	2.32792	519.948	8.1998
3 to 5 hrs.	191	2527.75	9.7594	1.09445	486.293	14.3516
More than 5 hrs.	60	721.01	2.7837	0.62196	484.352	19.0280

**Figure 9.17**  
**Extract of SPSS Computer Output for Performing Analyses with Student-Level Variables Using Plausible Values**

IDCNTY	CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Australia							
Population	3 Less than 1 Hr.	677	56672.35	34.10	2.09	531.88	10.96
	1 to 2 hrs.	823	73786.77	44.39	2.19	530.08	8.85
	3 to 5 hrs.	319	28046.65	16.87	1.45	526.98	7.28
	More than 5 hrs.	82	7704.60	4.64	.61	476.36	10.47
Austria							
Population	3 Less than 1 Hr.	681	25172.86	37.01	2.16	531.68	8.25
	1 to 2 hrs.	888	32046.58	47.12	1.75	517.65	4.99
	3 to 5 hrs.	287	9613.64	14.13	1.03	506.79	6.77
	More than 5 hrs.	40	1180.25	1.74	.42	480.06	15.42
Canada							
Population	3 Less than 1 Hr.	2023	99128.02	38.30	1.07	530.96	3.33
	1 to 2 hrs.	2240	114990.45	44.43	1.38	528.32	3.67
	3 to 5 hrs.	752	38069.53	14.71	1.27	512.25	5.70
	More than 5 hrs.	133	6644.14	2.57	.49	501.75	12.23
.							
.							
.							
Switzerland							
Population	3 Less than 1 Hr.	1936	34865.69	54.95	2.05	545.26	4.79
	1 to 2 hrs.	1069	22687.18	35.75	1.66	520.72	7.78
	3 to 5 hrs.	212	4918.94	7.75	.89	505.27	9.33
	More than 5 hrs.	32	980.72	1.55	.39	446.02	19.12
United States							
Population	3 Less than 1 Hr.	2162	883587.80	40.10	1.52	483.35	4.17
	1 to 2 hrs.	2197	868250.45	39.41	1.15	473.08	4.12
	3 to 5 hrs.	966	341157.85	15.48	.88	458.39	4.54
	More than 5 hrs.	363	110301.67	5.01	.44	424.33	7.86
Slovenia							
Population	3 Less than 1 Hr.	648	11282.19	43.56	2.56	520.08	9.24
	1 to 2 hrs.	660	11369.69	43.90	2.33	519.95	8.20
	3 to 5 hrs.	191	2527.75	9.76	1.09	486.29	14.35
	More than 5 hrs.	60	721.01	2.78	.62	484.35	19.03

## 9.6 Performing Analyses with Student-Level Variables Not Using Plausible Values

Many analyses of the TIMSS data can be undertaken using student-level data other than the plausible values. We have already presented an example in the previous section when explaining how to use the two macros provided with the data files. We now proceed to work out another example, including the invocation of the corresponding SAS and SPSS macro, in an analysis which does not include plausible values. Let us say we are interested in looking at the reports from the students in their final year of secondary school on the hours spent each day watching television and videos, and reports on the hours spent daily studying or doing homework. In this case we are

not interested in student achievement for which we would require plausible values but rather on the relationship between watching television and the numbers of hours studied.

To carry out this analysis we need to undertake several steps. After reviewing the codebooks and the questionnaire information, we identify the question asking about the number of hours students watch television (see Supplement 1). Students' responses to this question are captured in the variable CSBGDAY1, and this variable is found in the Student Background data file. Our next step is to review the documentation of national adaptations to the questionnaires to ensure that there were no deviations listed for this variable (see Supplement 2). If no changes were made we can continue with our analysis without any modifications. We also know from reviewing Supplement 1 that students were asked several questions about the number of hours they spent studying different subjects. These questions were summarized in a variable called CSDGSTUD (see Supplement 3) that is included in the database.

We then proceed to read from the SPSS system file or SAS data set that contains this variable, the derived variable that captures the number of hours spent studying or doing homework (CSDGSTUD), the sampling weight for the student (TOTWGTL4), the variables that contain the jackknife replication information (JKZONE and JKINDIC), and the variable containing the country identification code (IDCNTRY). In this analysis we will use the data for all countries in the database, although the same steps should be taken to analyze these variables within a single country or for a select group of countries.

The SAS code used to perform this analysis is presented in Figure 9.18, and the SPSS code is presented in Figure 9.19. A selection of the results obtained from these programs are displayed in Figure 9.20 and Figure 9.21. We have included as part of the programs the corresponding value labels and format statements so that the different categories or groups are labeled appropriately.

Notice that one of the steps in each of these programs was to select only those students in the sample who were assigned INTMSL4 = 1. These are those students who were administered the mathematics and science literacy test, our group of interest for this analysis. We also selected those who answered the question about the number of hours they watch television each day, and we recoded the variable so that there would be one category for those who reported watching television or video less than one hour per day.

Note also in this analysis that although we have used the data from all the countries, we selected from the system file or SAS data set only those variables that are relevant to our analysis. Although, in general, this analysis is quite feasible with a powerful desktop computer, the user needs to keep in mind that computing and storage requirements for these type of analysis are quite demanding so it is always prudent to select out only the variables that are necessary for an analysis.

In general, to perform analyses that involve percents and means on variables other than plausible values from the Student Background data files the user needs to do the following:

- Identify the variable or variables of interest in the student file and find out about any specific national adaptations to the variable.
- Retrieve the relevant variables from the data files, sampling weights, JRR replication information, and any other variables used in the selection of cases.
- Use the macro JACKGEN with the corresponding arguments and parameters.
- Print out the result file.

**Figure 9.18**

**SAS Control Statements for Performing Analyses with Student-Level Variables Not Using Plausible Values (EXAMPLE2.SAS)**

```

libname pop3 "{directory location}";

options nocenter;
* Read the variables from the student file;
data student;
    set pop3.csgall1
        (keep=idcntry idpop jkindic jkzone intmsl4 totwgtl4
          csbgday1 csdgstud);
    if intmsl4 = 1 ; * Select only Literacy Sample;
    if nmiss(csbgday1)=0 and nmiss(csdgstud)=0;
    if csbgday1=1 then csbgday1=2;

* Define the format for the variables used;
proc format library=work;
    value idpop 1 = "Population 1"
                2 = "Population 2"
                3 = "Population 3";
    value seetv 2 = "Less than 1 Hr." 3 = "1 to 2 hrs.      "
                4 = "3 to 5 hrs.      " 5 = "More than 5 hrs.";
    value country
036 = "Australia"      040 = "Austria"      056 = "Belgium (Fl)"
057 = "Belgium (Fr)"  100 = "Bulgaria"     124 = "Canada"
170 = "Colombia"      196 = "Cyprus"       200 = "Czech Republic"
208 = "Denmark"      826 = "England"     250 = "France"
280 = "Germany"      300 = "Greece"      344 = "Hong Kong"
348 = "Hungary"      352 = "Iceland"     364 = "Iran, Islamic Rep."
372 = "Ireland"      376 = "Israel"      380 = "Italy"
392 = "Japan"        410 = "Korea"       414 = "Kuwait"
428 = "Latvia (LSS)" 440 = "Lithuania"   528 = "Netherlands"
554 = "New Zealand" 578 = "Norway"      608 = "Philippines"
620 = "Portugal"     642 = "Romania"     643 = "Russian Federation"
827 = "Scotland"     702 = "Singapore"  201 = "Slovak Republic"
890 = "Slovenia"     717 = "South Africa" 724 = "Spain"
752 = "Sweden"       756 = "Switzerland" 764 = "Thailand"
840 = "United States";

* Now use the macro jackgen to get the results;
%include "jackgen.sas";

%jackgen(totwgtl4,jkzone,jkindic,75,idcntry idpop csbgday1,csdgstud,student);

proc print data=final noobs;
    by idcntry idpop;
    where idpop = 3;
    var csbgday1 N totwgtl4 pct pct_se mnx mnx_se ;
    format idcntry country. idpop idpop. csbgday1 seetv.;

```

**Figure 9.19****SPSS Control Statements for Performing Analyses with Student-Level Variables Not Using Plausible Values (EXAMPLE2.SPS)**

```

get file = "csgall1.sys"
  / keep=idcntry idstud idpop jkindic jkzone totwgtl4
    csbgday1 csdgstud intmsl4.

recode csbgday1 (1,2=2) (8,9=sysmis) (else=copy).

select if not(missing(csbgday1)) and intmsl4 = 1 and
  not(missing(csdgstud)).

value labels
  idpop 1 "Population 1"
        2 "Population 2"
        3 "Population 3" /
  csbgday1 2 "Less than 1 Hr." 3 "1 to 2 hrs."
           4 "3 to 5 hrs." 5 "More than 5 hrs." /
  id cntry
036 "Australia" 040 "Austria" 056 "Belgium (Fl)"
057 "Belgium (Fr)" 100 "Bulgaria" 124 "Canada"
170 "Colombia" 196 "Cyprus" 200 "Czech Republic"
208 "Denmark" 826 "England" 250 "France"
280 "Germany" 300 "Greece" 344 "Hong Kong"
348 "Hungary" 352 "Iceland" 364 "Iran, Islamic Rep."
372 "Ireland" 376 "Israel" 380 "Italy"
392 "Japan" 410 "Korea" 414 "Kuwait"
428 "Latvia (LSS)" 440 "Lithuania" 528 "Netherlands"
554 "New Zealand" 578 "Norway" 608 "Philippines"
620 "Portugal" 642 "Romania" 643 "Russian Federation"
827 "Scotland" 702 "Singapore" 201 "Slovak Republic"
890 "Slovenia" 717 "South Africa" 724 "Spain"
752 "Sweden" 756 "Switzerland" 764 "Thailand"
840 "United States".

* Now use the macro jackgen to get the results.
include "jackgen.sps".

jackgen cvar = idcntry idpop csbgday1 /
         dvar = csdgstud /
         njkr = 75 /
         jkz = jkzone /
         jki = jkindic /
         wgt = totwgtl4 / .

sort cases by idcntry idpop.

formats n (f5.0) totwgtl4 (f10.2).

report format=list automatic
  / var = csbgday1 (label) n totwgtl4 pct pct_se mnx mnx_se

  / break = idcntry idpop.

```



**Figure 9.20****Extract of SAS Computer Output for Performing Analyses with Student-Level Variables Not Using Plausible Values**

*COUNTRY ID*=Australia *POPULATION ID*=Population 3						
CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNX	MNX_SE
Less than 1 Hr.	650	54336.66	34.1294	2.04526	3.90726	0.16052
1 to 2 hrs.	793	70975.56	44.5805	2.14383	3.29747	0.20938
3 to 5 hrs.	307	26756.63	16.8061	1.48884	2.69753	0.18990
More than 5 hrs.	76	7138.78	4.4839	0.59396	1.63246	0.23806
*COUNTRY ID*=Austria *POPULATION ID*=Population 3						
CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNX	MNX_SE
Less than 1 Hr.	645	24233.46	37.2794	2.19183	2.11604	0.09942
1 to 2 hrs.	839	30657.87	47.1623	1.76584	1.99881	0.11157
3 to 5 hrs.	264	8978.56	13.8121	1.05207	1.76957	0.11100
More than 5 hrs.	38	1135.13	1.7462	0.43295	1.24869	0.34186
*COUNTRY ID*=Canada *POPULATION ID*=Population 3						
CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNX	MNX_SE
Less than 1 Hr.	1898	91888.29	38.1784	1.13757	2.89107	0.19463
1 to 2 hrs.	2083	106369.81	44.1953	1.46510	2.74880	0.12280
3 to 5 hrs.	707	36259.49	15.0654	1.34865	2.33941	0.20156
More than 5 hrs.	122	6163.52	2.5609	0.49995	2.97774	0.73527
.						
.						
.						
.						
.						
*COUNTRY ID*=Switzerland *POPULATION ID*=Population 3						
CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNX	MNX_SE
Less than 1 Hr.	1828	31390.46	54.9081	1.95935	2.09590	0.07387
1 to 2 hrs.	1004	20622.30	36.0725	1.51229	1.82188	0.10067
3 to 5 hrs.	201	4391.36	7.6814	0.93584	1.73611	0.25173
More than 5 hrs.	27	764.99	1.3381	0.37683	2.70538	0.73771
*COUNTRY ID*=United States *POPULATION ID*=Population 3						
CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNX	MNX_SE
Less than 1 Hr.	2090	853413.29	39.9273	1.51692	1.71364	0.08563
1 to 2 hrs.	2134	844955.57	39.5316	1.15117	1.77763	0.09183
3 to 5 hrs.	950	333677.37	15.6112	0.83980	1.69168	0.08920
More than 5 hrs.	349	105371.05	4.9298	0.44685	1.78200	0.19333
*COUNTRY ID*=Slovenia *POPULATION ID*=Population 3						
CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNX	MNX_SE
Less than 1 Hr.	579	10614.99	44.0946	2.67201	2.45069	0.13939
1 to 2 hrs.	585	10578.00	43.9410	2.49013	1.95378	0.13195
3 to 5 hrs.	165	2224.44	9.2403	0.98838	1.85671	0.19669
More than 5 hrs.	53	655.78	2.7241	0.64518	1.61176	0.47483

**Figure 9.21**  
**Extract of SPSS Computer Output for Performing Analyses with Student-Level Variables Not Using Plausible Values**

IDCNTRY	CSBGDAY1	N	TOTWGTL4	PCT	PCT_SE	MNX	MNX_SE
Australia							
Population	3 Less than 1 Hr.	650	54336.66	34.13	2.05	3.91	.16
	1 to 2 hrs.	793	70975.56	44.58	2.14	3.30	.21
	3 to 5 hrs.	307	26756.63	16.81	1.49	2.70	.19
	More than 5 hrs.	76	7138.78	4.48	.59	1.63	.24
Austria							
Population	3 Less than 1 Hr.	645	24233.46	37.28	2.19	2.12	.10
	1 to 2 hrs.	839	30657.87	47.16	1.77	2.00	.11
	3 to 5 hrs.	264	8978.56	13.81	1.05	1.77	.11
	More than 5 hrs.	38	1135.13	1.75	.43	1.25	.34
Canada							
Population	3 Less than 1 Hr.	1898	91888.29	38.18	1.14	2.89	.19
	1 to 2 hrs.	2083	106369.81	44.20	1.47	2.75	.12
	3 to 5 hrs.	707	36259.49	15.07	1.35	2.34	.20
	More than 5 hrs.	122	6163.52	2.56	.50	2.98	.74
.							
.							
.							
.							
Slovenia							
Population	3 Less than 1 Hr.	579	10614.99	44.09	2.67	2.45	.14
	1 to 2 hrs.	585	10578.00	43.94	2.49	1.95	.13
	3 to 5 hrs.	165	2224.44	9.24	.99	1.86	.20
	More than 5 hrs.	53	655.78	2.72	.65	1.61	.47
Switzerland							
Population	3 Less than 1 Hr.	1828	31390.46	54.91	1.96	2.10	.07
	1 to 2 hrs.	1004	20622.30	36.07	1.51	1.82	.10
	3 to 5 hrs.	201	4391.36	7.68	.94	1.74	.25
	More than 5 hrs.	27	764.99	1.34	.38	2.71	.74
United States							
Population	3 Less than 1 Hr.	2090	853413.29	39.93	1.52	1.71	.09
	1 to 2 hrs.	2134	844955.57	39.53	1.15	1.78	.09
	3 to 5 hrs.	950	333677.37	15.61	.84	1.69	.09
	More than 5 hrs.	349	105371.05	4.93	.45	1.78	.19

## 9.7 Performing Analyses with School-Level Variables

Although the students in the TIMSS samples were selected from within a sample of schools, the school sample was designed to optimize the resulting sample of students, rather than to give an optimal sample of schools. For this reason, it is preferable that school-level variables be analyzed as attributes of students, rather than as elements in their own right. In other words, the school-level data should not be analyzed to make statements about the number of schools with certain characteristics, but rather to make statements about the number of students attending schools with one characteristic or another. When school-level variables are analyzed, we recommend that the user merge the selected school-level variables with the student-level file, and then use the sampling and weight information contained in the student-level file to make the desired statements. The examples presented in this section describe how this can be accomplished using SAS or SPSS.

Lets us say that we want to find out the percent of students in their final year of secondary school that attend schools located in certain geographical area of the country (CCBGCOMM), and their average achievement in mathematics and science literacy. As in the previous example, the first step in our analysis is to locate the variables of interest in the specific codebook and file. We find the variable CCBGCOMM in the School Background file, and the student weights and plausible value scores in the Student Background file. We then proceed to review the documentation on national adaptations and discover, for example, that Australia has modified this variable slightly to fit their particular context. At this time we could proceed in one of two ways: we could exclude Australia from our analysis or we could label the variable accordingly so that we will not be making incorrect inferences about the specific groups. In the latter case, since we want to also explore the results for Australia we take the precaution of labeling the values for variable CCBGCOMM in a generic way before we proceed with the analysis.

After these considerations, we then proceed to read the School Background file and keep only the variables that are relevant to our analysis. In this case we keep the country identification (IDCNTRY) and school identification (IDSCHOOL). because these are the variables that will allow us to merge the school data to the student data, and also the variable of interest, in this case CCBGCOMM. When then read the variables we need from the student data file. First we read the identification of the country and the school (IDCNTRY and IDSCHOOL) which will allow us to merge the student data to the school data. We also select from this file the plausible values for mathematics and science literacy (CSLPV01, CSLPV02, CSLPV03, CSLPV04 and CSLPV05), the sampling weight for the student (TOTWGTL4), the variables that contain the jackknife replication information (JKZONE and JKINDIC), and the variable that will be used to select the sample of students in their final year of secondary school.

We then proceed to merge the school information with the student information using the variables IDCNTRY and IDSCHOOL as merge variables, and then use the macro JACKPV to obtain the corresponding percents of students within each group, and their mean achievement in mathematics and science literacy. If we were interested in looking at the mean of a variable other than the plausible values then we would use the macro JACKGEN. The computer code used to run this analysis in SAS and SPSS can be found in Figure 9.22 and Figure 9.23, and the results are shown in Figure 9.24 and Figure 9.25.

**Figure 9.22****SAS Control Statements for Performing Analyses with School-Level Variables Using Plausible Values (EXAMPLE3.SAS)**

```

libname pop3      "{directory location}";

options nocenter;

* Read the variables from the school file and sort them by the merge variables;
data school;
  set pop3.ccgall1
      (keep=idcntry idschool ccbgcomm);

proc sort data=school;
  by idcntry idschool;

* Read the variables from the student file and sort them by the merge variables;
data student;
  set pop3.csgall1
      (keep = idcntry idschool idpop jkzone jkindic intmsl4 totwgtl4
        cs1pv01--cs1pv05);

proc sort data=student;
  by idcntry idschool;

* Merge the student and school files together;
data merged;
  merge student school;
  by idcntry idschool;
  if nmiss(ccbgcomm)=0;

proc format library=work;
  value idpop 1 = "Population 1"
              2 = "Population 2"
              3 = "Population 3";

  value comm 1 = "Community Type 1"   2 = "Community Type 2"
              3 = "Community Type 3"   4 = "Community Type 4";

  value country
036 = "Australia"      040 = "Austria"      056 = "Belgium (Fl)"
057 = "Belgium (Fr)"  100 = "Bulgaria"     124 = "Canada"
170 = "Colombia"      196 = "Cyprus"       200 = "Czech Republic"
208 = "Denmark"      826 = "England"     250 = "France"
280 = "Germany"      300 = "Greece"      344 = "Hong Kong"
348 = "Hungary"      352 = "Iceland"     364 = "Iran, Islamic Rep."
372 = "Ireland"      376 = "Israel"      380 = "Italy"
392 = "Japan"        410 = "Korea"       414 = "Kuwait"
428 = "Latvia (LSS)" 440 = "Lithuania"   528 = "Netherlands"
554 = "New Zealand" 578 = "Norway"      608 = "Philippines"
620 = "Portugal"     642 = "Romania"     643 = "Russian Federation"
827 = "Scotland"     702 = "Singapore"  201 = "Slovak Republic"
890 = "Slovenia"     717 = "South Africa" 724 = "Spain"
752 = "Sweden"       756 = "Switzerland" 764 = "Thailand"
840 = "United States";

* Now use the macro JACKPV to get the results;
%include "jackpv.sas";

%jackpv (totwgtl4,jkzone,jkindic,75,idcntry idpop ccbgcomm,cs1pv0,5,merged);

proc print data=final noobs;
  where idpop = 3;
  by idcntry idpop;
  var ccbgcomm N totwgtl4 pct pct_se mnpv mnpv_se;

  format idcntry countryl. idpop idpop. ccbgcomm comm.;

```

**Figure 9.23****SPSS Control Statements for Performing Analyses with School-Level Variables Using Plausible Values (EXAMPLE3.SPS)**

```

get file = "ccgall1.sys"
  / keep = idcntry idschool ccbgcomm.

sort cases by idcntry idschool.

save outfile = school.

get file = "csgall1.sys"
  / keep = idcntry idschool idstud idpop jkzone jkindic
    intmsl4 totwgtl4 cslpv01 to cslpv05.

select if intmsl4=1.

sort cases by idcntry idschool.

save outfile = student.

* Now merge the two files.
match files
  / file=student
  / table=school
  / by idcntry idschool.

select if not(missing(ccbgcomm)).

save outfile = merged.

* Define the format for the variables used.
value labels
  idpop      1 "Population 1"  2 "Population 2"
             3 "Population 3"  /
  ccbgcomm   1 "Community Type 1"  2 "Community Type 2"
             3 "Community Type 3"  4 "Community Type 4" /
  idcntry
036 "Australia"      040 "Austria"      056 "Belgium (Fl)"
057 "Belgium (Fr)"  100 "Bulgaria"      124 "Canada"
170 "Colombia"      196 "Cyprus"        200 "Czech Republic"
208 "Denmark"       826 "England"      250 "France"
280 "Germany"       300 "Greece"       344 "Hong Kong "
348 "Hungary"       352 "Iceland"      364 "Iran, Islamic Rep."
372 "Ireland"       376 "Israel"       380 "Italy"
392 "Japan"         410 "Korea"        414 "Kuwait"
428 "Latvia (LSS)"  440 "Lithuania"    528 "Netherlands"
554 "New Zealand"  578 "Norway"       608 "Philippines"
620 "Portugal"     642 "Romania"      643 "Russian Federation"
827 "Scotland"     702 "Singapore"    201 "Slovak Republic"
890 "Slovenia"     717 "South Africa" 724 "Spain"
752 "Sweden"       756 "Switzerland" 764 "Thailand"
840 "United States" .

* Now use the macro to get the results.
include "jackpv.sps".

jackpv  cvar = idcntry idpop ccbgcomm /
        pvs  = cslpv01 to cslpv05 /
        npv  = 5 /
        njkr = 75 /
        jkz  = jkzone /
        jki  = jkindic /
        wgt  = totwgtl4 / .

sort cases by idcntry idpop.

formats n (f5.0) totwgtl4 (f10.2).

report format=list automatic
  / var   = ccbgcomm (label) n totwgtl4 pct pct_se mnpv mnpv_se
  / break = idcntry idpop.

```

**Figure 9.24**  
**Extract of SAS Computer Output for Performing Analyses with School-Level**  
**Variables Using Plausible Values**

*COUNTRY ID*=Australia *POPULATION ID*=Population 3						
CCBGCOMM	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Community Type 2	205	26037.28	16.0316	8.26704	505.577	4.7730
Community Type 3	813	68268.22	42.0339	6.36092	522.296	8.7826
Community Type 4	817	68106.85	41.9345	7.74910	533.811	21.3387
*COUNTRY ID*=Austria *POPULATION ID*=Population 3						
CCBGCOMM	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Community Type 2	191	10398.41	16.1045	3.28655	523.092	22.5036
Community Type 3	403	15758.48	24.4059	4.60452	515.899	13.3877
Community Type 4	1124	38411.40	59.4896	5.19142	525.108	6.3844
*COUNTRY ID*=Canada *POPULATION ID*=Population 3						
CCBGCOMM	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Community Type 1	242	4466.78	2.0900	0.69257	501.905	8.64908
Community Type 2	677	24054.12	11.2550	2.63010	519.462	5.26758
Community Type 3	963	49802.95	23.3030	3.30575	529.005	5.11759
Community Type 4	3017	135395.47	63.3520	3.75508	521.485	4.00409
.						
.						
.						
.						
.						
*COUNTRY ID*=United States *POPULATION ID*=Population 3						
CCBGCOMM	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Community Type 1	122	97059.85	5.1632	2.87786	441.696	19.3944
Community Type 2	1374	488475.82	25.9849	3.80045	471.964	5.1365
Community Type 3	1516	677170.44	36.0227	4.98287	484.969	5.0745
Community Type 4	1737	617137.48	32.8292	4.08738	469.115	7.2467
*COUNTRY ID*=Slovenia *POPULATION ID*=Population 3						
CCBGCOMM	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Community Type 1	35	217.70	2.4663	2.51723	492.974	2.3479
Community Type 3	290	2154.69	24.4100	9.05134	446.015	17.0364
Community Type 4	423	6454.70	73.1237	9.22420	521.665	26.2228

**Figure 9.25****Extract of SPSS Computer Output for Performing Analyses with School-Level Variables Using Plausible Values**

*COUNTRY	ID*	GEN\TYPE OF COMMUNITY	N	TOTWGTL4	PCT	PCT_SE	MNPV	MNPV_SE
Australia								
Population	3	Community Type 2	205	26037.28	16.03	8.27	505.58	4.77
		Community Type 3	813	68268.22	42.03	6.36	522.30	8.78
		Community Type 4	817	68106.85	41.93	7.75	533.81	21.34
Austria								
Population	3	Community Type 2	191	10398.41	16.10	3.29	523.09	22.50
		Community Type 3	403	15758.48	24.41	4.60	515.90	13.39
		Community Type 4	1124	38411.40	59.49	5.19	525.11	6.38
Canada								
Population	3	Community Type 1	242	4466.78	2.09	.69	501.90	8.65
		Community Type 2	677	24054.12	11.26	2.63	519.46	5.27
		Community Type 3	963	49802.95	23.30	3.31	529.00	5.12
		Community Type 4	3017	135395.47	63.35	3.76	521.49	4.00
.								
.								
.								
.								
United States								
Population	3	Community Type 1	122	97059.85	5.16	2.88	441.70	19.39
		Community Type 2	1374	488475.82	25.98	3.80	471.96	5.14
		Community Type 3	1516	677170.44	36.02	4.98	484.97	5.07
		Community Type 4	1737	617137.48	32.83	4.09	469.11	7.25
Slovenia								
Population	3	Community Type 1	35	217.70	2.47	.50	492.97	2.35
		Community Type 3	290	2154.69	24.41	9.05	446.01	17.04
		Community Type 4	423	6454.70	73.12	9.22	521.67	26.22

In summary, to perform analyses that involve the Student and School Background files, the user needs to do the following:

- Identify the variable or variables of interest in the student file and find out about any specific national adaptations to the variable.
- Retrieve the relevant variables from the student files, including the plausible values, sampling weights, JRR replication information and any other variables used in the selection of cases.
- Retrieve the relevant classification variable or variables from the school database.
- Merge the variables from the school database onto the student database using the variables IDCNTRY and IDSCHOOL.
- Use the macro JACKPV with the corresponding arguments and parameters.
- Print out the result file.

## 9.8 Scoring the Items

There were several types of items administered as part of the TIMSS tests. There were multiple-choice items, in which the student was asked to select one of four or five options as the correct response. These were administered as part of the written assessment. The responses to these items are coded with one digit. The codes used to represent the responses to these items are shown in Table 9.4 below.

**Table 9.4**  
**Definitions of Response Codes for the Multiple Choice Items in the Written Assessment Data Files**

Code	Description
1	Chose the first option (a)
2	Chose the second option (b)
3	Chose the third option (c)
4	Chose the fourth option (d)
5	Chose the fifth option (e)
6	Did not reach the item
7	Gave an invalid response (chose more than one of the options available)
8	The item was not administered
9	Did not respond to the item although the item was administered and was reached

There were also open-ended items where the students were asked to construct a response to a question, rather than choosing an answer from a list of options. The answers to these questions were coded by coders trained to use the two-digit scoring rubrics described in Chapter 4. The first digit of the two-digit code indicates the score given to the question, and the second digit in conjunction with the first provides diagnostic information on the type of answer given by the student. The codes used to represent the responses to these items are shown in Table 9.5 below.

**Table 9.5**  
**Definition of Response Codes for the Open-Ended Items in the Written Assessment**

Code	Description
30 to 39	Three points answer. Second digit provides diagnostic information.
20 to 29	Two points answer. Second digit provides diagnostic information.
10 to 19	One point answer. Second digit provides diagnostic information.
70 to 79	Zero point answer. Second digit provides diagnostic information.
90	Gave a response that could not be coded.
96	Did not reach the item.
98	The item was not administered.
99	Did not respond to the item although the item was administered and was reached.



The Written Assessment data files contained in the CD include information about the answer given to each item administered to a student or the code assigned by the coders to the student's response or report of a task. The user might want to work with these item data after they are recoded to the right-wrong format, in the case of multiple-choice items, or to the level of correctness in the case of the open-ended items. To this effect, we have included in the CD a set of macros in SPSS and SAS that will allow the user to recode the items from the written assessment to their right-wrong or correctness-level format. Each of these programs is called CSASCORE. When the user calls and executes these macros all the items in the Written Assessment file are scored. The programs convert the response option codes for multiple-choice items to dichotomous score levels (0 or 1) based on scoring keys. For the open-ended items the two-digit diagnostic codes will be converted to the corresponding correctness score level (3, 2, 1, 0) based on the value of the first digit, as described in Chapter 4. The filenames in SAS and SPSS have been kept constant, except for the last three characters of the file name.

When using these programs, the user must first consider the recoding scheme that is desired. For example, under certain circumstances the user might want to recode the not reached responses as incorrect (codes 6 and 96), whereas under other circumstances the user might want to recode these responses as not administered or invalid. In the case of TIMSS, not reached responses were recoded as not administered (and effectively as missing responses) for the purpose of calibrating the items when setting the international scale, but as incorrect for the purpose of calculating the scale scores for the individuals. By default, the scoring program provided with the database recodes the items coded as not reached and those left blank coded as incorrect responses.

To use these macros the user needs to include them as part of the SAS or SPSS programs used for the analysis. This is done by using the INCLUDE statement in the corresponding program. In the case of SAS, the scoring program code should be included as part of a DATA step that reads the items that are to be recoded. When using SPSS, the scoring program code should be included after the system file containing the item responses has been read into memory and becomes the working file. Both of these programs recode the items without creating a new file so, if the user want to preserve the original answers and codes assigned to the questions, then the file with the recoded item variables needs to be saved under a different file name. A copy of the macro that scores the items in SAS and SPSS and an example of how it is invoked is presented in Figure 9.21 and Figure 9.22.

When using SAS the macro the user needs to include the following command within the DATA step:

```
%include "CSASCORE.SAS";
```

When using SPSS the macro the user needs to include the following the GET FILE command using the include command as in

```
include "CSASCORE.SPS".
```

**Figure 9.26****SAS Control Code Used to Convert Cognitive Item Response Codes to Correctness-Score Levels (CSASCORE.SAS)**

```

%MACRO CSASCORE(ITEM,TYPE,RIGHT,NR,NA,OM,OTHER) ;
  %IF &TYPE = "MC" %THEN %DO;
    SCORE = 0 ;
    IF &ITEM = &RIGHT THEN SCORE = 1 ;
    IF &ITEM = &NR THEN SCORE = 0 ;
    IF &ITEM = &NA THEN SCORE = . ;
    IF &ITEM = &OM THEN SCORE = 0 ;
    IF &ITEM = &OTHER THEN SCORE = 0 ;
    &ITEM = SCORE;
  %END;
  %IF &TYPE = "SA" OR &TYPE = "EX" %THEN %DO;
    SCORE = 0 ;
    IF &ITEM >= 30 AND &ITEM < 40 THEN SCORE = 3 ;
    IF &ITEM >= 20 AND &ITEM < 30 THEN SCORE = 2 ;
    IF &ITEM >= 10 AND &ITEM < 20 THEN SCORE = 1 ;
    IF &ITEM >= 70 AND &ITEM < 80 THEN SCORE = 0 ;
    IF &ITEM = &NR THEN SCORE = 0 ;
    IF &ITEM = &NA THEN SCORE = . ;
    IF &ITEM = &OM THEN SCORE = 0 ;
    IF &ITEM = &OTHER THEN SCORE = 0 ;
    &ITEM = SCORE;
  %END;
%MEND CSASCORE;

%LET ARIGHT =
CSMGB06 CSMGB08 CSMGB09 CSMGB16 CSMGB19 CSMGC04 CSMGC10 CSMGC16
CSMGD07 CSMGD09 CSMGD13 CSMPE01 CSMPE09 CSMPE10 CSMPE11 CSMPE12
CSMPG05 CSMPG09 CSMPH03 CSMPH04 CSMMI02 CSMMI05 CSMMI07 CSMMJ01
CSMMJ11 CSMMK02 CSMMK05 CSMLL03 CSMLL09 CSMLL10 CSMLL12;

%LET BRIGHT =
CSMGA01 CSMGA03 CSMGB05 CSMGB11 CSMGB18 CSMGB20 CSMGB21 CSMGC02
CSMGC05 CSMGC07 CSMGC14 CSMGC15 CSMGD06 CSMGD08 CSMPE02 CSMPE07
CSMPF01 CSMPF04 CSMPF09 CSMPF10 CSMPG07 CSMPG10 CSMPH05 CSMPH07
CSMPH09 CSMMI04 CSMMJ08 CSMMJ09 CSMMJ12 CSMMJ14 CSMMK04 CSMMK07
CSMLL01 CSMLL05 CSMLL06;

%LET CRIGHT =
CSMGA04 CSMGA05 CSMGB02 CSMGB04 CSMGB07 CSMGB15 CSMGB23 CSMGB24
CSMGC03 CSMGC11 CSMGC17 CSMGD10 CSMGD14 CSMPPE03 CSMPPE05 CSMPF02
CSMPF03 CSMPF05 CSMPF08 CSMPG01 CSMPG04 CSMPH01 CSMPH06 CSMPH10
CSMMI09 CSMMI10 CSMMJ02 CSMMJ04 CSMMJ05 CSMMJ10 CSMMK03 CSMMK09
CSMLL10 CSMLL04 CSMLL07 CSMLL11;

%LET DRIGHT =
CSMGA02 CSMGB01 CSMGB10 CSMGB14 CSMGB17 CSMGB22 CSMGC06 CSMGC09
CSMGD01 CSMGD12 CSMPPE04 CSMPPE06 CSMPPE08 CSMPF06 CSMPF07 CSMPG03
CSMPG06 CSMPH02 CSMPH08 CSMPH11 CSMMI01 CSMMI06 CSMMI08 CSMMJ06
CSMMJ13 CSMMK06 CSMMK08 CSMMK11 CSMLL02;

%LET ERIGHT =
CSMGB03 CSMGC01 CSMGC08 CSMGD11 CSMPG08 CSMMI03 CSMMJ03 CSMMJ07
CSMMK01 CSMLL08;

%LET SHORTA =
CSSGA06A CSSGA06B CSSGA07 CSSGB12 CSSGB13 CSSGB25 CSSGB26 CSSGC12
CSSGC13 CSSGC18 CSSGC19 CSSGC20 CSSGD02 CSSGD03 CSSGD15A CSSGD15B
CSSGD16A CSSGD16B CSSGD17 CSSPF12 CSSPF13 CSSPF14 CSSPG11 CSSPG13
CSSPG14 CSSPG15 CSSPG17 CSSPH12 CSSPH13 CSSPH14 CSSPH15 CSSPH16
CSSPH17 CSSMJ15A CSSMJ15B CSSMJ16A CSSMJ16B CSSMJ17 CSSMJ19 CSSMK12
CSSMK13 CSSMK15 CSSMK17 CSSML13 CSSML14 CSSML15A CSSML15B CSSML18;
%LET EXTEND =
CSEGA08 CSEGA09A CSEGA09B CSEGA10 CSEGA11A CSEGA11B CSEGA11C CSEGA12
CSEGC21 CSEGD04 CSEGD05A CSEGD05B CSEPF15 CSEPF16 CSEPF17A CSEPF17B
CSEPG12 CSEPG16 CSEPG18 CSEPG19 CSEPH18 CSEPH19A CSEPH19B CSEMJ18
CSEMK14 CSEMK16 CSEMK18 CSEML16 CSEML17;

```

**Figure 9.26 (Continued)****SAS Control Code Used to Convert Cognitive Item Response Codes to Correctness-Score Levels (CSAScore.SAS)**

```
ARRAY ARIGHT &ARIGHT;
ARRAY ARIGHT &ARIGHT;
ARRAY ARIGHT &ARIGHT;
ARRAY ARIGHT &ARIGHT;
ARRAY ARIGHT &ARIGHT;
ARRAY SHORTA &SHORTA;
ARRAY EXTEND &EXTEND;

DO OVER ARIGHT; %CSAScore(ARIGHT,"MC",1, 6, 8, 9, 7); END;
DO OVER BRIGHT; %CSAScore(BRIGHT,"MC",2, 6, 8, 9, 7); END;
DO OVER CRIGHT; %CSAScore(CRIGHT,"MC",3, 6, 8, 9, 7); END;
DO OVER DRIGHT; %CSAScore(DRIGHT,"MC",4, 6, 8, 9, 7); END;
DO OVER ERIGHT; %CSAScore(ERIGHT,"MC",5, 6, 8, 9, 7); END;
DO OVER SHORTA; %CSAScore(SHORTA,"EX", ,96,98,99,90); END;
DO OVER EXTEND; %CSAScore(EXTEND,"EX", ,96,98,99,90); END;
```

**Figure 9.27****SPSS Control Code Used to Convert Cognitive Item Response Codes to Correctness-Score Levels (CSASCORE.SPS)**

```

SET MPRINT=ON.
DEFINE CSASCORE (Type = !charend('/') / Item = !charend('/') /
RIGHT = !charend('/') / nr = !charend('/') /
na = !charend('/') / om = !charend('/') /
other = !charend('/') ).

!If (!UPCASE(!Type)=MC) !Then
!Do !I !In(!Item).
Recode !I (!RIGHT =1)
(!nr =0)
(!na =sysmis)
(!om =0)
(!other =0)
(Else =0).
!DoEnd.
!IfEnd.

!If ((!UPCASE(!Type)=SA) !OR (!UPCASE(!Type)=EX)) !Then
!Do !I !In(!Item).
Recode !I (10 thru 19=1)
(20 thru 29=2)
(30 thru 39=3)
(70 thru 79=0)
(!nr = 0)
(!na = sysmis)
(!om = 0)
(!other = 0)
(Else = 0).
!DoEnd.
!IfEnd.
!enddefine.

CSASCORE Type = MC /
Item = CSMGB06 CSMGB08 CSMGB09 CSMGB16 CSMGB19 CSMGC04 CSMGC10 CSMGC16
CSMGD07 CSMGD09 CSMGD13 CSMPE01 CSMPE09 CSMPE10 CSMPP11 CSMPG02
CSMPG05 CSMPG09 CSMPH03 CSMPH04 CSMMI02 CSMMI05 CSMMI07 CSMMJ01
CSMMJ11 CSMMK02 CSMMK05 CSMLL03 CSMLL09 CSMLL10 CSMLL12/
RIGHT = 1 /
nr = 6 /
na = 8 /
om = 9 /
other = 7 .

CSASCORE Type = MC /
Item = CSMGA01 CSMGA03 CSMGB05 CSMGB11 CSMGB18 CSMGB20 CSMGB21 CSMGC02
CSMGC05 CSMGC07 CSMGC14 CSMGC15 CSMGD06 CSMGD08 CSMPE02 CSMPE07
CSMPF01 CSMPF04 CSMPF09 CSMPF10 CSMPG07 CSMPG10 CSMPH05 CSMPH07
CSMPH09 CSMMI04 CSMMJ08 CSMMJ09 CSMMJ12 CSMPJ14 CSMMK04 CSMMK07
CSMML01 CSMML05 CSMML06 /
RIGHT = 2 /
nr = 6 /
na = 8 /
om = 9 /
other = 7 .

CSASCORE Type = MC /
Item = CSMGA04 CSMGA05 CSMGB02 CSMGB04 CSMGB07 CSMGB15 CSMGB23 CSMGB24
CSMGC03 CSMGC11 CSMGC17 CSMGD10 CSMGD14 CSMPE03 CSMPE05 CSMPF02
CSMPF03 CSMPF05 CSMPF08 CSMPG01 CSMPG04 CSMPH01 CSMPH06 CSMPH10
CSMMI09 CSMMI10 CSMMJ02 CSMMJ04 CSMMJ05 CSMMJ10 CSMMK03 CSMMK09
CSMMK10 CSMLL04 CSMLL07 CSMLL11 /
RIGHT = 3 /
nr = 6 /
na = 8 /
om = 9 /
other = 7 .

```





---

## References Cited in Text

- Adams, R.J. and Gonzalez, E.J. (1996). TIMSS test design. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume I: Design and development*. Chestnut Hill, MA: Boston College.
- Adams, R.J., Wilson, M.R., and Wang, W.C. (1997). The multidimensional random coefficients multinomial logit. *Applied Psychological Measurement*.
- Adams, R.J., Wilson, M.R., and Wu, M.L. (1997). Multilevel item responses models: An approach to errors in variables regression. *Journal of Educational and Behavioral Statistics*, 22 (1), 46-75.
- Beaton, A.E. (1987). *Implementing the new design: The NAEP 1983-84 technical report*. Report No. 15-TR-20. Princeton, NJ: Educational Testing Service.
- Beaton, A.E., Martin, M.O., Mullis, I.V.S., Gonzalez, E.J., Smith, T.A., and Kelly, D.L. (1996a). *Science achievement in the middle school years: IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: Boston College.
- Beaton, A.E., Mullis, I.V.S., Martin, M.O. Gonzalez, E.J., Kelly, D.L., and Smith, T.A. (1996b). *Mathematics achievement in the middle school years: IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: Boston College.
- Bock, R.D., and Aitken, M. (1981). Marginal maximum likelihood estimation of item parameters: An application of the EM algorithm. *Psychometrika*, 46, 443-459.
- Dempster, A.P., Laird, N.M. and Rubin, D.B. (1977). Maximum likelihood from incomplete data via the EM algorithm. *Journal of the Royal Statistical Society, Series B*, 39, 1-38.
- Dumais, J. (1998). Implementation of the TIMSS sample design. In M.O. Martin and D.L. Kelly (Eds.) *TIMSS technical report, volume III: Implementation and analysis – final year of secondary school*. Chestnut Hill, MA: Boston College.
- Dumais, J and Foy, P. (1998). Calculation of sampling weights. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume III: Implementation and analysis – final year of secondary school*. Chestnut Hill, MA: Boston College.
- Foy, P., Rust, K., and Schleicher, A. (1996). Sample design. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume I: Design and development*. Chestnut Hill, MA: Boston College.
- Gonzalez, E.J. (1998). Reporting student achievement in mathematics and science. In M.O. Martin and D.L. Kelly (Eds.) *TIMSS technical report, volume III: Implementation and analysis – final year of secondary school*. Chestnut Hill, MA: Boston College.

- Gonzalez, E.J. and Foy, P. (1998). Estimation of sampling variability, design effects, and effective sample sizes. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume III: Implementation and analysis – final year of secondary school*. Chestnut Hill, MA: Boston College.
- Gonzalez, E.J. and Smith, T.A. (Eds.). (1997). *User guide for the TIMSS international database – primary and middle school years*. Chestnut Hill, MA: Boston College.
- Kelderman, H. and Rijkes, C.P.M. (1994). Loglinear multidimensional IRT models for polytomously scored items. *Psychometrika*, 59, 149-176.
- Lie, S., Taylor, A., and Harmon, M. (1996) Scoring techniques and criteria. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume I: Design and development*. Chestnut Hill, MA: Boston College.
- Martin, M.O. and Kelly, D.L., Eds. (1996). *TIMSS technical report, volume I: Design and development*. Chestnut Hill, MA: Boston College.
- Martin, M.O. and Kelly, D.L., Eds. (1997). *TIMSS technical report, volume II: Implementation and analysis – primary and middle school years*. Chestnut Hill, MA: Boston College.
- Martin, M.O. and Kelly, D.L., Eds. (1998). *TIMSS technical report, volume III: Implementation and analysis – final year of secondary school*. Chestnut Hill, MA: Boston College.
- Martin, M.O. and Mullis, I.V.S., Eds. (1996). *TIMSS: Quality assurance in data collection*. Chestnut Hill, MA: Boston College.
- Martin, M.O., Mullis, I.V.S., Beaton, A.E., Gonzalez, E.J., Smith, T.A., and Kelly, D.L. (1997). *Science achievement in the primary school years: IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: Boston College.
- Maxwell, B. (1996). Translation and cultural adaptation of the survey instruments. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume I: Design and development*. Chestnut Hill, MA: Boston College.
- Mislevy, R.J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56, 177-196.
- Mislevy, R.J., Beaton, A.E., Kaplan, B., and Sheehan, K.M. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, 29 (2), 133-161.
- Mislevy, R.J. and Sheehan, K.M. (1987). Marginal estimation procedures. In A.E. Beaton (Ed.), *Implementing the new design: The NAEP 1983-84 technical report*. Report No. 15-TR-20. Princeton, NJ: Educational Testing Service.
- Mislevy, R.J. and Sheehan, K.M. (1989). The role of collateral information about examinees in item parameter estimation. *Psychometrika*, 54(4), 661-679.



- 
- Mullis, I.V.S. and Smith, T.S. (1996). Quality control steps for free-response scoring. In M.O. Martin and I.V.S. Mullis (Eds.), *TIMSS: Quality assurance in data collection*. Chestnut Hill, MA: Boston College.
- Mullis, I.V.S., Jones, C.O., and Garden, R.A. (1996). Training sessions for free-response scoring and administration of the performance assessment. In M.O. Martin and I.V.S. Mullis (Eds.), *TIMSS: Quality assurance in data collection*. Chestnut Hill, MA: Boston College.
- Mullis, I.V.S., Martin, M.O., Beaton, A.E., Gonzalez, E.J., Kelly, D.L., and Smith, T.A. (1997). *Mathematics achievement in the primary school years IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: Boston College.
- Mullis, I.V.S., Martin, M.O., Beaton, A.E., Gonzalez, E.J., Kelly, D.L., and Smith, T.A. (1998). *Mathematics and science achievement in the final year of secondary school: IEA's Third International Mathematics and Science Study*. Chestnut Hill, MA: Boston College.
- Mullis, I.V.S. and Martin, M.O. (1998). Item analysis and review. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume III: Implementation and analysis-final year of secondary school*. Chestnut Hill, MA: Boston College.
- Robitaille, D.F., Ed. (1997). *National contexts for mathematics and science education: An encyclopedia of education systems participating in TIMSS*. Vancouver, B.C.: Pacific Educational Press.
- Robitaille, D.F., McKnight, C.C., Schmidt, W.H., Britton, E.D., Raizen, S.A., and Nicol, C. (1993). *TIMSS monograph no. 1 Curriculum frameworks for mathematics and science*. Vancouver, B.C.: Pacific Educational Press.
- Rubin, D.B. (1987). *Multiple imputation for non-response in surveys*. New York: John Wiley & Sons.
- Schmidt, W.H., McKnight, C.C., Valverde, G.A., Houang, R.T., and Wiley, D.E. (1997a). *Many visions, many aims: A cross-national investigation of curricular intentions in school mathematics*. Dordrecht, the Netherlands: Kluwer Academic Publishers.
- Schmidt, W.H., Raizen, S.A., Britton, E.D., Bianchi, L.J., and Wolfe, R.G. (1997b). *Many visions, many aims: A cross-national investigation of curricular intentions in school science*. Dordrecht, the Netherlands: Kluwer Academic Publishers.
- TIMSS (1995). *Guide to checking, coding, and entering the TIMSS data*. Chestnut Hill, MA: Boston College.
- TIMSS (1998). *TIMSS mathematics and science items – Released set for population 3 (final year of secondary school)*. Chestnut Hill, MA: Boston College.

- Wilson, M.R. (1992). The ordered partition model: An extension of the partial credit model. *Applied Psychological Measurement*, 16 (3).
- Wolter, K.M. (1985). *Introduction to variance estimation*. New York: Springer-Verlag.
- Wu, M.L., Adams, R.J., and Wilson, M. (1997). *Conquest: Generalized item response modelling software – Manual*. Melbourne: Australian Council for Educational Research.

---

# Acknowledgments

TIMSS was truly a collaborative effort among hundreds of individuals around the world. Staff from the national research centers, the international management, advisors, and funding agencies worked closely to design and implement the most ambitious study of international comparative achievement ever undertaken. TIMSS would not have been possible without the tireless efforts of all involved. Below, the individuals and organizations are acknowledged for their contributions. Given that implementing TIMSS has spanned more than seven years and involved so many people and organizations, this list may not pay heed to all who contributed throughout the life of the project. Any omission is inadvertent. TIMSS also acknowledges the students, teachers, and school principals who contributed their time and effort to the study. This report would not be possible without them.

## MANAGEMENT AND OPERATIONS

Since 1993, TIMSS has been directed by the International Study Center at Boston College in the United States. Prior to this, the study was coordinated by the International Coordinating Center at the University of British Columbia in Canada. Although the study was directed centrally by the International Study Center and its staff members implemented various parts of TIMSS, important activities also were carried out in centers around the world. The data were processed centrally by the IEA Data Processing Center in Hamburg, Germany. Statistics Canada was responsible for collecting and evaluating the sampling documentation from each country and for calculating the sampling weights. The Australian Council for Educational Research conducted the scaling of the achievement data.

### International Study Center (1993-)

Albert E. Beaton, International Study Director  
Michael O. Martin, Deputy International Study Director  
Ina V.S. Mullis, Co-Deputy International Study Director  
Eugenio J. Gonzalez, Director of Operations and Data Analysis  
Dana L. Kelly, Research Associate  
Teresa A. Smith, Research Associate  
Cheryl L. Flaherty, Research Associate  
Maryellen Harmon, Performance Assessment Coordinator  
Kelvin Gregory, Research Associate  
Robert Jin, Computer Programmer  
Ce Shen, Computer Programmer  
William J. Crowley, Fiscal Administrator  
Thomas M. Hoffmann, Publications Coordinator  
José Rafael Nieto, Senior Production Specialist  
Ann G.A. Tan, Conference Coordinator  
Mary C. Howard, Office Supervisor

### **International Study Center (continued)**

Diane Joyce, Secretary  
Joanne E. McCourt, Secretary  
Jennifer Boynton, Secretary  
Craig D. Hoyle, Graduate Assistant

### **International Coordinating Center (1991-93)**

David F. Robitaille, International Coordinator  
Robert A. Garden, Deputy International Coordinator  
Barry Anderson, Director of Operations  
Beverley Maxwell, Director of Data Management

### **Statistics Canada**

Pierre Foy, Senior Methodologist  
Suzelle Giroux, Senior Methodologist  
Jean Dumais, Senior Methodologist  
Nancy Darcovich, Senior Methodologist  
Marc Joncas, Senior Methodologist  
Laurie Reedman, Junior Methodologist  
Claudio Perez, Junior Methodologist

### **IEA Data Processing Center**

Michael Bruneforth, Senior Researcher (former)  
Jedidiah Harris, Research Assistant  
Dirk Hastedt, Senior Researcher  
Svenja Moeller, Research Assistant  
Knut Schwippert, Senior Researcher  
Heiko Sibberns, Senior Researcher  
Jockel Wolff, Research Assistant

### **Australian Council for Educational Research**

Raymond J. Adams, Principal Research Fellow  
Margaret Wu, Research Fellow  
Nikolai Volodin, Research Fellow  
David Roberts, Research Officer  
Greg Macaskill, Research Officer

## **IEA SECRETARIAT**

Tjeerd Plomp, Chairperson  
Hans Wagemaker, Executive Director  
Barbara Malak-Minkiewicz, Manager Membership Relations  
Leendert Dijkhuizen, Financial Officer  
Karin Baddane, Secretary

## **FUNDING AGENCIES**

Funding for the International Study Center was provided by the National Center for Education Statistics of the U.S. Department of Education, the U.S. National Science Foundation, and the International Association for the Evaluation for Educational Achievement. Eugene Owen and Lois Peak of the National Center for Education Statistics and Larry Suter of the National Science Foundation each played a crucial role in making TIMSS possible and for ensuring the quality of the study. Funding for the International Coordinating Center was provided by the Applied Research Branch of the Strategic Policy Group of the Canadian Ministry of Human Resources Development. This initial source of funding was vital in initiating the TIMSS project. Tjeerd Plomp, Chair of the IEA and of the TIMSS Steering Committee, has been a constant source of support throughout TIMSS. It should be noted that each country provided its own funding for the implementation of the study at the national level.

## **NATIONAL RESEARCH COORDINATORS**

The TIMSS National Research Coordinators and their staff had the enormous task of implementing the TIMSS design in their countries. This required obtaining funding for the project; participating in the development of the instruments and procedures; conducting field tests; participating in and conducting training sessions; translating the instruments and procedural manuals into the local language; selecting the sample of schools and students; working with the schools to arrange for the testing; arranging for data collection, coding, and data entry; preparing the data files for submission to the IEA Data Processing Center; contributing to the development of the international reports; and preparing national reports. The way in which the national centers operated and the resources that were available varied considerably across the TIMSS countries. In some countries, the tasks were conducted centrally, while in others, various components were subcontracted to other organizations. In some countries, resources were more than adequate, while in others, the national centers were operating with limited resources. Of course, across the life of the project, some NRCs have changed. This list attempts to include all past NRCs who served for a significant period of time as well as all the present NRCs. All of the TIMSS National Research Coordinators and their staff members are to be commended for their professionalism and their dedication in conducting all aspects of TIMSS.

**Argentina**

Carlos Mansilla  
Universidad del Chaco  
Av. Italia 350  
3500 Resistencia  
Chaco, Argentina

**Australia**

Jan Lokan  
Raymond Adams \*  
Australian Council for Educational Research  
19 Prospect Hill  
Private Bag 55  
Camberwell, Victoria 3124  
Australia

**Austria**

Guenter Haider  
Austrian IEA Research Centre  
Universität Salzburg  
Akademiestraße 26/2  
A-5020 Salzburg, Austria

**Belgium (Flemish)**

Christiane Brusselmans-Dehairs  
Rijksuniversiteit Ghent  
Vakgroep Onderwijskunde &  
The Ministry of Education  
Henri Dunantlaan 2  
B-9000 Ghent, Belgium

**Belgium (French)**

Georges Henry  
Christian Monseur  
Universite de Liège  
B32 Sart-Tilman  
4000 Liège 1, Belgium

**Bulgaria**

Kiril Bankov  
Foundation for Research, Communication,  
Education and Informatics  
Tzarigradsko Shausse 125, Bl. 5  
1113 Sofia, Bulgaria

**Canada**

Alan Taylor  
Applied Research & Evaluation Services  
University of British Columbia  
2125 Main Mall  
Vancouver, B.C. V6T 1Z4  
Canada

**Colombia**

Carlos Jairo Diaz  
Universidad del Valle  
Facultad de Ciencias  
Multitaller de Materiales Didacticos  
Ciudad Universitaria Meléndez  
Apartado Aereo 25360  
Cali, Colombia

**Cyprus**

Constantinos Papanastasiou  
Department of Education  
University of Cyprus  
Kallipoleos 75  
P.O. Box 537  
Nicosia 133, Cyprus

**Czech Republic**

Jana Strakova  
Vladislav Tomasek  
Institute for Information on Education  
Senovazne Nam. 26  
111 21 Praha 1, Czech Republic

**Denmark**

Peter Weng  
Peter Allerup  
Borge Prien\*  
The Danish National Institute for  
Educational Research  
28 Hermodsgade  
Dk-2200 Copenhagen N, Denmark

**England**

Wendy Keys  
Derek Foxman\*  
National Foundation for  
Educational Research  
The Mere, Upton Park  
Slough, Berkshire SL1 2DQ  
England

**France**

Anne Servant  
Ministère de l'Éducation  
Nationale 142, rue du Bac  
75007 Paris, France

Josette Le Coq\*  
Centre International d'Études  
Pédagogiques (CIEP)  
1 Avenue Léon Journault  
93211 Sèvres, France

\* Past National Research Coordinator

**Germany**

Rainer Lehmann  
Humboldt-Universitaet zu Berlin  
Institut fuer Allgemeine  
Erziehungswissenschaft  
Geschwister-Scholl-Str. 6  
10099 Berlin, Germany

Juergen Baumert  
Wilfried Bos  
Rainer Waterman  
Max-Planck Institute for Human  
Development and Education  
Lentzeallee 94  
14191 Berlin, Germany

Manfred Lehrke  
Universität Kiel  
IPN Olshausen Str. 62  
24098 Kiel, Germany

**Greece**

Georgia Kontogiannopoulou-Polydorides  
Department of Education (Nipiagogon)  
University of Athens  
Navarinou 13A, Neochimio  
Athens 10680, Greece

Joseph Solomon  
Department of Education  
University of Patras  
Patras 26500, Greece

**Hong Kong**

Frederick Leung  
Nancy Law  
The University of Hong Kong  
Department of Curriculum Studies Pokfulam  
Road, Hong Kong

**Hungary**

Péter Vari  
National Institute of Public Education  
Centre for Evaluation Studies  
Dorottya U. 8, P.O. Box 120  
1051 Budapest, Hungary

**Iceland**

Einar Gudmundsson  
Institute for Educational Research  
Department of Educational Testing  
and Measurement  
Surdgata 39  
101 Reykjavik, Iceland

**Indonesia**

Jahja Umar  
Ministry of Education and Culture  
Examination Development Center  
Jalan Gunung Sahari - 4  
Jakarta 10000, Indonesia

**Ireland**

Deirdre Stuart  
Michael Martin\*  
Educational Research Centre  
St. Patrick's College  
Drumcondra  
Dublin 9, Ireland

**Iran, Islamic Republic**

Ali Reza Kiamanesh  
Ministry of Education  
Center for Educational Research  
Iranshahr Shomali Avenue  
Teheran 15875, Iran

**Israel**

Pinchas Tamir  
The Hebrew University  
Israel Science Teaching Center  
Jerusalem 91904, Israel

Ruth Zuzovsky  
Tel Aviv University  
School of Education  
Ramat Aviv  
PO Box 39040  
Tel Aviv 69978, Israel

**Italy**

Anna Maria Caputo  
Ministero della Pubblica Istruzione  
Centro Europeo dell'Educazione  
Villa Falconieri  
00044 Frascati, Italy

**Japan**

Masao Miyake  
Eizo Nagasaki  
National Institute for Educational Research  
6-5-22 Shimomeguro  
Meguro-Ku, Tokyo 153, Japan

**Korea**

Jingyu Kim  
Hyung Im\*  
National Board of Educational Evaluation  
Evaluation Research Division  
Chungdam-2 Dong 15-1, Kangnam-Ku  
Seoul 135-102, Korea

**Kuwait**

Mansour Hussein  
Ministry of Education  
P. O. Box 7  
Safat 13001, Kuwait

**Latvia**

Andrejs Geske  
University of Latvia  
Faculty of Education & Psychology  
Jurmālas Gatve 74/76, Rm. 204a  
Riga, LV-1083, Latvia

**Lithuania**

Algirdas Zabulionis  
University of Vilnius  
Faculty of Mathematics  
Naugarduko 24  
2006 Vilnius, Lithuania

**Mexico**

Fernando Córdova Calderón  
Director de Evaluación de Políticas y  
Sistemas Educativos  
Netzahualcoyotl #127 2do Piso  
Colonia Centro  
Mexico 1, D.F., Mexico

**Netherlands**

Wilma Kuiper  
Klaas Bos  
University of Twente  
Faculty of Educational Science  
and Technology  
Department of Curriculum  
P.O. Box 217  
7500 AE Enschede, Netherlands

**New Zealand**

Megan Chamberlain  
Hans Wagemaker\*  
Steve May  
Ministry of Education  
Research and International Section  
P.O. Box 1666  
45-47 Pipitea Street  
Wellington, New Zealand

**Norway**

Svein Lie  
University of Oslo  
SLS Postboks 1099  
Blindern 0316  
Oslo 3, Norway

Gard Brekke  
Alf Andersensv 13  
3670 Notodden, Norway

**Philippines**

Milagros Ibe  
University of the Philippines  
Institute for Science and Mathematics  
Education Development  
Diliman, Quezon City  
Philippines

Ester Ogena  
Science Education Institute  
Department of Science and Technology  
Bicutan, Taguig  
Metro Manila 1604, Philippines

**Portugal**

Gertrudes Amaro  
Ministerio da Educacao  
Instituto de Inovação Educacional  
Rua Artilharia Um 105  
1070 Lisboa, Portugal

**Romania**

Gabriela Noveanu  
Institute for Educational Sciences  
Evaluation and Forecasting Division  
Str. Stirbei Voda 37  
70732-Bucharest, Romania



**Russian Federation**

Galina Kovalyova  
The Russian Academy of Education  
Institute of General Secondary School  
Ul. Pogodinskaya 8  
Moscow 119905, Russian Federation

**Scotland**

Brian Semple  
Scottish Office, Education &  
Industry Department  
Victoria Quay  
Edinburgh, E86 6QQ  
Scotland

**Singapore**

Wong Cheow Cher  
Chan Siew Eng\*  
Research and Evaluation Branch  
Block A Belvedere Building  
Ministry of Education  
Kay Siang Road  
Singapore 248922

**Slovak Republic**

Maria Berova  
Vladimir Burjan\*  
SPU-National Institute for Education  
Pluhova 8  
P.O. Box 26  
830 00 Bratislava  
Slovak Republic

**Slovenia**

Marjan Setinc  
Barbara Japelj  
Pedagoski Institut Pri Univerzi v Ljubljana  
Gerbiceva 62, P.O. Box 76  
61111 Ljubljana, Slovenia

**South Africa**

Sarah Howie  
Derek Gray\*  
Human Sciences Research Council  
134 Pretorius Street  
Private Bag X41  
Pretoria 0001, South Africa

**Spain**

José Antonio Lopez Varona  
Instituto Nacional de Calidad y Evaluación  
C/San Fernando del Jarama No. 14  
28071 Madrid, Spain

**Sweden**

Ingemar Wedman  
Anna Hofslagare  
Kjell Gisselberg\*  
Umeå University  
Department of Educational Measurement  
S-901 87 Umeå, Sweden

**Switzerland**

Erich Ramseier  
Amt Für Bildungsforschung der  
Erziehungsdirektion des Kantons Bern  
Sulgeneck Straße 70  
Ch-3005 Bern, Switzerland

**Thailand**

Suwaporn Semheng  
Institute for the Promotion of Teaching  
Science and Technology  
924 Sukhumvit Road  
Bangkok 10110, Thailand

**United States**

William Schmidt  
Michigan State University  
Department of Educational Psychology  
463 Erikson Hall  
East Lansing, MI 48824-1034  
United States

## **TIMSS ADVISORY COMMITTEES**

The TIMSS International Study Center was supported in its work by several advisory committees. The TIMSS International Steering Committee provided guidance to the International Study Director on policy issues and general direction of the study. The TIMSS Technical Advisory Committee provided guidance on issues related to design, sampling, instrument construction, analysis, and reporting, ensuring that the TIMSS methodologies and procedures were technically sound. The Subject Matter Advisory Committee ensured that current thinking in mathematics and science education were addressed by TIMSS, and was instrumental in the development of the TIMSS tests. The Free-Response Item Coding Committee developed the coding rubrics for the free-response items. The Performance Assessment Committee worked with the Performance Assessment Coordinator to develop the TIMSS performance assessment. The Quality Assurance Committee helped to develop the quality assurance program.

### **International Steering Committee**

Tjeerd Plomp (Chair), the Netherlands  
Lars Ingelstam, Sweden  
Daniel Levine, United States  
Senta Raizen, United States  
David Robitaille, Canada  
Toshio Sawada, Japan  
Benny Suprpto Brotosiswojo, Indonesia  
William Schmidt, United States

### **Technical Advisory Committee**

Raymond Adams, Australia  
Pierre Foy, Canada  
Andreas Schleicher, Germany  
William Schmidt, United States  
Trevor Williams, United States

### **Sampling Referee**

Keith Rust, United States

### **Subject Area Coordinators**

Robert Garden, New Zealand (Mathematics)  
Graham Orpwood, Canada (Science)

### **Special Mathematics Consultant**

Chancey Jones

### **Subject Matter Advisory Committee**

Svein Lie (Chair), Norway  
Antoine Bodin, France  
Peter Fensham, Australia  
Robert Garden, New Zealand  
Geoffrey Howson, England  
Curtis McKnight, United States  
Graham Orpwood, Canada  
Senta Raizen, United States  
David Robitaille, Canada  
Pinchas Tamir, Israel  
Alan Taylor, Canada  
Ken Travers, United States  
Theo Wubbels, the Netherlands

### **Free-Response Item Coding Committee**

Svein Lie (Chair), Norway  
Vladimir Burjan, Slovak Republic  
Kjell Gisselberg, Sweden  
Galina Kovalyova, Russian Federation  
Nancy Law, Hong Kong  
Josette Le Coq, France  
Jan Lokan, Australia  
Curtis McKnight, United States  
Graham Orpwood, Canada  
Senta Raizen, United States  
Alan Taylor, Canada  
Peter Weng, Denmark  
Algirdas Zabulionis, Lithuania

### **Mathematics and Science Literacy Item Task Force**

Henk Schuring, The Netherlands  
John Lindsey, Australia  
Curtis McKnight, United States  
Graham Orpwood, Canada  
Ken Travers, United States  
A.I. Weinzwieg, Germany

## **Performance Assessment Committee**

Derek Foxman, England  
Robert Garden, New Zealand  
Per Morten Kind, Norway  
Svein Lie, Norway  
Jan Lokan, Australia  
Graham Orpwood, Canada

## **Quality Control Committee**

Jules Goodison, United States  
Hans Pelgrum, The Netherlands  
Ken Ross, Australia

## **Editorial Committee**

David F. Robitaille (Chair), Canada  
Albert Beaton, International Study Director  
Paul Black, England  
Svein Lie, Norway  
Rev. Ben Nebres, Philippines  
Judith Torney-Purta, United States  
Ken Travers, United States  
Theo Wubbels, the Netherlands