

International Association for the Evaluation of Educational Achievement

User Guide for the TIMSS International Database
Primary and Middle School Years

(Population 1 and Population 2)
Data Collected in 1995

Edited by

Eugenio J. Gonzalez
Teresa A. Smith

with contributions by

Heiko Jungclaus
Dirk Hastedt
Dana L. Kelly
Ina V.S. Mullis
Michael O. Martin
Knut Schwippert
Jens Brockmann
Ray Adams
Pierre Foy
Ce Shen

September 1997



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

© 1997 International Association for the Evaluation of Educational Achievement (IEA).

User Guide for the TIMSS International Database – Primary and Middle School Years
1995 Assessment / Edited by Eugenio J. Gonzalez and Teresa A. Smith

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

For more information about TIMSS contact the TIMSS International Study Center

TIMSS International Study Center
Campion Hall 323- CSTEPP
School of Education
Boston College
Chestnut Hill, MA 02167
United States

email: timss@bc.edu
<http://www.cstepp.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

Chapter 1: Overview of TIMSS and the Database User Guide.....	1-1
1.1 Overview of the International Database.....	1-1
1.2 Overview of TIMSS.....	1-3
1.3 TIMSS International Reports.....	1-5
1.4 Contents of the Database.....	1-5
1.5 Contents of the User Guide.....	1-6
1.6 Management and Operations of TIMSS.....	1-8
1.7 Additional Resources.....	1-8
Chapter 2: TIMSS Instruments and Booklet Design.....	2-1
2.1 Introduction.....	2-1
2.2 The TIMSS Mathematics and Science Content.....	2-1
2.3 The TIMSS Items.....	2-3
2.4 Organization of the Test Booklets.....	2-4
2.5 Performance Assessment.....	2-8
2.6 Release Status for TIMSS Test Items, Performance Tasks, and Background Questionnaires.....	2-10
2.7 Questionnaires.....	2-10
Chapter 3: Sampling and Sampling Weights.....	3-1
3.1 The Target Populations.....	3-1
3.2 School Sample Selection.....	3-5
3.3 Classroom and Student Sampling.....	3-6
3.4 Performance Assessment Subsampling.....	3-6
3.5 Response Rates.....	3-6
3.5.1 School-Level Response Rates.....	3-7
3.5.2 Student-Level Response Rates.....	3-7
3.5.3 Overall Response Rates.....	3-8
3.6 Compliance with Sampling Guidelines.....	3-8
3.7 Sampling Weights.....	3-11
3.8 Weight Variables Included in the Student Data Files.....	3-14
3.9 Weight Variables Included in the Student-Teacher Linkage Files.....	3-16
3.10 Weight Variables Included in the School Data Files.....	3-17
Chapter 4: Data Collection, Materials Processing, Scoring, and Database Creation.....	4-1
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.....	4-10

Chapter 5: TIMSS Scaling Procedures.....	5-1
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
5.8.1 Drawing The International Calibration Sample	5-7
5.8.2 Standardizing the International Scale Scores.....	5-7
Chapter 6: Student Achievement Scores.....	6-1
6.1 Achievement Scores in the Student Files	6-1
6.2 Achievement Scores in the School Background Files	6-9
Chapter 7: Content and Format of Database Files	7-1
7.1 Introduction	7-1
7.2 Data Files.....	7-2
7.2.1 Background Files	7-4
7.2.1.1 Student Background File	7-4
7.2.1.2 Teacher Background File	7-4
7.2.1.3 School Background File.....	7-5
7.2.1.4 Identification Variables	7-5
7.2.1.5 Achievement Scores	7-7
7.2.1.6 Linking and Tracking Variables	7-8
7.2.1.7 International Background Variables	7-10
7.2.1.8 Variables Derived from Student, Teacher, and School Background Data	7-12
7.2.1.9 Sampling Variables.....	7-15
7.2.2 Assessment Files	7-16
7.2.2.1 Written Assessment Files	7-16
7.2.2.2 Performance Assessment Files.....	7-16
7.2.2.3 Cognitive Item Variable Names	7-16
7.2.2.4 Performance Assessment Task Identifications	7-17
7.2.2.5 Cognitive Item Response Code Values	7-18
7.2.2.6 Analysis By Mathematics and Science Content Area Reporting Categories.....	7-18
7.2.2.7 Release Status of TIMSS Test Items and Performance Tasks.....	7-23
7.2.2.8 Other Variables in the Student Assessment Files	7-23
7.2.2.9 School Performance Assessment Files	7-23
7.2.3 Coding Reliability Files.....	7-24
7.2.4 Student-Teacher Linkage Files.....	7-28
7.2.5 Missing Codes in the International Data Files	7-29
7.2.6 National Data Issues Affecting the Usage of International Data Files	7-31

7.3	Codebook Files	7-33
7.3.1	Accessing the Codebook Files.....	7-33
7.3.2	Using the Codebooks	7-34
7.4	Program Files	7-38
7.5	Data Almanacs	7-39
7.6	Test-Curriculum Matching Analysis Data Files	7-42
Chapter 8: Estimating Sampling Variance		8-1
8.1	Computing Error Variance Using the JRR Method.....	8-1
8.2	Construction of Sampling Zones for Sampling Variance Estimation	8-2
8.3	Computing the JRR Replicate Weights	8-3
Chapter 9: Performing Analyses with the TIMSS Data: Some Examples		9-1
9.1	Contents of the CDs	9-4
9.2	Creating SAS Data Sets and SPSS System Files.....	9-5
9.3	Computing JRR Replicate Weights and Sampling Variance Using SPSS and SAS.....	9-8
9.3.1	SAS Macro for Computing Mean and Percents with Corresponding Standard Errors (JACK.SAS)	9-8
9.3.2	SPSS Macro for Computing Mean and Percents with Corresponding Standard Errors (JACK.SPS) ...	9-16
9.4	Performing Analyses with Student-Level Variables	9-21
9.5	Performing Analyses with Teacher-Level Variables	9-27
9.6	Performing Analyses with School-Level Variables	9-33
9.7	Scoring the Items.....	9-39

Tables and Figures

Table 1.1	Countries Participating in TIMSS at Population 1 and 2 (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 Content Area Reporting Categories	2-3
Table 2.2	Distribution of Item Types Across Clusters – Population 1	2-5
Table 2.3	Ordering of Item Clusters Within Population 1 Booklets	2-6
Table 2.4	Distribution of Item Types Across Clusters – Population 2	2-7
Table 2.5	Ordering of Clusters Within Population 2 Booklets	2-8
Table 2.6	Assignment of Performance Assessment Tasks to Stations	2-9
Table 2.7	Assignment of Students to Stations in the Performance Assessment	2-9
Table 2.8	Countries Administering the Specialized and Non-Specialized Versions of the Population 2 Student Questionnaire	2-11
Table 3.1	Grades Tested in TIMSS – Population 1	3-2
Table 3.2	Grades Tested in TIMSS – Population 2	3-3
Figure 3.1	Relationship Between the Desired Populations and Exclusions	3-4
Figure 3.2	Countries Grouped for Reporting of Achievement According to Compliance with Guidelines for Sample Implementation and Participation Rates – Population 1 Written Assessment	3-9
Figure 3.3	Countries Grouped for Reporting of Achievement According to Compliance with Guidelines for Sample Implementation and Participation Rates – Population 2 Written Assessment	3-10
Figure 3.4	Countries Grouped for Reporting of Achievement According to Compliance with Guidelines for Sample Implementation and Participation Rates – Performance Assessment	3-11
Table 3.3	Sample Information for TIMSS Population 1 Countries	3-12
Table 3.4	Sample Information for TIMSS Population 2 Countries	3-13
Figure 4.1	Example Coding Guide for Short-Answer Mathematics Item	4-3
Figure 4.2	Example Coding Guide for Extended-Response 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	4-8
Table 6.1	Descriptive Statistics for the International Mathematics Achievement Scores for Population 1 (Variable: AIMATSCR)	6-5
Table 6.2	Descriptive Statistics for the International Science Achievement Scores for Population 1 (Variable: AISCISCR)	6-6
Table 6.3	Descriptive Statistics for the International Mathematics Achievement Scores for Population 2 (Variable: BIMATSCR)	6-7
Table 6.4	Descriptive Statistics for the International Science Achievement Scores for Population 2 (Variable: BISCISCR)	6-8
Table 7.1	TIMSS Population 1 and Population 2 Data Files	7-2
Table 7.2	Country Identification and Inclusion Status in Population 1 and Population 2 Data Files	7-3
Table 7.3	Background Questionnaire Item Field Location Format Conventions	7-11
Table 7.4	International Background Variable Naming Conventions	7-12
Table 7.5	International Report Table/Figure Location Reference Definition for Derived Variables	7-13
Table 7.6	Variable Name Definitions for the Written Assessment and Performance Assessment Items	7-17

Table 7.7	Classification of Population 1 Items into Mathematics Content Area Reporting Categories.....	7-19
Table 7.8	Classification of Population 1 Items into Science Content Area Reporting Categories.....	7-20
Table 7.9	Classification of Population 2 Items into Mathematics Content Area Reporting Categories.....	7-21
Table 7.10	Classification of Population 2 Items into Science Content Area Reporting Categories.....	7-22
Table 7.11	Recodes Made to Free-Response Item Codes in the Written Assessment and Performance Assessment Items.....	7-26
Table 7.12	Population 1 and Population 2 Codebook Files.....	7-33
Table 7.13	File Structure of Machine-Readable Codebook Files.....	7-34
Figure 7.1	Example Printout of a Codebook Page	7-36
Table 7.14	Population 1 and Population 2 Program Files	7-38
Table 7.15	Data Almanac Files	7-39
Figure 7.2	Example Data Almanac Display for Categorical Variable.....	7-40
Figure 7.3	Example Data Almanac Display for Continuous Variable.....	7-41
Figure 9.1	Sample Table for Student-Level Analysis Taken From the TIMSS International Report “Mathematics Achievement in the Middle School Years”	9-2
Figure 9.2	Sample Table for Teacher-Level Analysis Taken From the TIMSS International Report “Mathematics Achievement in the Middle School Years”	9-3
Table 9.1	Three-letter Extension Used to Identify the Files Contained in the CD	9-4
Figure 9.3	Extract from SAS Control Code for Creating a Student Background SAS Data Set	9-6
Figure 9.4	Extract from SPSS Control Code for Creating a Student Background SPSS Data Set.....	9-7
Figure 9.5	SAS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACK.SAS)	9-9
Table 9.2	Number of Replicate Weights Needed for Computing the JRR Error Variance Estimate... ..	9-12
Figure 9.6	SAS Control Code and Extract of Output File for Using the Macro JACK.SAS	9-15
Figure 9.7	SPSS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACK.SPS)	9-16
Figure 9.8	SPSS Control Code and Extract of Output File for Using the Macro JACK.SPS	9-21
Figure 9.9	SAS Control Statements for Performing Analyses with Student-Level Variables (EXAMPLE1.SAS)	9-23
Figure 9.10	SPSS Control Statements for Performing Analyses with Student-Level Variables (EXAMPLE1.SPS)	9-24
Figure 9.11	Extract of SAS Computer Output for Performing Analyses with Student-Level Variables (EXAMPLE 1)	9-25
Figure 9.12	Extract of SPSS Computer Output for Performing Analyses with Student-Level Variables (EXAMPLE 1)	9-26
Figure 9.13	SAS Control Statements for Performing Analyses with Teacher-Level Variables (EXAMPLE2.SAS)	9-28
Figure 9.14	SPSS Control Statements for Performing Analyses with Teacher-Level Variables (EXAMPLE2.SPS)	9-29
Figure 9.15	Extract of SAS Computer Output for Performing Analyses with Teacher-Level Variables (EXAMPLE 2)	9-31
Figure 9.16	Extract of SPSS Computer Output for Performing Analyses with Teacher-Level Variables (EXAMPLE 2)	9-32

Figure 9.17 SAS Control Statements for Performing Analyses with School-Level Variables (EXAMPLE3.SAS)	9-35
Figure 9.18 SPSS Control Statements for Performing Analyses with School-Level Variables (EXAMPLE3.SPS)	9-36
Figure 9.19 Extract of SAS Computer Output for Performing Analyses with School-Level Variables (EXAMPLE 3)	9-37
Figure 9.20 Extract of SPSS Computer Output for Performing Analyses with School-Level Variables (EXAMPLE 3)	9-38
Table 9.3 Definitions of Response Codes for the Multiple Choice Items in the Written Assessment Data Files	9-39
Table 9.4 Definition of Response Codes for the Open-Ended Items in the Written Assessment and Performance Assessment Data Files	9-40
Figure 9.21 Extracted Sections of SAS Control Code Used to Convert Cognitive Item Response Codes to Correctness-Score Levels	9-42
Figure 9.22 Extracted Sections of SPSS Control Code Used to Convert Cognitive Item Response Codes to Correctness-Score Levels	9-43

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 Primary and Middle School Years (TIMSS Populations 1 and 2). The database, provided on two compact disks, contains achievement data (written test and performance assessment) and student, teacher, and school background data collected in 42 countries in 1995. Table 1.1 lists, for each of Populations 1 and 2, the countries for which written assessment and performance assessment data are included in the International Database. Each of these countries gave the IEA permission to release its national data.

The TIMSS International Database contains the following, for each country for which internationally comparable data are available.

- Mathematics and science proficiency scale scores
- Students' responses to cognitive mathematics and science items
- Students' responses to hands-on performance tasks
- Students' background questionnaire data
- Mathematics and science teacher background questionnaire data
- School background questionnaire data
- Test-curriculum matching analysis data
- 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 500 files on the two compact disks 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 four supplements to the 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 reports.

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.

Table 1.1
Countries Participating in TIMSS at Population 1 and 2 (Data Included in Database)

Population 1		Population 2	
Written Assessment	Performance Assessment	Written Assessment	Performance Assessment
Australia	Australia	Australia	Australia
Austria	Canada	Austria	Canada
Canada	Cyprus	Belgium*	Colombia
Cyprus	Hong Kong	Bulgaria	Cyprus
Czech Republic	Iran, Islamic Republic	Canada	Czech Republic
England	Israel	Colombia	England
Greece	New Zealand	Cyprus	Hong Kong
Hong Kong	Portugal	Czech Republic	Iran, Islamic Rep.
Hungary	Slovenia	Denmark	Israel
Iceland	United States	England	Netherlands
Iran, Islamic Republic		France	New Zealand
Ireland		Germany	Norway
Israel		Greece	Portugal
Japan		Hong Kong	Romania
Korea		Hungary	Scotland
Kuwait		Iceland	Singapore
Latvia		Iran, Islamic Republic	Slovenia
Netherlands		Ireland	Spain
New Zealand		Israel	Sweden
Norway		Japan	Switzerland
Portugal		Korea	United States
Scotland		Kuwait	
Singapore		Latvia	
Slovenia		Lithuania	
Thailand		Netherlands	
United States		New Zealand	
		Norway	
		Philippines	
		Portugal	
		Romania	
		Russian Federation	
		Scotland	
		Singapore	
		Slovak Republic	
		Slovenia	
		South Africa	
		Spain	
		Sweden	
		Switzerland	
		Thailand	
		United States	

*The Flemish and French education systems in Belgium participated separately

1.2 Overview of TIMSS

The Third International Mathematics and Science Study (TIMSS) was conducted in 1995 across more than 40 countries.¹ 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 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. In an effort to extend what had been learned from previous studies and provide contextual and explanatory information, TIMSS expanded beyond the already substantial task of measuring achievement in two subject areas by also including a thorough investigation of curriculum and how it is delivered in classrooms around the world. In addition, extending the work of previous IEA studies, TIMSS included a performance assessment.

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

¹ Countries on a Southern Hemisphere school schedule - Australia, Korea, New Zealand, and Singapore - tested students in September - November 1994. All other countries tested students in 1995.

volumes: *Many Visions, Many Aims: A Cross-National Investigation of Curricular Intentions in School Mathematics* (Schmidt et al., 1997) and *Many Visions, Many Aims: A Cross-National Investigation of Curricular Intentions in School Science* (Schmidt et al., 1998).

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.

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 taking advanced courses in mathematics and students taking 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 collected from the assessment of students in their final year of secondary school (Population 3) will be released in a separate database.

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, following that, 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 International Database 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

The International Database, provided in two compact disks, includes more than 3000 variables in more than 500 files. One disk contains Population 1 data and the other disk contains Population 2 data. The files included on each disk are briefly described below.

Data Files. These files include the written assessment data, performance assessment data, background questionnaire data, coding reliability data, information to link students and teachers, 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 unweighted summary statistics for each participating country for each variable in the background questionnaires.

Test-Curriculum Matching Analysis Files. These files contain data collected for the TIMSS test-curriculum matching analysis.

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 CDs. 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 read-through (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 lower and upper grades of Populations 1 and 2; the performance assessment administered to subsamples of the upper-grade students in Populations 1 and 2; and the student, teacher, 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.

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–Population 1

This supplement contains the international versions of the student, teacher, and school background questionnaires for Population 1 and tables that map each question to a variable in the database.

Supplement 2 - International Versions of the Background Questionnaires–Population 2

This supplement contains the international versions of the student, teacher, and school background questionnaires for Population 2 and tables that map each question to a variable in the database.

Supplement 3 - Documentation of National Adaptations of the International Background Questionnaire Items

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

Supplement 4 - Documentation of Derived Variables Based on Student and Teacher Background Questionnaire Items

The TIMSS international reports included a number of variables derived from questions in the student and teacher questionnaires. 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 managed 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* (Martin and Kelly, 1997)

Chapter 2

TIMSS Instruments and Booklet Design

2.1 Introduction

TIMSS used several types of instruments to collect data about students, teachers, and schools. Each assessed student received a test booklet containing cognitive items in mathematics and science along with a separate background questionnaire. Subsamples of students participating in the written assessment also participated in a performance assessment in which they completed hands-on mathematics and science tasks. Teacher questionnaires were given to the mathematics and science teachers of the assessed students. A school questionnaire was distributed to each participating school and completed by the school principal or headmaster.

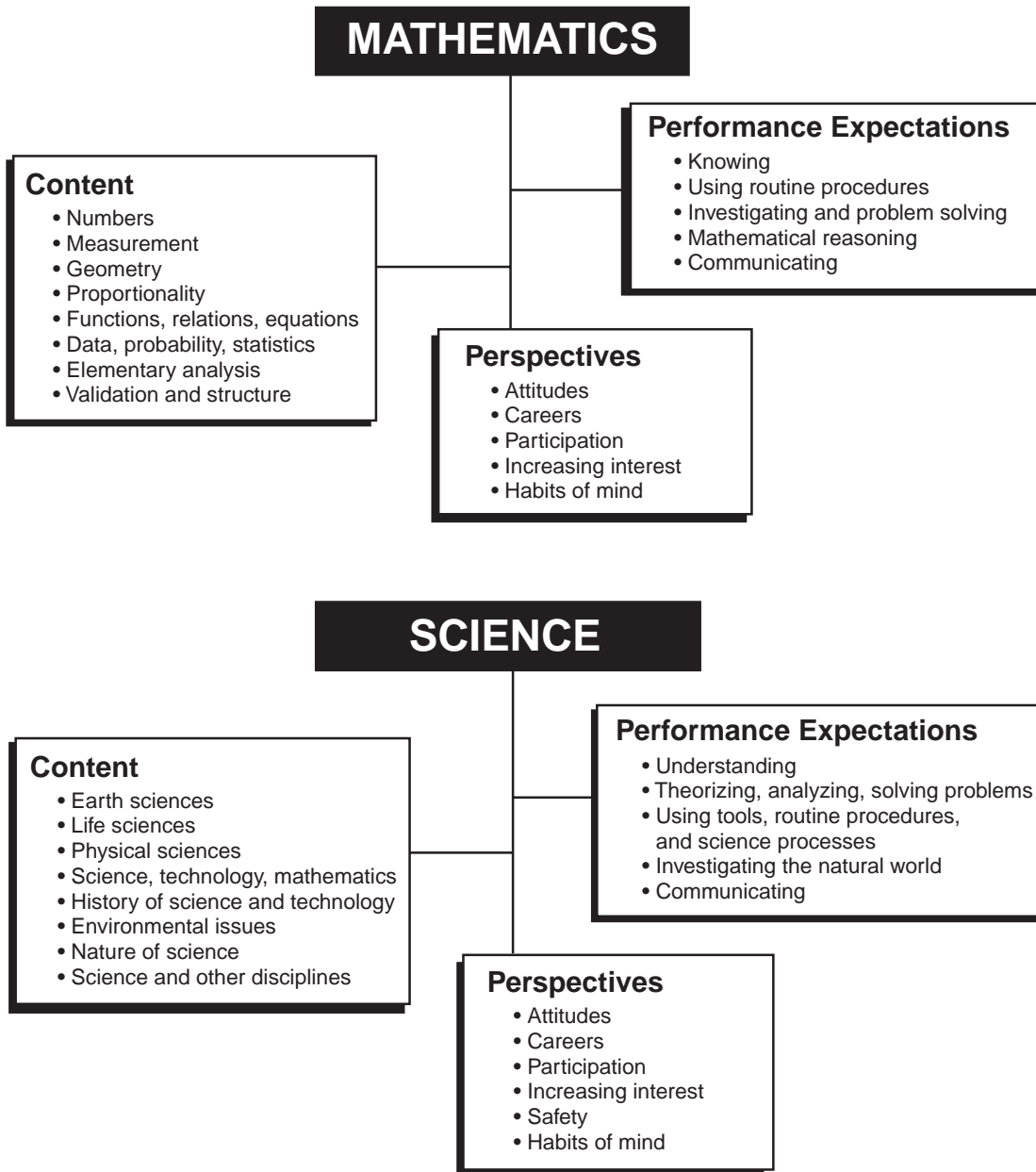
This chapter describes the content and organization of the assessment instruments for Populations 1 and 2.

2.2 The TIMSS 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 mathematics and science content “reporting categories.” In the analysis, each item in the TIMSS test was included in one reporting category based on its principal content code. Table 2.1 presents the reporting categories for the mathematics and science content areas used in the international reports. The classification of items into each mathematics and science reporting category are shown in Tables 7.7 through 7.10 in Chapter 7.

Table 2.1
Mathematics and Science Content Area Reporting Categories

	Mathematics	Science
Population 1	Whole numbers Fractions and proportionality Measurement, estimation, and number sense Data representation, analysis, and probability Geometry Patterns, relations, and functions	Earth science Life science Physical science Environmental issues and the nature of science
Population 2	Fractions and number sense Geometry Algebra Data representation, analysis, and probability Measurement Proportionality	Earth science Life science Physics Chemistry Environmental issues and the nature of science

2.3 The TIMSS Items

The task of putting together the achievement item pools for the different TIMSS student populations 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’ curriculum.

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).

At Population 1, there were 102 mathematics items, including 79 multiple-choice items, 15 short-answer items, and 8 extended-response items. The science test contained 97 items, of which 13 were classified as requiring short answers and 10 as requiring more extended responses. In all, there is a total pool of 235 unique testing minutes in Population 1, 118 for mathematics and 117 for science.

At Population 2, the overall pool of cognitive items contained 151 mathematics items, including 125 multiple-choice items, 19 short-answer items, and 7 extended-response items. There were 135 science items, including 102 multiple-choice items and 33 free-response items. Population 2 contained a total of 396 unique testing minutes, 198 for mathematics and 198 for science.

2.4 Organization of the Test Booklets

At each population, the test items were allocated to 26 different clusters labeled A through Z. Also, at each population, the 26 clusters were assembled into eight booklets. Each student completed one booklet. At Population 1 students were given 64 minutes to complete their booklets. At Population 2, students were given 90 minutes. The organization of the clusters and booklets is summarized below, first for Population 1 and then for Population 2.¹

At Population 1, the clusters were either 9 or 10 minutes in length, as shown in Table 2.2. The core cluster, labeled A, composed of five mathematics and five science multiple-choice items, was included in all booklets. Focus clusters, labeled B through H, appeared in at least three booklets, so that the items were answered by a relatively large fraction (three-eighths) of the student sample in each country. The breadth clusters, largely containing multiple-choice items, appeared in only one booklet. The breadth clusters are labeled I through M for mathematics and N through R for science. The free-response clusters were each assigned to two booklets, so that items statistics of reasonable accuracy would be available. These clusters are labeled S through V for mathematics and W through Z for science.

¹ The organization of the test design is fully documented in Adams and Gonzalez (1996).

Table 2.2
Distribution of Item Types Across Clusters - Population 1

Cluster Type	Cluster Label	Number Mathematics Items			Number Science Items		
		Multiple Choice	Short Answer	Extended Response	Multiple Choice	Short Answer	Extended Response
Core (10 minutes)	A	5	-	-	5	-	-
Focus (9 minutes)	B	5	-	-	4	-	-
	C	4	-	-	5	-	-
	D	5	-	-	4	-	-
	E	4	-	-	4	1	-
	F	5	-	-	4	-	-
	G	4	-	-	5	-	-
	H	5	-	-	3	1	-
Breadth (Mathematics) (9 minutes)	I	9	-	-			
	J	9	-	-			
	K	9	-	-			
	L	8	1	-			
Breadth (Science) (9 minutes)	M	7	2	-			
	N				9	-	-
	O				7	2	-
	P				8	1	-
Mathematics Free-Response (9 minutes)	Q				7	2	-
	R				8	1	-
	S	-	3	2			
	T	-	3	2			
Science Free-Response (9 minutes)	U	-	3	2			
	V	-	3	2			
	W				-	3	2
	X				1	2	2
	Y				-	-	3
	Z				-	-	3

Table 2.3
Ordering of Item Clusters Within Population 1 Booklets

Cluster Order	Booklet							
	1	2	3	4	5	6	7	8
1st	B	C	D	E	F	G	H	B
2nd	A	A	A	A	A	A	A	A
3rd	C	D	E	F	G	H	B	R
4th	S	W	T	X	U	Y	V	Z
Break								
5th	E	F	G	H	B	C	D	I
6th	J	N	K	O	L	P	M	Q
7th	T	X	U	Y	V	Z	W	S

The Population 1 test booklets were designed to be administered in two consecutive testing sessions with a 15-20 minute break between the sessions. The order of the clusters within the Population 1 booklets is shown in Table 2.3. All booklets contain mathematics and science items. The core cluster appears in the second position in all booklets. The rotation design used to assign cluster B through H to booklets 1 through 7 allows the estimation of all item covariances for the items in cluster A through H. Each of the focus clusters occurs once in the first, third, and fifth positions in booklets 1 through 7. There are free-response clusters in Part 1 as well as in Part 2 of each test booklet (fourth and seventh cluster in each booklet). Booklet 8, which contains three breadth clusters, serves primarily to increase the content coverage of the tests.

Table 2.4
Distribution of Item Types Across Clusters - Population 2

Cluster Type	Cluster Label	Number Mathematics Items			Number Science Items		
		Multiple Choice	Short Answer	Extended Response	Multiple Choice	Short Answer	Extended Response
Core (12 Minutes)	A	6	-	-	6	-	-
Focus (12 minutes)	B	6	-	-	6	-	-
	C	6	-	-	6	-	-
	D	6	-	-	6	-	-
	E	6	-	-	6	-	-
	F	6	-	-	6	-	-
	G	6	-	-	6	-	-
	H	6	-	-	6	-	-
Breadth (Mathematics and Science) (22 minutes)	I	7	2	-	9	1	-
	J	7	2	-	7	2	-
	K	7	2	-	8	2	-
	L	9	1	-	6	-	1
	M	7	2	-	2	2	1
	N	7	2	-	8	2	-
	O	7	2	-	4	4	-
	P	9	1	-	3	4	-
	Q	9	1	-	5	3	-
R	7	2	-	2	-	1	
Mathematics Free-Response (10 minutes)	S	-	-	2			
	T	-	-	2			
	U	-	-	2			
	V	-	2	1			
Science Free-Response (10 minutes)	W				-	-	2
	X				-	-	2
	Y				-	-	2
	Z				-	-	2

The booklet design for Population 2 is very similar to that for Population 1. Of the 26 clusters in Population 2, eight take 12 minutes, ten take 22 minutes, and eight take 10 minutes. The core cluster (cluster A), comprising six mathematics and six science multiple-choice items, appears in the second position in every booklet. The seven focus clusters appear in at least three booklets, and the ten breadth clusters appears in only one booklet. The eight free-response clusters, each containing 10 minutes of short-answer and extended-response items, were each assigned to two booklets. Tables 2.4 and 2.5 show the number of items in each cluster and the assignment of clusters to booklets, respectively.

Table 2.5
Ordering of Clusters Within Population 2 Booklets

Cluster Order	Booklet							
	1	2	3	4	5	6	7	8
1st	B	C	D	E	F	G	H	B
2nd	A	A	A	A	A	A	A	A
3rd	C	D	E	F	G	H	B	Q
4th	S	W	T	X	U	Y	V	
Break								
5th	E	F	G	H	B	C	D	R
6th	I	J	K	L	M	N	O	P
7th	T	X	U	Y	V	Z	W	

2.5 Performance Assessment

The TIMSS performance assessment was administered at Populations 1 and 2 to a subsample of students in the upper grades that participated in the written assessment (Harmon and Kelly, 1996; Harmon et al., 1997). The *performance tasks* permitted students to demonstrate their ability to make, record, and communicate observations; to take measurements or collect experimental data and present them systematically; to design and conduct a scientific investigation; or to solve certain types of problems. A set of 13 such “hands-on” activities was developed and used with subsamples of students at fourth and eighth grades. Eleven of the tasks were either identical or similar across populations, and two tasks were different. Of these two, one task was administered to the Population 1 (fourth graders) and one was administered to Population 2 (eighth graders).

The 12 tasks administered at each population were presented at nine different stations. Each station required about 30 minutes working time. Each student was assigned to three stations by a sequence number, for a total testing time of 90 minutes. Because the complete circuit of nine stations occupies nine students, students participating in the performance assessment were sampled in sets of nine. However, the complete rotation of students required two sets of 9, or 18 students, to assure that each task was paired with each other task at least once. Taken together, Tables 2.6 and 2.7 show the stations each student visited and the tasks completed according to the rotation assignment (either Rotation 1 or Rotation 2) and sequence number.

Table 2.6
Assignment of Performance Assessment Tasks to Stations

Station	Task
A	S1 Pulse M1 Dice
B	S2 Magnets M2 Calculator
C	G1 Shadows
D	S3 Batteries M3 Folding and Cutting
E	S4 Rubber Band
F	M5 Packaging
G	S5 or S6 Solutions (Population 2) Containers (Population 1)
H	M4 Around the Bend
I	G2 Plasticine

Table 2.7
Assignment of Students to Stations in the Performance Assessment

Student Sequence Number	Rotation 1 Stations	Rotation 2 Stations
1	A, B, C	A, B, E
2	B, E, D	B, D, G
3	C, F, E	C, A, D
4	D, G, H	D, E, F
5	E, A, G	E, I, H
6	F, H, B	F, H, A
7	G, I, F	G, F, I
8	H, C, I	H, G, C
9	I, D, A	I, C, B

2.6 Release Status for TIMSS Test Items, Performance Tasks, and Background Questionnaires

Some TIMSS items have been released for unrestricted public use. At Populations 1 and 2, all items in clusters I through Z are classified as public release and are available to secondary users (TIMSS, 1996a; 1996b; 1997a; 1997b). Information about how to obtain the released item sets is provided at the end of this chapter. All of the performance assessment tasks are contained in the performance assessment international report (Harmon et al., 1997). All student, teacher, and school questionnaires are also classified as public release and are available to secondary users in Supplements 1 and 2 of this User Guide.

2.7 Questionnaires

The *student questionnaires* were separate instruments 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 more time was required in about 60% of the sessions (about 20 minutes extra).

The student questionnaires asked about students' demographics and home environment, including academic activities outside of school, people living in the home, parental education (only at Population 2), 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. The final sections of the questionnaires asked about classroom experiences in mathematics and science.

In general, the structure and content of the student questionnaires were similar across Populations 1 and 2. However, for many questions, the response options were reduced in the Population 1 version from four to three categories. At Population 2, there were two versions of the student questionnaires. One version was for use in countries teaching general or integrated science (non-specialized version) and the other was for use in systems where students take courses in specific sciences such as biology, chemistry, earth science, or physics (specialized version). Slightly more than half the countries used the general questionnaire and slightly less than half used the questionnaire developed for use with specific science curriculum areas. Table 2.8 presents the countries that administered the specialized and non-specialized versions.

Table 2.8**Countries Administering the Specialized and Non-Specialized Versions of the Population 2 Student Questionnaire**

Non-Specialized Version (Science as an Integrated Subject)	Specialized Version (Science as Separate Subjects)
Australia	Belgium (Flemish)
Austria	Belgium (French)
Canada	Czech Republic
Colombia	Denmark
Cyprus	France
England	Germany
Hong Kong	Greece
Iran	Hungary
Ireland	Iceland
Israel	Latvia
Japan	Lithuania
Korea	Netherlands
Kuwait	Portugal
New Zealand	Romania
Norway	Russian Federation
Philippines	Slovak Republic
Scotland	Slovenia
Singapore	Sweden (Upper-Grade)
South Africa	
Spain	
Sweden (Lower-Grade)	
Switzerland	
Thailand	
United States	

The *teacher questionnaires* for Population 2 addressed four major areas: teachers' background, instructional practices, students' opportunity to learn, and teachers' pedagogic beliefs. There are separate questionnaires for teachers of mathematics and of science. Since most Population 1 teachers teach all subjects, a single teacher questionnaire was developed to address both mathematics and science. So as not to overburden the teachers, the classroom practices questions in the Population 1 teacher questionnaire pertain mostly to mathematics. However, teachers also were asked about how they spend their time in school and the atmosphere in their schools (e.g., teaching loads, collaboration policies, responsibilities for decision-making, and the availability of resources).

The teacher questionnaires were designed to provide information about the teachers of the students sampled in TIMSS. The teachers who completed the TIMSS questionnaires do not constitute a sample from any definable population of teachers. Rather, they represent the teachers of a national sample of students.

The *school questionnaires* for each population sought information about the school's community, staff, students, curriculum and programs of study, and instructional resources and time. At Populations 1 and 2, the school questionnaires also ask about the number of years students are taught by the same teacher. A school questionnaire was to be completed by the principal, headmaster, or other administrator of each school that participated in TIMSS.

TIMSS Released Item Sets

The TIMSS released items are available in four volumes:

TIMSS Mathematics Items – Released Set for Population 1 (Third and Fourth Grades)

TIMSS Science Items – Released Set for Population 1 (Third and Fourth Grades)

TIMSS Mathematics Items – Released Set for Population 2 (Seventh and Eighth Grades)

TIMSS Science Items – Released Set for Population 2 (Seventh and Eighth Grades)

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

The TIMSS released items also are available on the World Wide Web at:
wwwcsteep.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 Foy (1997a). The weighting procedures are described in Foy (1997b).

3.1 The Target Populations

The selection of valid and efficient samples is crucial to the quality and success of an international comparative study such as TIMSS. For TIMSS, National Research Coordinators worked on all phases of sampling with staff from Statistics Canada. In consultation with the TIMSS Sampling Referee (Keith Rust, WESTAT, Inc.), staff from Statistics Canada reviewed the national documentation on 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 national samples.

For Populations 1 and 2, the International Desired Populations for all countries were defined as follows.

Population 1. All students enrolled in the two adjacent grades that contain the largest proportion of 9-year-olds at the time of testing.

Population 2. All students enrolled in the two adjacent grades that contain the largest proportion of 13-year-olds at the time of testing.

Tables 3.1 and 3.2 show the grades tested in each country that participated in TIMSS. This information is captured in the variable IDGRADE in the student data files.

Table 3.1
Grades Tested in TIMSS – Population 1

Country	Lower Grade		Upper Grade	
	Country's Name for Lower Grade	Years of Formal Schooling Including Lower Grade ¹	Country's Name for Upper Grade	Years of Formal Schooling Including Upper Grade ¹
² Australia	3 or 4	3 or 4	4 or 5	4 or 5
Austria	3	3	4	4
Canada	3	3	4	4
Cyprus	3	3	4	4
Czech Republic	3	3	4	4
England	Year 4	4	Year 5	5
Greece	3	3	4	4
Hong Kong	Primary 3	3	Primary 4	4
Hungary	3	3	4	4
Iceland	3	3	4	4
Iran, Islamic Rep.	3	3	4	4
Ireland	3rd Class	3	4th Class	4
Israel	–	–	4	4
Japan	3	3	4	4
Korea	3rd Grade	3	4th Grade	4
Kuwait	–	–	5	5
Latvia	3	3	4	4
³ Netherlands	5	3	6	4
⁴ New Zealand	Standard 2	3.5–4.5	Standard 3	4.5–5.5
Norway	2	2	3	3
Portugal	3	3	4	4
Scotland	Year 4	4	Year 5	5
Singapore	Primary 3	3	Primary 4	4
Slovenia	3	3	4	4
Thailand	Primary 3	3	Primary 4	4
United States	3	3	4	4

¹ Years of schooling based on the number of years children in the grade level have been in formal schooling, beginning with primary education (International Standard Classification of Education Level 1). Does not include preprimary education.

² Australia: Each state/territory has its own policy regarding age of entry to primary school. In 4 of the 8 states/territories students were sampled from grades 3 and 4; in the other four states/territories students were sampled from grades 4 and 5.

³ In the Netherlands kindergarten is integrated with primary education. Grade-counting starts at age 4 (formerly kindergarten 1). Formal schooling in reading, writing, and arithmetic starts in grade 3, age 6.

⁴ New Zealand: The majority of students begin primary school on or near their 5th birthday so the "years of formal schooling" vary.

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95. Information provided by TIMSS National Research Coord

Table 3.2
Grades Tested in TIMSS – Population 2

Country	Lower Grade		Upper Grade	
	Country's Name for Lower Grade	Years of Formal Schooling Including Lower Grade ¹	Country's Name for Upper Grade	Years of Formal Schooling Including Upper Grade ¹
² Australia	7 or 8	7 or 8	8 or 9	8 or 9
Austria	3. Klasse	7	4. Klasse	8
Belgium (Fl)	1A	7	2A & 2P	8
Belgium (Fr)	1A	7	2A & 2P	8
Bulgaria	7	7	8	8
Canada	7	7	8	8
Colombia	7	7	8	8
³ Cyprus	7	7	8	8
Czech Republic	7	7	8	8
Denmark	6	6	7	7
England	Year 8	8	Year 9	9
France	5ème	7	4ème (90%) or 4ème Technologique (10%)	8
Germany	7	7	8	8
Greece	Secondary 1	7	Secondary 2	8
Hong Kong	Secondary 1	7	Secondary 2	8
Hungary	7	7	8	8
Iceland	7	7	8	8
Iran, Islamic Rep.	7	7	8	8
Ireland	1st Year	7	2nd Year	8
Israel	–	–	8	8
Japan	1st Grade Lower Secondary	7	2nd Grade Lower Secondary	8
Korea, Republic of	1st Grade Middle School	7	2nd Grade Middle School	8
Kuwait	–	–	9	9
Latvia	7	7	8	8
Lithuania	7	7	8	8
Netherlands	Secondary 1	7	Secondary 2	8
⁴ New Zealand	Form 2	7.5 - 8.5	Form 3	8.5 - 9.5
³ Norway	6	6	7	7
³ Philippines	Grade 6 Elementary	6	1st Year High School	7
Portugal	Grade 7	7	Grade 8	8
Romania	7	7	8	8
⁵ Russian Federation	7	6 or 7	8	7 or 8
Scotland	Secondary 1	8	Secondary 2	9
Singapore	Secondary 1	7	Secondary 2	8
Slovak Republic	7	7	8	8
Slovenia	7	7	8	8
Spain	7 EGB	7	8 EGB	8
³ South Africa	Standard 5	7	Standard 6	8
³ Sweden	6	6	7	7
³ Switzerland (German)	6	6	7	7
(French and Italian)	7	7	8	8
Thailand	Secondary 1	7	Secondary 2	8
United States	7	7	8	8

¹ Years of schooling based on the number of years children in the grade level have been in formal schooling, beginning with primary education (International Standard Classification of Education Level 1). Does not include preprimary education.

² Australia: Each state/territory has its own policy regarding age of entry to primary school. In 4 of the 8 states/territories students were sampled from grades 7 and 8; in the other four states/territories students were sampled from grades 8 and 9.

³ Indicates that there is a system-split between the lower and upper grades. In Cyprus, system-split occurs only in the large or city schools. In Switzerland there is a system-split in 14 of 26 cantons.

⁴ New Zealand: The majority of students begin primary school on or near their 5th birthday so the "years of formal schooling" vary.

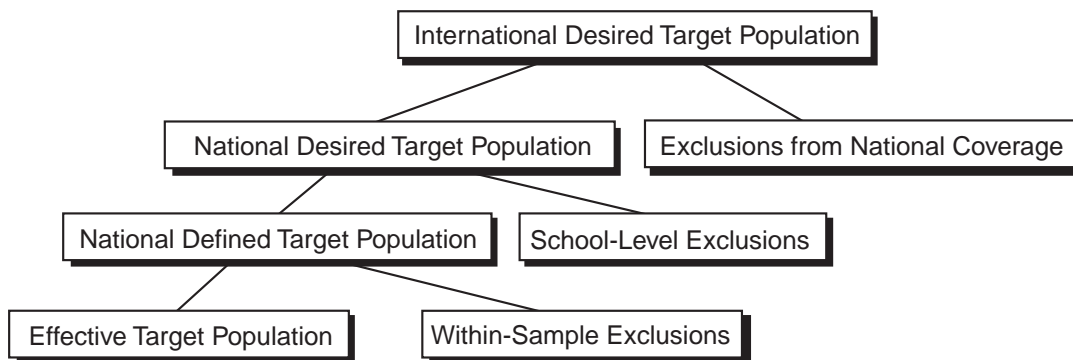
⁵ Russian Federation: 70% of students in the seventh grade have had 6 years of formal schooling; 70% in the eighth grade have had 7 years of formal schooling.

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95. Information provided by TIMSS National Research Coordinators.

TIMSS used a grade-based definition of the target populations. In a few cases, TIMSS components were administered only to the upper grade of these populations (i.e., the performance assessment was conducted at the upper grade and some background questions were asked of the upper-grade students only). However, two adjacent grades were chosen to ensure extensive coverage of the same age cohort for most countries, thereby increasing the likelihood of producing useful age-based comparisons in addition to the grade-based analyses.

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 Populations as a basis, participating countries had to operationally define their populations 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. Thus, the Nationally Desired Population sometimes differed from the International Desired Population.

The basic sample design used by TIMSS is generally referred to as a two-stage stratified cluster sample design. The first stage consisted of a sample of schools, which may be stratified; the second stage consisted of samples of classrooms from each eligible target grade in sampled schools. In some countries, a third stage consisted of sampling students within classrooms. Exclusions could occur at the school level, student level, or both. TIMSS participants were expected to keep such exclusions to no more than 10% of the national desired populations. Nevertheless, the National Defined Populations generally were subsets of the desired populations.

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.

3.2 School Sample Selection

In the first stage of sampling, representative samples of schools were selected from sampling frames (comprehensive lists of all eligible students). The TIMSS standard for sampling precision required that all population samples have an effective sample size of at least 400 students for the main criterion variables. To meet the standard, at least 150 schools were to be selected per target population. However, the clustering effect of sampling classrooms rather than students was also considered in determining the overall sample size for TIMSS. Because the magnitude of the clustering effect is determined by the size of the cluster and the intraclass correlation, TIMSS produced sample-design tables showing the number of schools to sample for a range of intraclass correlations and minimum-cluster-size values. Some countries needed to sample more than 150 schools. Countries, however, were asked to sample 150 schools even if the estimated number of schools to sample fell below 150. Information about design effect and effective sample size can be found in Gonzalez and Foy (1997).

The sample-selection method used for first-stage sampling was based on a systematic probability-proportional-to-size (PPS) technique. The schools in each explicit stratum (e.g., geographical region, public/private, etc.) were listed in order of the implicit stratification variables, and then further sorted according to their measure of size (MOS). Of course, the stratification variables differed from country to country. Small schools were handled either through explicit stratification or through the use of pseudo-schools. 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.

It was sometimes the case that a sampled school was unable to participate in the assessment. In such cases, this originally sampled school needed to be replaced by *replacement schools*. 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. Using either explicit or implicit stratification variables and ordering of the school sampling frame by size ensured that any original sampled school's replacement would have similar characteristics.

3.3 Classroom and Student Sampling

In the second sampling stage, classrooms of students were sampled. Generally, in each school, one classroom was sampled from each target grade, although some countries opted to sample two classrooms at the upper grade in order to be able to conduct special analyses. Most participants tested all students in selected classrooms, and in these instances the classrooms were selected with equal probabilities. The few participants who chose to subsample students within selected classrooms sampled classrooms with PPS.

As an optional third sampling stage, participants with particularly large classrooms in their schools could decide to subsample a fixed number of students per selected classroom. This was done using a simple random sampling method whereby all students in a sampled classroom were assigned equal selection probabilities.

3.4 Performance Assessment Subsampling

For the performance assessment, TIMSS participants were to sample at least 50 schools from those already selected for the written assessment, and from each school a sample of either 9 or 18 upper-grade students already selected for the written assessment. This yielded a sample of about 450 students in the upper grade of each population (eighth and fourth grades in most countries) in each country. For the performance assessment, in the interest of ensuring the quality of administration, countries could exclude additional schools if the schools had fewer than nine students in the upper grade or if the schools were in a remote region. The exclusion rate for the performance assessment sample was not to exceed 25% of the national desired population.

3.5 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.

3.5.1 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. School-level response rates were computed and reported by grade weighted and unweighted, 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 (π_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 are computed in a similar way, where all school ratios are 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 π_i was set proportional to MOS_i within each explicit stratum, it is generally the case that weighted and unweighted rates are similar.

3.5.2 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. Student-level response rates were computed and reported by grade, weighted and unweighted. The general formula for computing student-level response rates is shown in the following equation:

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

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.

3.5.3 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. Weighted overall response rates were computed and reported by grade, both with and without replacement schools.

Response rates and school and student sample sizes for the written and performance assessments are available in Mullis et al., (1997); Martin et al., (1997); Beaton et al., (1996a); Beaton et al., (1996b); and Harmon et al., (1997).

3.6 Compliance with Sampling Guidelines

Figures 3.2 and 3.3, for Populations 1 and 2, respectively, indicate the degree to which countries complied with the TIMSS sampling guidelines. Countries were considered to have met the TIMSS guidelines if they 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, and if they complied with the TIMSS guidelines for grade selection and classroom sampling. Countries that met these guidelines are shown in the top panels of Figures 3.2 and 3.3. Countries that met the guidelines only after including replacement schools are identified.

Countries not reaching at least 50% school participation without the use of replacement schools, or that failed to reach the sampling participation standard even with the inclusion of replacement schools, are shown in the second panels. To provide a better curricular match, several countries elected to test students in the seventh and eighth grades (the two grades tested by most countries), even though that meant not testing the two grades with the most age-eligible students. This led to the students in these countries being somewhat older than those in the other countries, and they are shown in a separate panel. For a variety of reasons, some countries did not comply with the guidelines for sampling classrooms. They also are shown in a separate section, as are the countries that had unapproved classroom sampling procedures as well as other departures from the guidelines. Finally, at Population 2, the Philippines had unapproved sampling procedures at the school level and so sampling weights could not be computed. Therefore, data for the Philippines are unweighted.¹

The performance assessment subsamples also were reviewed on the basis of their quality and adherence to the international standards. Figure 3.4 indicates the degree to which countries' performance assessment samples met the standards. The sample of schools and students for the performance assessment was a subsample of schools and students that participated in the main written assessment. Consequently, the characteristics of each country's performance assessment sample reflect the quality of the sampling for the written assessment and compliance with the guidelines for the performance assessment sampling. Due to unapproved sampling procedures at the school level, the performance assessment data for Israel at both Population 1 and Population 2 are unweighted.

¹ This is effectively implemented by assigning a weight of 1 to all students in the sample for the Philippines.

Figure 3. 2
Countries Grouped for Reporting of Achievement According to Compliance with
Guidelines for Sample Implementation and Participation Rates – Population 1
Written Assessment

Fourth Grade		Third Grade	
Countries satisfying guidelines for sample participation rates, grade selection and sampling procedures			
Canada	Norway	Canada	Norway
Cyprus	Portugal	Cyprus	Portugal
Czech Republic	[†] Scotland	Czech Republic	Scotland
¹² England	Singapore	¹² England	Singapore
Greece	United States	Greece	United States
Hong Kong		Hong Kong	
Iceland		Iceland	
Iran, Islamic Rep.		Iran, Islamic Rep.	
Ireland		Ireland	
Japan		Japan	
Korea		Korea	
New Zealand		New Zealand	
Countries not satisfying guidelines for sample participation			
Australia		Australia	
Austria		Austria	
¹ Latvia (LSS)		¹ Latvia (LSS)	
Netherlands		Netherlands	
		Scotland	
Countries not meeting age/grade specifications (high percentage of older students)			
Slovenia		Slovenia	
Countries with unapproved sampling procedures at the classroom level			
Hungary		Hungary	
Countries with unapproved sampling procedures at classroom level and not meeting other guidelines			
¹ Israel			
Kuwait		Thailand	
Thailand			

[†] Met guidelines for sample participation rates only after replacement schools were included.

¹ National Desired Population does not cover all of International Desired Population. Because coverage falls below 65%, Latvia is annotated LSS for Latvian Speaking Schools only.

² National Defined Population covers less than 90 percent of National Desired Population.

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

Figure 3. 3
Countries Grouped for Reporting of Achievement According to Compliance with
Guidelines for Sample Implementation and Participation Rates – Population 2
Written Assessment

Eighth Grade		Seventh Grade	
Countries satisfying guidelines for sample participation rates, grade selection and sampling procedures			
[†] Belgium (Fl) Canada Cyprus Czech Republic ^{†2} England France Hong Kong Hungary Iceland Iran, Islamic Rep. Ireland Japan Korea	[†] Latvia (LSS) [†] Lithuania New Zealand Norway Portugal Russian Federation Singapore Slovak Republic Spain Sweden [†] Switzerland [†] United States	[†] Belgium (Fr) [†] Belgium (Fl) Canada ^{†2} Cyprus Czech Republic England France Hong Kong Hungary Iceland Iran, Islamic Rep. Ireland Japan Korea	[†] Latvia (LSS) [†] Lithuania New Zealand Norway Portugal Russian Federation [†] Scotland Singapore Slovak Republic Spain Sweden [†] Switzerland [†] United States
Countries not satisfying guidelines for sample participation			
Australia Austria Belgium Bulgaria Netherlands Scotland		Australia Austria Bulgaria Netherlands	
Countries not meeting age/grade specifications (high percentage of older students)			
Colombia ^{†1} Germany Romania Slovenia		Colombia ^{†1} Germany Romania Slovenia	
Countries with unapproved sampling procedures at the classroom level			
Denmark Greece Thailand		Denmark Greece [†] South Africa Thailand	
Countries with unapproved sampling procedures at classroom level and not meeting other guidelines			
[†] Israel Kuwait South Africa			
Countries with unapproved sampling procedures at school level			
³ Phillipines		³ Phillipines	

[†] Met guidelines for sample participation rates only after replacement schools were included.

¹ National Desired Population does not cover all of International Desired Population. Because coverage falls below 65%, Latvia is annotated LSS for Latvian Speaking Schools only.

² National Defined Population covers less than 90 percent of National Desired Population.

³ TIMSS was unable to compute sampling weights for the Phillipines.

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

Figure 3. 4
Countries Grouped for Reporting of Achievement According to Compliance with Guidelines for Sample Implementation and Participation Rates – Performance Assessment

Eighth Grade	Fourth Grade
Countries satisfying guidelines for sample participation rates, grade selection and sampling procedures	
Canada Cyprus Czech Republic Iran, Islamic Republic New Zealand Norway Portugal [†] Scotland Singapore Spain Sweden ^{†1} Switzerland	Canada Cyprus Iran, Islamic Republic ^{†3} New Zealand Portugal
Countries not satisfying guidelines for sample participation	
Australia ² England Netherlands United States	Australia Hong Kong United States
Countries not meeting age/grade specifications (high percentage of older students)	
Colombia ³ Romania Slovenia	Slovenia
Countries with small sample sizes	
Hong Kong	
Countries with unapproved sampling procedures	
Israel	Israel

[†] Met guidelines for sample participation rates only after replacement schools were included.

¹ National Desired Population does not cover all of International Desired Population. German-speaking car

² National Defined Population covers less than 90 percent of National Desired Population before allowing for performance assessment sample exclusions.

³ School-level exclusions for performance assessment exceed 25% of National Desired Population.

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

3.7 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.²

The students within each country were selected using probability sampling. This means that the probability of each student being selected as part of the sample is known. 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. As is the case in TIMSS, the sum of the sampling weights for a sample is an estimate of the size of the population of students within the country in the sampled grades. The sampling weights must be used whenever population estimates are required. The use of the appropriate sampling weights ensures that the different subgroups that constitute the sample are properly and proportionally represented in the computation of population estimates.

Tables 3.3 and 3.4 present the sample sizes and the estimate of the population size (sum of the weights) for each participating country in Populations 1 and 2, respectively.

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

Table 3.3
Sample Information for TIMSS Population 1 Countries

COUNTRY	Third Grade			Fourth Grade		
	Sample Size	Sum of Weights	Maximum Jackknifing Zones	Sample Size	Sum of Weights	Maximum Jackknifing Zones
Australia	4741	232333.3	74	6507	237328.1	74
Austria	2526	86043.5	68	2645	91390.7	68
Canada	7594	371166.0	75	8408	389159.8	75
Cyprus	3308	9740.3	74	3376	9995.4	74
Czech Republic	3256	116051.5	73	3268	120406.0	73
England	3056	531682.1	67	3126	534922.4	67
Greece	2955	98999.8	75	3053	106181.0	75
Hong Kong	4396	83846.6	62	4411	89901.4	62
Hungary	3038	116778.6	75	3006	117228.1	75
Iceland	1698	3734.6	75	1809	3739.3	75
Iran, Islamic Rep.	3361	1391859.1	75	3385	1433314.5	75
Ireland	2889	58502.8	73	2873	60496.9	73
Israel	.	.	.	2351	66967.0	44
Japan	4306	1388749.3	74	4306	1438464.9	74
Korea	2777	607007.1	75	2812	615004.3	75
Kuwait	.	.	.	4318	24071.0	75
Latvia (LSS)	2054	15120.7	59	2216	18882.5	59
Netherlands	2790	171561.2	52	2524	173406.9	52
New Zealand	2504	48385.9	75	2421	52254.3	75
Norway	2219	49035.8	70	2257	49896.4	70
Portugal	2650	114774.8	72	2853	133186.2	72
Scotland	3132	59393.4	65	3301	59053.7	65
Singapore	7030	41904.0	75	7139	41244.0	75
Slovenia	2521	27453.4	61	2566	27685.1	61
Thailand	2870	883764.8	75	2992	864525.4	75
United States	3819	3643393.3	59	7296	3563795.3	59

Table 3.4
Sample Information for TIMSS Population 2 Countries

Country	Seventh Grade			Eighth Grade		
	Sample Size	Sum of Weights	Maximum Jackknifing Zones	Sample Size	Sum of Weights	Maximum Jackknifing Zones
Australia	5599	238294.2	74	7253	231349.2	74
Austria	3013	89592.6	65	2773	86739.4	66
Belgium (Fl)	2768	64177.2	71	2894	75069.0	71
Belgium (Fr)	2292	49897.7	60	2591	59269.8	60
Bulgaria	1798	140979.0	52	1973	147094.0	58
Canada	8219	377731.5	75	8362	377425.8	75
Colombia	2655	619462.0	71	2649	527145.4	71
Cyprus	2929	10033.0	55	2923	9347.0	55
Czech Republic	3345	152491.8	75	3327	152493.7	75
Denmark	2073	44980.0	75	2297	54172.1	75
England	1803	465457.5	64	1776	485280.1	64
France	3016	860657.2	67	2998	815509.8	68
Germany	2893	742346.4	69	2870	726087.5	69
Greece	3931	130221.7	75	3990	121910.7	75
Hong Kong	3413	88590.8	43	3339	88573.7	43
Hungary	3066	118726.6	75	2912	112436.1	75
Iceland	1957	4212.4	75	1773	4234.2	75
Iran, Islamic Rep.	3735	1052795.3	75	3694	935093.1	75
Ireland	3127	68476.8	66	3076	67643.8	66
Israel	.	.	.	1415	60584.3	23
Japan	5130	1562417.6	75	5141	1641941.4	75
Korea	2907	798409.3	75	2920	810403.8	75
Kuwait	.	.	.	1655	13093.0	36
Latvia (LSS)	2567	17041.1	64	2409	15414.0	64
Lithuania	2531	36551.0	73	2525	39700.0	73
Netherlands	2097	175419.1	48	1987	191662.7	48
New Zealand	3184	48507.6	75	3683	51133.3	75
Norway	2469	51165.0	72	3267	50223.8	74
Portugal	3362	146882.0	71	3391	137459.0	71
Romania	3746	295348.5	72	3725	296533.6	72
Russian Federation	4138	2168163.5	41	4022	2004792.2	41
Scotland	2913	61938.0	64	2863	64637.6	64
Singapore	3641	36181.0	69	4644	36538.5	69
Slovak Republic	3600	83074.1	73	3501	79766.4	73
Slovenia	2898	28048.9	61	2708	26010.7	61
South Africa	5301	649180.0	66	4491	766333.6	66
Spain	3741	549032.2	75	3855	547113.6	75
Sweden	2831	96493.9	75	4075	98193.1	60
Switzerland	4085	66681.1	75	4855	69732.5	75
Thailand	5810	680225.3	74	5833	657748.2	74
United States	3886	3156846.8	55	7087	3188296.6	55

3.8 Weight Variables Included in the Student Data Files³

There are several sampling weight variables included in the student data files. Some of these variables capture different aspects of the sampling process, and others constitute the sampling weights themselves. For the purpose of consistency, the variable names across both populations are identical so the explanation that follows applies to the Population 1 and Population 2 data files. The variables named in this section are included in the Student Background and the Performance Assessment data files. Their values will vary across these files even for the same students because the probability of selection for a student for each of these samples was different. The performance assessment sample was selected as a sub-sample of the written assessment sample and therefore the values for the different weighting factors and their adjustments are computed separately for the students within each file. The meaning and interpretation of the weights in each of the files remains the same. The weighting factors and their adjustments 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***

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***

This is the inverse of the probability of selection of the classroom within the school. In most cases the value of this variable is an integer, but it could take other values when more than one classroom is selected in the school. Although most countries selected classrooms within schools, this was not always the case. When a country selected students within the school without first selecting a specific classroom, or when there was only one classroom at the target grade, the value of this variable is set to 1 for all the students in the school. Since it was expected that only one classroom would be selected per grade within each school, there was no need to compute an adjustment factor for the classroom weight.

WGTFAC3 ***Student Weighting Factor***

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

³ For the user not familiar with the data files included in the TIMSS International Database we recommend reading Chapter 7 before proceeding with sections 3.8, 3.9, and 3.10.

WGTADJ3 Student Weighting Adjustment

This is an adjustment applied to the variable WGTADJ3 to account for non-participating students in the selected school and/or classroom. If we were to multiply the variables WGTADJ2, WGTADJ3, and WGTADJ3 and add 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 to compute a student's overall sampling weight. A two-stage sampling design was used in TIMSS: schools were first selected from a national list of schools; in the second stage classrooms were selected within these schools. Some countries used a third stage in which students were selected within classrooms. We compute the probability for selecting an individual student as the product of three independent events: selecting the school, the classroom, and the student. To obtain the probability of selection for an individual student we need to multiply three selection probabilities – school, classroom, and student – and their respective adjustment factors. The resulting product of these three probabilities gives us the individual probability of selection for the student. Inverting this probability give us the sampling weight for the student. The same result is achieved by multiplying the different weights of the different selection stages (school, classroom, student).

Three versions of the students' sampling weight are provided in the user database. All three give the same figures for statistics such as means and proportions, but vary for statistics such as totals and population sizes. Each one has particular advantages in certain circumstances.

TOTWGT Total Student Weight

This is obtained by simply multiplying the variables WGTADJ1, WGTADJ2, WGTADJ3, and WGTADJ3 for the student. The sum of these weights within a sample provides an estimate of the size of the population. Although this is a commonly used sampling weight, it sometimes adds 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 the population across all countries, using the variable TOTWGT as our weight variable will lead each country to contribute proportionally to its population size, with the large countries counting more than small countries. Although this might be desirable in some circumstances (e.g., when computing the 75th percentile for mathematics achievement for students around the world), 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 (TOTWGT). 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 obtained from each of these proportional transformations of the weights will be exactly the same. To this effect, two other weight variables are computed and included in the student data files. Each of these is computed for a specific purpose and will yield exactly the same results within each country, but will have some desirable properties when estimates across countries are computed or significance tests performed.

SENWGT *Senate Weight*

This variable is computed as

$$\text{TOTWGT} \left(\frac{1000}{\sum \text{TOTWGT}} \right)$$

within each country. The transformation of the weights will be different within each country, but in the end, the sum of the variable SENWGT within each country will add up to 1,000.⁴ The variable SENWGT, within each country, is proportional to TOTWGT by the ratio of 1,000 divided by the size of the population. These sampling weights can be used when international estimates are sought and the user wants to have each country contribute the same amount to the international estimate. When this variable is used as the sampling weight for international estimates, the contribution of each country is the same, regardless of the size of the population.

HOUWGT *House Weight*

This variable is computed as

$$\text{TOTWGT} \left(\frac{n}{\sum \text{TOTWGT}} \right)$$

within each country. The transformation of the weights will be different within each country, but in the end, the sum of the variables HOUWGT within each country will add up to the sample size for that country. The variable HOUWGT is proportional to TOTWGT 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 is used as the weighting variable. When performing analyses using such software, we recommend using the variable HOUWGT as the weight variable. Because of the clustering effect in most TIMSS samples, it may also be desirable to apply a correction factor such as a design effect to the HOUWGT variable.

3.9 Weight Variables Included in the Student-Teacher Linkage Files

The individual student sampling weights generally should be used when the user wants to obtain estimates at the student level. The exception is when student and teacher data are to be analyzed together. In this case, a separate set of weights have been computed to account for the fact that a student could have more than one mathematics or science teacher. These weight variables are included in the Student-Teacher Linkage file and are listed below.

⁴ In Israel and Kuwait only one grade was tested and the SENWGT adds to 500.

MATWGT

This weight is computed by dividing the sampling weight for the student by the number of mathematics teachers that the student has. This weight should be used whenever the user wants to obtain estimates regarding students and their mathematics teachers.

SCIWGT

This weight is computed by dividing the sampling weight for the student by the number of science teachers that the student has. This weight should be used whenever the user wants to obtain estimates regarding students and their science teachers.

TCHWGT

This weight is computed by dividing the sampling weight for the student by the number of mathematics and science teachers that the student has. This weight should be used whenever the user wants to obtain estimates regarding students and their mathematics and science teachers combined.

The Student-Teacher Linkage file also includes variables that indicate the number of mathematics teachers, the number of science teachers, and the number of mathematics and science teachers the student has.

3.10 Weight Variables Included in the School Data Files

The TIMSS samples are samples of students within countries. Although they are made up of a sample of schools within the countries, the samples of schools are selected so that the sampling of students, rather than the sampling of schools, is optimized. In other words, the samples are selected to make statements about the students in the country rather than about the schools in the country. To this effect, several weight variables are included in the school files. These variables 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.

STOTWGTL *Sum of the Student Weights for the Lower Grade*

STOSWGTU *Sum of the Student Weights for the Upper Grade*

STOTWGTE *Sum of the Student Weights for the Extra Grade⁵*

These variables are the sum of the weights of the students in the corresponding grade level within the school. If there are no students at this grade level, then the variable is 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.

⁵ Students in the extra grade are those students above the upper grade in the TIMSS sample. Sweden and Switzerland tested students at this grade level and those data are included in the database.

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. The four countries on a Southern Hemisphere school schedule (Australia, Korea, New Zealand, and Singapore) tested in September through November of 1994, which was the end of the school year in the Southern Hemisphere. The remaining countries tested their students at the end of the 1994-95 school year, most often in May and June of 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 designated School Coordinators. They also were responsible for arranging for the return of the testing materials from the school sites 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 for Populations 1 and 2 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, classrooms, and students, and ensure proper linkage among 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 once 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 administered the teacher questionnaires, 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.

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 and teachers, 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 and all of the performance assessment 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 mathematics item is shown below in Figure 4.1. In this guide, students received one point for a correct answer and then there were multiple diagnostic codes for incorrect answers.

The rubrics for more complicated items were correspondingly more complicated, having categories for full and partial credit. As shown in Figure 4.2, on both parts of this extended 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 Mathematics Item

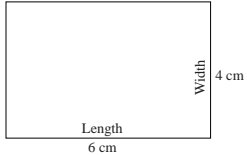
L16. Find x if $10x - 15 = 5x + 20$

Answer: _____

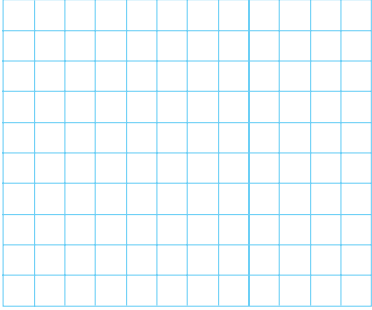
Code	Response
Correct Response	
10	7
Incorrect Response	
70	1 OR 2.33.. OR 3
71	Other incorrect numeric answers.
72	Any expression or equation containing x .
79	Other incorrect
Nonresponse	
90	Crossed out/erased, illegible, or impossible to interpret.
99	BLANK

Figure 4. 2
Example Coding Guide for Extended-Response Mathematics Item

U2.



a. In the space below, draw a new rectangle whose length is one and one half times the length of the rectangle above, and whose width is half the width of the rectangle above. Show the length and width of the new rectangle in centimeters on the figure.



b. What is the ratio of the area of the new rectangle to the area of the first one?

Show your work.

Note: There is no distinction made between responses with and without units.

A: Codes for Drawing

Code	Response
Correct Response	
20	9 cm and 2 cm. Correct drawing shown.
Partial Response	
10	9 cm and 2 cm. Drawing is incorrect or missing.
11	The length and/or width is not given or is incorrect. Correct drawing is shown.
Incorrect Response	
70	15 cm and 2 cm. Explicitly written or implicit from the drawing.
71	7.5 cm and 2 cm. Explicitly written or implicit from the drawing.
72	3 cm and 2 cm. Explicitly written or implicit from the drawing.
73	2 cm width and a length equal to any other numbers except those given above. Explicitly written or implicit from the drawing.
74	9 cm length and a width equal to any other numbers than those given above. Explicitly written or implicit from the drawing.
79	Other incorrect
Nonresponse	
90	Crossed out/erased, illegible, or impossible to interpret.
99	BLANK

Figure 4.2 (Continued)

B: Codes for ratio and areas

Code	Response
Correct Response	
20	3:4, 3/4 or equivalent. The areas are 18 cm ² and 24 cm ² . (No work needs to be shown. Also, areas do not need to be mentioned if the ratio is consistent with the areas of the given rectangle and the rectangle the student has drawn in U-2a)
21	The ratio is NOT 3:4 but areas and ratio of part (b) are consistent with response in part (a).
Partial Response	
10	4:3 or equivalent. (Ratio is reversed.) The areas are 18 cm ² and 24 cm ² .
11	An incorrect ratio or no ratio is given. The areas are 18 cm ² and 24 cm ² .
12	The difference between the areas, 6, is given instead of the ratio. The areas are 18 cm ² and 24 cm ² .
13	Areas are NOT 18 cm ² and 24 cm ² but are consistent with response in part a) and an incorrect ratio or no ratio is given.
14	Areas are NOT 24 cm ² and 18 cm ² but are consistent with response in part a) and a difference consistent with those areas is given instead of the ratio.
Incorrect Response	
70	Focuses exclusively on the ratios of lengths and widths between the given rectangle and the new rectangle. No areas are shown.
79	Other incorrect.
Nonresponse	
90	Crossed out/erased, illegible, or impossible to interpret.
99	BLANK

To meet the goal of reliable scoring, TIMSS had an ambitious training program. The training sessions were designed to assist representatives of the national centers who were then responsible for training personnel in their respective 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 2, a very high percentage of exact agreement was observed on the mathematics items for all countries, with averages across items for the correctness score ranging from 97% to 100% and an overall average of 99% across all 26 countries. The corresponding average across all countries for diagnostic code agreement was 95%, with a range of 89% to 99%. At Population 1, the average agreement for correctness on the mathematics items was very similar – 97% overall for the 16 countries, ranging from 88% to 100%. For diagnostic codes, the overall average was 93% and the range was from 79% to 99%.

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
Mathematics						
Population 1	97	88	100	93	79	99
Population 2	99	97	100	95	89	99
Science						
Population 1	94	83	99	87	72	98
Population 2	95	88	100	87	73	98

Note: Overall averages are based on all countries providing coding reliability data: Population 1 data are based on 16 countries. Population 2 data are based on 26 countries.

While the percentages of exact agreement for science items were somewhat lower than for the mathematics items, they were still quite good. For the 26 countries at Population 2, the averages across items for the correctness score ranged from 88% to 100%, and the overall average across the 26 countries was 95%. The overall average for diagnostic score agreement was 87%, with a range of 73% to 98%. At Population 1, the average agreement for correctness was 94%, ranging from 83% to 99%. The average exact agreement on diagnostic codes was 87% overall, with a range of 72% to 90%.

TIMSS also conducted a special cross-country coding reliability study at Population 2 in which 39 scorers from 21 of the participating countries coded common sets of student responses. In this study, an overall average percentage of exact agreement of 92% for correctness scores and 80% for diagnostic codes was obtained.

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 more than 40 participating countries, 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 and performance assessment tasks were originally developed in English and then translated by the participating TIMSS countries into more than 30 languages. 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, 1997). 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. 2
List of Deleted Cognitive Items

Country	Item	Subject	Variable Name
Population 2 - Written Assessment			
All	M09	Mathematics	BSMMM09
Austria	M05	Mathematics	BSMMM05
Belgium (FR)	N05	Science	BSMSN05
Colombia	O12	Science	BSMSO12
Cyprus	F03	Science	BSMSF03
	H01	Science	BSMSH01
	J01	Science	BSMSJ01
	J07	Science	BSMSJ07
	O11	Science	BSMSO11
	Q15	Science	BSMSQ15
Denmark	N16	Mathematics	BSMMN16
France	N05	Science	BSMSN05
	Q16	Science	BSMSQ16
Greece	V01	Mathematics	BSSMV01
Hungary	L17	Mathematics	BSMML17
	Q09	Mathematics	BSMMQ09
	Z01B	Science	BSESZ01
	Z01C	Science	BSESZ01
Iceland	O04	Mathematics	BSMMO04
	Q17	Science	BSSSQ17
Iran	F04	Science	BSMSF04
	N09	Science	BSMSN09
Israel	B05	Science	BSMSB05
Japan	J11	Mathematics	BSMMJ11
	N17	Mathematics	BSMMN17
	U02A	Mathematics	BSEMU02
	U02B	Mathematics	BSEMU02
Kuwait	D10	Mathematics	BSMMD10
	F07	Mathematics	BSMMF07
	I01	Mathematics	BSMMI01
	J05	Science	BSMSJ05
	J18	Mathematics	BSMMJ18
	O04	Mathematics	BSMMO04
Latvia	E12	Science	BSMSE12
Lithuania	F11	Mathematics	BSMMF11
Philippines	I03	Mathematics	BSMMI03
	J02	Science	BSMSJ02
Romania	A03	Mathematics	BSMMA03
Russian Federation	I19	Science	BSMSI19
Singapore	U02A	Mathematics	BSEMU02
	U02B	Mathematics	BSEMU02
Slovak Republic	K10	Science	BSSSK10
Thailand	K04	Mathematics	BSMMK04
	P04	Science	BSMSP04
	Q16	Science	BSMSQ16
United States	I13	Science	BSMSI13

Table 4.2
(Continued)

Country	Item	Subject	Variable Name
Population 1 - Written Assessment			
Cyprus	B04	Science	ASMSB04
England	T01	Mathematics	ASEMT01
Hong Kong	J05	Mathematics	ASMMJ05
Hungary	D01	Science	ASMSD01
	T04	Mathematics	ASEMT04
Iceland	B03	Science	ASMSB03
Iran	K04	Mathematics	ASMMK04
Japan	G03	Mathematics	ASMMG03
	J03	Mathematics	ASMMJ03
	L03	Mathematics	ASMML03
	L06	Mathematics	ASMML06
	P06	Science	ASMSP06
Korea	A01	Mathematics	ASMM A01
	R08	Science	ASMSR08
Latvia	G01	Mathematics	ASMMG01
	G02	Mathematics	ASMMG02
Thailand	B04	Science	ASMSB04
	U04	Mathematics	ASSMU04
	X01	Science	ASESX01
Populations 1 and 2 - Performance Assessment			
Colombia	Task M4 Item 6 (Pop2)	Mathematics	BSPM46
Cyprus	Task M2 Item 5 (Pop1)	Mathematics	ASPM25
Cyprus	Task S6 Item 1A (Pop1)	Science	ASPS61A

4.5.2 Background Questionnaire Items

As was the case with the test instruments, the international versions of the student, teacher, and school background questionnaires were also developed in English and then translated into other languages by the TIMSS countries. While the intent of TIMSS is to provide internationally comparable data for all variables, there are many contextual differences among countries so that the international version of the questions are 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 (Supplements 1 and 2), 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 items that were used for reporting the contextual data in the international reports. 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 3.¹

¹ 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}_{jk} , ($j = 1, \dots, I; k = 1, \dots, K_i$), 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 θ 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

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

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

where Ω 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 , $\mathbf{B}_i = (\mathbf{b}_{i1}, \mathbf{b}_{i2}, \dots, \mathbf{b}_{iD})^T$, 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, ξ , 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

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

And for a response vector we have:

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

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

The difference between the unidimensional model and the multidimensional model is that the ability parameter is a scalar, θ , in the former, and a D by one-column vector, θ , 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, θ . The complete definition of the TIMSS model, therefore, requires the specification of a density, $f_{\theta}(\theta; \alpha)$, for the latent variable, θ . We use \mathbf{a} to symbolize a set of parameters that characterize the distribution of θ . 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 m and variance s^2 . That is:

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

$$\text{or equivalently} \quad \theta = \mu + E \quad (9)$$

$$\text{where} \quad E \sim N(0, \sigma^2).$$

Adams, Wilson, and Wu (1997) discuss how a natural extension of (8) is to replace the mean, m , 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 s^2 so that (10) is equivalent to

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

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

The TIMSS scaling model takes the generalization one step further by applying it to the vector valued θ rather than the scalar valued θ resulting in the multivariate population model

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

where $\boldsymbol{\gamma}$ is a $u \times d$ matrix of regression coefficients, $\boldsymbol{\Sigma}$ 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 $\boldsymbol{\gamma}$, $\boldsymbol{\Sigma}$, and \mathbf{x} .

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 γ , Σ , 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:

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

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

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

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

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

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 θ_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, $\{\varphi_{nm}\}_{m=1}^M$ 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

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

At the same time the values

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

are calculated, so that we obtain the set of pairs $(\varphi_{nm}, p_{mn}/\mathfrak{S})_{m=1}^M$ that can be used as an approximation to the posterior density (23); and the probability that φ_{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, $\varphi_{n_{i_0}}$, 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 was undertaken using a subsample of students drawn from the samples of the participating countries. These samples were called the *international calibration* samples. In a second step the model was fitted separately to the data for each country within the item parameters fixed at values estimated in the first step.

There were three principal reasons for using an international calibration sample for estimating international item parameters. First, it seemed unnecessary to estimate parameters using the complete data set; second, drawing equal sized subsamples from each country for inclusion in the international calibration sample ensured that each country was given equal weight in the estimation of the international parameters; and third, the drawing of appropriately weighted samples meant that weighting would not be necessary in the international scaling runs.

5.8.1 Drawing The International Calibration Sample

For each target population, samples of 600 tested students were selected from the database for each participating country. This generally led to roughly equal samples from each target grade. For Israel, where only the upper grade was tested, the sample size was reduced to 300 tested students. The sampled students were selected using a probability-proportionality-to size systematic selection method. The overall sampling weights were used as measures of size for this purpose. This resulted in equal selection probabilities, within national samples, for the students in the calibration samples. The Populations 1 and 2 international calibration samples contained 14,700 and 23,100 students respectively.

5.8.2 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 these scores 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, with the exception of the countries that tested only one grade. Countries that tested only one grade had only half the contribution of

the remaining countries. The contribution of the students from each grade within each country was proportional to the number of students at each grade level within the country. The transformation applied to the logit scores was

$$S_{ijk} = 500 + 100 * \left(\frac{\theta_{ijk} - \bar{\theta}_j}{SD_{\theta_j}} \right)$$

where S_{ijk} is the standardized scale score for student i , in plausible value j , in country k ; θ_{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_{θ_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 exactly 500. This is because one or two countries were not included in the computation of $\bar{\theta}_j$ since the scaling of their data was not completed.

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 50, the score was recoded to 50, therefore making 50 the lowest score on the transformed scale. This happened in very few cases across the sample. The highest transformed scores did not exceed 1,000 points, so the transformed values were left untouched at the upper end of the distribution.

Five plausible values were drawn for each 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. In the TIMSS international reports, plausible values were used only to compute total scores in mathematics and science. The reliability of these total scores, as reflected in the intercorrelations between the five plausible values, was found to be sufficiently high that the imputation error component could be ignored. For the purpose of reporting international achievement, therefore, only the first plausible value was used. However, all five values are provided in the International Database.

Chapter 6

Student Achievement Scores

The TIMSS international database contains several student-level achievement scores. These scores were computed at different stages of the study 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 they can be used by users of the database. For identification purposes, the first letter for the variable name identifies the population for which the score was computed. The scores computed for Population 1 have the letter A as the first character in their name and scores for Population 2 have the letter B as the first character. This convention was followed with other background and derived variables and with the files included in the database.

6.1 Achievement Scores in the Student Files¹

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

ASMSCPT	Number of raw score points obtained on the mathematics items – Population 1
ASSSCPT	Number of raw score points obtained on the science items – Population 1
BMSCPT	Number of raw score points obtained on the mathematics items – Population 2
BSSSCPT	Number of raw score points obtained on the science items – Population 2

After the items were recoded 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 receive 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 determines the number of score points assigned to an open-ended item. A description of the algorithm used to score the items can be found in Chapter 9 in the section “Scoring the Items.”

Although these scores can be used to compare students’ performances on the same booklet, they should not be used to compare students’ performances across different booklets. Different booklets contain different numbers of items for each subject, and the specific items contained in one booklet had varying difficulties. It is recommended that these scores be used only to verify whether the items have been recoded correctly when a user decides to recode the items to their level of correctness. Raw scores can be found in the Student Background data files and in the Written Assessment data files.

¹ For the user not familiar with the data files included in the International Database we recommend reading Chapter 7 before proceeding with sections 6.1 and 6.2.

ASMSTDR	Standardized mathematics raw score – Population 1
ASSSTDR	Standardized science raw score – Population 1
BSMSTDR	Standardized mathematics raw score – Population 2
BSSSTDR	Standardized science raw score – Population 2

Because of the difficulty in making any comparisons across the test booklets using only the number of raw score points obtained on a set of items, raw scores were standardized by booklet to provide a simple score which could be used in comparisons across booklets in preliminary analyses. The standardized score was computed so that the weighted mean score within each booklet was equal to 50, and the weighted standard deviation was equal to 10.

These standardized raw scores were used in the initial item analysis for computing the discrimination coefficients for each of the items in the test. This initial item analysis was conducted prior to scaling the test items. The standardized raw scores can be found in the Student Background data files and in the Written Assessment data files.

ASMNRSC	National Rasch Mathematics Score (ML) – Population 1
ASSNRSC	National Rasch Science Score (ML) – Population 1
BSMNRSC	National Rasch Mathematics Score (ML) – Population 2
BSSNRSC	National Rasch Science Score (ML) – Population 2

The national Rasch scores were also designed for preliminary analyses. These provided a basic Rasch score for preliminary analyses within countries, but could not be used for international comparisons since each country has been assigned the same mean score. The national Rasch scores were computed by standardizing mathematics and science logit scores to have a weighted mean of 150 and a standard deviation of 10 within each country. The logit scores were computed using the Quest Rasch analysis software; Quest provides maximum likelihood (ML) estimates of a scaled score, based on the Rasch model, for the performance of the students on a set of items. The computation 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 for each country using all available item responses for the country and centering the item difficulty around zero. When computing the item difficulties, responses marked as "not reached" were treated as items that were not administered. This avoids giving inflated item difficulties to the items located at the end of the test in cases where students systematically do not reach the end of the test. These item difficulties were then used to compute logit scores for each student.

When computing the student logit scores the responses marked as "not reached" were treated as incorrect responses. 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. Since 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, respectively. These logit scores were then standardized to have a weighted mean of 150 and a standard deviation of 10.

The 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.

ASMPV01-ASMPV05	Mathematics Plausible Value 1 to Plausible Value 5 – Population 1
ASSPV01-ASSPV05	Science Plausible Value 1 to Plausible Value 5 – Population 1
BSMPV01-BSMPV05	Mathematics Plausible Value 1 to Plausible Value 5 – Population 2
BSSPV01-BSSPV05	Science Plausible Value 1 to Plausible Value 5 – Population 2

As described in chapter 5, TIMSS made use of multiple imputation or “plausible values” methodology to provide estimates of student proficiency in mathematics and science. Because of the error involved in the imputation process, TIMSS produced not one but five imputed values for each student in mathematics and science. 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 did not allow for all the items to be administered to each student. A plausible value is an estimate of how the individual student would have performed on a test that included all possible 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 improve the estimation of population parameters. They were developed during the analysis of the 1983-84 NAEP data in order to improve estimates of population distributions. 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. 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 from replication to replication may be influenced by unreliability in the achievement measures, and considered to be suspect. In the TIMSS international reports, the reliability of the achievement measures, as reflected in the inter correlations between the five plausible values was found to be sufficiently high that the imputation error could be ignored. For the purpose of reporting international achievement, therefore, only the first plausible value was used. However, all five values are provided in the International Database for use by other analysts. The plausible values are included in the Student Background data files and in the Student-Teacher Linkage files.

AIMATEAP	International Mathematics Score (EAP) – Population 1
AISCIEAP	International Science Score (EAP) – Population 1
BIMATEAP	International Mathematics Score (EAP) – Population 2
BISCIEAP	International Science Score (EAP) – Population 2

International expected a posteriori (EAP) scores are the average of the distribution from which the corresponding plausible values are drawn. Although the average of an individual's distribution of plausible values may be a better estimate of the individual's proficiency than a single plausible value, in general it will not produce consistent population estimates or estimates of their error variance. The EAP scores can be used if individual scores are necessary, but should not be used to obtain population estimates of proficiency. These scores are included in the Student Background, the student Written Assessment, and the student Performance Assessment files.

AIMATSCR	International Mathematics Achievement Score – Population 1
AISCISCR	International Science Achievement Score – Population 1
BIMATSCR	International Mathematics Achievement Score – Population 2
BISCISCR	International Science Achievement Score – Population 2

These are the international mathematics and science achievement scores used to report achievement at the international level. They correspond to the first plausible value in each subject area. 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. International proficiency scores are included in the Student Background, the student Written Assessment, the student Performance Assessment files, and the Student-Teacher Linkage files. The mean of these scores within each school, by grade level, are also included in the School Background data files. Tables 6.1, 6.2, 6.3, and 6.4 present weighted statistics for these scores within each participating country.

Table 6.1**Descriptive Statistics for the International Mathematics Achievement Scores for Population 1 (Variable: AIMATSCR)**

Country	Lower Grade				Upper Grade			
	Mean	Standard Deviation	Minimum	Maximum	Mean	Standard	Minimum	Maximum
Australia	483.82	90.1	172.95	798.72	546.69	91.7	166.48	875.19
Austria	487.02	82.9	209.90	764.08	559.25	78.8	281.67	804.63
Canada	469.45	78.2	208.37	738.88	532.13	83.7	235.42	811.92
Cyprus	430.42	77.4	99.75	724.48	502.42	86.4	182.30	786.59
Czech Republic	497.19	82.8	171.76	812.04	567.09	86.3	260.14	859.06
England	456.50	87.4	134.13	801.16	512.70	91.2	221.63	842.78
Greece	428.05	85.2	150.01	690.19	491.90	89.9	147.10	787.93
Hong Kong	523.95	72.5	199.78	740.00	586.64	79.0	215.81	837.71
Hungary	476.10	89.3	99.86	753.77	548.36	88.1	232.46	852.64
Iceland	410.11	67.2	198.48	651.12	473.75	72.3	230.84	737.05
Iran, Islamic Rep.	377.98	65.6	107.41	622.65	428.50	69.1	185.25	678.57
Ireland	475.77	81.0	230.00	719.64	549.94	85.3	247.73	786.11
Israel	531.40	84.6	249.70	832.55
Japan	537.91	75.3	182.47	767.02	596.83	81.2	307.27	869.28
Korea	560.90	70.2	304.21	798.87	610.70	73.9	313.10	871.17
Kuwait	400.16	66.8	170.31	679.28
Latvia (LSS)	463.29	80.9	200.57	766.15	525.38	84.9	224.30	851.19
Netherlands	492.88	64.9	231.78	688.67	576.66	70.5	339.47	839.41
New Zealand	439.53	82.3	125.33	690.97	498.72	89.6	137.19	822.00
Norway	421.30	71.5	161.52	656.86	501.87	74.1	217.24	717.43
Portugal	425.26	85.4	70.51	676.83	475.42	80.3	175.98	734.25
Scotland	458.03	79.5	183.19	745.03	520.41	89.4	198.64	844.11
Singapore	552.08	99.9	178.79	876.85	624.95	104.2	208.88	944.04
Slovenia	487.64	77.3	208.50	762.90	552.41	82.4	275.41	806.96
Thailand	444.32	71.2	175.21	657.46	490.19	69.5	260.72	693.78
United States	479.80	81.9	189.54	739.31	544.57	85.1	175.26	816.43

SOURCE: IEA Third International Mathematics and Science Study (TIMSS) 1994-95.

Table 6.2
Descriptive Statistics for the International Science Achievement Scores for
Population 1 (Variable: AISCISCR)

Country	Lower Grade				Upper Grade			
	Mean	Standard Deviation	Minimum	Maximum	Mean	Standard	Minimum	Maximum
Australia	510.24	97.8	109.91	852.73	562.55	93.3	168.80	911.57
Austria	504.55	87.6	119.53	748.09	564.75	79.8	277.16	807.41
Canada	490.42	88.1	132.18	791.18	549.26	85.9	160.12	862.82
Cyprus	414.72	73.1	94.56	681.38	475.44	75.7	215.03	717.97
Czech Republic	493.67	84.6	201.74	786.70	556.51	81.2	224.10	843.97
England	499.23	100.6	77.98	796.94	551.46	96.0	219.91	857.75
Greece	445.87	82.5	146.54	703.33	497.18	83.0	102.13	744.22
Hong Kong	481.56	73.5	176.38	703.72	532.98	77.8	225.17	824.38
Hungary	464.42	88.8	165.66	725.85	531.59	80.7	223.82	826.10
Iceland	435.38	82.1	164.43	681.32	504.74	84.9	213.72	800.97
Iran, Islamic Rep.	356.16	76.0	50.00	658.64	416.50	74.5	170.54	678.76
Ireland	479.11	87.8	146.67	755.30	539.48	84.9	230.41	786.25
Israel	504.82	86.3	184.34	787.68
Japan	521.78	72.6	224.27	754.71	573.61	72.8	253.39	800.76
Korea	552.92	71.4	273.99	814.67	596.90	68.1	332.21	832.19
Kuwait	401.29	85.2	57.95	744.01
Latvia (LSS)	465.25	82.6	228.76	778.61	512.17	83.8	234.55	871.62
Netherlands	498.84	63.4	132.77	747.38	556.66	65.7	293.48	801.08
New Zealand	473.05	99.6	94.41	793.59	531.03	97.0	217.08	867.40
Norway	450.28	89.8	104.47	749.08	530.27	86.2	176.79	788.49
Portugal	422.99	95.6	50.00	702.77	479.79	84.4	92.77	741.33
Scotland	483.85	95.0	110.84	813.50	535.62	93.4	195.12	816.46
Singapore	487.74	98.8	105.30	812.96	546.69	97.2	186.63	887.39
Slovenia	486.94	78.0	207.33	740.59	545.68	76.0	226.89	775.09
Thailand	432.56	77.5	140.87	704.27	472.86	70.8	171.15	736.01
United States	511.23	93.8	124.71	780.52	565.48	95.0	117.45	882.36

SOURCE: IEA Third International Mathematics and Science Study (TIMSS) 1994-95.

Table 6.3

Descriptive Statistics for the International Mathematics Achievement Scores for Population 2 (Variable: BIMATSCR)

Country	Lower Grade				Upper Grade			
	Mean	Standard Deviation	Minimum	Maximum	Mean	Standard	Minimum	Maximum
Australia	497.90	91.9	209.10	833.42	529.63	98.2	183.17	863.47
Austria	509.17	85.2	180.91	823.31	539.43	92.0	210.89	853.29
Belgium (Fl)	557.62	76.7	303.11	802.15	565.18	91.8	307.12	863.33
Belgium (Fr)	507.14	78.0	208.03	744.33	526.26	86.2	222.72	815.95
Bulgaria	513.80	103.3	178.70	814.00	539.66	110.4	216.59	916.03
Canada	493.99	80.0	188.74	785.69	527.24	86.3	222.14	819.10
Colombia	368.51	63.0	141.82	588.92	384.76	64.2	155.00	605.36
Cyprus	445.67	82.1	190.75	741.01	473.59	87.7	165.24	769.21
Czech Republic	523.39	89.3	183.86	845.02	563.75	93.7	277.26	887.44
Denmark	464.85	77.7	161.66	703.72	502.29	83.7	248.96	795.24
England	476.15	89.9	138.71	801.00	505.73	93.0	244.42	831.36
France	492.19	73.9	220.14	755.28	537.83	76.0	312.55	801.00
Germany	484.38	85.1	161.29	792.75	509.16	89.6	184.60	816.06
Greece	439.89	85.4	121.62	744.68	483.90	88.3	165.39	788.45
Hong Kong	563.60	99.2	200.30	914.83	588.02	100.9	219.51	939.24
Hungary	501.80	90.7	163.97	822.51	537.26	93.0	200.79	859.32
Iceland	459.43	67.8	234.71	712.27	486.78	76.0	263.28	740.84
Iran, Islamic Rep.	400.95	56.9	224.88	601.82	428.33	59.3	220.79	631.53
Ireland	499.69	86.8	217.36	758.57	527.40	92.5	245.13	843.90
Israel	521.59	92.0	184.13	842.65
Japan	571.09	96.0	259.88	920.00	604.77	101.9	244.70	891.33
Korea	577.06	104.5	186.96	889.86	607.38	108.8	279.07	987.44
Kuwait	392.18	57.7	176.20	596.44
Latvia (LSS)	461.59	76.5	218.93	741.47	493.36	82.1	200.63	774.17
Lithuania	428.19	75.2	141.60	702.94	477.23	80.2	190.89	752.23
Netherlands	515.98	78.9	201.36	751.23	540.99	88.9	229.74	841.28
New Zealand	471.70	86.8	144.17	784.26	507.80	90.3	229.76	820.50
Norway	460.66	76.0	172.39	740.18	503.29	83.9	210.57	782.49
Portugal	423.15	59.7	199.12	640.48	454.45	63.8	261.59	671.89
Romania	454.42	84.2	182.73	759.93	481.55	89.2	224.48	787.16
Russian Federation	500.93	88.2	173.70	818.45	535.47	91.9	203.13	852.58
Scotland	462.89	81.7	221.01	705.39	498.46	87.4	234.26	795.92
Singapore	601.04	93.2	269.85	917.08	643.30	88.2	315.41	957.95
Slovak Republic	507.84	85.1	179.74	820.82	547.11	92.1	219.41	860.50
Slovenia	498.17	81.9	232.68	796.55	540.80	87.8	284.99	838.63
South Africa	347.51	63.4	110.23	576.66	354.14	65.3	149.68	583.28
Spain	447.97	69.5	180.18	703.79	487.35	73.5	266.48	743.53
Sweden	477.47	76.9	164.46	776.91	518.64	85.3	205.78	818.23
Switzerland	505.55	75.4	241.46	800.77	545.44	87.6	240.06	841.72
Thailand	494.77	78.6	192.48	731.30	522.37	85.8	219.60	810.88
United States	475.68	89.3	143.07	792.40	499.76	90.9	172.03	816.67

SOURCE: IEA Third International Mathematics and Science Study (TIMSS) 1994-95.

Table 6.4
Descriptive Statistics for the International Science Achievement Scores for
Population 2 (Variable: BISCISCR)

Country	Lower Grade				Upper Grade			
	Mean	Standard Deviation	Minimum	Maximum	Mean	Standard	Minimum	Maximum
Australia	504.37	102.6	180.52	869.99	544.60	106.5	163.23	910.31
Austria	518.76	94.0	163.74	856.53	557.68	98.2	204.41	897.20
Belgium (Fl)	528.67	73.1	283.47	802.84	550.28	81.1	311.38	824.10
Belgium (Fr)	442.05	78.6	140.23	679.01	470.57	85.5	170.06	702.27
Bulgaria	530.85	103.0	226.14	907.46	564.83	110.8	170.28	940.45
Canada	499.19	89.7	162.34	821.00	530.89	93.0	193.94	852.60
Colombia	387.48	72.2	115.89	648.38	411.09	75.5	138.97	671.46
Cyprus	419.92	87.0	95.88	666.34	462.57	88.5	143.53	706.24
Czech Republic	532.93	81.8	220.04	776.61	573.94	87.0	261.59	872.36
Denmark	439.02	86.3	112.66	683.34	478.30	88.0	157.58	790.49
England	512.02	101.1	208.51	811.17	552.07	105.8	230.34	916.19
France	451.46	74.2	173.64	712.07	497.65	76.8	221.46	759.89
Germany	499.50	95.6	135.63	846.32	531.33	101.4	168.84	879.53
Greece	448.65	87.4	128.24	753.26	497.35	85.0	182.48	802.98
Hong Kong	495.26	86.4	174.27	801.59	522.12	88.9	201.45	773.11
Hungary	517.90	91.4	183.04	837.61	553.68	90.0	217.89	872.46
Iceland	461.96	75.1	173.54	737.66	493.55	79.0	205.19	769.31
Iran, Islamic Rep.	436.33	71.6	175.26	688.21	469.74	72.6	206.79	719.75
Ireland	495.21	91.0	148.15	823.85	537.78	95.6	196.09	866.90
Israel	524.48	103.7	142.97	889.63
Japan	531.04	86.2	209.11	839.82	570.98	90.0	247.92	822.67
Korea	535.02	91.8	198.12	858.62	564.87	93.7	227.39	887.88
Kuwait	429.57	73.9	158.20	687.41
Latvia (LSS)	434.91	78.0	141.27	652.04	484.76	81.2	191.58	715.31
Lithuania	403.06	79.5	112.16	682.08	476.45	81.0	184.62	754.54
Netherlands	517.25	79.0	218.16	803.21	560.07	85.0	266.54	847.36
New Zealand	480.96	96.5	173.38	826.02	525.45	99.8	164.41	871.56
Norway	483.19	84.8	168.05	724.14	527.19	87.3	210.69	828.46
Portugal	427.93	71.5	165.14	682.23	479.63	73.8	218.01	735.10
Romania	451.58	100.0	85.02	805.22	486.05	101.7	116.79	836.99
Russian Federation	483.96	94.3	132.89	820.14	538.08	95.3	187.59	874.85
Scotland	468.14	93.7	116.03	747.61	517.18	99.8	162.60	856.16
Singapore	544.70	100.2	184.41	888.64	607.33	95.4	251.63	950.77
Slovak Republic	509.67	85.0	188.49	821.07	544.41	92.0	218.76	855.94
Slovenia	529.93	85.9	267.64	772.02	560.05	87.7	239.91	865.46
South Africa	317.14	92.0	50.00	656.22	325.90	98.8	50.00	664.24
Spain	477.15	79.9	193.70	753.34	517.05	77.9	228.37	792.09
Sweden	488.43	84.3	158.51	805.67	535.40	90.3	204.14	851.31
Switzerland	483.66	81.9	164.47	734.92	521.74	90.9	198.90	829.91
Thailand	492.90	69.1	229.51	744.08	525.43	72.3	263.18	777.75
United States	508.24	105.0	121.88	876.38	534.45	105.7	148.11	902.61

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

6.2 Achievement Scores in the School Background Files

Although the achievement scores are computed at the individual level, several summary variables of achievement are included in the School Background data files. These correspond to the average achievement scores within the school, in mathematics and science, by grade level. The variables are identified below.

ACLGMAT	Mean Mathematics Achievement at the Lower Grade – Population 1
ACLGSCI	Mean Science Achievement at the Lower Grade – Population 1
ACUGMAT	Mean Mathematics Achievement at the Upper Grade – Population 1
ACUGSCI	Mean Science Achievement at the Upper Grade – Population 1
BCLGMAT	Mean Mathematics Achievement at the Lower Grade – Population 2
BCLGSCI	Mean Science Achievement at the Lower Grade – Population 2
BCUGMAT	Mean Mathematics Achievement at the Upper Grade – Population 2
BCUGSCI	Mean Science Achievement at the Upper Grade – Population 2
BCEGMAT	Mean Mathematics Achievement at the Extra Grade – Population 2
BCEGSCI	Mean Science Achievement at the Extra Grade – Population 2

Note that, although most countries tested across two grades per the target population definition, Switzerland and Sweden opted to test students in an extra grade at Population 2, and therefore the variables *BCEGMAT* and *BCEGSCI* are set to missing in all but these two countries.

Chapter 7

Content and Format of Database Files

7.1 Introduction

This chapter describes the content and format of the TIMSS International Database for Population 1 and Population 2. 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.¹ 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 4 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, teacher, and school background questionnaires. To investigate the match between the TIMSS achievement tests 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. 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, teachers, and schools.

¹ 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 and II* (Martin and Kelly, 1996; 1997) and *TIMSS: Quality Assurance in Data Collection* (Martin and Mullis, 1996).

7.2 Data Files

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

- Background Files
- Assessment Files
- Coding Reliability Files
- Student-Teacher Linkage Files

These files and the variables contained in each are described in Sections 7.2.1 through 7.2.4.

Data files are provided for each country that participated in TIMSS at Population 1 and Population 2 and for which internationally comparable data are available. At Population 1 there are nine different data file types and at Population 2 there are ten different file types, reflecting achievement data collected from students (both for the written assessment and performance assessment); background data collected from students, teachers, and schools; student-teacher linkage information; 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 1 and Population 2 data files provided.

Table 7.1
TIMSS Population 1 and Population 2 Data Files

File Type	Population 2 File Name	Population 1 File Name
Student Written Assessment File	BSA<Country>1.DAT	ASA<Country>1.DAT
Student Background File	BSG<Country>1.DAT	ASG<Country>1.DAT
Teacher Background File(s)	BTM<Country>1.DAT - Mathematics BTS<Country>1.DAT - Science	ATG<Country>1.DAT
School Background File	BCG<Country>1.DAT	ACG<Country>1.DAT
Student - Teacher Linkage File	BLG<Country>1.DAT	ALG<Country>1.DAT
Student Written Assessment Reliability File	BSR<Country>1.DAT	ASR<Country>1.DAT
Student Performance Assessment File	BSP<Country>1.DAT	ASP<Country>1.DAT
Student Performance Assessment Reliability File	BSQ<Country>1.DAT	ASQ<Country>1.DAT
School Performance Assessment File	BCT<Country>1.DAT	ACT<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 and the available files 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
Country Identification and Inclusion Status in Population 1 and Population 2 Data Files

Country	Abbreviation	Numeric Code	Inclusion in Population 1 Data Files									Inclusion in Population 2 Data Files								
			A S G	A S A	A T G	A L G	A C G	A C T	A S P	A S Q	A S R	B S G	B S A	B T M	B T S	B L G	B C G	B C T	B S P	B S Q
Australia	AUS	036	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓
Austria	AUT	040	✓	✓	✓	✓	✓													
Belgium (Fi)	BFL	056																		✓
Belgium (Fr)	BFR	057																		
Bulgaria	BGR	100										✓								✓
Canada	CAN	124	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Colombia	COL	170										✓	✓	✓	✓	✓	✓	✓	✓	✓
Cyprus	CYP	196	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓	✓		✓		
Czech Republic	CSK	200	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Denmark	DNK	208										✓	✓	✓	✓	✓				
England	GBR	826	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓		✓		✓
France	FRA	250										✓	✓	✓	✓	✓				✓
Germany	DEU	280										✓	✓	✓	✓	✓				✓
Greece	GRC	300	✓	✓	✓	✓	✓					✓	✓	✓	✓	✓				
Hong Kong	HKG	344	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Hungary	HUN	348	✓	✓	✓	✓	✓					✓	✓	✓	✓	✓				
Iceland	ISL	352	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓				✓
Iran, Islamic Rep.	IRN	364	✓	✓	✓	✓	✓		✓		✓	✓	✓	✓	✓	✓		✓		✓
Ireland	IRL	372	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓				✓
Israel	ISR	376	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓	✓		✓		
Japan	JPN	392	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓				✓
Korea	KOR	410	✓	✓	✓	✓	✓					✓	✓	✓	✓	✓				
Kuwait	KWT	414	✓	✓	✓	✓	✓					✓	✓	✓	✓	✓				
Latvia	LVA	428	✓	✓	✓	✓	✓					✓	✓	✓	✓	✓				
Lithuania	LTU	440										✓	✓	✓	✓	✓				
Netherlands	NLD	528	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
New Zealand	NZL	554	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓		✓
Norway	NOR	578	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓		✓	✓	✓
Philippines	PHL	608										✓	✓	✓	✓	✓				
Portugal	PRT	620	✓	✓	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Romania	ROM	642										✓	✓	✓	✓	✓	✓	✓		
Russian Federation	RUS	643										✓	✓	✓	✓	✓				✓
Scotland	SCO	827	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓	✓	✓	✓	✓
Singapore	SGP	702	✓	✓	✓	✓	✓				✓	✓	✓	✓	✓	✓		✓	✓	✓
Slovak Republic	SLV	201										✓	✓	✓	✓	✓				✓
Slovenia	SVN	890	✓	✓	✓	✓	✓		✓			✓	✓	✓	✓	✓		✓		
South Africa	ZAF	717										✓	✓	✓	✓	✓				
Spain	ESP	724										✓	✓	✓	✓	✓		✓	✓	✓
Sweden	SWE	752										✓	✓	✓	✓	✓		✓	✓	✓
Switzerland	CHE	756										✓	✓	✓	✓	✓	✓	✓	✓	✓
Thailand	THA	764	✓	✓	✓	✓	✓					✓	✓	✓	✓	✓				
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 three different types of TIMSS background files – student, teacher, 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. One version of the questionnaire was administered in Population 1 and two versions were administered in Population 2. In Population 2, one version was tailored for educational systems where science is taught as an integrated subject (non-specialized version). The second version was tailored for educational systems where the sciences (biology, earth science, physics, and chemistry) are taught separately (specialized version). Table 2.8 in Chapter 2 shows which countries administered the specialized and non-specialized versions of the questionnaire. Although most countries chose to use one version of the questionnaire, some countries opted to use both. The variable ITQUEST identifies the version of the questionnaire that was administered to the students in Population 2. For students who were administered the non-specialized version, all questions that were given only in the specialized version were coded as not administered. For students who were assigned a specialized version of the questionnaire, all questions that were asked only in the non-specialized version were coded as not administered. The Student Background files also contain a series of identification variables, link variables, sampling variables, achievement variables, and the derived variables that were used for the creation of the international reports.

7.2.1.2 Teacher Background File

The mathematics and science teachers of the students who were sampled for TIMSS were administered at least one questionnaire with questions pertaining to their background and their teaching practices in the classes of the sampled students. Each teacher was asked to respond to a questionnaire for each class taught that contained sampled students. The Teacher Background files contain one record for each of the classes taught by either a mathematics or a science teacher. In some cases, although the teacher was to respond to more than one questionnaire, responses to only one were obtained. In these cases, there were as many records entered in the teacher file as classes were taught by the teacher, and the background information from the complete questionnaire was entered into these teacher records. In Population 1 a single questionnaire was administered since both mathematics and science are usually taught by the same teacher at this age level; the responses to this questionnaire can be found in one file. There were two questionnaires administered in Population 2 – one for the mathematics teachers and one for the science teachers. The data from these questionnaires are found in separate files. Variable names for questions asked in both questionnaires are the same.

In the Teacher Background data files each teacher was assigned a unique identification number (IDTEACH) and a Teacher Link Number (IDLINK) that is specific to the class taught by the teacher and to which the information in the data record corresponds. The IDTEACH and IDLINK combination uniquely identifies a teacher teaching one specific class. So, for example, students linked to teachers identified by the same IDTEACH but different IDLINK are taught by the same teacher but in different classes. The Teacher Background files cannot be merged directly onto the

student data files and they do not contain sampling information or achievement scores. It is important to note that the Teacher Background data files do not constitute a representative sample of teachers in a country, but consist rather of the teachers who teach a representative sample of students. The teacher data should therefore be analyzed only in conjunction with the Student-Teacher Linkage file. The Teacher Background data files contain a series of other identification variables, link variables, and the derived variables that were used for the creation of the international reports

7.2.1.3 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. That file also contains a series of identification variables, link variables, sampling variables, and achievement variables. The school data files can be merged with the student data files by using the country and school identification variables.

7.2.1.4 Identification Variables

In all background files, several identification variables are included that provide information used to identify students, teachers, 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, Teacher, and School Background Files

IDCOUNTRY	Three-digit country identification code (see Table 7.2 for 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 (1 or 2).
IDSTRAT	This variable identifies the sampling stratum within the country. This variable was used in some but not all countries to identify specific strata within the population for that country.
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 IDCOUNTRY and IDSCHOOL combination.

Additional Variables Included in the Student Background Files

IDSTUD	Identification number that uniquely identifies each student in the country sampled. The variable IDSTUD is a hierarchical identification number. It is formed by the combination of the variables IDSCHOOL and IDCLASS, followed by a two-digit sequential number within each classroom. Students can be uniquely identified in the database by the combination of IDSTUD and IDCOUNTRY.
--------	--

IDCLASS	Identification number that uniquely identifies the sampled class within the school.
IDBOOK	Identifies the specific test booklet (1-8) that was administered to the student.
IDGRADER	Indicates whether the student was selected from the upper or the lower grade of the target population.
IDGRADE	Indicates the actual grade denomination within the country.

Additional Variables Included in the Teacher Background Files

IDTEACH	Identification number that uniquely identifies the selected teacher within the school. It is a hierarchical identification number formed by the combination IDSCHOOL and a two-digit sequential number within each school. This variable is unique to each teacher within each country but is not unique in the teacher file.
IDLINK	This variable uniquely identifies the class for which the teacher answered the questionnaire. The combination of variables IDCNTRY, IDTEACH, and IDLINK uniquely identifies a teacher-class combination in the database.
IDGRADE	The target grades for which the teacher answered the questionnaire.
IDGRADER	Indicates the TIMSS grade level associated with each teacher (1 = lower grade; 2 = upper grade)
IDSUBJECT	The subject(s) taught by the teacher (mathematics, science or both).

Additional Variable Included in the School Background Files

IDGRADER	Indicates the TIMSS grade levels contained in each school (1 = lower-grade school only; 2 = upper-grade school only; 3 = both lower and upper grades)
----------	---

In the Student Background file, the IDSTUD variable provides a unique identification number to identify each student within each country. Since teachers may teach more than one class, the IDTEACH and IDLINK combinations in the Teacher Background files provide a unique identification for each teacher teaching a specific class. Teacher background variables are linked to appropriate students using the Student-Teacher Linkage file described in Section 7.2.4. The variable IDSCHOOL, contained in all three 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 or teachers.

In the Teacher Background data file, the identification variable IDSUBJECT may be used to identify teachers as either mathematics only, science only, or mathematics/science teachers. For Population 2, separate background files are provided for mathematics and science teachers. At Population 1, a single questionnaire was given, and only one Teacher Background file is

provided, including all Population 1 teachers. In some countries, however, a portion or even all of the students were taught mathematics and science by different teachers. Therefore, the IDSUBJECT variable was used to identify the appropriate teachers to include in background data tables in the international mathematics and science reports for Population 1.

7.2.1.5 Achievement Scores

Student-Level Achievement Score Variables

Variables representing students' proficiency scores in mathematics and science are discussed in Chapter 6. These variables were used in the reporting of international achievement results in the international reports and are provided in the Population 1 and Population 2 Student Background files under the following variable names:

AIMATSCR	Population 1 International Mathematics Score
AISCISCR	Population 1 International Science Score
BIMATSCR	Population 2 International Mathematics Score
BISCISCR	Population 2 International Science Score

The following additional achievement variables described in Chapter 6 are also provided in the Student Background files:

ASMPV01 - ASMPV05	Population 1 - Mathematics Plausible Values (1-5)
ASSPV01 - ASSPV05	Population 1 - Science Plausible Values (1-5)
BSMPV01 - BSMPV05	Population 2 - Mathematics Plausible Values (1-5)
BSSPV01 - BSSPV05	Population 2 - Science Plausible Values (1-5)
AIMATEAP	Population 1 - International Mathematics Score (EAP)
AISCIEAP	Population 1 - International Science Score (EAP)
BIMATEAP	Population 2 - International Mathematics Score (EAP)
BISCIEAP	Population 2 - International Science Score (EAP)
ASMNRSC	Population 1 - National Rasch Mathematics Score (ML)
ASSNRSC	Population 1 - National Rasch Science Score (ML)
BSMNRSC	Population 2 - National Rasch Mathematics Score (ML)
BSSNRSC	Population 2 - National Rasch Science Score (ML)
ASMSCPT	Population 1 - Number of Raw Score Points in Mathematics
ASSSCPT	Population 1 - Number of Raw Score Points in Science

BMSMCP	Population 2 - Number of Raw Score Points in Mathematics
BSSMCP	Population 2 - Number of Raw Score Points in Science

School-Level Achievement Score Variables

In addition to these student-level scores, the following variables containing the mean of the overall mathematics and science scale scores for the lower and upper grades of each school are provided in the School Background files:

ACLMAT	Population 1 - Mean Overall Mathematics Achievement for Lower Grade
ACUMAT	Population 1 - Mean Overall Mathematics Achievement for Upper Grade
ACLSCI	Population 1 - Mean Overall Science Achievement for Lower Grade
ACUSCI	Population 1 - Mean Overall Science Achievement for Upper Grade
BCLMAT	Population 2 - Mean Overall Mathematics Achievement for Lower Grade
BCUMAT	Population 2 - Mean Overall Mathematics Achievement for Upper Grade
BCLSCI	Population 2 - Mean Overall Science Achievement for Lower Grade
BCUSCI	Population 2 - Mean Overall Science Achievement for Upper Grade

7.2.1.6 Linking and Tracking Variables

Information about students, teachers, and schools provided on the survey tracking forms² is included in linking or tracking variables. These variables have prefixes of IL or IT. Some of the important linking and tracking variables are listed below.³

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

² Survey Tracking forms are listings of students, teachers, or schools used for sampling and administration purposes.

³ Some additional linking and tracking variables included in the international background files but not listed above are indicated in the codebooks described in Section 7.3.

ITLANG	Language of testing for each student. Data are included for six 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 Spain: 1 = Castellano; 3 = Catalan; 4 = Gallego; 8 = Valenciano Switzerland: 1 = German; 2 = French; 3 = Italian Romania: 348 = Hungarian; 642 = Romanian
ITPART	Participation status variable indicating whether each student participated in any TIMSS session; only those students with ITPART equal to 3 (participated in a TIMSS session) are included in the Student Background files
ITPART1, ITPART2, ITPART3	Separate participation status variables for students related to the specific sessions as applicable: first-half testing session, second-half testing session, and questionnaire session
ILRELIAB	Linking variable indicating the inclusion status of each student in the reliability file containing double-coded free-response items (see Section 7.2.3)

The tracking information available regarding students' gender (ITSEX) or date of birth (ITBIRTHM, ITBIRTHY) was compared to the students' reports obtained from the background questionnaires. If the student data were missing, data in the corresponding tracking variables were substituted. Also, if a student's gender or date of birth reported in the background variables differed from the tracking information, in Population 2 the tracking information was replaced by the background data, and in Population 1, the background data were replaced by the tracking information.

Variables Included in the Teacher Background Files

ILCLASS1, ILCLASS2, ILCLASS3	Linking variables containing class identification numbers for classes to which teachers are linked
ILMATH, ILSCI	Sampling status of each teacher as a mathematics teacher or a science teacher
ITQUEST	Questionnaire status indicating whether the teacher completed the mathematics or science teacher questionnaire
ITPART	Participation status variable indicating whether the student participated in questionnaire session

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

7.2.1.7 International Background Variables

General Information

International background variables obtained from the student, teacher, 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, teacher, and school background questionnaire items, see Chapter 2. Copies of the international versions of the questionnaires are provided in Supplements 1 and 2. More information regarding how national adaptations of background questionnaire items were handled in creating the TIMSS background data files is provided in Supplement 3.

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, teachers, and schools contain the international background variable names, descriptive labels, response code definitions, formats, and field locations corresponding to each questionnaire item.

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. The questionnaire item numbers associated with each variable are indicated by field locations according to the formats given in Table 7.3. A set of tables mapping each background variable to its location in the questionnaire is included in Supplement 1 and Supplement 2.

Table 7.3
Background Questionnaire Item Field Location Format Conventions

Population	Questionnaire	Field Location Format	Example Question
Population 1	Student Questionnaire	SQ1-***	SQ1-2 Are you a boy or a girl?
Population 1	Teacher Questionnaire	TQ1-***	TQ1-9C Hours/week planning lessons
Population 1	School Questionnaire	SCQ1-***	SCQ1-4A How many full-time teachers
Population 2	Student Questionnaire (variables in both SQ2 and SQ2(s) forms: general and math-related questions)	SQ2-***	SQ2-25C How often do you have a test in mathematics
Population 2	Student Questionnaire (variables only in SQ2 form: questions related to integrated or general science)	SQ2G-***	SQ2G-31C How often do you have a test in science
Population 2	Student Questionnaire (variables in only the SQ2(s) form: questions related to specific science subject areas)	SQ2S-***	SQ2S-31C How often do you have a test in biology
Population 2	Mathematics Teacher Questionnaire	TQM2A*** TQM2B*** TQM2C*** TQM2D***	TQM2A17A Familiarity with National Curriculum Guide for Mathematics
Population 2	Science Teacher Questionnaire	TQS2A*** TQS2B*** TQS2C*** TQS2D***	TQS2A17A Familiarity with National Curriculum Guide for Science
Population 2	School Questionnaire	SCQ2-***	SCQ2-1 In what type of community is school located

International Background Variable Names

The naming system for the background variables permits the determination of the population and questionnaire based on 7- or 8-digit codes according to the general definitions given in Table 7.4.

Table 7.4
International Background Variable Naming Conventions

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

For example:

BSBGEDUM = Population 2 students' reports of mother's highest level of education

BTBSCLTM = Population 2 science teachers' reports of minutes in last science lesson

ACBMTEAW = Population 1 schools' reports of how many hours during the school week teachers have for teaching mathematics

7.2.1.8 Variables Derived from Student, Teacher, and School Background Data

General Information

In addition to the background variables contained in the student, teacher, and school questionnaires, a number of derived variables based on the student and teacher background variables were computed for use in the international reports. These derived variables, many of which use combined response options or data from more than one item, are also included in the International Database for use in secondary analyses. There are also several derived variables based on school background variables that are included in the international school background file, although these were not presented in the international reports.

Use of Derived Variables in the International Reports

The derived variables may, in general, be used to reproduce the values shown in the international report tables, applying the appropriate teacher or student filters and weights.⁴ Some exceptions are noted in Supplement 4, which lists all derived variables used to produce the international reports along with their corresponding international table and figure reference location. The order of variables in the index and in the Student Background file codebooks is based on the referenced table or figure in the international report. The nomenclature used to indicate the international table or figure 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 7.5

International Report Table/Figure Location Reference Definition for Derived Variables

Character Position	Definition	Value(s)
1	Population	A = Population 1 B = Population 2
2-3	International Report Type	MR = Mathematics Report SR = Science Report
4-9	Table/Figure Reference	T**.**(a,b,..) = Table Number F**.**(a,b,..) = Figure Number (column/section) [Chapter.Number(column/section if applicable)]

For example: BMRT5.9a = Population 2 Math Report Table 5.9, first major section

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 7.4). In the case of the derived variables, however, the third character is a D (derived variable) instead of B (background variable). For example, BSDGEDUP is a derived variable including Population 2 students' reports of parents' education.

⁴ Additional information about the methods used to report student and teacher background data may be found in (Martin and Kelly, 1996).

Information About How the Reported Derived Variables Were Computed

Descriptions of each of the derived variables based on student and teacher background data, their associated international table/figure references, and analysis notes indicating how they were computed using the associated student or teacher source variables are provided in Supplement 4 for the following four types of derived variables:

Section 1	Population 1 - Derived variables based on student background data
Section 2	Population 1 - Derived variables based on teacher background data
Section 3	Population 2 - Derived variables based on student background data
Section 4	Population 2 - Derived variables based on teacher background data

Each section of Supplement 4 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 deviations were handled in the international derived variables is also summarized in Supplement 4 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 and teacher codebook files, which are described in Section 7.3.

Derived Variables Based on School Background Data

In addition to the derived reporting variables based on the student and teacher background data, several additional derived variables are included in the school background files for Population 1 and Population 2. These variables are listed below with the questionnaire locations of the school background variables used to obtain them (SCQ1-**, SCQ2-** = Population 1 or Population 2 School Questionnaire Item Number).

ACDGTENR/BCDGTENR	Total school enrollment (sum of boys and girls reported in SCQ1-16A/SCQ2-17A)
ACDGLTER/BCDGLTER	Total lower-grade enrollment (sum of boys and girls reported in SCQ1-16E/SCQ2-17E)
ACDGLTRT/BCDGLTRT	Total students repeating lower grade (sum of boys and girls reported in SCQ1-16F/SCQ2-17F)
ACDMLTER/BCDMLTER	Total lower-grade students studying mathematics (sum of boys and girls reported in SCQ1-16I/SCQ2-17I)
ACDSLTER/BCDSLTER	Total lower-grade students studying science (sum of boys and girls reported in SCQ1-16J/SCQ2-17J)
ACDGUTER/BCDGUTER	Total upper-grade enrollment (sum of boys and girls reported in SCQ1-16K/SCQ2-17K)

ACDGUTRT/BCDGUTRT	Total students repeating upper grade (sum of boys and girls reported in SCQ1-16L/SCQ2-17L)
ACDMUTER/BCDMUTER	Total upper-grade students studying mathematics (sum of boys and girls reported in SCQ1-16O/SCQ2-17O)
ACDSUTER/BCDSUTER	Total lower-grade students studying science (sum of boys and girls reported in SCQ1-16P/SCQ2-17P)

7.2.1.9 Sampling Variables

The following variables related to the samples selected in each country are included in the student and school background files for use in obtaining the national population estimates presented in the international reports:

TOTWGT	Total student weight
HOUWGT	House weight
SENWGT	Senate weight
WGTFAC1	School weighting factor
WGTAJ1	School weighting adjustment
WGTFAC2	Class weighting factor
WGTFAC3	Student weighting factor
WGTAJ3	Student weighting adjustment
JKZONE	Jackknife zone
JKINDIC	Replicate code
WGTFAC1	School weighting factor
WGTAJ1	School weighting adjustment
SCHWGT	School-level weight
STOTWGTL	Sum of student weights - lower grade
STOTWGTU	Sum of student weights - upper grade
STOTWGTE	Sum of student weights - extra grade

Detailed descriptions of the sampling weights and jackknife variables and how they are used in computing the population estimates and standard errors presented in the international reports are described in Chapter 3 and Chapter 8. In addition, Chapter 9 includes sample analyses using these variables.

7.2.2 Assessment Files

Student assessment files contain the student response data for the individual cognitive items in the TIMSS written assessment and the performance assessment. Four student assessment files for each country are contained in the TIMSS International Database:

- Student written assessment file - Population 1
- Student written assessment file - Population 2
- Student performance assessment file - Population 1
- Student performance assessment file - Population 2

7.2.2.1 Written Assessment Files

Students who participated in TIMSS were administered one of eight test booklets with questions in mathematics and science. Some of these questions were multiple-choice questions and some were open-ended. The responses to the open-ended questions were coded using a two-digit coding system (described in Chapter 4). The written assessment data files contain the answers to the multiple-choice questions and the codes assigned by the coders to the student responses. Since under the TIMSS test design a student received only a fraction of the total test item pool, the variables for the items that were not included in the version of the test that was administered to the student are coded as not administered. 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 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.

7.2.2.2 Performance Assessment Files

A subset of the students who participated in the TIMSS written assessment were also sampled to participate in a performance assessment. These students were presented with three to four performance assessment tasks, each of which asked them conduct some activities and respond to a set of questions. The responses to the performance assessment tasks were also coded by trained coders using a two-digit system. The performance assessment files contain the codes assigned by the coders to the student responses. They also contain identification variables, sampling variables, and achievement variables. The achievement variables included in the performance assessment data files are those from the written assessment. The data contained in these files can be linked to the Student Background data files and the Written Assessment data files by using the variables IDCNTRY and IDSTUD.

7.2.2.3 Cognitive Item Variable Names

The cognitive item variable names are based on 7-digit alphanumeric codes according to the general definitions given in Table 7.6.

Table 7. 6
Variable Name Definitions for the Written Assessment and Performance Assessment Items

Character Position	Definition	Value(s)	
		Written Assessment	Performance Assessment
1	Population	A = Population 1 B = Population 2	A = Population 1 B = Population 2
2	Student Variable	S	S
3	Item Format	M = Multiple Choice S = Short Answer E = Extended Response	P = Performance Assessment
4	Subject Area	M = Mathematics S = Science	M = Mathematics S = Science G = Math/Science Combined
5	Cluster or Task Location	A,B,...,Z	1 - 6 (number of task within each subject category)
6-7	Item Number	01-99	1, 2, 3A, 3B, etc.(item number within task)

For example:

ASMMI01 = Population 1 written assessment mathematics multiple-choice item number I01

BSPS11A = Population 2 performance assessment science task 1 (S1), item 1A

7.2.2.4 Performance Assessment Task Identifications

The performance assessment tasks are listed in Chapter 2 (Table 2.7) which indicates the task identification number and task titles.⁵ Complete descriptions of the tasks administered to students in each population and the items contained in each may be found in the international report describing the performance assessment results (Harmon et al., 1997).

⁵ The task identification information is also indicated in the variable labels in the performance assessment codebooks.

7.2.2.5 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 written assessment items and the performance assessment 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 or performance assessment task they completed, a 'Not Administered' code is given to all items that were not included in the test booklet or performance assessment tasks assigned to each student.⁶

7.2.2.6 Analysis By Mathematics and Science Content Area Reporting Categories

The TIMSS cognitive items measured student achievement in different areas within mathematics and the sciences, and student performance was presented in the international reports in several different content area reporting categories. In order to permit secondary analyses based on these content areas, the classification of TIMSS written assessment items into each of the Population 1 and Population 2 mathematics and science content area reporting categories is presented in Tables 7.7 through 7.10. Achievement scale scores were not computed for each of these content area reporting categories.

⁶ See Chapter 2 for a discussion of the design of the TIMSS written assessment and performance assessment.

Table 7.7
Classification of Population 1 Items into Mathematics Content Area Reporting Categories

Items in Content Area Reporting Category					
Whole Numbers (25 items)	Fractions and Proportionality (21 items)	Measurement, Estimation, and Number Sense (20 items)	Data Representation, Analysis, and Probability (12 items)	Geometry (14 items)	Patterns, Relations, and Functions (10 items)
A03	A01	A05	B05	B06	G04
A04	A02	B07	C01	D07	H08
B08	B09	C02	D05	F09	I07
C04	C03	D06	F05	H05	J05
E02	D08	E01	J03	I01	K03
F07	D09	E03	K04	I06	K06
G03	E04	F08	L01	J01	L04
H07	F06	G02	L02	J02	L09
H09	G01	H06	M01	K01	M09
I03	I02	J06	M02	K08	U04
I04	I05	J08	S01	L03	
I09	I08	K05	T01A	L05	
J04	J07	K07	T01B	M04	
J09	K09	L06		T05	
K02	M05	L08			
L07	S03	M07			
M03	S04	S05			
M06	T04A	T03			
M08	T04B	U01			
S02	U02	V05			
T02	U03A				
U05	U03B				
V02	U03C				
V03	V01				
V04A					
V04B					

Table 7.8
Classification of Population 1 Items into Science Content Area Reporting Categories

Items in Content Area Reporting Category					
Earth Science	Life Science		Physical Science		Environmental Issues and the Nature of Science
(17 items)	(41 items)		(30 items)		(9 items)
A08	A09	P08	A06	Q09	A07
A10	B03	P09	B01	R01	C07
B02	B04	Q01	C05	R05	E08
C09	C06	Q02	D01	R08	O06
D04	C08	Q05	E05	R09	P06
E06	D02	Q06	F04	W01	Q07
E09	D03	R03	G09	X01	R02
F02	E07	R04	H01	Z03	W05A
F03	F01	R06	N04		W05B
G07	G05	R07	N07		X03
G08	G06	W02	N08		
H03	H02	W03	N09		
N01	H04	W04	O01		
O04	N02	X02	O05		
O09	N03	X04	O08		
Y01	N05	X05	P03		
Z01A	N06	Y02A	P04		
Z01B	O02	Y02B	P05		
	O03	Y03A	P07		
	O07	Y03B	Q03		
	P01	Z02	Q04		
	P02		Q08		

Table 7.9
Classification of Population 2 Items into Mathematics Content Area Reporting Categories

Items in Content Area Reporting Category							
Fractions and Number Sense		Algebra		Measurement	Geometry	Data Representation, Analysis, and Probability*	Proportionality
(51 items)		(27 items)		(18 items)	(23 items)	(20 items)	(11 items)
A01	L09	A02	L16	A03	A05	A06	A04
B09	L17	B12	N13	C01	B11	B07	B08
B10	M04	C05	O07	D11	C03	C02	D08
C04	M08	D10	P10	E06	D07	E01	F07
C06	N11	E05	P15	F10	E02	F08	G04
D09	N14	F11	Q01	G02	G03	G01	L14
D12	N16	G06	Q02	I03	I08	H07	M06
E03	N17	H10	Q07	J10	J11	H11	Q05
E04	N19	H12	R09	K05	J15	I09	R14
F09	O02	I01	R11	L12	J16	J13	T02A
F12	O04	I04	S01A	M01	K03	K07	T02B
G05	O09	J18	S01B	N15	K08	L10	V03
H08	P12	K04	T01A	O06	L15	M03	
H09	P13	L11	T01B	P11	M02	N18	
I02	P14	L13		Q03	M05	O01	
I05	P16			S02A	M07	O05	
I06	Q06			S02B	N12	P17	
I07	Q08			S02C	O03	Q04	
J12	Q09			U02A	O08	R08	
J14	R06			U02B	P08	V02	
J17	R07			V04	P09		
K01	R12				Q10		
K02	R13				R10		
K06	U01A						
K09	U01B						
L08	V01						

*Item M09 was excluded from Population 2 data files due to problematic item statistics.

Table 7.10
Classification of Population 2 Items into Science Content Area Reporting Categories

Items in Content Area Reporting Category						
Earth Science	Life Science		Physics		Chemistry	Environmental Issues and the Nature of Science
(22 items)	(40 items)		(40 items)		(19 items)	(14 items)
A12	A07	K18	A08	L07	A09	A11
B01	B04	L02	A10	M12	C10	C11
B05	C08	L03	B02	M14	F06	F04
C07	D05	L05	B03	N08	G10	G12
D03	D06	L06	B06	N10	H06	I12
E09	E08	M11	C09	O10	J03	I13
E12	E10	N02	C12	O13	J04	I15
F05	F01	N04	D01	P01	J06	I18
G11	F03	N06	D02	P02	J08	K19
H03	G08	O16	D04	P05	M10	N01
H04	G09	O17	E07	Q12	M13	N03
I17	H01	P04	E11	Q13	N07	N05
J01	H02	P06	F02	Q18	N09	P07
K15	I10	Q17	G07	R01	O11	Z02A
O12	I11	R03	H05	R02	O15	Z02B
O14	I14	X01	I16	Y01	Q14	
P03	I19	X02A	J05	Y02	Q15	
Q11	J02	X02B	K10		R05	
Q16	J07		K13		Z01A	
R04	J09		K14		Z01B	
W01A	K11		K17		Z01C	
W01B	K12		L01			
W02	K16		L04			

7.2.2.7 Release Status of TIMSS Test Items and Performance Tasks

To aid in the interpretation of the student cognitive item response data, a large number of TIMSS items have been released to the public along with their scoring guides. For both Populations 1 and 2, written assessment items in Clusters I through Z and the entire set of performance assessment tasks and coding guides are available. See Chapter 2 for information on how to obtain the released item sets.

7.2.2.8 Other Variables in the Student Assessment Files

In addition to the written assessment and performance 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
- Linking and tracking variables⁷
- Sampling 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.

7.2.2.9 School Performance Assessment Files

The School Performance Assessment files for Population 1 and Population 2 contain school-level information relevant to the administration of the performance assessment. In these files, there is one record for each school that participated in the performance assessment in each country, containing identification variables required to link to other files. In addition, a number of tracking variables are included with specific information provided by the performance assessment administrators in each school. The variables included in the School Performance Assessment files are the following:

ITROTAT	Rotation scheme used
ITPAS1	Adequacy of room
ITPAS2	Time allotted per station
ITPAS3	Adequacy of equipment and material
ITPAS4	Missing information from manual
ITPAS5	Clearness of instructions

⁷ For the performance assessment files, additional tracking variables are included related to student participation in the performance assessment, such as to which rotation and sequence number each student was assigned (ITROTAT, ITSEQ) and students' participation status for each performance assessment task (ITPARTM1-ITPARTM5; ITPARTS1-ITPARTS5; ITPARTG1,ITPARTG2). See Harmon and Kelly (1996) for a description of the performance assessment administration procedures.

ITPAS6	Difficulty of performance tasks
ITPAS7	Students' attitude towards the performance assessment
ITPAS8	Attitude of staff towards the performance assessment

The codebooks for School Performance Assessment files include a complete list of all of these variables as well as their response option definitions.

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. Four coding reliability files for each country are included in the TIMSS International Database:

- Student Written Assessment Coding Reliability File - Population 1
- Student Written Assessment Coding Reliability File - Population 2
- Student Performance Assessment Coding Reliability File - Population 1
- Student Performance Assessment Coding Reliability File - Population 2

Each of these files contains one record for each student included in the reliability sample for the written assessment or the performance assessment, as described in Chapter 4. For each free-response item in the written or performance 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 reports.

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 Original Code variables shown in Table 7.6. The three reliability variables may be differentiated by the second character in the variable name:

- Original Code Variable: Second Character = S (e.g. ASSSR01, BSPM42)
- Second Code Variable: Second Character = R (e.g. ARSSR01, BRPM42)
- Agreement Code Variable: Second Character = I (e.g. AISSR01, BIPM42)

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 are identical to those contained in the Student Written Assessment or Student Performance Assessment files. In some 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 (see Martin and Mullis, 1996). The recoded diagnostic code values were used in computing the achievement scores reflected in the international reports. Table 7.11 lists the recodings made to the Population 1 and Population 2 written assessment and performance assessment items. These recodes are reflected in the Written Assessment and Performance Assessment data files.

For the items indicated in Table 7.11, the response codes in the Student Written Assessment or Student Performance Assessment files reflect the recoded values. 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.

Table 7.11
Recodes Made to Free-Response Item Codes in the Written Assessment and Performance Assessment Items

	Item	Variable	Recodes	Comment
General	All Items		37, 38 → 39 27, 28 → 29 17, 18 → 19 77, 78 → 79	Country-specific diagnostic codes recoded to 'other' categories within the score level.
	K10	BSMMK08	71 → 70	Training team found it difficult to distinguish between the 70 and 71 codes; both codes combined in 70.
Population 2 - Written Assessment Items	L04	BSESL04	20 → 10 21 → 11 29 → 19 10 → 74 11 → 75 12 → 76 19 → 79	Only 20s have positive point-biserial correlation; change to 1-point item codes.
	M11	BSESM11	10, 11, 12, 13 → 71 20, 21, 22, 23, 24, 25 → 72 30 → 10 31 → 11	Only 30s have positive point-biserial correlation; change to 1-point item codes.
	Y01	BSESY01	20 → 10 21 → 11 22 → 12 29 → 19 10 → 73 11 → 74 19 → 75	Only 20s have positive point-biserial correlation; change to 1-point item codes.
	Y02	BSESY02	21 → 19	Typographical error in category 21 in coding guide.
	J03	BSSSJ03	19 → 10	Typographical error in coding guide.
	M12	BSSSM12	19 → 10	Typographical error in coding guide.
	O14	BSES014	20 → 10 29 → 19 10 → 72 11 → 73 19 → 74	Only 20s have positive point-biserial correlation.
	Q18	BSSSQ18	19 → 10 29 → 20	Typographical error in coding guide.
	L16	BSSML16	19 → 10	Typographical error in coding guide.
	M06	BSSMM06	19 → 10	Typographical error in coding guide.
	M08	BSSMM08	19 → 10	Typographical error in coding guide.
	Q10	BSSMQ10	19 → 10	Typographical error in coding guide.
	R13	BSSMR13	74 → 79	Typographical error in code 74 (28 instead of 280); leaves gap in 7* diagnostic codes.
	S01A	BSEMS01A	19 → 10	Typographical error in coding guide.
	S02A	BSEMS02A	19 → 10	Typographical error in coding guide.
	T01A	BSEMT01A	29 → 20	Typographical error in coding guide.
	T02A	BSEMT02A	19 → 10	Typographical error in coding guide.
	U01A	BSEMU01A	19 → 10	Typographical error in coding guide.
	U02A	BSEMU02A	19 → 10 29 → 20	Typographical error in coding guide.
	U02B	BSEMU02B	19 → 10 29 → 20	Typographical error in coding guide.

Table 7.11
(Continued)

	Item	Variable	Recodes	Comment
Population 1 - Written Assessment Items	T04A	ASEMT04A	20 → 10 29 → 19 10 → 72 11 → 73	Only 20s have positive point-biserial correlation.
	T04B	ASEMT04B	20 → 10 29 → 19 10 → 72 11 → 73	Only 20s have positive point-biserial correlation.
	V04A	ASEMV04A	30 → 20 20 → 12 21 → 13	Differentiation between 30s, 20s, and 10s not clear.
	Y01	ASESY01	20 → 10 29 → 19 10 → 72 11 → 73 19 → 74	Only 20s have positive point-biserial correlation.
	Z02	ASESZ02	30 → 10 31 → 11 20 → 71 29 → 72 10, 11, 12, 13 → 73	Only 30s have positive point-biserial correlation.
Performance Assessment Items	Task M2 (Calculator) Item 5 (Population 2)	BSPM25	11 → 12	Error in coding guide: valid codes listed as 10,12, 19 (no code 11). Recoded 11 codes used in some countries.
	Task M5 (Packaging) Item 1 (Populations 2 & 1)	BSPM51 ASPM51	30 → 22 31 → 23	Two versions of task used across countries: original asked for 2 OR 3 boxes; revised asked for 3. Item changed to 2-point value for report tables; changed codes for 3 correct boxes (30,31) to 2-point codes (22,23).
	Task S5 (Solutions) Item 2A (Population 2)	BSPS52A	99 → 98	Administrator notes not coded consistently across countries; invalid 99 codes (blank) used in several countries recoded to not administered. Item omitted from report table but kept in data file.
	Task S5 (Solutions) Item 4 (Population 2)	BSPS54	10 → 21	Coding guide revised based on reports of problematic scoring during training development.
	Task S6 (Containers) Item 1A (Population 1)	ASPS61A	99 → 98	Administrator notes not coded consistently across countries; invalid 99 codes (blank) used in several countries recoded to not administered.

Other Variables in the Reliability Files

In addition to the coding reliability variables, the reliability files also include identification variables to aid in case identification. Some tracking variables are also included that were used in conducting of 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).⁸

7.2.4 Student-Teacher Linkage Files

General Information

The Student-Teacher Linkage files for Population 1 and Population 2 contain information required to link the student and teacher files and to compute appropriately weighted teacher-level data using the student as the unit of analysis. Example analyses using these data files are described in Chapter 9.

The Student-Teacher Linkage files contain one entry per student-teacher linkage combination in the data. In many cases, students are linked to more than one mathematics and/or science teacher, and in these cases there will be one record for each student-teacher link. This is particularly true in Population 2, where the majority of students have separate teachers for science and mathematics. In addition, in some countries many students may also have more than one teacher for each of the two subject areas. For instance, if three teachers are linked to a student, there are three entries in the file corresponding to that student.

Variables Required for Computing Teacher-Level Data

The following important variables required to compute appropriately weighted teacher-level data are included in the linkage files.

NMTEACH	Number of mathematics teachers for each student
NSTEACH	Number of science teachers for each student
NTEACH	Number of total teachers for each student (mathematics and science)
MATWGT	Adjusted weight for mathematics teacher data
SCIWGT	Adjusted weight for science teacher data
TCHWGT	Adjusted weight for total teachers (mathematics and science)

The MATWGT, SCIWGT, and TCHWGT variables contain mathematics, science, and total teacher weights adjusted for the total number of teachers of each type for each student. For example, if a student has three mathematics teachers, the adjusted weight for each mathematics teacher (MATWGT) will be equal to one-third so that each mathematics teacher contributes equally to the teacher-based data for students.

⁸ See Mullis and Smith (1996) for more detailed information about the coding reliability procedures and results.

Other Variables in the Student-Teacher Linkage Files

The linkage files also contain the identification variables required to identify cases and link the student and teacher files. In addition some tracking and achievement score variables are also included. The codebooks for Student-Teacher Linkage files contain a complete list of all variables included.

7.2.5 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.⁹ 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, teacher, 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 and performance 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 or the performance assessment, uninterpretable student responses were given a code 90, which is distinguished from the code 99 given to items that were left blank.

⁹ 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).

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.

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 eight rotated booklets used in the TIMSS study 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 session - If a student was not present for a particular testing session, then all variables referring to that session have been coded to "Not administered". However, if a student participated in a session and did not answer any of the items, these questions have been 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, teacher, 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 removed for a country.

Background questionnaire items were omitted - Questions in the student, teacher, or school background questionnaires that were considered not applicable in some countries were not included in their questionnaires.

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 3 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 or performance assessment items, a Not Reached Item Code value of 96 is used.

7.2.6 National Data Issues Affecting the Use of International Data Files

In some cases, resources were not available to resolve database issues for specific countries in time for either the release of the international reports or the production of the international data files. As a result, some international data are modified or not available for some countries. These general database issues are documented below.

Australia

Information made available after publication of the Population 1 international reports required that the Population 1 sampling weights for Australia be recomputed. The adjusted sampling weights are included in the International Database. As a consequence, any computations using these new weights may be slightly different from those shown in the international report tables for Population 1.

Bulgaria

The student, teacher, and school background data submitted by Bulgaria were not deemed internationally comparable and are thus not included in the International Database.

Israel

Due to the use of unapproved school sampling procedures for the performance assessment, the performance assessment results for Israel in the international report are based on unweighted data at both the fourth and eighth grades (Harmon et al., 1997). Thus, the sampling weights have been set to 1 for all cases in the performance assessment file for Israel.

New Zealand

There are some incorrect Population 1 student-teacher linkages for students in multi-grade classes containing both lower- and upper-grade students. In some of these classes, different teachers for mathematics and/or science are linked to the students in the other grade included in their composite class. As a result, some students in these classes were given not only the correct linkages to their own teacher(s) but also some incorrect linkages to teacher(s) of students in the other grade in their class. The student-teacher linkage problems were discovered after the production of the international reports and too late to correct the database. The errors mostly affect the lower-grade students and are expected to account for less than 10% of students.

Philippines

The teacher and school data submitted by the Philippines were not deemed internationally comparable and thus are not included in the International Database.

Due to the use of unapproved school sampling procedures, the results presented in the international reports for the Philippines reflect unweighted data. Consequently, the sampling weights have been set to 1 for all cases in the files for the Philippines.

South Africa

The teacher and school data submitted by South Africa were not deemed internationally comparable and thus are not included in the International Database.

Thailand

Information made available after publication of the Population 2 international reports required that the sampling weights for Thailand be recomputed. The adjusted sampling weights are included in the International Database. As a consequence, any computations using these new weights may be slightly different from those shown in the international report tables for Population 2.

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. One codebook file is provided for each of the data files listed in Table 7.12.

Table 7.12
Population 1 and Population 2 Codebook Files

Data File	Population 2 Codebook Name	Population 1 Codebook Name
Student Written Assessment File	BSACODE.*	ASACODE.*
Student Background File	BSGCODE.*	ASGCODE.*
Teacher Background File(s)	BTMCODE.* - Mathematics BTSCODE.* - Science	ATGCODE.*
School Background File	BCGCODE.*	ACGCODE.*
Student - Teacher Linkage File	BLGCODE.*	ALGCODE.*
Student Written Assessment Reliability File	BSRCODE.*	ASRCODE.*
Student Performance Assessment File	BSPCODE.*	ASPCODE.*
Student Performance Assessment Reliability File	BSQCODE.*	ASQCODE.*
School Performance Assessment File	BCTCODE.*	ACTCODE.*

Codebook files are available in two different formats, differentiated by the file extension:

- ASCII Text Format (* = CDT)
- Machine-Readable ASCII Format (* = 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 7.13 describes the structure of the machine-readable codebook files.

Table 7.13
File Structure of Machine-Readable Codebook Files

Field Name	Location
Variable name	1-8
Question location	10-20
Starting location in data file	22-24
Ending location in data file	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 the section for each codebook type. The major sections of each codebook type are as follows:

Student, Teacher, and School Background File Codebooks

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

Student Written and Performance Assessment File Codebooks

- Identification Variables
- Tracking/Linking Variables
- Cognitive Item Variables (in order by item within clusters or performance assessment tasks)
- Sampling Variables
- Score Variables

School Performance Assessment File Codebooks

- Identification Variables
- Tracking/Linking Variables

Student-Teacher Linkage File Codebooks

- Identification Variables
- Sampling Variables
- Score Variables
- Teacher Linking/Weighting Variables

Reliability File Codebooks

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

An example printout of a page from the codebook for the student background data (BSGCODE.*) is shown in Figure 7.1. The files are as follows:

Figure 7.1
Example Printout of a Codebook Page

```
Codebook, Date: 24.09.97 File: asgcode.SDB
-----
Var.Question  Variable Label      Code  Option  Location/Format
No.           Name
-----
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      7-9 / N 3.0
                999   missing
                998   not admin.
-----
5 SCHOOL       *SCHOOL ID*     VALUE  School ID       10-15 / N 6.0
-----
6 CLASS        *CLASS ID*      VALUE  Class ID        16-23 / N 8.0
-----
7 STUD         *STUDENT ID*    VALUE  Student ID      24-33 / N 10.0
-----
8 GRADE        *GRADE*         VALUE  Grade           34-35 / N 2.0
-----
9 EXCLUD      *INDICATOR FOR EXCLUDED STUDENTS*
                1      student is functionally disabled
                2      student is educable mentally retarded
                3      student unable to speak or read language
                4      <nationally defined reason>
                9      missing
                36 / C 1.0
-----
```

Variable Number: The first column (Var. No.) 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.

Variable Name: The third column (Variable Name) 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 (Label) 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 (Code) contains the codes used for variable responses. For variables where numeric data are supplied in response 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 7.2.5 are also included for either numerical or categorical variables. For example, for multiple-choice cognitive items, the code options are a,b,c,d,e, while for the free-response cognitive items, the code options are the two-digit numerical codes described in Chapter 4.

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

XX-XX / N(C) X

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

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 convert cognitive item response codes to the score values used in the computation of international scores, with one file provided for the written-assessment items and one for the performance assessment items. 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 7.14 lists all program files provided.

Table 7.14
Population 1 and Population 2 Program Files

Data Files	Population 2 File Name	Population 1 File Name
I. Data Access Control Files		
Control Files for the Student Written Assessment File	BSACONTR.*	ASACONTR.*
Control Files for Student Background File	BSGCONTR.*	ASGCONTR.*
Control Files for Teacher Background File(s)	BTMCONTR.* - Mathematics BTSCONTR.* - Science	ATGCONTR.*
Control Files for School Background File	BCGCONTR.*	ACGCONTR.*
Control Files for Student - Teacher Linkage File	BLGCONTR.*	ALGCONTR.*
Control Files for Student Written Assessment Reliability File	BSRCONTR.*	ASRCONTR.*
Control Files for Student Performance Assessment File	BSPCONTR.*	ASPCONTR.*
Control Files for Student Performance Assessment Reliability File	BSQCONTR.*	ASQCONTR.*
Control Files for School Performance Assessment File	BCTCONTR.*	ACTCONTR.*
II. Jackknife Statistics Program Files		
	JACK.*	JACK.*
III. Scoring Program Files		
Scoring Program Files for Written-Assessment Cognitive Items	BSASCORE.*	ASASCORE.*
Scoring Program for Performance Assessment Items	BSPSCORE.*	ASPSCORE.*

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

7.5 Data Almanacs

Data almanacs are included for all student, teacher, and school background variables. The data almanacs are text files that display unweighted summary statistics, by grade, for each participating country on each variable included in the background questionnaires administered to students, teachers, and school administrators or principals. The data almanac files corresponding to each questionnaire type are listed in Table 7.15.

Table 7.15
Data Almanac Files

Almanac File	Contents
ASGALM.LIS	Student background data - Population 1
ATGALM.LIS	Teacher background data - Population 1
ACGALM.LIS	School background data - Population 1
BSGALM.LIS	Student background data - Population 2
BTMALM.LIS	Mathematics teacher background data - Population 2
BTSALM.LIS	Science teacher background data - Population 2
BCGALM.LIS	School background data - Population 2

There are two types of displays in the data almanacs, depending on whether the item is a categorical variable or a continuous variable. The display for categorical variables includes the sample size, the count of students who were not administered the question, the count of students choosing each of the options on the question, and the count of students who did not choose any of the valid options to the questions. In cases where the question did not apply, the count of students to whom the question did not apply is also presented in the almanac. An example of a categorical variable almanac display is shown in Figure 7.2.

Figure 7.2

Example Data Almanac Display for Categorical Variable

Third International Mathematics and Science Study
 Student Background Variables - Population 1
 Unweighted data almanac for the TIMSS International Database

September 22, 1997

Question: Are you a girl or a boy? (ASBGSEX)
 Location: S01-2

R:*GRADE INDICATOR*=Lower grade

Country	Cases	GEN\STUDENT'S SEX			OMIT
		1.GIRL	2.BOY	NOT ADM.	
Australia	4741	2335	2280	126	.
Austria	2526	1261	1243	10	12
Canada	7594	3660	3708	83	143
Cyprus	3308	1640	1632	17	19
Czech Republic	3256	1646	1600	10	.
England	3056	1507	1478	71	.
Greece	2955	1441	1499	15	.
Hong Kong	4396	1967	2406	8	15
Hungary	3038	1492	1456	54	36
Iceland	1698	836	819	43	.
Iran, Islamic Rep.	3361	1722	1595	44	.
Ireland	2889	1348	1506	35	.
Japan	4306	2105	2189	12	.
Korea	2777	1323	1451	3	.
Latvia (LSS)	2054	1027	991	35	1
Netherlands	2790	1379	1391	19	1
New Zealand	2504	1287	1210	6	1
Norway	2219	1063	1101	7	48
Portugal	2650	1286	1359	5	.
Scotland	3132	1523	1488	121	.
Singapore	7030	3378	3639	7	6
Slovenia	2521	1228	1282	11	.
Thailand	2870	1433	1415	21	1
United States	3819	1857	1962	.	.

A second type of data display is used for continuous variables. This includes the sample size, the count of students who were not administered the question, the count of those who did not respond, the count of those to whom the question did not apply, the mean, mode, minimum, maximum, and the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles. An example of a continuous variable almanac display is shown in Figure 7.3.

September 22, 1997 5

Figure 7.3
Example Data Almanac Display for Continuous Variable

Third International Mathematics and Science Study
 Student Background Variables - Population 1
 Unweighted data almanac for the TIMSS International Database

Question: How old were you when you came to <country>? (ASBGRN2)
 Location: SQ1-3B

R:**GRADE INDICATOR**=Lower grade

Country	Cases	Not Adm.	Omit	Not Appl.	Mean	Min	P10	Q1	Median	Q3	P90	Max
Australia	4741	126	59	4208	4.4	0.0	1.0	2.0	4.0	6.0	8.0	9.0
Austria	2526	11	46	2188	5.0	1.0	1.0	3.0	5.0	7.0	8.0	10.0
Canada	7594	231	132	6579	4.3	1.0	1.0	2.0	4.0	6.0	8.0	9.0
Cyprus	3308	22	140	2808	4.5	1.0	1.0	2.0	4.0	6.0	9.0	13.0
Czech Republic	3256	24	36	3150	4.1	1.0	1.0	2.0	4.0	5.0	8.0	9.0
England	3056	71	70	2727	4.0	1.0	1.0	1.0	4.0	6.0	8.0	9.0
Greece	2955	33	115	2580	4.5	1.0	1.0	3.0	4.0	6.0	8.0	11.0
Hong Kong	4396	42	74	3483	4.8	0.0	1.0	2.0	4.0	8.0	9.0	13.0
Hungary	3038	54	152	2740	4.3	1.0	1.0	1.5	4.0	6.5	9.0	11.0
Iceland	1698	48	41	1459	4.1	0.0	1.0	1.0	4.0	7.0	8.0	10.0
Iran, Islamic Rep.	3361	44	300	2602	3.9	1.0	1.0	1.0	3.0	6.0	9.0	12.0
Ireland	2889	35	44	2699	4.5	1.0	1.0	2.0	4.0	6.0	8.0	10.0
Japan	4306	4306	0
Korea	2777	3	78	2696
Latvia (LSS)	2054	38	55	1897	4.3	1.0	1.0	2.0	4.0	6.0	9.0	10.0
Netherlands	2790	225	40	2218	1.9	0.0	0.0	0.0	0.0	4.0	6.0	12.0
New Zealand	2504	5	59	2191	3.9	1.0	1.0	2.0	3.0	6.0	8.0	9.0
Norway	2219	23	36	2059	3.2	1.0	1.0	1.0	3.0	5.0	6.0	9.0
Portugal	2650	59	39	2378	4.4	0.0	1.0	2.0	4.0	7.0	8.0	11.0
Scotland	3132	121	80	2627	4.3	1.0	2.0	3.0	4.0	6.0	7.0	9.0
Singapore	7030	13	80	6344	5.7	1.0	2.0	3.0	6.0	8.0	9.0	13.0
Slovenia	2521	13	22	2399	3.9	0.0	1.0	2.0	4.0	6.0	7.0	10.0
Thailand	2870	22	36	2812
United States	3819	31	110	3243	3.7	1.0	1.0	1.0	3.0	5.0	7.0	10.0

7.6 Test-Curriculum Matching Analysis Data Files

To investigate the match of the TIMSS 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 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 B of Beaton et al. (1996a); Beaton et al. (1996b); Mullis et al. (1997); and Martin et al. (1997). The selection of the items by each country, by grade, is included as part of the International Database.

There are four files that contain the item selection by each country, at each grade level. These files are located in the subdirectory called TCMA in the corresponding CD. The four files are:

TCMAMP1.CSV	Test-Curriculum Matching Analysis Population 1 mathematics item selection
TCMASP1.CSV	Test-Curriculum Matching Analysis Population 1 science item selection
TCMAMP2.CSV	Test-Curriculum Matching Analysis Population 2 mathematics item selection
TCMASP2.CSV	Test-Curriculum Matching Analysis Population 2 science item selection

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 for inclusion (indicated with a "1") or exclusion (indicated with a "0") of the items for a specific grade level. For more information on the TCMA see Beaton and Gonzalez (1997).

Chapter 8

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 (Walter, 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 Computing Error 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:

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

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.

8.2 Construction of Sampling Zones for Sampling Variance Estimation

An important step in applying the JRR techniques 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 schools 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.

This 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. In the Performance Assessment data files there is a maximum of 42 sampling zones in Population 2 and 39 in Population 1. This variable is included in the Student Background, the student Written Assessment, the Student-Teacher Linkage, and the student Performance Assessment data files. For each individual student, this variable is identical in the first three files, but differs in the Performance Assessment data files because the performance assessment sample is a sub-sample of the written assessment sample.

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, the student written assessment, the student teacher linkage, and the student performance assessment data files. For each individual student, this variable is identical in the first three files, but differs in the performance assessment data files because the performance assessment sample is a sub-sample of the written assessment sample.

8.3 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 double 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 the chapter on sampling and sampling weights.

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 $k_{hi}=1$ otherwise.

In TIMSS, a total of 75 replicate weights were computed regardless of the number of actual zones within each 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.

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 and scores discussed in previous chapters. It also provides details on some SPSS and SAS code to conduct such analyses, and the results of these analyses. 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 for each subgroup, and the corresponding standard errors for the percent and mean statistics. The analyses presented in this chapter, based on student and teacher data, replicate analyses that are included in the TIMSS mathematics international report. Two tables from the international report (*Mathematics Achievement in the Middle School Years*), shown in Figure 9.1 and Figure 9.2, are replicated in Examples 1 and 2 in this chapter. The user is welcomed to compare the results from these analysis 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 reports.¹

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.

¹ Documentation regarding the computational methods used to obtain any derived variables included in the international reports is presented in Supplement 4.

Figure 9.1
Sample Table for Student-Level Analysis Taken From the TIMSS International Report
“Mathematics Achievement in the Middle School Years”

CHAPTER 4

Table 4.10

Students' Reports on the Hours Spent Each Day Watching Television and Videos
Mathematics - Upper Grade (Eighth Grade*)

Country	Less than 1 Hour		1 to 2 Hours		3 to 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>	24 (0.9)	539 (6.0)	41 (0.8)	539 (4.1)	27 (0.8)	528 (3.8)	9 (0.6)	487 (5.5)
<i>Austria</i>	25 (1.4)	540 (5.4)	53 (1.1)	546 (4.2)	17 (1.0)	539 (5.2)	5 (0.6)	497 (8.6)
Belgium (Fl)	24 (1.2)	580 (6.7)	52 (1.2)	575 (6.2)	19 (1.0)	535 (7.1)	5 (0.5)	514 (12.1)
<i>Belgium (Fr)</i>	33 (1.3)	536 (4.2)	44 (1.8)	536 (4.9)	17 (1.3)	522 (4.0)	6 (1.0)	445 (9.0)
Canada	22 (0.7)	522 (2.9)	46 (0.8)	534 (3.5)	25 (0.7)	532 (3.0)	7 (0.6)	504 (5.2)
<i>Colombia</i>	31 (1.5)	384 (4.9)	39 (1.2)	397 (3.3)	20 (1.2)	391 (5.2)	11 (1.0)	374 (5.3)
Cyprus	25 (1.1)	466 (4.4)	45 (1.1)	486 (2.7)	21 (0.8)	479 (3.7)	9 (0.7)	441 (5.7)
Czech Republic	15 (0.8)	556 (7.5)	45 (1.2)	575 (6.2)	31 (1.2)	562 (4.3)	9 (0.8)	531 (8.9)
<i>Denmark</i>	28 (1.1)	499 (3.9)	42 (1.2)	507 (4.0)	22 (1.0)	510 (4.5)	8 (0.7)	488 (6.0)
England	20 (1.3)	500 (8.1)	37 (1.2)	515 (3.9)	31 (1.2)	516 (3.7)	11 (0.9)	481 (6.1)
France	42 (1.3)	546 (3.9)	45 (1.1)	539 (2.9)	9 (0.7)	532 (5.5)	4 (0.5)	494 (10.8)
<i>Germany</i>	31 (1.0)	510 (6.2)	47 (1.1)	517 (4.5)	16 (0.8)	511 (5.9)	6 (0.6)	467 (7.4)
<i>Greece</i>	32 (0.9)	486 (3.5)	42 (0.7)	489 (3.7)	17 (0.7)	486 (4.9)	9 (0.5)	470 (5.7)
Hong Kong	22 (0.9)	582 (7.7)	39 (0.9)	599 (6.8)	28 (1.0)	599 (6.5)	11 (0.8)	556 (9.1)
<i>Hungary</i>	11 (0.7)	550 (6.2)	41 (1.1)	552 (4.0)	33 (0.9)	537 (3.9)	15 (1.0)	496 (5.2)
Iceland	24 (1.3)	475 (7.4)	47 (1.3)	494 (4.5)	22 (1.2)	498 (5.7)	7 (0.8)	473 (11.8)
Iran, Islamic Rep.	32 (1.3)	421 (3.1)	46 (0.9)	434 (2.9)	17 (0.9)	438 (4.1)	5 (0.6)	425 (7.9)
Ireland	20 (0.8)	517 (6.4)	51 (1.1)	539 (5.2)	23 (0.8)	531 (5.3)	5 (0.5)	486 (8.5)
<i>Israel</i>	9 (1.4)	506 (17.0)	33 (2.1)	536 (7.0)	44 (1.7)	525 (5.4)	14 (1.2)	505 (7.8)
Japan	9 (0.5)	606 (5.7)	53 (0.9)	615 (2.1)	30 (0.8)	596 (3.4)	9 (0.5)	569 (5.1)
Korea	32 (1.0)	612 (4.6)	40 (1.0)	618 (3.4)	20 (0.8)	595 (5.3)	7 (0.6)	570 (6.9)
<i>Kuwait</i>	39 (1.7)	386 (2.9)	38 (1.3)	398 (3.3)	14 (1.2)	400 (3.8)	9 (0.8)	384 (4.1)
Latvia (LSS)	16 (1.0)	474 (4.4)	44 (1.1)	500 (3.7)	29 (1.2)	509 (4.2)	10 (0.7)	475 (5.1)
Lithuania	12 (0.7)	469 (6.2)	44 (1.3)	480 (4.6)	32 (1.2)	483 (4.0)	12 (0.9)	472 (5.8)
<i>Netherlands</i>	17 (1.8)	544 (14.0)	47 (1.7)	556 (7.0)	27 (1.5)	529 (6.3)	9 (0.9)	496 (7.3)
New Zealand	24 (1.0)	506 (6.4)	38 (0.9)	521 (4.8)	26 (0.9)	510 (4.7)	12 (0.8)	474 (5.7)
Norway	15 (0.7)	508 (4.2)	48 (1.0)	509 (2.5)	30 (1.0)	503 (3.7)	7 (0.4)	470 (6.0)
Portugal	27 (1.0)	450 (3.3)	48 (0.9)	458 (2.9)	20 (0.8)	460 (3.3)	5 (0.5)	440 (5.3)
<i>Romania</i>	38 (1.4)	475 (5.6)	39 (1.2)	489 (5.5)	16 (0.9)	495 (5.6)	8 (0.7)	470 (7.7)
Russian Federation	12 (1.0)	515 (6.9)	42 (1.4)	538 (5.9)	32 (1.0)	547 (4.8)	14 (0.9)	535 (7.5)
<i>Scotland</i>	15 (0.7)	488 (7.2)	43 (1.0)	504 (6.9)	31 (1.0)	508 (5.9)	11 (0.7)	472 (4.8)
Singapore	7 (0.6)	657 (7.2)	50 (1.1)	650 (5.2)	37 (1.2)	636 (5.2)	6 (0.5)	619 (8.6)
Slovak Republic	14 (0.7)	561 (7.4)	47 (1.0)	550 (3.5)	28 (0.9)	547 (4.1)	11 (0.8)	523 (5.6)
<i>Slovenia</i>	23 (1.1)	546 (4.1)	54 (1.1)	541 (3.4)	19 (0.9)	540 (4.7)	4 (0.4)	518 (9.9)
Spain	33 (1.2)	481 (3.0)	46 (1.0)	494 (2.4)	17 (0.8)	489 (3.9)	4 (0.5)	464 (5.1)
Sweden	16 (0.7)	518 (4.9)	51 (0.9)	528 (3.3)	27 (0.8)	514 (3.7)	6 (0.5)	478 (5.5)
Switzerland	45 (1.5)	556 (4.1)	44 (1.3)	543 (3.2)	9 (0.7)	528 (6.6)	2 (0.2)	~ ~
<i>Thailand</i>	28 (1.4)	510 (4.7)	46 (1.0)	524 (6.4)	19 (1.1)	540 (7.3)	8 (0.7)	521 (6.9)
United States	22 (0.8)	504 (5.7)	40 (0.9)	513 (5.1)	25 (0.6)	501 (4.2)	13 (1.0)	461 (4.6)

*Eighth grade in most countries; see Table 2 for more information about the grades tested in each country.

() Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.

Countries shown in italics did not satisfy one or more guidelines for sample participation rates, age/grade specifications, or classroom

sampling procedures (see Figure A.3). Background data for Bulgaria and South Africa are unavailable.

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

A tilde (~) indicates insufficient data to report achievement.

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

Figure 9.2
Sample Table for Teacher-Level Analysis Taken From the TIMSS International Report
“Mathematics Achievement in the Middle School Years”

C H A P T E R 9

Table 5.3

Teachers' Reports on Their Years of Teaching Experience
Mathematics - Upper Grade (Eighth Grade*)

Country	0 - 5 Years		6-10 Years		11-20 Years		More than 20 Years	
	Percent of Students	Mean Achievement	Percent of Students	Mean Achievement	Percent of Students	Mean Achievement	Percent of Students	Mean Achievement
<i>Australia</i>	18 (2.3)	517 (8.5)	19 (2.6)	528 (11.6)	35 (2.7)	540 (8.5)	28 (2.6)	533 (8.5)
<i>Austria</i>	r 7 (2.3)	516 (19.7)	13 (2.5)	546 (9.5)	51 (4.0)	554 (6.7)	28 (3.6)	549 (8.8)
Belgium (Fl)	10 (2.8)	556 (17.9)	9 (2.2)	590 (14.5)	32 (4.8)	554 (13.4)	49 (4.9)	575 (10.6)
<i>Belgium (Fr)</i>	s 8 (3.2)	536 (12.3)	8 (2.3)	528 (13.8)	31 (5.2)	558 (7.0)	54 (4.8)	543 (6.4)
Canada	17 (2.6)	527 (6.7)	15 (2.9)	527 (5.0)	22 (3.6)	526 (7.6)	46 (3.8)	528 (3.8)
<i>Colombia</i>	18 (3.0)	409 (7.7)	22 (5.0)	375 (11.7)	27 (4.3)	385 (6.0)	33 (4.2)	385 (5.0)
<i>Cyprus</i>	r 30 (4.6)	474 (4.6)	19 (4.3)	474 (7.6)	25 (5.0)	467 (6.4)	26 (4.7)	471 (5.5)
Czech Republic	12 (3.1)	566 (17.7)	9 (1.9)	538 (8.6)	17 (4.1)	584 (11.4)	62 (4.7)	562 (5.7)
<i>Denmark</i>	4 (1.9)	487 (2.6)	4 (2.0)	493 (14.4)	47 (4.9)	504 (3.3)	45 (4.8)	508 (4.4)
England	s 19 (2.5)	522 (10.8)	11 (2.1)	518 (13.5)	39 (3.5)	512 (8.1)	31 (3.0)	515 (11.3)
France	11 (2.5)	539 (8.1)	11 (3.1)	529 (10.2)	26 (4.6)	540 (8.8)	52 (4.3)	538 (5.4)
<i>Germany</i>	s 10 (2.2)	534 (14.5)	14 (4.3)	471 (12.1)	32 (5.1)	521 (10.6)	44 (5.5)	516 (9.3)
<i>Greece</i>	16 (3.1)	464 (7.2)	20 (3.4)	469 (5.3)	47 (4.3)	490 (3.5)	17 (4.4)	503 (11.9)
Hong Kong	53 (5.9)	585 (9.7)	14 (3.3)	606 (16.3)	18 (4.2)	574 (19.2)	15 (3.9)	596 (19.8)
Hungary	13 (2.9)	530 (12.7)	10 (2.8)	510 (7.4)	38 (4.1)	537 (5.6)	38 (4.1)	547 (5.2)
<i>Iceland</i>	r 19 (5.1)	478 (5.3)	14 (3.8)	480 (8.5)	33 (7.1)	492 (7.3)	35 (7.7)	496 (10.6)
Iran, Islamic Rep.	38 (4.5)	417 (3.7)	24 (4.8)	437 (3.8)	24 (4.3)	433 (3.2)	14 (3.0)	440 (4.8)
Ireland	13 (3.0)	513 (16.3)	18 (3.5)	512 (12.5)	42 (4.5)	535 (8.4)	28 (4.5)	523 (10.0)
<i>Israel</i>	r 16 (6.1)	490 (9.1)	12 (4.3)	555 (15.9)	45 (7.4)	510 (8.3)	27 (7.4)	548 (13.7)
Japan	19 (3.3)	606 (5.0)	25 (3.5)	607 (4.3)	36 (3.8)	598 (3.5)	19 (2.9)	614 (4.0)
Korea	28 (3.5)	610 (4.7)	29 (3.9)	622 (5.6)	23 (3.7)	597 (5.6)	20 (3.1)	606 (5.5)
<i>Kuwait</i>	r 30 (6.7)	397 (3.3)	33 (5.5)	388 (3.4)	31 (7.0)	388 (4.1)	6 (4.1)	418 (8.5)
Latvia (LSS)	12 (3.4)	496 (7.0)	16 (3.4)	482 (8.8)	38 (5.0)	496 (5.5)	34 (5.1)	490 (5.8)
<i>Lithuania</i>	r 5 (1.8)	455 (9.2)	15 (3.3)	465 (11.0)	33 (4.2)	482 (8.4)	47 (4.3)	481 (5.2)
<i>Netherlands</i>	13 (3.6)	530 (19.5)	21 (3.6)	525 (10.2)	42 (5.3)	548 (17.8)	24 (4.0)	556 (9.3)
New Zealand	17 (3.1)	497 (7.5)	28 (4.0)	515 (7.9)	34 (4.1)	517 (9.2)	20 (3.4)	487 (9.4)
Norway	12 (2.7)	499 (10.7)	10 (2.5)	500 (6.1)	35 (4.0)	508 (4.0)	43 (4.6)	503 (3.4)
Portugal	51 (4.7)	449 (3.0)	16 (3.1)	447 (5.4)	27 (3.9)	462 (4.3)	6 (2.3)	477 (8.6)
<i>Romania</i>	10 (2.3)	452 (14.2)	15 (3.1)	466 (9.9)	14 (3.1)	496 (12.8)	61 (4.2)	486 (5.7)
Russian Federation	16 (3.7)	541 (25.2)	14 (2.5)	532 (9.7)	29 (4.0)	526 (7.1)	41 (5.0)	538 (6.6)
<i>Scotland</i>	17 (3.4)	483 (9.2)	12 (3.2)	484 (14.3)	42 (4.4)	496 (8.5)	29 (4.3)	507 (12.3)
Singapore	30 (4.5)	617 (9.4)	11 (2.8)	658 (14.0)	11 (3.0)	664 (13.4)	48 (4.6)	652 (7.0)
Slovak Republic	6 (1.9)	556 (13.3)	15 (3.3)	531 (8.5)	21 (3.5)	539 (8.2)	58 (4.5)	553 (4.6)
<i>Slovenia</i>	r 4 (1.9)	537 (23.2)	19 (4.0)	533 (6.0)	55 (5.0)	542 (5.5)	22 (3.8)	550 (6.2)
Spain	3 (0.8)	472 (17.7)	8 (2.4)	487 (7.6)	39 (4.3)	488 (3.8)	50 (4.3)	488 (3.1)
Sweden	16 (2.4)	529 (7.1)	15 (2.8)	512 (9.5)	26 (3.1)	518 (6.2)	44 (4.1)	520 (4.4)
Switzerland	14 (3.3)	540 (10.1)	6 (1.8)	545 (19.0)	37 (4.6)	549 (8.4)	42 (4.9)	548 (7.4)
<i>Thailand</i>	s 48 (6.6)	517 (8.9)	12 (2.6)	499 (9.3)	35 (6.2)	540 (10.9)	5 (3.4)	615 (17.7)
United States	25 (3.4)	484 (6.3)	14 (2.7)	488 (9.8)	25 (3.2)	501 (7.3)	36 (3.3)	513 (7.5)

*Eighth grade in most countries; see Table 2 for more information about the grades tested in each country.
 Countries shown in italics did not satisfy one or more guidelines for sample participation rates, age/grade specifications, or classroom sampling procedures (see Figure A.3). Background data for Bulgaria and South Africa are unavailable.
 Because population coverage falls below 65%, Latvia is annotated LSS for Latvian Speaking Schools only.
 () Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.
 An "r" indicates teacher response data available for 70-84% of students. An "s" indicates teacher response data available for 50-69% of students.

SOURCE: IEA Third International Mathematics and Science Study (TIMSS), 1994-95.

9.1 Contents of the CDs

There are two CDs that accompany this User Guide – one CD containing the Population 1 data and one containing the Population 2 data. Each CD has the following internal file structure:

- A main directory identifying the population (POP 1 or POP 2).
- 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

The directory names within each 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
.LIS	Almanac
.CDT	Codebook in printout 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 7.14. 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:

JACK.SAS and JACK.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. These macros also generate replicate weights and compute JRR sampling variance for the percent and mean estimates. Although, in general, the continuous variable of choice will be one of the achievement scores in either mathematics or science, this continuous variable can be any other continuous variable in the file. How to use each of these macro programs is described later in this chapter.

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 only in the CD for Population 2, although the same examples can be easily adapted to perform the same analyses with the Population 1 data.

ASASCORE.SAS, BSASCORE.SAS, ASPSCORE.SAS, BSPSCORE.SAS
ASASCORE.SPS, BSASCORE.SPS, ASPSCORE.SPS, BSPSCORE.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.

The files beginning with the letters ASA and BSA can be used to recode the items from the written assessment. Those beginning with the letter ASP and BSP can be used to recode the items from the performance assessment.

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 9.3, and for creating the file in SPSS in Figure 9.4. 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 should 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.

Figure 9.3**Extract from SAS Control Code for Creating a Student Background SAS Data Set**

```

LIBNAME libdat "{dataset_library}" ;
LIBNAME library "{format_library}" ;
FILENAME rawinp "bsg<country>.dat" ;

PROC FORMAT LIBRARY=library ;
.
.
DATA libdat.bsg<country> ;
.
.{INSERT VALUE FORMATS HERE}
.
INFILE rawinp LRECL=1024 END=eof LENGTH=inline LINE=ln MISSEVER ;
*LENGTH STATEMENTS FOR ALL VARIABLES ;
LENGTH DEFAULT=5 ;
LENGTH VERSION 3 ;
.
.{INSERT VARIABLE LENGTH STATEMENTS HERE}
.
*MISSING CATEGORIES ;
MISSING A B R N ;
INPUT VERSION 1 - 2
.
.{INSERT VARIABLE NAME, LOCATION, AND FORMATS HERE}
.
*Assign value labels to variables ;
FORMAT IDPOP ASG3F. ;
.
.{INSERT ASSIGNMENT OF FORMATS TO VARIABLES HERE}
.
*Assign labels to variables ;
LABEL
VERSION = "*FILE VERSION*"
.
{INSERT VARIABLE LABELS HERE}
.
*RECODE MISSING VALUES
Select(IDSTRAT ) ;
when(998) IDSTRAT =.A ;
when(999) IDSTRAT =. ;
otherwise;
end;
.
.{INSERT RECODES FOR MISSING VALUES HERE}
.

```

Figure 9.4

Extract from SPSS Control Code for Creating a Student Background SPSS Data Set

```
*NOTE: Assignment of variable names.
data list file = bsg<country>1.dat
/      VERSION  1      - 2
.
.{INSERT VARIABLE NAME FORMAT AND LOCATION HERE}
.
*NOTE: Assignment of variable labels.
variable labels      VERSION  "*FILE VERSION*"
.
.{INSERT VARIABLE LABELS HERE}
.
VALUE LABELS
      IDPOP          1      "population 1"
                   2      "population 2"
                   3      "population 3"
/
.
.{INSERT VARIABLE LABELS HERE}
.
*RECODE of sysmis-values.
RECODE      IDSTRAT  (999          = sysmis).
.
.{INSERT RECODES OF MISSING VALUES HERE}
.
*RECODE of other missing-values.
MISSING value  IDSTRAT  (998 ).
.
.{DEFINE MISSING VALUES HERE}
.
SAVE OUTFILE = bsg<country>.sys.
```

9.3 Computing JRR Replicate Weights and Sampling Variance 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, 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.

9.3.1 SAS Macro for Computing Mean and Percents with Corresponding Standard Errors (JACK.SAS)

The CD containing the TIMSS International Database contains a SAS macro program called JACK.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 JACK.SAS is presented in Figure 9.5.

Figure 9.5**SAS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACK.SAS)**

```

%macro jack (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))));

* Now create the data file with the replicate weights;
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;
  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;                        * for the sum of X;

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;
  by k;
  maxtype2 = maxtype-1;

* Select the record which has the totals for &cvar2 and save to a file;
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 9.5 continued

```

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 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) ; * Compute the mean for each replicate;
    else rmn(i) = . ;
    rp(i) = 100 * (rwt(i) / trwt(i)) ; * Compute the proportion of each replicate;
    if rmn(i) <> . then
      mn_jvar = mn_jvar + (rmn(i) - mnx )**2; * Compute JK_VAR for Mean ;
      p_jvar = p_jvar + (rp(i) - pct )**2; * Compute JK_VAR for P ;
    if rwt(i) > 0 then
      n_jvar = n_jvar + (&wgt - rwt(i))**2; * Compute JK_VAR for N ;
  end;
  mnx_se = sqrt(mn_jvar ) ;
  pct_se = sqrt(p_jvar ) ;
  n_se = sqrt(n_jvar ) ;
run;

%mend jack;

```

The user needs to know some basic SAS macro language in order to use JACK.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 "{directory_location}jack.sas";
```

where {directory_location} points to the specific drive and directory where the macro JACK.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. Generally TOTWGT when using the student files, or MATWGT, SCIWGT, or TCHWGT when using the teacher files.

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, school, and teacher background data, and 42 for the performance assessment data. The user working with the data for only one country should set the NJKR argument to as many replicates as there were in the 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.

Table 9.2
Number of Replicate Weights Needed for Computing the JRR Error Variance Estimate

Country	3rd Grade	4th Grade	7th Grade	8th Grade
Australia	74	74	74	74
Austria	68	68	65	66
Belgium (Fl)	-	-	71	71
Belgium (Fr)	-	-	60	60
Bulgaria	-	-	52	58
Canada	75	75	75	75
Colombia	-	-	71	71
Cyprus	74	74	55	55
Czech Republic	73	73	75	75
Denmark	-	-	75	75
England	67	67	64	64
France	-	-	67	68
Germany	-	-	69	69
Greece	75	75	75	75
Hong Kong	62	62	43	43
Hungary	75	75	75	75
Iceland	75	75	75	75
Iran, Islamic Rep.	75	75	75	75
Ireland	73	73	66	66
Israel	-	44	-	23
Japan	74	74	75	75
Korea	75	75	75	75
Kuwait	-	75	-	36
Latvia (LSS)	59	59	64	64
Lithuania	-	-	73	73
Netherlands	52	52	48	48
New Zealand	75	75	75	75
Norway	70	70	72	74
Portugal	72	72	71	71
Romania	-	-	72	72
Russian Federation	-	-	41	41
Scotland	65	65	64	64
Singapore	75	75	69	69
Slovak Republic	-	-	73	73
Slovenia	61	61	61	61
South Africa	-	-	66	66
Spain	-	-	75	75
Sweden	-	-	75	60
Switzerland	-	-	75	75
Thailand	75	75	74	74
United States	59	59	55	55

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 IDCNTRY.

DVAR

This is the variable for which means are to be computed. Only one variable has to be listed here. If the user wants to find out, for example, the results in mathematics and science, then the macro needs to be invoked separately to generate each table. Although in most cases the continuous variable of interest will be an achievement variable, this can actually be any other continuous variable in the data file.

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:

```
%jack(TOTWGT ,JKZONE,JKINDIC,75,IDCNTRY IDGRADER ITSEX, BIMATSCR, BSGALL1);
```

it will compute the mean mathematics achievement, and its corresponding standard error, for boys and girls, by grade, within each country, using the variable TOTWGT as the sampling weight. It will also compute the percent of boys and girls by grade within the country, and its corresponding standard error. The data will be read from the data set BSGALL1, 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 would be three variables in the WORK.FINAL data set. These would be IDCNTRY, IDGRADER, 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 our example this variable is called TOTWGT .

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 our example, we would obtain 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 then 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 9.6. In this example the macro computes the percent of boys and girls, by grade and by country, and their mean achievement in mathematics. The listing presented in Figure 9.6 is interpreted in the following way. The first line shows that there were 3039 students in the sample for IDCNTRY=36 (Australia), in IDGRADER=1 (Seventh grade), and who had ITSEX=1 (Girls). It is estimated that there are 123649 seventh-grade girls in Australia, their mean mathematics score is 500, with a standard error of 4.3. We can also tell from this line of data that it is estimated that 52 percent of the seventh graders in Australia are girls, and the standard error of this percent is 2.3. The second line shows the same information, but for the seventh-grade boys (ITSEX=2). It is estimated that there are 114646 seventh-grade boys in Australia, their mean mathematics score is 495, with a standard error of 5.2. We can also tell from this line of data that it is estimated that 48 percent of the seventh graders in Australia are boys, and the standard error of this percent is 2.3.

Figure 9.6**SAS Control Code and Extract of Output File for Using the Macro JACK.SAS**

```

options nocenter;
* Read the variables from the student file;
data student;
  set bsgall1
      (keep=idcntry idgrader jkindic jkzone totwgt itsex bimatscr);
  where ITSEX >0;

%include "jack.sas";

%jack (totwgt,jkzone,jkindic,75,idcntry idgrader,itsex,bimatscr,student);

proc print data=final noobs;

```

IDCNTRY	IDGRADER	ITSEX	N	TOTWGT	MNX	PCT	MNX_SE	PCT_SE
36	1	1	3039	123648.61	500.451	51.8891	4.32185	2.32468
36	1	2	2560	114645.55	495.142	48.1109	5.20426	2.32468
36	2	1	3722	114806.50	531.990	49.6409	4.58168	2.13471
36	2	2	3529	116467.28	527.357	50.3591	5.10432	2.13471
40	1	1	1545	45945.00	508.568	53.2220	3.27647	1.63905
40	1	2	1358	40382.01	509.963	46.7780	4.58177	1.63905
40	2	1	1321	42574.48	535.625	50.0342	4.54757	1.79216
40	2	2	1385	42516.32	543.581	49.9658	3.23439	1.79216
56	1	1	1344	31717.08	558.544	49.4211	4.71205	2.76051
56	1	2	1424	32460.16	556.717	50.5789	4.47097	2.76051
56	2	1	1437	37480.16	567.222	49.9276	7.43010	3.70983
56	2	2	1457	37588.84	563.139	50.0724	8.78875	3.70983
.
.
.
.
.
840	1	1	1976	1587912.47	473.310	50.3006	5.74234	1.20563
840	1	2	1910	1568934.35	478.069	49.6994	5.74972	1.20563
840	2	1	3561	1574689.22	497.464	49.3897	4.47104	0.77280
840	2	2	3526	1613607.42	501.996	50.6103	5.21162	0.77280
890	1	1	1486	14410.14	495.847	51.3907	3.22189	0.76075
890	1	2	1411	13630.25	500.643	48.6093	3.48741	0.76075
890	2	1	1381	13341.15	536.878	51.3542	3.29050	0.90855
890	2	2	1324	12637.53	544.918	48.6458	3.76508	0.90855

9.3.2 SPSS Macro for Computing Mean and Percents with Corresponding Standard Errors (JACK.SPS)

The CD containing the TIMSS International Database also contains an SPSS macro program called JACK.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 JACK.SPS also computes the JRR error estimate for these means and percents. The control code for the macro JACK.SPS is presented in Figure 9.7.

Figure 9.7
SPSS Macro for Computing Mean and Percents with Corresponding JRR Standard Errors (JACK.SPS)

```
set mprint=on.

define jack      (cvar = !charend('/') /
                 dvar = !charend('/') /
                 njkr = !charend('/') /
                 jkz  = !charend('/') /
                 jki  = !charend('/') /
                 wgt  = !charend('/') ) .

weight off.

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

!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)
```

Figure 9.7 continued

```

aggregate outfile = tmpjckA
  / break = !cvar2a
  /      !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)).

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.

!if (!head(!cvar1)=!head(!cvar2a)) !then
save outfile = jackfn1.
!else
add files
  / file = jackfn1
  / file = *.
save outfile = jackfn1.
!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.

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 needs to 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 needs to be called in every batch. If the user is using SPSS interactively then 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. Once the macro is included in a specific session the

word “JACK” 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 "{directory_location}jack.sps".
```

where {directory_location} points to the specific drive and directory where the macro JACK.SPS can be found. The macro requires that several argument be submitted when it is invoked. These parameters are:

WGT

The sampling weight to be used in the analysis. Generally TOTWGT when using the student files, or MATWGT, SCIWGT, or TCHWGT when using the teacher files.

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, school, and teacher background data, and 42 for the performance assessment data. The user working with the data for only one country should set the NJKR argument to as many replicates there were in the country (see Table 9.2 for the maximum number of replicates by country). 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 IDCNTRY.

DVAR

This is the variable for which means are to be computed. Only one variable has to be listed here. If the user wants to examine, for example, results in mathematics and science, then the macro needs to be invoked separately to generate each table. Although in most cases the continuous variable of interest will be an achievement variable, this can actually be any other continuous variable.

Unlike the SAS macro (JACK.SAS), the JACK 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 = BSGALL1.
jack cvar = IDCNTRY IDGRADER ITSEX /
      dvar = BIMATSCR/
      jkz  = JKZONE /
      jki  = JKINDIC /
      njkr = 75 /
      WGT  = TOTWGT / .
```

it will compute the mean mathematics achievement, and its corresponding standard error, for boys and girls, by grade, within each country, using the variable TOTWGT as the sampling weight. It will also compute the percent of boys and girls, by grade, within the country, and its corresponding standard error. The data would be read from the system file BSGALL1.

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 is kept in the resulting file. In our example above there would be three variables in the resulting system file. These would be IDCNTRY, IDGRADER, 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.

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 MNX 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 our example, we would obtain 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 then 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 9.8. In this example the macro will compute the percent of boys and girls, by grade and by country, and their mean achievement in mathematics. The listing presented in Figure 9.8 is interpreted in the following way. The first line shows the results for the students with IDCNTY=36 (Australia), in IDGRADER=1 (Seventh grade), and who had ITSEX=1 (Girls). It is estimated that there are 123649 seventh-grade girls in Australia, their mean mathematics score is 500, with a standard error of 4.3. We can also tell from this line of data that it is estimated that 52 percent of the seventh graders in Australia are girls, and the standard error of this percent is 2.3. The second line shows the same information, but for the seventh-grade boys (ITSEX=2). It is estimated that there are 114646 seventh grade boys in Australia, their mean mathematics score is 495, with a standard error of 5.2. We can also tell from this line of data that it is estimated that 48 percent of the seventh graders in Australia are boys, and the standard error of this percent is 2.3.

Figure 9.8**SPSS Control Code and Extract of Output File for Using the Macro JACK.SPS**

```

get file = "bsgall1.sys"
  / keep=idcntry idstud idgrader jkindic jkzone totwgt itsex bimatscr .

select if not(missing(itsex)).

* Now use the macro to get the results.
include "jack.sps".

jack  cvar = idcntry idgrader itsex /
      dvar = bimatscr /
      njkr = 75 /
      jkz = jkzone /
      jki = jkindic /
      wgt = totwgt / .

sort cases by idcntry idgrader.

list var = all.

```

IDCNTRY	IDGRADER	ITSEX	N	TOTWGT	MNX	PCT	MNX_SE	PCT_SE
36	1	1	3039	123648.61	500.451	51.8891	4.32185	2.32468
36	1	2	2560	114645.55	495.142	48.1109	5.20426	2.32468
36	2	1	3722	114806.50	531.990	49.6409	4.58168	2.13471
36	2	2	3529	116467.28	527.357	50.3591	5.10432	2.13471
40	1	1	1545	45945.00	508.568	53.2220	3.27647	1.63905
40	1	2	1358	40382.01	509.963	46.7780	4.58177	1.63905
40	2	1	1321	42574.48	535.625	50.0342	4.54757	1.79216
40	2	2	1385	42516.32	543.581	49.9658	3.23439	1.79216
56	1	1	1344	31717.08	558.544	49.4211	4.71205	2.76051
56	1	2	1424	32460.16	556.717	50.5789	4.47097	2.76051
56	2	1	1437	37480.16	567.222	49.9276	7.43010	3.70983
56	2	2	1457	37588.84	563.139	50.0724	8.78875	3.70983
.
.
.
.
.
840	1	1	1976	1587912.47	473.310	50.3006	5.74234	1.20563
840	1	2	1910	1568934.35	478.069	49.6994	5.74972	1.20563
840	2	1	3561	1574689.22	497.464	49.3897	4.47104	0.77280
840	2	2	3526	1613607.42	501.996	50.6103	5.21162	0.77280
890	1	1	1486	14410.14	495.847	51.3907	3.22189	0.76075
890	1	2	1411	13630.25	500.643	48.6093	3.48741	0.76075
890	2	1	1381	13341.15	536.878	51.3542	3.29050	0.90855
890	2	2	1324	12637.53	544.918	48.6458	3.76508	0.90855

9.4 Performing Analyses with Student-Level Variables

Many analyses of the TIMSS data can be undertaken using student-level data. We have already presented one 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 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 looking at the eighth graders' reports on the hours spent each day watching television and videos, and their achievement in mathematics. These are the results that are presented in Figure 9.1 earlier in this chapter.

To do this we need to undertake several steps. After reviewing the codebooks and the questionnaire information we find out if the students were asked a question about the number of hours they watch television (see Supplement 2). The data collected from this variable are captured in the variable BSBGDAY1, 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 3). 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 international mathematics achievement score (BIMATSCR), the sampling weight for the student (TOTWGT), the variables that contain the jackknife replication information (JKZONE and JKINDIC), the variable that will be used to select the eighth graders from the data file (IDGRADER), 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 exact same steps need to be taken if the user wants to examine 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 9.9, and the SPSS code is presented in Figure 9.10. A selection of the results obtained from these programs are displayed in Figure 9.11 and Figure 9.12. 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 eighth grade who chose one of the five options presented in the question, 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.

In general, to perform analyses such as those in figures 9.9 and 9.10 using 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, including the achievement score, sampling weights, JRR replication information, and any other variables used in the selection of cases.
- Use the macro JACK with the corresponding arguments and parameters.
- Print out the result file.

Figure 9.9**SAS Control Statements for Performing Analyses with Student-Level Variables
(EXAMPLE1.SAS)**

```

options nocenter;
* Read the variables from the student file;
data student;
  set bsgall1
      (keep=idcntry idgrader jkindic jkzone totwgt bsbgday1 bimatscr);
  where idgrader=2 and nmiss(bsbgday1)=0; * Select only 8th graders;
  if bsbgday1=1 then bsbgday1=2;

* Define the format for the variables used;
proc format library=work;
  value grade 1 = "Seventh Grade " 2 = "Eighth Grade ";

  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 JACK to get the results;
%include "jack.sas";

%jack (totwgt,jkzone,jkindic,75,idcntry idgrader bsbgday1,bimatscr,student);

proc print data=final noobs;
  by idcntry idgrader;
  where idgrader = 2;
  var bsbgday1 N totwgt mnx mnx_se pct pct_se;
  format idcntry country. idgrader grade. bsbgday1 seetv.;

```

Figure 9.10**SPSS Control Statements for Performing Analyses with Student-Level Variables
(EXAMPLE1.SPS)**

```

get file = "bsgall1.sys"
  / keep=idcntry idstud idgrader jkindic jkzone totwgt
    bsbgday1 bimatscr .

recode bsbgday1 (1,2=2) (8,9=sysmis) (else=copy).

select if not(missing(bsbgday1)).

value labels
  idgrader 1 "Seventh Grade " 2 "Eighth Grade " /
  bsbgday1 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 JACK to get the results.
include "jack.sps".

jack  cvar = idcntry idgrader bsbgday1 /
      dvar = bimatscr /
      njkr = 75 /
      jkz = jkzone /
      jki = jkindic /
      wgt = totwgt /

sort cases by idcntry idgrader.

temporary.
select if idgrader = 2.
report format=list automatic
  / var = bsbgday1 (label) totwgt mnx mnx_se pct pct_se
  / break = idcntry idgrader.

```

Figure 9.11**Extract of SAS Computer Output for Performing Analyses with Student-Level Variables (EXAMPLE 1)**

COUNTRY ID=Australia *GRADE*=Eighth Grade

BSBGDAY1	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Less than 1 Hr.	1651	52718.94	538.883	5.99951	23.5254	0.91886
1 to 2 hrs.	2943	92637.64	538.578	4.12722	41.3388	0.84451
3 to 5 hrs.	1880	59663.57	527.545	3.80272	26.6244	0.84064
More than 5 hrs.	563	19073.65	487.008	5.47335	8.5115	0.61293

COUNTRY ID=Austria *GRADE*=Eighth Grade

BSBGDAY1	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Less than 1 Hr.	673	21112.51	539.820	5.42470	25.2880	1.39738
1 to 2 hrs.	1402	44420.93	545.816	4.20850	53.2063	1.06455
3 to 5 hrs.	469	14158.96	539.469	5.19961	16.9592	0.97982
More than 5 hrs.	118	3795.76	496.845	8.63972	4.5465	0.57168

COUNTRY ID=Belgium (Fl) *GRADE*=Eighth Grade

BSBGDAY1	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Less than 1 Hr.	757	18055.42	579.948	6.6674	24.3543	1.24276
1 to 2 hrs.	1517	38846.19	575.118	6.1529	52.3983	1.17633
3 to 5 hrs.	478	13808.20	534.656	7.1230	18.6254	1.03135
More than 5 hrs.	112	3426.62	513.901	12.0735	4.6220	0.53665

COUNTRY ID=Belgium (Fr) *GRADE*=Eighth Grade

BSBGDAY1	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Less than 1 Hr.	887	19212.63	536.028	4.20418	33.3444	1.26106
1 to 2 hrs.	1124	25151.10	535.669	4.87625	43.6509	1.77399
3 to 5 hrs.	394	9549.30	522.292	3.98052	16.5732	1.32403
More than 5 hrs.	119	3705.76	445.040	9.01330	6.4315	0.98084

.
.

.

.

.

COUNTRY ID=United States *GRADE*=Eighth Grade

BSBGDAY1	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Less than 1 Hr.	1488	687980.70	503.686	5.72799	22.1953	0.75219
1 to 2 hrs.	2668	1239197.59	512.935	5.10220	39.9783	0.90504
3 to 5 hrs.	1727	767004.30	501.228	4.22001	24.7447	0.58673
More than 5 hrs.	1001	405491.75	460.767	4.57929	13.0818	0.95070

COUNTRY ID=Slovenia *GRADE*=Eighth Grade

BSBGDAY1	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Less than 1 Hr	614	5838.91	546.071	4.11421	22.6437	1.06473
1 to 2 hrs.	1442	13852.82	540.938	3.36504	53.7221	1.09384
3 to 5 hrs.	516	4992.78	540.094	4.65821	19.3623	0.86774
More than 5 hrs.	112	1101.55	517.721	9.92666	4.2719	0.42142

Figure 9.12
Extract of SPSS Computer Output for Performing Analyses with Student-Level Variables (EXAMPLE 1)

COUNTRY ID	GEN\OUTSIDE SCHL\WATCH TY OR VIDEOS	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Australia						
Eighth Grade	Less than 1 Hr.	52718.94	538.88	6.00	23.53	.92
	1 to 2 hrs.	92637.64	538.58	4.13	41.34	.84
	3 to 5 hrs.	59663.57	527.54	3.80	26.62	.84
	More than 5 hrs.	19073.65	487.01	5.47	8.51	.61
Austria						
Eighth Grade	Less than 1 Hr.	21112.51	539.82	5.42	25.29	1.40
	1 to 2 hrs.	44420.93	545.82	4.21	53.21	1.06
	3 to 5 hrs.	14158.96	539.47	5.20	16.96	.98
	More than 5 hrs.	3795.76	496.85	8.64	4.55	.57
Belgium (Fl)						
Eighth Grade	Less than 1 Hr.	18055.42	579.95	6.67	24.35	1.24
	1 to 2 hrs.	38846.19	575.12	6.15	52.40	1.18
	3 to 5 hrs.	13808.20	534.66	7.12	18.63	1.03
	More than 5 hrs.	3426.62	513.90	12.07	4.62	.54
Belgium (Fr)						
Eighth Grade	Less than 1 Hr.	19212.63	536.03	4.20	33.34	1.26
	1 to 2 hrs.	25151.10	535.67	4.88	43.65	1.77
	3 to 5 hrs.	9549.30	522.29	3.98	16.57	1.32
	More than 5 hrs.	3705.76	445.04	9.01	6.43	.98
.						
.						
.						
.						
United States						
Eighth Grade	Less than 1 Hr.	687980.70	503.69	5.73	22.20	.75
	1 to 2 hrs.	1239797.59	512.94	5.10	39.98	.91
	3 to 5 hrs.	767004.30	501.23	4.22	24.74	.59
	More than 5 hrs.	405491.75	460.77	4.58	13.08	.95
Slovenia						
Eighth Grade	Less than 1 Hr.	5838.91	546.07	4.11	22.64	1.06
	1 to 2 hrs.	13852.82	540.94	3.37	53.72	1.09
	3 to 5 hrs.	4992.78	540.09	4.66	19.36	.87
	More than 5 hrs.	1101.55	517.72	9.93	4.27	.42

9.5 Performing Analyses with Teacher-Level Variables

When analyzing the teacher data, it is first necessary to link the students with their corresponding teachers. Each student record in the Student Background data file can contain a link to as many as six different teachers in the Teacher Background data file. To facilitate the linking between students and their teachers in the teacher file, the Student-Teacher Linkage file was created and is part of the International Database. This file is called ALG<COUNTRY>1 in Population 1 and BLG<COUNTRY>1 in Population 2. The Student-Teacher Linkage file contains one record for each student-by-teacher combination, with the corresponding identification variables. Each record also contains the number of mathematics and science teachers for the student and a set of weights that can be used when conducting analyses with these data. In order to simplify the merging process for analyses that link teacher variables to student achievement, student achievement scores, sampling weights, and JRR replication information have been added to the Student-Teacher Linkage file. For such analyses it is necessary only to merge the Teacher Background file with the Student-Teacher Linkage file. For analyses linking teacher variables to other student variables, it is necessary also to merge the Student Background files to the other two.

Conducting analyses with the teacher data requires some extra steps that are not required when analyzing the student or school background data. The code in SAS and SPSS for the example that follows is presented in Figures 9.13 and Figure 9.14. For our example, we want to find out about the years of experience of the mathematics teachers that teach the eighth-grade students in each of the TIMSS countries. In particular, we want to find out what percent of eighth-grade students are taught by teachers of specific groups based on their years of experience, and the mean achievement in mathematics of the students taught by those teachers. These are the results that are presented in Figure 9.2 earlier in this chapter.

The TIMSS teacher files do not contain representative samples of teachers within a country. Rather, these are the teachers for a representative sample of students within a country. Therefore, it is appropriate that statements about the teachers be made only in terms of how many students are taught by teachers of one kind or another, and not in terms of how many teachers in the country do one thing or another.

Figure 9.13**SAS Control Statements for Performing Analyses with Teacher-Level Variables
(EXAMPLE2.SAS)**

```

options nocenter;

* Read the variables from the Mathematics Teacher file and sort by merge variables;
data teacher;
  set btmall1 (keep=idcntry idteach idlink btbgtaug);
  if btbgtaug >= 0 & btbgtaug <= 5 then btbgtaug = 1;
  if btbgtaug > 5 & btbgtaug <= 10 then btbgtaug = 2;
  if btbgtaug > 10 & btbgtaug <= 20 then btbgtaug = 3;
  if btbgtaug > 20 & btbgtaug <= 90 then btbgtaug = 4;

proc sort data=teacher;
  by idcntry idteach idlink;

* Read the variables from the student teacher link file and sort by merge variables;
data studteac;
  set blgall1
    (keep=idcntry idteach idlink idgrader jkindic jkzone matwgt bimatscr);
proc sort data=studteac;
  by idcntry idteach idlink;

* Now merge the two files;
data merged;
  merge studteac teacher;
  by idcntry idteach idlink;
  if matwgt > 0 and nmiss(btbgtaug) = 0;

* Define the format for the variables used;
proc format library=work;
  value grade 1 = "Seventh Grade" 2 = "Eighth Grade";

  value exper 1 = "0 to 5 years" 2 = "6 to 10 years"
    3 = "11 to 20 years" 4 = "Over 20 years ";

  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 JACK to get the results;
%include "jack.sas";

%jack (matwgt,jkzone,jkindic,75,idcntry idgrader btbgtaug,bimatscr,merged);

proc print data=final noobs;
  by idcntry idgrader;
  where idgrader=2;
  var btbgtaug N matwgt mnx mnx_se pct pct_se;
  format idcntry country. btbgtaug exper. idgrader grade.;

```

Figure 9.14**SPSS Control Statements for Performing Analyses with Teacher-Level Variables (EXAMPLE2.SPS)**

```

get file = "btmall1.sys"
  / keep = idcntry idteach idlink btbgtaug.

recode btbgtaug ( 0 thru 5 = 1) ( 6 thru 10 = 2)
               (10 thru 20 = 3) (20 thru 90 = 4)
               (98,99 = sysmis).

sort cases by idcntry idteach idlink.

save outfile = teacher.

get file = "blgall1.sys"
  / keep = idcntry idteach idlink idgrader jkindic jkzone matwgt bimatscr.

select if matwgt > 0 and idgrader=2.

sort cases by idcntry idteach idlink.

save outfile= studteac.

* Now merge the two files.
match files
  / file=studteac
  / table=teacher
  / by idcntry idteach idlink.

select if not(missing(btbgtaug)).

save outfile = merged.

* Define the format for the variables used.
value labels
idgrader 1 "Seventh Grade " 2 "Eighth Grade" /
btbgtaug 1 " 0 to 5 years " 2 " 6 to 10 years"
          3 "11 to 20 years " 4 "Over 20 years " /

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 "jack.sps".

jack  cvar      = idcntry idgrader btbgtaug /
      dvar      = bimatscr /
      njkr      = 75 /
      jkz       = jkzone /
      jki       = jkindic /
      wgt       = matwgt /

sort cases by idcntry idgrader.

temporary.
select if idgrader = 2.
report format=list automatic
  / var = btbgtaug (label) matwgt mnx mnx_se pct pct_se
  / break = idcntry idgrader.

```


As before, we first proceed to identify the variables relevant to the analysis in the corresponding files, and review the documentation on the specific national adaptations to the questions of interest (Supplements 1, 2, 3). Since we are using teacher-level variables we need to look into the teacher file and the Student-Teacher Linkage files to find the variables. From the mathematics teacher file we extract the variable that contains the information on the mathematics teachers' years of experience (BTBGTAUG), the variable that identifies the country (IDCNTRY) and the two variables that will allow us to link the teacher information to the student data (IDTEACH and IDLINK).

In Population 2 there is one teacher file for the mathematics teachers and a second teacher file for the science teachers. If the user wants to look only at mathematics teachers, then the user will need to use the mathematics teacher file (BTM<COUNTRY>1); if the interest is in the science teachers then the user will need to use the science teacher file (BTS<COUNTRY>1); but if the interest is in the mathematics and science teachers combined, both these files need to be combined by appending or adding one file to the other. In doing so it is important to keep in mind that although there are variables in common between these two files, most of them are not. On the other hand, there is only one file for the Population 1 teachers where both mathematics and science teacher variables are found (ATG<COUNTRY>1). In each population there is only one student-teacher-link file where the sampling and achievement information is found.

In our example, our teacher variable of interest (years of teaching experience) is a continuous variable. However, we want to categorize the teachers into 4 groups (0 to 5 years experience, 6 to 10 years experience, 11 to 20 years experience, and More than 20 years experience). While reading the Teacher file we use commands in SAS or SPSS to collapse the different values into four categories and we label them accordingly. We then proceed to read the necessary information from the Student-Teacher Linkage file. From this file we keep the country identification (IDCNTRY) and the two variables that will allow us to link the student information to the teacher data (IDTEACH and IDLINK). We also keep the variable that indicates the grade for the student (IDGRADER), the mathematics achievement score (BIMATSCR), and the information necessary to compute the replicate weights for estimating the JRR standard error. We select the variable that will give us the correct weight for mathematics teacher variable (MATWGT). If the user is interested in looking at the science teachers, then the weight variable that should be selected is SCIWGT, and TCHWGT if the user is interested in analyzing both mathematics and science teachers combined.

The two files are then merged or matched into one file that will then be used with the JACK macro. These two files will be merged using the variables IDCNTRY, IDTEACH, and IDLINK. The combination of values for these three variables is unique within the teacher data, but is repeated in the Student-Teacher Linkage file as many times as the specific teacher teaches students in a class. After the files are merged, the macro JACK is used and the results can be printed. Selected results from this analysis are shown in Figure 9.15 and Figure 9.16.

Figure 9.15**Extract of SAS Computer Output for Performing Analyses with Teacher-Level Variables (EXAMPLE 2)**

COUNTRY ID=Australia *GRADE*=Eighth Grade

BTBGTAUG	N	MATWGT	MXN	MXN_SE	PCT	PCT_SE
0 to 5 years	1174	35747.90	517.042	8.5001	17.6582	2.31566
6 to 10 years	1164	39265.42	528.252	11.5941	19.3957	2.55903
11 to 20 years	2205	71548.18	541.140	8.3675	35.3423	2.74924
Over 20 years	1859	55882.22	533.099	8.4781	27.6038	2.60588

COUNTRY ID=Austria *GRADE*=Eighth Grade

BTBGTAUG	N	MATWGT	MXN	MXN_SE	PCT	PCT_SE
0 to 5 years	159	4573.10	515.596	19.7319	7.1599	2.31322
6 to 10 years	298	8429.89	545.782	9.4934	13.1983	2.47861
11 to 20 years	1182	32826.23	553.934	6.6800	51.3945	4.00150
Over 20 years	500	18041.87	549.280	8.7719	28.2473	3.64019

COUNTRY ID=Belgium (Fl) *GRADE*=Eighth Grade

BTBGTAUG	N	MATWGT	MXN	MXN_SE	PCT	PCT_SE
0 to 5 years	281	6915.80	555.760	17.9455	9.5878	2.75864
6 to 10 years	306	6351.73	589.843	14.5214	8.8058	2.21509
11 to 20 years	864	23216.69	553.862	13.4302	32.1866	4.75228
Over 20 years	1357	35647.25	574.548	10.5636	49.4198	4.90806

COUNTRY ID=Belgium (Fr) *GRADE*=Eighth Grade

BTBGTAUG	N	MATWGT	MXN	MXN_SE	PCT	PCT_SE
0 to 5 years	159	3051.78	536.164	12.2866	8.1705	3.15008
6 to 10 years	168	2844.35	528.197	13.7981	7.6151	2.31007
11 to 20 years	565	11466.42	558.394	6.9786	30.6989	5.20517
Over 20 years	976	19988.70	543.097	6.3910	53.5155	4.84059

.
 .
 .
 .
 .

COUNTRY ID=United States *GRADE*=Eighth Grade

BTBGTAUG	N	MATWGT	MXN	MXN_SE	PCT	PCT_SE
0 to 5 years	1642	684804.90	483.554	6.34204	24.6560	3.42971
6 to 10 years	950	383442.47	488.370	9.75550	13.8056	2.69418
11 to 20 years	1408	706610.59	500.842	7.29526	25.4411	3.24706
Over 20 years	2107	1002577.37	513.453	7.51101	36.0972	3.31137

COUNTRY ID=Slovenia *GRADE*=Eighth Grade

BTBGTAUG	N	MATWGT	MXN	MXN_SE	PCT	PCT_SE
0 to 5 years	83	839.82	536.528	23.2197	3.8502	1.89562
6 to 10 years	435	4158.14	532.594	5.9679	19.0633	4.02549
11 to 20 years	1214	12025.34	541.671	5.4834	55.1310	4.97920
Over 20 years	542	4789.01	550.436	6.1702	21.9555	3.82312

Figure 9.16
Extract of SPSS Computer Output for Performing Analyses with Teacher-Level Variables
(EXAMPLE 2)

COUNTRY ID	GEN\YEARS BEEN TEACHING	MATWGT	MNX	MNX_SE	PCT	PCT_SE
Australia						
Eighth Grade	0 to 5 years	35747.90	517.04	8.50	17.66	2.32
	6 to 10 years	39265.42	528.25	11.59	19.40	2.56
	11 to 20 years	71548.18	541.14	8.37	35.34	2.75
	Over 20 years	55882.22	533.10	8.48	27.60	2.61
Austria						
Eighth Grade	0 to 5 years	4573.10	515.60	19.73	7.16	2.31
	6 to 10 years	8429.89	545.78	9.49	13.20	2.48
	11 to 20 years	32826.23	553.93	6.68	51.39	4.00
	Over 20 years	18041.87	549.28	8.77	28.25	3.64
Belgium (Fl)						
Eighth Grade	0 to 5 years	6915.80	555.76	17.95	9.59	2.76
	6 to 10 years	6351.73	589.84	14.52	8.81	2.22
	11 to 20 years	23216.69	553.86	13.43	32.19	4.75
	Over 20 years	35647.25	574.55	10.56	49.42	4.91
Belgium (Fr)						
Eighth Grade	0 to 5 years	3051.78	536.16	12.29	8.17	3.15
	6 to 10 years	2844.35	528.20	13.80	7.62	2.31
	11 to 20 years	11466.42	558.39	6.98	30.70	5.21
	Over 20 years	19988.70	543.10	6.39	53.52	4.84
.						
.						
.						
.						
.						
United States						
Eighth Grade	0 to 5 years	684804.90	483.55	6.34	24.66	3.43
	6 to 10 years	383442.47	488.37	9.76	13.81	2.69
	11 to 20 years	706610.59	500.84	7.30	25.44	3.25
	Over 20 years	1002577.37	513.45	7.51	36.10	3.31
Slovenia						
Eighth Grade	0 to 5 years	839.82	536.53	23.22	3.85	1.90
	6 to 10 years	4158.14	532.59	5.97	19.06	4.03
	11 to 20 years	12025.34	541.67	5.48	55.13	4.98
	Over 20 years	4789.01	550.44	6.17	21.96	3.82

In summary, to perform analyses such as those in Figures 9.15 and 9.16 using the Student and Teacher Background data files the user needs to do the following:

- Identify the variable or variables of interest in the corresponding teacher file and find out about any specific national adaptations to the variable.
- Retrieve the relevant variable or variables from the teacher data files. There is one teacher file for the teachers of the Population 1 students, and two for the teachers of the Population 2 students. If the user is interested in looking at both mathematics and science teachers combined for the Population 2 students, then the files for these teachers need to be added or appended to each other.
- Retrieve the relevant variables from the Student-Teacher Linkage file. This includes the identification information for the country and teacher (IDCNTRY, IDTEACH, and IDLINK), the achievement score, JRR replication information, and the sampling weight. If the analysis is to be based on mathematics teachers only, then the weight variable to use is MATWGT. If the analysis is to be based on the science teachers only, then the weight variable to be used is SCIWGT. If the analysis is to be based on the science and mathematics teachers combined, then the weight variable to be used is TCHWGT.
- Merge the variables from the teacher data files into the Student-Teacher Linkage files using the variables IDCNTRY, IDTEACH and IDLINK.
- Use the macro JACK with the corresponding arguments and parameters.
- Print out the result file.

9.6 Performing Analyses with School-Level Variables

Although the students in the TIMSS samples were selected from within a sample of schools, the school-level variables should only be analyzed with respect to the number of students attending schools of one type or another. 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 eighth graders that attend schools located in a certain geographical area of the country (BCBGCOMM), and their average achievement in mathematics. 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 BCBGCOMM in the School Background file, and the student weights and scores in the Student Background file. We then proceed to review the documentation on national adaptations and discover 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 BCBGCOMM 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). We keep these variables because these are the variables that will allow us to merge the school data to the student data. We also keep from the School Background file the variable of interest, in this case BCBGCOMM. We then read the variables of interest 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 variable the international mathematics achievement score (BIMATSCR), the sampling weight for the student (TOTWGT), the variables that contain the jackknife replication information (JKZONE and JKINDIC), and the variable that will be used to select the eighth graders from the data file (IDGRADER).

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 JACK to obtain the corresponding percents of students within each group, and their mean achievement in mathematics. The computer code used to run this analysis in SAS and SPSS can be found in Figure 9.17 and Figure 9.18, and the results are shown in Figure 9.19 and Figure 9.20.

Figure 9.17**SAS Control Statements for Performing Analyses with School-Level Variables (EXAMPLE3.SAS)**

```

options nocenter;

* Read the variables from the school file and sort them by the merge variables;
data school;
  set bcgall1
    (keep=idcntry idschool bcbgcomm);

proc sort data=school;
  by idcntry idschool;

* Read the variables from the student file and sort them by the merge variables;
data student;
  set bsgall1
    (keep=idcntry idschool idgrader jkindic jkzone
      totwgt bimatscr );

proc sort data=student;
  by idcntry idschool;

* Merge the student and school files together;
data merged;
  merge student school;
  by idcntry idschool;
  if nmiss(bcbgcomm)=0;

proc format library=work;
  value grade 1 = "Seventh Grade"      " 2 = "Eighth Grade"      ;

  value comm 1 = "Community Type 1"    " 2 = "Community Type2"
  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 JACK to get the results;
%include "jack.sas";

%jack (totwgt,jkzone,jkindic,75,idcntry idgrader bcbgcomm,bimatscr,merged);

proc print data=final noobs;
  where idgrader = 2;
  by idcntry idgrader;
  var bcbgcomm N totwgt mnx mnx_se pct pct_se;
  format idcntry country1. idgrader grade. bcbgcomm comm.;

```

Figure 9.18**SPSS Control Statements for Performing Analyses with School-Level Variables (EXAMPLE3.SPS)**

```

get file = "bcgall1.sys"
  / keep = idcntry idschool bcbgcomm.

sort cases by idcntry idschool.

save outfile = school.

get file = "bsgall1.sys"
  / keep = idcntry idschool idstud idgrader jkindic jkzone
    totwgt bimatscr .

select if idgrader = 2.

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(bcbgcomm)).

execute.

* Define the format for the variables used.
value labels
idgrader 1 "Seventh Grade" 2 "Eighth Grade" /
bcbgcomm 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 "jack.sps".

jack cvar = idcntry idgrader bcbgcomm /
     dvar = bimatscr /
     njkr = 75 /
     jkz = jkzone /
     jki = jkindic /
     wgt = totwgt /

sort cases by idcntry idgrader.

temporary.
select if idgrader = 2.
report format=list automatic
  / var = bcbgcomm (label) totwgt mnx mnx_se pct pct_se
  / break = idcntry idgrader.

```

Figure 9.19**Extract of SAS Computer Output for Performing Analyses with School-Level Variables (EXAMPLE 3)**

COUNTRY ID=Australia *GRADE*=Eighth Grade

BCBGCOMM	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Community Type 2	866	33600.29	512.906	11.3257	17.1558	3.71297
Community Type 3	3184	98242.75	531.360	7.2504	50.1613	4.36371
Community Type 4	2177	64010.75	528.813	7.7368	32.6829	4.41223

COUNTRY ID=Austria *GRADE*=Eighth Grade

BCBGCOMM	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Community Type 1	40	303.99	605.685	2.1439	0.3818	0.27798
Community Type 2	908	38258.26	526.105	5.0042	48.0505	3.65878
Community Type 3	287	7440.07	537.116	20.9813	9.3444	2.77784
Community Type 4	1212	33618.58	551.566	5.2165	42.2233	4.01114

COUNTRY ID=Belgium (Fl) *GRADE*=Eighth Grade

BCBGCOMM	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Community Type 2	898	22954.95	562.497	7.2354	33.5920	4.81714
Community Type 3	729	16854.23	584.657	7.7854	24.6642	4.42650
Community Type 4	996	28525.47	553.321	13.9251	41.7438	6.06839

COUNTRY ID=Belgium (Fr) *GRADE*=Eighth Grade

BCBGCOMM	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Community Type 2	518	10906.33	515.153	18.8832	23.3082	5.02550
Community Type 3	357	10109.43	518.489	10.3821	21.6051	4.90721
Community Type 4	1252	25776.11	539.305	5.9846	55.0867	5.44131

.
.
.
.
.
.

COUNTRY ID=United States *GRADE*=Eighth Grade

BCBGCOMM	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Community Type 1	132	90271.12	466.310	21.1273	3.2568	1.65686
Community Type 2	961	587437.27	509.643	6.0996	21.1936	3.38785
Community Type 3	1784	898597.98	513.360	6.4977	32.4196	4.10878
Community Type 4	3116	1195464.02	498.183	7.2497	43.1300	4.40161

COUNTRY ID=Slovenia *GRADE*=Eighth Grade

BCBGCOMM	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Community Type 1	163	1630.53	537.019	18.1296	7.4994	2.34076
Community Type 2	520	5435.88	521.496	6.8803	25.0016	2.92850
Community Type 3	492	4860.04	559.011	6.1974	22.3531	4.22613
Community Type 4	1092	9815.66	541.874	4.3193	45.1459	3.81019

Figure 9.20**Extract of SPSS Computer Output for Performing Analyses with School-Level Variables (EXAMPLE 3)**

COUNTRY ID	GEN\TYPE OF COMMUNITY	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Australia						
Eighth Grade	Community Type 2	33600.29	512.91	11.33	17.16	3.71
	Community Type 3	98242.75	531.36	7.25	50.16	4.36
	Community Type 4	64010.75	528.81	7.74	32.68	4.41
Austria						
Eighth Grade	Community Type 1	303.99	605.69	2.14	.38	.28
	Community Type 2	38258.26	526.10	5.00	48.05	3.66
	Community Type 3	7440.07	537.12	20.98	9.34	2.78
	Community Type 4	33618.58	551.57	5.22	42.22	4.01
Belgium (Fl)						
Eighth Grade	Community Type 2	22954.95	562.50	7.24	33.59	4.82
	Community Type 3	16854.23	584.66	7.79	24.66	4.43
	Community Type 4	28525.47	553.32	13.93	41.74	6.07
Belgium (Fr)						
Eighth Grade	Community Type 2	10906.33	515.15	18.88	23.31	5.03
	Community Type 3	10109.43	518.49	10.38	21.61	4.91
	Community Type 4	25776.11	539.30	5.98	55.09	5.44
.						
.						
.						
.						
.						
United States						
Eighth Grade	Community Type 1	90271.12	466.31	21.13	3.26	1.66
	Community Type 2	587437.27	509.64	6.10	21.19	3.39
	Community Type 3	898598.98	513.36	6.50	32.42	4.11
	Community Type 4	1195464.02	498.18	7.25	43.13	4.40
Slovenia						
Eighth Grade	Community Type 1	1630.53	537.02	18.13	7.50	2.34
	Community Type 2	5435.88	521.50	6.88	25.00	2.93
	Community Type 3	4860.04	559.01	6.20	22.35	4.23
	Community Type 4	9815.66	541.87	4.32	45.15	3.81

In summary, to perform analyses such as those in Figures 9.17 and 9.18 using 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 achievement score, 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 IDCNTY and IDSCHOOL.
- Use the macro JACK with the corresponding arguments and parameters.
- Print out the result file.

9.7 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 9.3 below.

Table 9.3
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 or perform and report on a specific task, 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 specific answer given by the student. These types of response codes were used for the free-response items administered as part of the written assessment and for the items in the performance assessment tasks. The codes used to represent the responses to these items are shown in Table 9.4 below.

Table 9.4
Definition of Response Codes for the Open-Ended Items in the Written Assessment and Performance Assessment Data Files

Code	Description
30 to 39	Three-point answer. Second digit provides diagnostic information.
20 to 29	Two-point 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 scored.
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 Performance Assessment and 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 and performance assessment tasks. To this effect, we have included in the CD a set of programs in SPSS and SAS that will allow the user to recode the items from the written assessment and from the performance assessment to their right-wrong or correctness-level format. Each of these programs contains a macro called SCOREIT and the necessary call to this macro so that all the items in the corresponding file are scored. These programs will 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. As defined previously in Table 7.14, four files are included to provide control code to perform the recodes of the test items in the written and performance assessment in Populations 1 and 2:

- Written Assessment Files (ASASCORE, BSASCORE)
- Performed Assessment Files (ASPCORE, BSPSCORE)

When using this code, 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 the not-reached responses were then recoded as incorrect when scoring the item for the individual countries, and 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 onto themselves 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.

Figure 9.21
Extracted Sections of SAS Control Code Used to Convert Cognitive Item Response Codes to Correctness-Score Levels

```

%MACRO SCOREIT (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 SCOREIT;

%LET ARIGHT =
BSMMA01 BSMSA10 BSMSB05 BSMSB06 BSMMB08 BSMMB09 BSMMB12 BSMMC01
BSMMC03 BSMMC06 BSMSC11 BSMSD06 BSMMD09 BSMMD10 BSMSSE10 BSMSF02
BSMSF03 BSMSF05 BSMMF11 BSMSG07 BSMSH01 BSMSH03 BSMSH04 BSMSH06
BSMMI07 BSMMI09 BSMSI13 BSMSI15 BSMSI16 BSMSI19 BSMSJ02 BSMMJ15
BSMMJ16 BSMMK03 BSMSK11 BSMSK16 BSMSK18 BSMLL15 BSMMM02 BSMMM03
BSMMM04 BSMSM13 BSMSN06 BSMMN16 BSMMN17 BSMSO12 BSMMQ05 BSMSQ11
BSMSQ15 BSMSR01 BSMMR06 BSMMR11;

%LET SHORTA =
BSSMI04 BSSMI06 BSSSI18 BSSSJ03 BSSSJ09 BSSMJ12 BSSMJ13 BSSMK02
BSSMK05 BSSSK10 BSSSK19 BSSML16 BSSMM06 BSSMM08 BSSSM12 BSSSM14
BSSSN07 BSSSN10 BSSMN13 BSSMN19 BSSMO06 BSSMO09 BSSSO10 BSSSO16
BSSSP07 BSSSP02 BSSSP03 BSSSP05 BSSSP06 BSSMP16 BSSMQ10 BSSSQ12
BSSSQ17 BSSSQ18 BSSSR04 BSSSR05 BSSMR13 BSSMR14 BSSMV01 BSSMV04;

ARRAY ARIGHT &ARIGHT;
ARRAY EXTEND &SHORTA;

DO OVER ARIGHT; %SCOREIT (ARIGHT,"MC",1, 6, 8, 9, 7); END;
DO OVER EXTEND; %SCOREIT (SHORTA,"SA", ,96,98,99,90); END;

```

Figure 9.22**Extracted Sections of SPSS Control Code Used to Convert Cognitive Item Response Codes to Correctness-Score Levels**

```

SET MPRINT=ON.
DEFINE SCOREIT (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.

SCOREIT Type = MC /
Item = BSMMMA01 BSMSA10 BSMSB05 BSMSB06 BSMMB08 BSMMB09
BSMMB12 BSMMC01 BSMMC03 BSMMC06 BSMSC11 BSMSD06
BSMMD09 BSMD10 BSMSSE10 BSMSF02 BSMSF03 BSMSF05
BSMMF11 BSMSG07 BSMSH01 BSMSH03 BSMSH04 BSMSH06
BSMMI07 BSMMI09 BSMSI13 BSMSI15 BSMSI16 BSMSI19
BSMSJ02 BSMMJ15 BSMMJ16 BSMMK03 BSMSK11 BSMSK16
BSMSK18 BSML15 BSMMM02 BSMMM03 BSMMM04 BSMSM13
BSMSN06 BSMMN16 BSMMN17 BSMSO12 BSMMQ05 BSMSQ11
BSMSQ15 BSMSR01 BSMMR06 BSMMR11 /
RIGHT = 1 /
nr = 6 /
na = 8 /
om = 9 /
other = 7 .

SCOREIT Type = SA /
Item = BSSMI04 BSSMI06 BSSSI18 BSSSJ03 BSSSJ09 BSSMJ12
BSSMJ13 BSSMK02 BSSMK05 BSSSK10 BSSSK19 BSSML16
BSSMM06 BSSMM08 BSSSM12 BSSSM14 BSSSN07 BSSSN10
BSSMN13 BSSMN19 BSSMO06 BSSMO09 BSSSO10 BSSSO16
BSSSO17 BSSSP02 BSSSP03 BSSSP05 BSSSP06 BSSMP16
BSSMQ10 BSSSQ12 BSSSQ17 BSSSQ18 BSSSR04 BSSSR05
BSSMR13 BSSMR14 BSSMV01 BSSMV04 /
RIGHT = 0 /
nr = 96 /
na = 98 /
om = 99 /
other = 90 .

```

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. and Gonzalez, E.J. (1997). Test-curriculum matching analysis. In M.O. Martin and D.L. Kelly (Eds.) *TIMSS technical report, volume II: Implementation and analysis*. Chestnut Hill, MA: Boston College.
- 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.
- Foy, P. (1997a). Implementation of the TIMSS sample design. In M.O. Martin and D.L. Kelly (Eds.) *TIMSS technical report, volume II: Implementation and analysis*. Chestnut Hill, MA: Boston College.
- Foy, P. (1997b). Calculation of sampling weights. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume II: Implementation and analysis*. 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. and Foy, P. (1997). Estimation of sampling variability, design effects, and effective sample sizes. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume II: Implementation and analysis*. Chestnut Hill, MA: Boston College.
- Harmon, M. and Kelly, D.L. (1996). Performance assessment. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume I: Design and development*. Chestnut Hill, MA: Boston College.
- Harmon, M., Smith, T.A., Martin, M.O., Kelly, D.L., Beaton, A.E., Mullis, I.V.S., Gonzalez, E.J., and Orpwood, G. (1997). *Performance assessment in IEA's Third International Mathematics and Science Study*. 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*. 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. and Martin, M.O. (1997). Item analysis and review. In M.O. Martin and D.L. Kelly (Eds.), *TIMSS technical report, volume II: Implementation and analysis*. 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. (1997). *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. (1998). *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 (1996a). *TIMSS mathematics items – Released set for population 2 (seventh and eighth grades)*. Chestnut Hill, MA: Boston College.
- TIMSS (1996b). *TIMSS science items – Released set for population 2 (seventh and eighth grades)*. Chestnut Hill, MA: Boston College.
- TIMSS (1997a). *TIMSS mathematics items – Released set for population 1 (third and fourth grades)*. Chestnut Hill, MA: Boston College.
- TIMSS (1997b). *TIMSS science items – Released set for population 1 (third and fourth grades)*. 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. The TIMSS performance assessment was an integral part of the study and one that required a great deal of additional resources and effort for all involved in that component. Below, the individuals and organizations are acknowledged for their contributions to TIMSS. 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.

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
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

International Study Center (continued)

Mary C. Howard, Office Supervisor
Diane Joyce, Secretary
Joanne E. McCourt, Secretary
Kathleen A. Haley, Graduate Assistant
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

Jens Brockmann, Senior Researcher
Michael Bruneforth, Senior Researcher (former)
Jedidiah Harris, Research Assistant
Dirk Hastedt, Senior Researcher
Heiko Jungclaus, Senior Researcher
Svenja Moeller, Research Assistant
Knut Schwippert, 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

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 to initiate the TIMSS project. Tjeerd Plomp, Chair of the IEA and of the TIMSS International 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. This list only includes information for those countries for which data are included in the International Database.

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
Université 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 CY-1789, 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

Max-Planck Institute for Human
Development and Education
Lentzeallee 94
D-14195 Berlin, Germany

Manfred Lehrke

Universität Kiel
IPN
Olshausen Str. 62
24098 Kiel, Germany

Greece

Georgia Kontogiannopoulou-Polydorides
Department of Education
Tmima Nipiagogon
University of Athens
Navarinou 13, Neochimio
Athens 106 80, 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.F. 701/420
1051 Budapest, Hungary

Iceland

Einar Gudmundsson
Institute for Educational Research
Department of Educational Testing
and Measurement
Surdgata 39
101 Reykjavik, Iceland

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

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
Research Division
15-1 Chungdam-2 dong, 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 Jurmalas
gatve 74/76, Rm. 204A
Riga, LV-1083, Latvia

Lithuania

Algirdas Zabulionis
National Examination Centre
Ministry of Education & Science
M. Katkaus 44
2006 Vilnius, Lithuania

Netherlands

Wilmad Kuiper
Anja Knuver
Klaas Bos
University of Twente
Faculty of Educational Science
and Technology
Department of Curriculum
P.O. Box 217
7500 AE Enschede, Netherlands

New Zealand

Hans Wagemaker
Megan Chamberlain
Steve May
Robyn Caygill
Ministry of Education
Research Section
Private 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
Center for IEA Studies
Educational Research Institute
Gerbiceva 62, P.O. Box 2976
61111 Ljubljana, Slovenia

South Africa

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
28002 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 International Study Center was supported in its work by several advisory committees. The 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

Performance Assessment Committee

Derek Foxman, England
Robert Garden, New Zealand
Per Morten Kind, Norway
Svein Lie, Norway
Jan Lokan, Australia
Graham Orpwood, Canada

Quality Assurance 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