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

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- CSTEEP
School of Education
Boston College
Chestnut Hill, MA 02167
United States

email: timss@bc.edu
http://wwwcsteep.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 implemenation of TIMSS.

Printed and bound in the United States.

Contents

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 are95.(zuance)Tj T* essessmee8. Elists,95.(s 1 and 2, thgave95.(zuEA for each of)Tj T* 0.007 Tessmissie, po

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

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

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.

Figure 2.1
The Major Categories of the TIMSS Curriculum Frameworks

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	-			
	M	7	2	-			
Breadth (Science) (9 minutes)	N				9	-	-
	O				7	2	-
	P				8	1	-
	Q				7	2	-
	R				8	1	-
Mathematics Free-Response (9 minutes)	S	-	3	2			
	T	-	3	2			
	U	-	3	2			
	V	-	3	2			
Science Free-Response (9 minutes)	W				-	3	2
	X				1	2	2
	Y				-	-	3
	Z				-	-	3

CHAPTER 2

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	Solutions (Population 2)
	S6	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

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).

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.

CHAPTER 3

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

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:

$$R \quad stu \quad \frac{p}{p\mathring{a}partpart}$$

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] T[illegible] P[illegible]

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
†2 England	Singapore	†2 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**Fourth Grade**	

Ths86337(C8 34647(Ths[05-satisfyinKonu484ry

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)	[1] Latvia (LSS)	† Belgium (Fr)	[1] Latvia (LSS)
Canada	[1] Lithuania	† Belgium (Fl)	[1] Lithuania
Cyprus	New Zealand	Canada	New Zealand
Czech Republic	Norway	†[2] Cyprus	Norway
†[2] England	Portugal	Czech Republic	Portugal
France	Russian Federation	England	Russian Federation
Hong Kong	Singapore	France	† Scotland
Hungary	Slovak Republic	Hong Kong	Singapore
Iceland	Spain	Hungary	Slovak Republic
Iran, Islamic Rep.	Sweden	Iceland	Spain
Ireland	[1] Switzerland	Iran, Islamic Rep.	Sweden
Japan	† United States	Ireland	[1] Switzerland
Korea		Japan	† United States
		Korea	
Countries not satisfying guidelines for sample participation			
Australia		Australia	
Austria		Austria	
Belgium		Bulgaria	
Bulgaria		Netherlands	
Netherlands			
Scotland			
Countries not meeting age/grade specifications (high percentage of older students)			
Colombia		Colombia	
†[1] Germany		†[1] Germany	
Romania		Romania	
Slovenia		Slovenia	
Countries with unapproved sampling procedures at the classroom level			
Denmark		Denmark	
Greece		Greece	
Thailand		[1] South Africa	
		Thailand	
Countries with unapproved sampling procedures at classroom level and not meeting other guidelines			
[1] Israel			
Kuwait			
South Africa			
Countries with unapproved sampling procedures at school level			
[3] Phillipines		[3] Phillipinesof older studenser d7i0 0 3tevel	
		Netherlands	

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	Canada
Cyprus	Cyprus
Czech Republic	Iran, Islamic Republic
Iran, Islamic Republic	†3 New Zealand
New Zealand	Portugal
Norway	
Portugal	
† Scotland	
Singapore	
Spain	
Sweden	
†1 Switzerland	
Countries not satisfying guidelines for sample participation	
Australia	Australia
2 England	Hong Kong
Netherlands	United States
United States	
Countries not meeting age/grade specifications (high percentage of older students)	
Colombia	Slovenia
3 Romania	
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

Table 3.3
Sample Information for TIMSS Population 1 Countries

COUNTRY	Third Grade			Fourth Grade		
	Sample Size	Sum of Weights	Maximum Jacknifing Zones	Sample Size	Sum of Weights	Maximum Jacknifing 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 Jacknifing Zones	Sample Size	Sum of Weights	Maximum Jacknifing 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 Files[3]

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.

WGTADJ3 Student Weighting Adjustment

This is an adjustment applied to the variable WGTFAC3 to account for non-participating students in the selected school and/or classroom. If we were to multiply the variables WGTFAC2, WGTFAC3, and WGTADJ3 and 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 WGTFAC1,WGTADJ1, WGTFAC2, WGTFAC3, 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

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

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 desireable 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] I [illegible] K [illegible] SEN GT [illegible] 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

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.

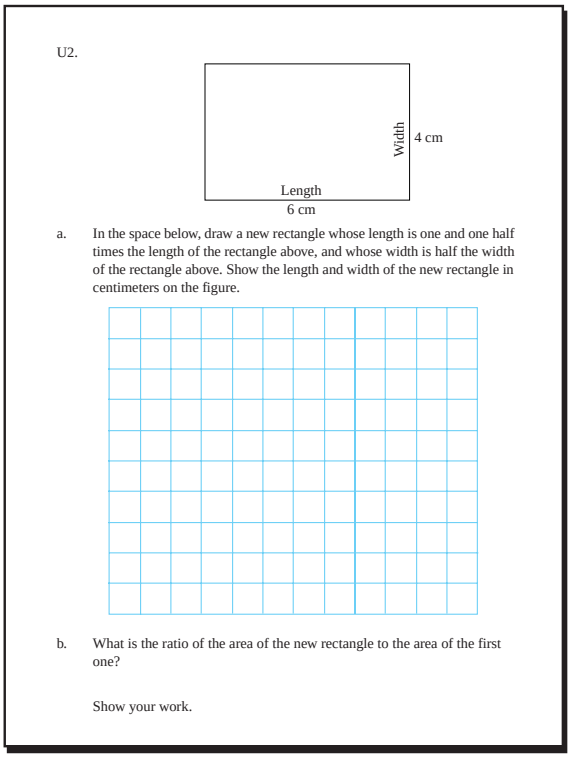
Figure 4. 1
Example Coding Guide for Short-Answer Mathematics Item

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



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

A: Codes for Drawin

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

Table 4.2
(Continued)

Country	Item	Subject	Variable Name
Population 1 - Written Assessment			
Cyprus	B04	Science	ASMSB04
England	TA G Science 6		ASMSB04 e emF15c 43910881 m 192.5 660MA01 [(Sc-8.81)-ula49Engl8(

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] T NRC

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 $i=1,...,I$ with each item admitting $K_i + 1$ response alternatives $k = 0,1,...,K_i$. Use the vector valued random variable, $\mathbf{X}_i = \left(X_{i1}, X_{i2}, \ldots, X_{iK_i}\right)^{\prime}$,

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

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 $\mathbf{X}' = (\mathbf{X}_1', \mathbf{X}_2', ..., \mathbf{X}_I')$, 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 $\mathbf{x}^T = \left(x_1, x_2, \ldots, x_p\right)$ 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 $\mathbf{a}_{jk}$, $\left(j = 1, \ldots, I;\ k = 1, \ldots K_i\right)$, each of length p, which can be collected to form a design matrix $\mathbf{A}' = \left(\mathbf{a}_{11}, \mathbf{a}_{12}, \ldots, \mathbf{a}_{1K_1}, \mathbf{a}_{21}, \ldots, \mathbf{a}_{2K_2}, \ldots, \mathbf{a}_{iK_i}\right)$. 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 $\mathbf{b}^T = \left(b_{11}, b_{12}, \ldots, b_{1K_1}, b_{21}, b_{22}, \ldots, b_{2K_2}, \ldots, b_{iK_i}\right)$. (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 q be the latent variable the item response probability model is written as:

$$\Pr\left(\mathbf{X}_{ij} = 1; \mathbf{A}, \mathbf{b}, \mathbf{x} \mid q\right) = \frac{\exp\left(b_{ij}q + \mathbf{a}_{ij}^{T}\mathbf{x}\right)}{\sum_{k=1}^{K_i} \exp\left(b_{ik}q + \mathbf{a}_{ij}^{T}\mathbf{x}\right)}, \quad (2)$$

and a response vector probability model as

$$f\left(\mathbf{x}; \mathbf{x} \mid q\right) = \gamma\left(q, \mathbf{x}\right)\exp\left[\mathbf{x}^{T}\left(\mathbf{b}q + \mathbf{A}\mathbf{x}\right)\right], \quad (3)$$

with $$\gamma\left(q, \mathbf{x}\right) = \left[\sum_{\mathbf{z} \in W} \exp\left[\mathbf{z}^{T}\left(\mathbf{b}q + \mathbf{A}\mathbf{x}\right)\right]\right]^{-1}, \quad (4)$$

where W is the set of all possible response vectors.

5.3 The Multidimensional R5 Ts Tc . 14 Tw 5. 4oefcf8 2455 4.8 im 12 1 Tf 1. 62 1

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 $\boldsymbol{\alpha}$ 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.

5 Tm 1.5,

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,

$$f_x(\mathbf{x};\xi,\gamma,\Sigma) = \int_\theta f_x(\mathbf{x};\xi \mid \theta)\, f_\theta(\theta;\gamma,\Sigma)\, d\theta \,, \tag{13}$$

and it follows that the likelihood is,

$$L = \prod_{n=1}^{N} f_x(\mathbf{x}_n;\xi,\gamma,\Sigma), \tag{14}$$

where N is the total number of sampled students.

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

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

provides the following system of likelihood equations:

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

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

and $\hat{\Sigma} = \frac{1}{N}\sum_{n=1}^{N}(\bar{\theta}_n - \hat{\gamma}\mathbf{W}_n)(\bar{\theta}_n - \hat{\gamma}\mathbf{W}_n)^T + \frac{1}{N}\sum_{n=1}^{N}\int_{\theta_n}(\theta_n - \bar{\theta}_n)(\theta_n - \bar{\theta}_n)^T h_\theta(\theta_n;\mathbf{W}_n,\hat{\gamma},\Sigma \mid \mathbf{x}_n)\, d\theta_n$,

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

$$\theta_n^{EAP} = \sum_{r=1}^{P} Q_r \, h_Q\left(Q_r; \mathbf{W}_n, \hat{\xi}, \hat{\gamma}, \hat{\Sigma} \mid \mathbf{x}_n\right). \qquad (21)$$

Variance estimates for these predictions were estimated using

$$\mathrm{var}\left(\theta_n^{EAP}\right) = \sum_{r}^{P} \left(Q_r - \theta_n^{EAP}\right)^2 h_Q\left(Q_r; \mathbf{W}_n, \hat{\xi}, \hat{\gamma}, \hat{\Sigma} \mid \mathbf{x}_n\right). \qquad (22)$$

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

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

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 qual weigh2Sght in t secSE in ttTj 0 -1.2727 TD 0.00 u1 9al cali0 7 drawing 1 Tc

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 Files[1]

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
BSMSCPT	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.

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 say e- take into a omparisons across cou4 Tw (po4D -0.)[ountrieem 0.ssmn

Table 6.1

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

Table 6.2

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

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	222.14 75 -1.5 TD -0.024 Tc (Bu60Ti2220 TD6.75	
	Belgium8507.14		141.82	588.m 56		98..26 68Ti2220 TD6.7413.5( )-5( )-222.78		819 3
			Belgium8					
			Cana (Irel -0JTc [(507.14	)-2R	)-89	180.91	8232 Tc3	

Table 6.4
Descriptive Statistics for the International Science Achievement Scores for

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
C a F r a Da a a F

7.1 I

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] o mo e detailed info mation about data ent y, data p oce ing, data cleaning, data management, databa e t uctu e, a well a analy i and epo ting of the TIMSS data, ee the *TIMSS Technical Repo t, Volume I and II (Ma tin and Kelly, 1996;1997)* and *TIMSS: Quality A u ance in Data Collection (Ma tin and Mulli, 1996).*

Table 7.2

7.2.1 Background Files

There are three different types of TIMSS background files – student, teacher, and school.

7.2.1.1 Student Background Files

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 Files

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 International 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 Identifying Teacher, and School Background File

IDCNTRY	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 IDCNTRY and IDSCHOOL combination.

Additional Variables Identifying Student Background File

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 IDCNTRY.

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 in the Teacher Background File

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)

IDSUBJCT The subject(s) taught by the teacher (mathematics, science or both).

Additional Variables in the School Background File

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 IDSUBJCT 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

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 in the Teacher Background File

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 School Background File

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.

International Background Variable Questionnaire Location

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 Files and Corresponding Variables

Population	Questionnaire	File Name 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 Convention

Case Position	Description	Value		
		Student Value	Teacher Value	School Value
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 teaches' 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.

Using Derived Variables in Data Analysis

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 Table/Figure Location Reference Derived Variables

Chapter Number	Derived Variable(s)

International Achievement Reporting Derived Variables

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/BCDGTENR	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
WGTADJ1	School weighting adjustment
WGTFAC2	Class weighting factor
WGTFAC3	Student weighting factor
WGTADJ3	Student weighting adjustment
JKZONE	Jackknife zone
JKINDIC	Replicate code
WGTFAC1	School weighting factor
WGTADJ1	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 Written and Performance Assessment Items

Character Position	Definition	Value (Written Assessment)	Value (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 Identification

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 Achievement Data Scale 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.

Ta 7.7
Ca a P a 11 Ma a C A a R Ca

I C A a R Ca					
W N (25)	F a a P a (21)	M a , E a , a N S (20)	Da a R a , A a , a P a (12)	G (14)	Pa , R a , a F. (10)
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	

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

Item Content Area Reporting Category					
Earth Science (17 items)	Life Science (41 items)		Physical Science (30 items)		Environmental Issues and the Nature of Science (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

Mathematics Content Area Reporting Category						
Fractions and Number Sense (51 items)		Algebra (27 items)		Measurement (18 items)	Geometry (23 items)	Data Representation, Analysis, and Probability* (20 items)	Proportionality (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.

7.2.2.7 Released TIMSS Item Data

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 Student 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 variables[7]
- 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] To 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.

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 Codes

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
Recoding Map for Country-Specific Diagnostic Codes in Part A

	Item	Variable	Recoding	Comment
Grade 4	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.
Part 1-Written	K10	BSMMK08	71 → 70	Training team found it difficult to distinguish between the 70 and 71 codes; both codes combined in 70.
	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.

Other Variables in 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

Content

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

Other Variables in the Student-Teacher Linkage File

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.

Trackiseode [illegible]

7.2.5 Missing Cases in 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 or Invalid 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:
value * (iallvalid conn0v.speon3nt8value gi.9s also contain ppglue p3 TD eu(unreadabacpec0togpon Trackdeon3nt8va

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

NA (a R C (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.

NR a I r C (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 Na (a Da a I A U I a (a Da a F (

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 C F

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.

Ta 7.12
P a 1 a P a 2 C F

Da a F	P a 2 C Na	P a 1 C Na
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 A C F

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.

P F a (*.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

S. W[illegible] a P[illegible]a A[illegible] F[illegible]C

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

S P[illegible]a A[illegible] F[illegible]C

- Identification Variables
- Tracking/Linking Variables

S. -T a L[illegible]a/ F[illegible]C

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

R [illegible]F[illegible]C

- 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:

Example Printout o

Codebook, Date: 24.09.97 File: asgcode.SDB

Var. Question No.	Variable Name	Label	Code	Option	Location/Format
1 VERSION	VERSION	*FILE VERSION*	VALUE	Version number	1-2 / C 2.0
2 COUNTRY	IDCNTRY	*COUNTRY ID*	VALUE	Country ID	3-5 / N 3.0
3 POP	IDPOP	*POPULATION ID*	1	population 1	6 / C 1.0
			2	population 2	
			3	population 3	
4 STRAT	IDSTRAT	*STRATUM ID*	VALUE	Stratum ID	7-9 / N 3.0
			999	missing	
			998	not admin.	
5 SCHOOL	IDSCHOOL	*SCHOOL ID*	VALUE	School ID	10-15 / N 6.0
6 CLASS	IDCLASS	*CLASS ID*	VALUE	Class ID	16-23 / N 8.0
7 STUD	IDSTUD	*STUDENT ID*	VALUE	Student ID	24-33 / N 10.0
8 GRADE	IDGRADE	*GRADE*	VALUE	Grade	34-35 / N 2.0
9 EXCLUD	IDEXCLUD	*INDICATOR FOR EXCLUDED STUDENTS*			36 / C 1.0
			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	

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 File	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.*	AS
Control Fiw 91-c8e		

7.5 Data Arrays

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

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

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 eb0ud4t code to conduct such analyses, eb0uthe results of these analyses. Although the analyses presented here are simple in nature, they are designed to familiarize the user with the different files eb0utheir 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, eb0uthe corresponding steb0ard errors for the percent eb0 mean stetistics. The analyses presented in this chapter, based on student eb0uteacher 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 eb0uFigure 9.2, are replicated in Examples 1 eb0u2 in this chapter. The user is welcomed to compare the results from these analysis to the tables in the reports, eb0uis encouraged to practice analyzing the TIMSS data by trying to replicate some of the tables that are presented in the international reports.[1]

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

116

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

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

Figure 9.5 continued

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:

where 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.***C V D R***

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 ExtRiYct of Output File for Using the Macro JACK.AS

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)

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:

where 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:

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

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

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.

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

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

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

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

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

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)

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

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

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 IDCNTRY 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 (ASPSCORE, 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 in 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  BSMSE10  BSMSF02
 BSMSF03  BSMSF05  BSMMF11  BSMSG07  BSMSH01  BSMSH03  BSMSH04  BSMSH06
 BSMMI07  BSMMI09  BSMSI13  BSMSI15  BSMSI16  BSMSI19  BSMSJ02  BSMMJ15
 BSMMJ16  BSMMK03  BSMSK11  BSMSK16  BSMSK18  BSMML15  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
 BSSSO17  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

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.

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

International Study Center (continued)

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'Education Nationale
142, rue du Bac
75007 Paris, France

Josette Le Coq*
Centre International d'Etudes 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

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

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

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

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