13

Chapter 13

Reporting TIMSS 2003 Questionnaire Data

María José Ramírez and Alka Arora

13.1 Overview

The purpose of TIMSS is to provide information that policymakers, curriculum specialists, and researchers can use to understand better the performance of their educational systems. With this aim, TIMSS collects data on hundreds of contextual variables from nationally representative samples of students, their science and mathematics teachers, and their schools. Once the data are collected, one of the major challenges for TIMSS is reporting this vast array of information in a useful and meaningful way. The challenge is to focus on the most important educational contexts, inputs, and processes without overburdening the audiences with unmanageable amounts of information. TIMSS strives to report educational indicators that are easy to understand and interpret by policymakers and school personnel.

This chapter documents the analysis and reporting procedures used for the background questionnaire data in producing the TIMSS 2003 International Reports in mathematics and science. It provides an overview of the consensus process used to develop the report outlines and prototype exhibits; explains how single- and multiple-item indicators from the student, teacher, and school data were developed and computed; describes methods used by TIMSS to compute these indicators; and details the analysis and reporting of curriculum data. The final section explains how the data are displayed in the exhibits, and addresses issues regarding the unit of analysis, trend data, and response rates.

13.2 General Procedures

As described in Chapter 3, TIMSS 2003 used four types of questionnaires at both the fourth and eighth grades to gather information at various levels of the educational system:

- Student Questionnaire (separate versions for general/integrated science countries and separate science countries at eighth grade)
- Teacher Questionnaire (separate versions for mathematics and science at eighth grade)
- School Questionnaire
- Curriculum Questionnaire (separate versions for mathematics and science at both eighth and fourth grades)

The TIMSS & PIRLS International Study Center (ISC) at Boston College produced data almanacs summarizing the basic data from the student, teacher, and school questionnaires. For each participating country, these almanacs presented descriptive statistics for each question (variable) in the survey instruments. The statistics included the percentages of students checking each response option for categorical and ordinal data, as well as means, standards deviations, and percentile scores for continuous data. The almanacs were distributed periodically to the National Research Coordinators (NRCs) for review. Each time, a new data version was provided with more cases and updated cleaning rules and corrections implemented.

The ISC began working on the analysis of background data in May 2003. The main steps involved in this process were as follows. First, the TIMSS 2003 questionnaires were reviewed in the light of the contextual framework (see Chapter 3) to identify major conceptual categories or constructs that would enable a better understanding of the participating countries' educational systems and a fuller interpretation of their students' achievement in mathematics and science. Second, an outline describing the chapters and exhibits to be included in the TIMSS 2003 International Reports was prepared. Third, questions that could be used to measure the constructs of interest were identified, and extensive exploratory data analysis was conducted to decide what information to show and how to display it in each of the exhibits of the International Reports.

At the time the ISC started working on the reporting of data from the background questionnaires, data from the countries that operated with the southern hemisphere schedule were available for preliminary analyses.[1] These countries – Australia, Botswana, Chile, Malaysia, New Zealand, Singapore, and South Africa – provided data from some 40,000 students covering the entire spectrum of achievement on the TIMSS 2003 assess-

1 Countries that used the southern hemisphere schedule collected their data during September-November 2002, approximately six months earlier than countries using the northern hemisphere schedule.

ment and representing great cultural diversity. The preliminary analyses used background data from the Student Questionnaire (general version), Mathematics Teacher Questionnaire, and School Questionnaire from the TIMSS 2003 eighth-grade population.

As a first step, staff at the ISC reviewed the data thoroughly to ensure its quality. Descriptive analyses were run for each country separately, as well as for all the countries together. Statistics showing total number of cases, response rates, mean scores, standard deviations, and minimum and maximum scores were computed. For open-ended questions, ranges of valid responses were defined. When there were questions about the data, the national versions of the questionnaires were reviewed, and in some cases the NRC was contacted for further clarifications. As a result of this data review, the IEA Data Processing Center (DPC) in Hamburg implemented a number of revisions to the data cleaning rules.

Several preliminary versions of the indicators were developed and reviewed at the ISC. As explained in the following section, TIMSS 2003 used three methods for reporting background data: the direct reporting method (for single-item indicators), the scale method, and the combination of responses method (for multiple-item indicators). At this exploratory stage, all the analyses were run on unweighted data, using the first plausible value for mathematics as a criterion.[2] All the programming at this stage was done using SPSS version 11.5 (SPSS Inc., Chicago IL).

Once there was a clearer idea about how to combine the data into multiple-item indicators, the analyses were extended and adapted to the TIMSS 2003 fourth-grade population as well as to the science-specific instruments – Student Questionnaire (integrated science), Student Questionnaire (separate science subjects) and Science Teacher Questionnaire. All the indicators were reviewed for their effectiveness in providing information about educational contexts in the participating countries. Starting in October 2003, data from the northern hemisphere countries became available and was included in the analyses. The suitability of the preliminary indicators was checked again for these additional countries, and changes in the measures were made as necessary.

For each exhibit (table or figure) in the International Reports, analysis notes were created to document how the data were to be analyzed. These notes identified the source questions used to gather the data, explained how the data were processed before reporting, and described how the data would be displayed in the exhibits. The analysis notes also served as directions for programming the analyses in SAS version 9.0 (SAS Institute Inc., Cary NC),

2 See Chapters 11 and 12 for more information on plausible values.

the software used by TIMSS in implementing the data analysis. The exhibits in the International Reports were produced in SAS using all five plausible values in the TIMSS 2003 dataset, and standard errors were computed using the jackknife procedure (see Chapter 12). Based on the analysis notes also, the graphic production staff at the ISC designed and prepared prototype exhibits to display the background information.

Representatives from the participating countries reviewed the outlines for the International Reports, the proposed exhibits and indicators, and the analysis notes at the seventh NRC meeting held in Cape Town, South Africa, in November 2003. At that time, although data were available for just a few countries, they were useful in providing a sense of how the complex exhibits would look. NRCs approved the report outlines and almost all the proposed indicators; revisions were required in some exhibits based on suggestions for improvements from NRCs.

In January 2004, the ISC posted to its website revised Chapter 4 (Mathematics/Science Student Background) exhibits for the NRCs to review. Weighted data from 45 countries at the eighth-grade and 22 countries at the fourth-grade were available at that time. In March 2004, a revised version of the exhibits in Chapter 5 (Mathematics/Science Curriculum), Chapter 6 (Teachers of Mathematics/Science), Chapter 7 (Instruction in Mathematics/Science), and Chapter 8 (Mathematics/Science School Context) were posted to the ISC website, together with updated analyses notes. NRCs reviewed their national data and informed the ISC about any problems or anomalies that required further attention. In the meantime, staff at the ISC continued checking the data. All analyses were conducted in SAS, and repeated independently in SPSS to ensure that the same results were obtained.

The penultimate version of the TIMSS background exhibits was presented at the eighth NRC meeting held in Santiago, Chile, in June 2004. Country representatives reviewed their data and approved the exhibits for the International Reports. In a few cases, changes in the exhibits' format and type of information displayed were requested. NRCs informed the ISC about any questionable results that required further examination. After the meeting, staff at the ISC made final revisions to the exhibits.

Once the final exhibits of the background chapters were available, the companion text for those chapters was written. The background chapters with final exhibits and draft text were posted to the ISC website from August 16-30, 2004. NRCs reviewed the text and shared their comments with the ISC.

13.3 Methods for Reporting Background Data

This section describes the specific methods used to report TIMSS 2003 questionnaire data: the direct reporting method (for single-item indicators); scale method and combination of responses method (for multiple-item indicators).

13.3.1 Direct Reporting Method

Direct reporting was the simplest method used by TIMSS to report background data. The direct reporting method simply used the response categories in the questionnaires as reporting categories in the exhibits in the International Reports. In some cases, slight modifications were introduced: some response categories were collapsed, or were presented in a different order. Although the direct reporting method had the advantage of simplicity, it would have been impossible to report the vast amount of information collected by TIMSS in this way. Some data reduction was required, necessitating the use of more sophisticated approaches, as described below.

13.3.2 Methods for Computing Multiple-Item Indicators

Around one-fourth of the exhibits in the TIMSS 2003 International Reports were multiple-item indicators (derived variables) that combined data from several questions in the TIMSS 2003 questionnaires. Multiple-item indicators were used with complex constructs, such as the teacher's emphasis on mathematics homework, or school climate. Because the source items making up a multiple-item indicator target different facets of the construct, these measures can provide a more global and thorough picture of the phenomenon being studied than can single variables. Multiple-item indicators also have the advantage of providing more reliable measures of the construct, since random errors tend to cancel out when data are combined from different sources (see DeVellis, 1991; Spector, 1992).

Multiple-item indicators maximize the information that can be preserved in the presence of missing data. TIMSS required that at least two-thirds of the component questions have valid responses before computing an index. For instance, if an index was based on five questions, this rule allowed for one missing response only.

The starting point for creating a multiple-item indicator was to identify the questions in the TIMSS 2003 questionnaires that were related to the construct of interest. In some cases, these source questions were all sub-items of a more general question, and all had the same format. In other cases, the source questions came from different parts of the questionnaires, and did not share the same format. Depending upon the construct of interest and the item formats, TIMSS used two different methods to create derived variables: the scale method and the combination of responses method.

13.3.2.1 Scale Method

The "scale method" was used when the construct of interest had an underlying quantitative continuum. For example, schools can have a better or a worse climate for learning, or students can have higher or lower self-confidence in learning science. The scale method also required that all the questions (items) have the same number of response categories. These conditions allowed data to be combined from several items into one underlying scale while retaining the original metric of the items.

Before combining data from different questions, TIMSS gathered evidence that the source questions had the expected relationship with the achievement scores. For instance, it was expected that students who agreed with a statement such as "I usually do well in mathematics" would have higher mathematics scores than students who disagreed with the statement. Descriptive statistics, analysis of variance (ANOVA), and eta-squared (η^2) were useful in assessing whether the expected relationships held true (see Hinkle, Wiersma & Jurs, 1998, pp. 565-569; Pedhazur, 1997, pp. 355, 505-507).

Questions addressing a construct were expected to be correlated in the data. Chi-square (χ^2) and Spearman's rank order correlation coefficient were used to measure the association between pairs of categorical or ordinal items. Principal component analysis (PCA) was used to identify questions related to a common construct. Building on these analyses, new variables (components) were created that accounted for most of the variance in the source items.

Once there was enough evidence that a set of questions or items was measuring the construct of interest, TIMSS examined the reliability of a scale made up from these items. Cronbach's alpha (α) was used to measure the internal consistency of these scales; item-total correlations (or point-biserial correlations) were used to identify questions that did not cluster together with the others.

Using the scale method, TIMSS computed index scores by averaging the numerical values associated with each response option. This procedure had the advantage of preserving the original scale categories, thus allowing for a straightforward interpretation of the index scores. The TIMSS 2003 questionnaires made extensive use of the 4-point Likert scale format, with "*strongly agree*" coded 4, "*agree*" coded 3, "*disagree*" coded 2, and "*strongly disagree*" coded 1. Before averaging the scores associated with the responses,

responses were recoded as necessary, with items coded so that high scores were associated with the response category indicating higher levels of the attribute being measured.

Whenever the scale method was used to create an index, TIMSS classified the students into three levels: high, medium, and low. In the International Reports, these derived variables are referred to as indices. To classify the cases into three groups, two cutoff points were established. Three main criteria were used in setting the cutoff points. First, the high level of the index should correspond to conditions or activities generally associated with good educational practice or high academic achievement. Second, there should be a reasonably even distribution of students across the three index levels. Third, the scale categories should be about the same size.

Once the cutoff points were defined, a critical step was to check the overall quality of the indices. Indices were intended to discriminate among students with high and low achievement. The extent of the association with achievement was measured using eta-squared (η^2). This was computed for each country separately and for all the countries together. Only indices that discriminated reasonably well in most of the participating countries were included in the International Reports.

Line graphs plotting mean achievement by index level also were useful in checking the hypothesized positive association between index levels and achievement scores. The slope of the line joining the means served as an indicator of how well the index discriminated among students with different achievement levels. The steeper the line the greater were the differences between the average achievement scores of one index level and the next.

13.3.2.2 Combination of Responses Method

TIMSS also made extensive use of the "combination of responses method" to construct indices. Cases were classified into the high, medium, or low level of an index depending upon the combination of responses provided to the source items. For example, in the index of *Good School and Class Attendance*, cases were classified into the high index level if the three source items (arriving late at school, absenteeism, and skipping classes) were reported to be *not a problem*. Cases went to the low index level when two or more behaviors were reported to be a *serious problem* or two behaviors were reported to be a *minor problem* and the third a *serious problem*. The medium level included all other combinations of responses.

In addition to constructing indices, the combination of responses method also was used to construct some specific derived variables. An example is students' *Use of Computer*. Students were asked if they use a computer "at home," "at school," "at a library," "at a friend's house," "at an Internet cafe," or "elsewhere." The reporting categories for this derived variable were "use computer both home and at school," "use computer at home but not at school," "use computer at school but not at home," "use computer only at places other than home and school," and "do not use computer at all."

13.3.2.3 Summary of Derived Variables in the TIMSS 2003 International Reports

The TIMSS 2003 International Reports in mathematics and science each present some 60 exhibits with background information, providing data on some 250 indicators. The mathematics report presents data on 17 derived variables and the science report on 16; each report includes 11 indices. Exhibits 13.1 and 13.2 list the indices computed for the TIMSS 2003 International Reports in mathematics and science, respectively. Exhibit 13.3 lists the other derived variables presented in the mathematics and science reports. The name of the indicators, the label used to identify them in the International Reports and database, the mathematics or science exhibit where the data are reported, and the analysis method used to compute the data are provided.

13.4 Analysis of Curriculum Data

The Mathematics and Science Curriculum Questionnaires were used to collect information about the intended curriculum in each participating country. The NRC for each country, with the help of curriculum specialists, completed curriculum questionnaires for the grade assessed (fourth grade and/or eighth grade). Chapter 5 in the TIMSS 2003 International Reports combined data from the Curriculum Questionnaires and the Teacher Questionnaire to inform about both the intended and implemented Mathematics and Science curricula in the participating countries. The following information was presented:

- Existence of a national curriculum, the year it was introduced, and whether it was under revision
- Methods used to support and monitor curriculum implementation
- Use of public examinations and grades tested
- Instructional time intended for mathematics and science
- Differentiation of curriculum for students with different levels of ability
- Emphasis on different approaches and processes in the intended curriculum (e.g., knowing facts, understanding concepts)
- Coverage of the TIMSS 2003 topics in the intended and implemented curriculum
- Science subjects offered through the eighth grade (science only)

Exhibit 13.1 **Summary Indices in the TIMSS 2003 International Mathematics Report**

Index	Analysis Method
Exhibit 4.7 Index of Time Students Spend Doing Mathematics Homework (TMH)	Index based on students' reports on the frequency and amount of mathematics homework they are given. High level indicates more than 30 minutes of mathematics homework assigned 3-4 times a week. Low level indicates no more than 30 minutes of mathematics homework no more than twice a week. Medium level includes all other possible combinations of responses.
Exhibit 4.9 Index of Students' Self-Confidence in Learning Mathematics (SCM)	Index based on students' responses to four statements about mathematics: 1) I usually do well in mathematics; 2) Mathematics is more difficult for me than for many of my classmates (Reversed); 3) Mathematics is not one of my strengths (Reversed); 4) I learn things quickly in mathematics. Average is computed across the four items based on a 4-point scale: 1. Agree a lot; 2. Agree a little; 3. Disagree a little; 4. Disagree a lot. Students agreeing a little or a lot on average across the four statements are assigned to the high level. Students disagreeing a little or a lot on average are assigned to the low level. All other students are assigned to the middle level.
Exhibit 4.10 Index of Students' Valuing Mathematics (SVM) (Grade 8 only)	Index based on students' responses to seven statements about mathematics: 1) I would like to take more mathematics in school; 2) I enjoy learning mathematics; 3) I think learning mathematics will help me in my daily life; 4) I need mathematics to learn other school subjects; 5) I need to do well in mathematics to get into the university of my choice; 6) I would like a job that involved using mathematics; 7) I need to do well in mathematics to get the job I want. Average is computed across the seven items based on a 4-point scale: 1. Agree a lot; 2. Agree a little; 3. Disagree a little; 4. Disagree a lot. Students agreeing a little or a lot on average across the seven statements are assigned to the high level. Students disagreeing a little or a lot on average are assigned to the low level. All other students are assigned to the middle level.
Exhibit 7.2 Index of Teachers' Reports on Teaching Mathematics Classes with Few or No Limitations on Instruction due to Student Factors (MCFL) (Grade 8 only)	Index based on teachers' responses to six statements about student factors limiting mathematics instruction: 1) Students with different academic abilities; 2) Students who come from a wide range of backgrounds; 3) Students with special needs; 4) Uninterested students; 5) Low morale among students; 6) Disruptive students. Average is computed across the six statements based on a 4-point scale: 1. Not at all/Not applicable; 2. A little; 3. Some; 4. A lot. High level indicates average is less than or equal to 2. Medium level indicates average is greater than 2 and less than 3. Low level indicates average is greater than or equal to 3.
Exhibit 7.13 Index of Teachers' Emphasis on Mathematics Homework (EMH)	Index based on teachers' responses to two questions about how often they usually assign mathematics homework and how many minutes of mathematics homework they usually assign. High level indicates the assignment of more than 30 minutes of homework about half of the lessons or more. Low level indicates no assignment or the assignment of less than 30 minutes of homework about half of the lessons or less. Medium level includes all other possible combinations of responses.
Exhibit 8.3 Index of Availability of School Resources for Mathematics Instruction (ASRMI)	Index based on principals' average response to five questions about shortages that affect general capacity to provide instruction: instructional materials (e.g., textbook); budget for supplies (e.g., paper, pencils); school buildings and grounds; heating/cooling and lighting systems; and instructional space (e.g., classrooms); and the average response to five questions about shortages that affect mathematics instruction: computers for mathematics instruction; computer software for mathematics instruction; calculators for mathematics instruction; library materials relevant to mathematics instruction; and audio-visual resources for mathematics instruction. Average is computed based on a 4-point scale: 1. None; 2. A little; 3. Some; 4. A lot. High level indicates that both shortages are on average lower than 2. Low level indicates that both shortages are on average greater than or equal to 3. Medium level includes all other possible combinations of responses.

Exhibit 13.1 **Summary Indices in the TIMSS 2003 International Mathematics Report**
(...Continued)

Index	Analysis Method
Exhibit 8.4 Index of Principals' Perception of School Climate (PPSC)	Index based on principals' responses to eight questions about their schools: teachers' job satisfaction; teachers' understanding of the school's curricular goals; teachers' degree of success in implementing the school's curriculum; teachers' expectations for student achievement; parental support for student achievement; parental involvement in school activities; students' regard for school property; and students' desire to do well in school. Average is computed based on a 5-point scale: 1. Very high; 2. High; 3. Medium; 4. Low; 5. Very low. High level indicates average is less than or equal to 2. Medium level indicates that average is greater than 2 and less or equal to 3. Low level indicates average is greater than 3.
Exhibit 8.5 Index of Mathematics Teachers' Perception of School Climate (TPSC)	Index based on teachers' responses to eight questions about their schools: teachers' job satisfaction; teachers' understanding of the school's curricular goals; teachers' degree of success in implementing the school's curriculum; teachers' expectations for student achievement; parental support for student achievement; parental involvement in school activities; students' regard for school property; and students' desire to do well in school. Average is computed based on a 5-point scale: 1. Very high; 2. High; 3. Medium; 4. Low; 5. Very low. High level indicates average is less than or equal to 2. Medium level indicates that average is greater than 2 and less or equal to 3. Low level indicates average is greater than 3.
Exhibit 8.6 Index of Good School and Class Attendance (GSCA)	Index based on principals' responses to three questions about the seriousness of attendance problems in the school: arriving late at school; absenteeism (i.e., unjustified absences); and skipping class. High level indicates that all three behaviors either never occur or are reported not to be a problem. Low level indicates that two or more behaviors are reported to be a serious problem, or two behaviors are reported to be minor problems and the third a serious problem. Medium level includes all other possible combinations of responses.
Exhibit 8.7 Index of Mathematics Teachers' Perception of Safety in the Schools (TPSS)	Index based on teachers' responses to three statements about their schools: this school is located in a safe neighborhood; I feel safe at this school; this school's security policies and practices are sufficient. High level indicates that the teacher agrees a lot or agrees to all three statements. Low level indicates that teacher disagrees or disagrees a lot to all three statements. Medium level includes all other combinations of responses.
Exhibit 8.8 Index of Students' Perception of Being Safe in the Schools (SPBSS)	Index based on students' responses to five statements about things that happened in their schools in the last month (1 = yes, 2 = no): something of mine was stolen; I was hit or hurt by other student(s) (e.g., shoving, hitting, kicking); I was made to do things that I didn't want to do by other students; I was made fun of or called names; I was left out of activities by other students. High level indicates that the student answered NO to all five statements. Low level indicates that the student answered YES to three or more statements. Medium level includes all other possible combinations of responses.

Note: Detailed information about the computation of indices can be found in the TIMSS 2003 User Guide.

Exhibit 13.2 **Summary Indices in the TIMSS 2003 International Science Report**

Index	Analysis Method
Exhibit 4.7 Index of Time Students Spend Doing Science Homework (TSH)	Index based on students' reports on the frequency and amount of science homework they are given. High level indicates more than 30 minutes of science homework assigned 3-4 times a week. Low level indicates no more than 30 minutes of science homework no more than twice a week. Medium level includes all other possible combinations of responses.
Exhibit 4.9 Index of Students' Self-Confidence in Learning Science (SCS)	Index based on students' responses to four statements about science: 1) I usually do well in science; 2) Science is more difficult for me than for many of my classmates (Reversed); 3) Science is not one of my strengths (Reversed); 4) I learn things quickly in science. Average is computed across the four items based on a 4-point scale: 1. Agree a lot; 2. Agree a little; 3. Disagree a little; 4. Disagree a lot. Students agreeing a little or a lot on average across the four statements are assigned to the high level. Students disagreeing a little or a lot on average are assigned to the low level. All other students are assigned to the middle level.
Exhibit 4.10 Index of Students' Valuing Sciences (SVS) (Grade 8 only)	Index based on students' responses to seven statements about science: 1) I would like to take more science in school; 2) I enjoy learning science; 3) I think learning science will help me in my daily life; 4) I need science to learn other school subjects; 5) I need to do well in science to get into the university of my choice; 6) I would like a job that involved using science; 7) I need to do well in science to get the job I want. Average is computed across the seven items based on a 4-point scale: 1. Agree a lot; 2. Agree a little; 3. Disagree a little; 4. Disagree a lot. Students agreeing a little or a lot on average across the seven statements are assigned to the high level. Students disagreeing a little or a lot on average are assigned to the low level. All other students are assigned to the middle level.
Exhibit 7.2 Index of Teachers' Reports on Teaching Science Classes with Few or No Limitations on Instruction due to Student Factors (SCFL) (Grade 8 only)	Index based on teachers' responses to six statements about student factors limiting science instruction: 1) Students with different academic abilities; 2) Students who come from a wide range of backgrounds; 3) Students with special needs; 4) Uninterested students; 5) Low morale among students; 6) Disruptive students. Average is computed across the six statements based on a 4-point scale: 1. Not at all/Not applicable; 2. A little; 3. Some; 4. A lot. High level indicates average is less than or equal to 2. Medium level indicates average is greater than 2 and less than 3. Low level indicates average is greater than or equal to 3.
Exhibit 7.10 Index of Teachers' Emphasis on Science Homework (ESH)	Index based on teachers' responses to two questions about how often they usually assign science homework and how many minutes of science homework they usually assign. High level indicates the assignment of more than 30 minutes of homework about half of the lessons or more. Low level indicates no assignment or the assignment of less than 30 minutes of homework about half of the lessons or less. Medium level includes all other possible combinations of responses.
Exhibit 8.3 Index of Availability of School Resources for Science Instruction (ASRSI)	Index based on principals' average response to five questions about shortages that affect general capacity to provide instruction: instructional materials (e.g., textbook); budget for supplies (e.g., paper, pencils); school buildings and grounds; heating/cooling and lighting systems; and instructional space (e.g., classrooms); and the average response to six questions about shortages that affect science instruction: science laboratory equipment and materials; computers for science instruction; computer software for science instruction; calculators for science instruction; library materials relevant to science instruction; and audio-visual resources for science instruction. Average is computed based on a 4-point scale: 1. None; 2. A little; 3. Some; 4. A lot. High level indicates that both shortages are on average lower than 2. Low level indicates that both shortages are on average greater than or equal to 3. Medium level includes all other possible combinations of responses.

Exhibit 13.2 **Summary Indices in the TIMSS 2003 International Science Report**
(…Continued)

Index	Analysis Method
Exhibit 8.4 Index of Principals' Perception of School Climate (PPSC)	Index based on principals' responses to eight questions about their schools: teachers' job satisfaction; teachers' understanding of the school's curricular goals; teachers' degree of success in implementing the school's curriculum; teachers' expectations for student achievement; parental support for student achievement; parental involvement in school activities; students' regard for school property; and students' desire to do well in school. Average is computed based on a 5-point scale: 1. Very high; 2. High; 3. Medium; 4. Low; 5. Very low. High level indicates average is less than or equal to 2. Medium level indicates that average is greater than 2 and less or equal to 3. Low level indicates average is greater than 3.
Exhibit 8.5 Index of Science Teachers' Perception of School Climate (TPSC)	Index based on teachers' responses to eight questions about their schools: teachers' job satisfaction; teachers' understanding of the school's curricular goals; teachers' degree of success in implementing the school's curriculum; teachers' expectations for student achievement; parental support for student achievement; parental involvement in school activities; students' regard for school property; and students' desire to do well in school. Average is computed based on a 5-point scale: 1. Very high; 2. High; 3. Medium; 4. Low; 5. Very low. High level indicates average is less than or equal to 2. Medium level indicates that average is greater than 2 and less or equal to 3. Low level indicates average is greater than 3.
Exhibit 8.6 Index of Good School and Class Attendance (GSCA)	Index based on principals' responses to three questions about the seriousness of attendance problems in the school: arriving late at school; absenteeism (i.e., unjustified absences); and skipping class. High level indicates that all three behaviors either never occur or are reported not to be a problem. Low level indicates that two or more behaviors are reported to be a serious problem, or two behaviors are reported to be minor problems and the third a serious problem. Medium level includes all other possible combinations of responses.
Exhibit 8.7 Index of Science Teachers' Perception of Safety in the Schools (TPSS)	Index based on teachers' responses to three statements about their schools: this school is located in a safe neighborhood; I feel safe at this school; this school's security policies and practices are sufficient. High level indicates that the teacher agrees a lot or agrees to all three statements. Low level indicates that teacher disagrees or disagrees a lot to all three statements. Medium level includes all other combinations of responses.
Exhibit 8.8 Index of Students' Perception of Being Safe in the Schools (SPBSS)	Index based on students' responses to five statements about things that happened in their schools in the last month (1 = yes, 2 = no): something of mine was stolen; I was hit or hurt by other student(s) (e.g., shoving, hitting, kicking); I was made to do things that I didn't want to do by other students; I was made fun of or called names; I was left out of activities by other students. High level indicates that the student answered NO to all five statements. Low level indicates that the student answered YES to three or more statements. Medium level includes all other possible combinations of responses.

Note: Detailed information about the computation of indices can be found in the TIMSS 2003 User Guide

Exhibit 13.3 **Summary of Derived Variables Other than Indices in the TIMSS 2003 International Mathematics and Science Reports**

Derived Variable	Analysis Method
Exhibit 4.1 Highest Level of Education of Either Parent (Grade 8 only)	Derived variable based on students' responses to the highest level of education of mother and father. Cases classified in four categories: 1. Finished University or Equivalent or Higher 2. Finished Post-secondary Vocational/Technical Education but Not University 3. Finished Upper Secondary Schooling 4. Finished Lower Secondary Schooling 5. No More Than Primary Schooling
Exhibit 4.2 Students' Educational Aspirations Relative to Parents' Educational Level (Grade 8 only)	Derived variable based on students' responses to the highest level of education of mother and father, and students' expectations for further education. Cases were classified in four categories: 1. Finish University and Either Parent Went to University or Equivalent 2. Finish University but Neither Parent Went to University Equivalent 3. Not Finish University Regardless of Parents' Education 4. Do Not Know Regardless of Parents' Education
Exhibit 4.6 Use of Computer	Derived variable based on students' responses to where do they use a computer. Cases were classified in five categories: 1. Use Computer Both at Home and at School 2. Use Computer at Home but Not at School 3. Use Computer at School but Not at Home 4. Use Computer only at places other than home and school 5. Do Not Use Computer at All
Exhibit 6.5 Preparation to Teach Mathematics (Grade 4 only)*	Derived variable based on teachers' responses to main area of study during post-secondary education, and main area in specialization. Cases were classified in five categories: 1. Primary/Elementary Education with a Major or Specialization in Mathematics 2. Primary/Elementary Education with a Major or Specialization in Science but Not in Mathematics 3. Mathematics or Science Major or Specialization without a Major in Primary/Elementary Education 4. Primary/Elementary Education without a Major or Specialization in Mathematics or Science 5. Other
Exhibit 6.5 Preparation to Teach Science (Grade 4 only)*	Derived variable based on teachers' responses to main area of study during post-secondary education, and main area in specialization. Cases were classified in five categories: 1. Primary/Elementary Education with a Major or Specialization in Mathematics 2. Primary/Elementary Education with a Major or Specialization in Science but Not in Mathematics 3. Mathematics or Science Major or Specialization without a Major in Primary/Elementary Education 4. Primary/Elementary Education without a Major or Specialization in Mathematics or Science 5. Other

Note: Detailed information about the computation of indices can be found in the TIMSS 2003 User Guide
* At grade 8, "Preparation to teach" was reported using the direct reporting method.

In general, information from the curriculum questionnaires was directly reported in the exhibits. The information extracted from these questionnaires is mostly textual and qualitative in nature. In the case of quantitative information, descriptive statistics were provided. NRCs reviewed and approved the display of the curriculum information at the seventh NRC meeting. At that time, exhibits with data were available only for the mathematics curriculum at the eighth grade. After that meeting, ISC staff implemented the suggested changes to the curriculum exhibits, and completed them for both grades and subjects. Given the qualitative nature of the curriculum data, extensive follow-up and data cleaning were required. From January to June 2004, ISC staff carefully reviewed the curriculum data and asked NRCs to provide missing data, correct inconsistent data, and clarify questionable data. The final version of the curriculum exhibits was presented and approved at the eighth NRC meeting, when any lingering questions about the curriculum data were resolved.

13.5 Display of Background Data

TIMSS 2003 results were reported separately by subject area, with the mathematics and science results appearing in separate reports. Final exhibits with background data were organized into chapters 4 through 8 in the International Reports (the first three chapters reported achievement data). Chapter 4 reported data on students' characteristics, Chapter 5 on the curriculum, Chapter 6 on teachers' characteristics, Chapter 7 on instructional practices, and Chapter 8 on the schools.

It is important to note that in the data reported in the exhibits the student was always the unit of analysis, even when information from the teacher or school questionnaire was reported. In general, the exhibits presented the percentage of students having certain characteristics, or the percentage of students whose teachers or schools have various characteristics. For example, the International Reports give the percentage of students taught by teachers having a teaching certificate. This approach is consistent with the main goal of TIMSS, which is to inform about students' educational contexts and performance. The percentages in the exhibits were often accompanied by the students' mean achievement (mathematics or science). Information for each country was presented in individual rows, with the international average for all the participating countries (mean of countries' means) displayed separately. In general, where only one variable with several categories was reported in an exhibit, countries were displayed in rank order based on one of the categories, and where more than one variable was reported, countries were displayed in alphabetical order.

Whenever possible and relevant, the International Reports included trend data from 1995 (fourth and eighth grades) and 1999 (eighth grade only). Significant differences between the percentages of students having a given trait in each cycle were indicated. In other exhibits, data were displayed separately for boys and girls, and significant differences were also indicated.

In the science report, eighth grade background information was reported separately for the integrated science countries and for the separate science countries. The integrated science countries were reported in a "General/ Integrated Science" panel. The separate science countries were reported in four different panels: Biology, Earth Science, Chemistry, and Physics.

The exhibits in the International Reports contained special notations regarding response rates for the background variables. Although in general there were high response rates, some indicators and some countries had less than acceptable response rates. Since the student was the unit of analysis, the notation used in the International Reports always reflected the percentage of students for whom the responses from students, teachers, or schools were available. The following special notations were used to convey information about response rates in the exhibits in the International Reports:

- For a country where student, teacher, or school responses were available for 70 to 85 percent of the students, an "r" appeared next to the data for that country.
- Where student, teacher, or school responses were available for 50 to 69 percent of the students, an "s" appeared next to the data for that country.
- Where student, teacher, or school responses were available for less than 50 percent of the students, "x" replaced the data.
- Where the percentage of students in a particular category was less than two percent, achievement data were not reported in that category; the data were replaced by a tilde (~).
- Where data were not comparable for all respondents in a country, a dash (–) was used in place of data in all of the affected columns.[3]

3 A dash usually indicates that a background question was not administered in a country, but could also be due to translation problems or to the administration of a question that was determined to be not internationally comparable. In the exhibits based on the separate science subjects, the inclusion of dashes for specific countries is by design and reflects the specific science subjects not included in each country.

References

DeVellis, R. (1991). *Scale development, theory and applications*. Sage Publications on Applied Social Research Method Series Vol. 26. Newburry Park, California: Sage Publications, Inc.

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1998). *Applied statistics for the behavioral sciences* (4th ed.). Boston: Houghton Mifflin Company.

Pedhazur, E. (1997). *Multiple regression in behavioral research* (3rd ed.). Boston: Harcourt Brace College Publishers.

SAS Institute (2002). *SAS system for Windows* (version 9). Cary, NC: SAS Institute.

Spector, P. (1992). *Summated rating scale construction, an introduction*. Sage University Papers Series on Quantitative Applications in the Social Sciences, Series No. 07-082. Beverly Hills, California: Sage University Press.

SPSS Inc. (2002). *SPSS for Windows* (version 11.5). CHicago, IL: SPSS Inc.