6

Chapter 6

TIMSS 2003 Survey Operations Procedures

Juliane Barth, Eugenio J. Gonzalez, and Oliver Neuschmidt

6.1 Overview

The TIMSS 2003 data collection in each country was a very demanding exercise, with test administration at two grade levels in at least 150 schools, and with questionnaires for students, mathematics and science teachers, and school principals. To conduct the data collection successfully called for close cooperation between the National Research Coordinator (NRC) and school personnel – principals and teachers – and students. The first part of this chapter describes the field operations for collecting the data, including the responsibilities of the NRC, the procedure for sampling classrooms within schools and tracking students and teachers, and the steps involved in administering the achievement tests and background questionnaires. The second part describes the activities involved in preparing the data files at national centers, particularly the procedures for scoring the constructed-response items, creating and checking data files for achievement test and questionnaire responses, and dispatching the completed data files to the IEA Data Processing Center (DPC) in Hamburg, Germany.

6.2 TIMSS 2003 Field Operations

The TIMSS 2003 field operations were developed jointly by the TIMSS & PIRLS International Study Center at Boston College, the IEA Data Processing Center, and Statistics Canada. They were based on procedures used successfully in TIMSS 1995, TIMSS 1999, and other IEA studies, and were refined on the basis of TIMSS 2003 field-test experience.

6.2.1 Responsibilities of the National Research Coordinator

In conducting field operations in each country, the National Research Coordinator was the key person. The NRC had ultimate responsibility for collecting the data for the TIMSS assessment according to internationally agreed-upon procedures and preparing the data according to international specifications. NRC responsibilities in other areas, including sampling schools and translating the achievement tests and questionnaires, have been outlined in earlier chapters of this report.[1] This section focuses on NRC activities with regard to administering the assessment in participating schools. Specifically, it describes the procedures for sampling classes within schools, for tracking classes, teachers, and students in the sampled schools, and for organizing the administration of the achievement tests and questionnaires.

6.2.2 Documentation and Software

NRCs were provided with a comprehensive set of procedural manuals detailing all aspects of the data collection.

- The *TIMSS 2003 Survey Operations Manual* (TIMSS, 2002a) was the essential handbook of the National Research Coordinator, and described in detail all of the activities and responsibilities of the NRC, from the moment the TIMSS instruments arrived at the national center to the moment the checked and verified data files and accompanying documentation were submitted to the IEA Data Processing Center.
- The *TIMSS 2003 School Sampling Manual* (TIMSS, 2001) defined the TIMSS 2003 target populations and sampling goals and described the procedures for the sampling of schools.
- The *TIMSS 2003 School Coordinator Manual* (TIMSS, 2002b) described the activities of the School Coordinator – the person in the school responsible for organizing the TIMSS test administration – from the time the testing materials arrived at the school to the time the completed materials were returned to the national TIMSS center.
- The *TIMSS 2003 Test Administrator Manual* (TIMSS, 2002c) described in detail the procedures for administering the TIMSS tests and questionnaires, from the beginning of the test administration to the return of the testing materials to the School Coordinator.
- The *TIMSS 2003 Scoring Guides for Mathematics and Science Constructed-Response Items* (TIMSS, 2002d; TIMSS, 2002e) contained instructions for scoring the short-answer and extended-response test items.
- The *Manual for Entering the TIMSS 2003 Data* (TIMSS, 2002f) provided the NRCs with instructions for coding, entering, and verifying the data.

1 See Chapter 5 for information about sampling schools, and Chapter 4 for details of the translation task.

- The *TIMSS 2003 National Quality Control Observer's Manual* (TIMSS, 2002g) provided instructions for conducting classroom observations during data collection in a sample of participating schools.

Additionally, six software packages were supplied by the IEA Data Processing Center to assist NRCs with the data collection:

- The within-school sampling software (WinW3S) is a computer program that helps NRCs randomly sample the TIMSS class or classes in each sampled school; prepare the survey tracking forms that keep track of sampled students, classes, and teachers; and assign test booklets to students. The software stores all tracking information in an MS-Access database so that it can be used later in constructing sampling weights and in verifying the integrity of the sampling procedure.
- The DataEntryManager for Windows (WinDEM), is a computer program developed by IEA to enable national center staff to capture all of the TIMSS data through keyboard data entry and to perform a range of validity checks on the keyed data. The WinDEM database includes codebooks for each of the TIMSS 2003 test booklets and questionnaires, providing all information necessary to produce data files for each instrument in a standard international format.
- The WinLink program allows NRCs to check the correspondence between the tracking information stored in the WinW3S database and the student, teacher, and school information keyed into the WinDEM files. Using this program, for example, NRCs can check that each student listed on the student tracking form has a corresponding data record in the student achievement and student questionnaire WinDEM files.
- The Data Correction Software (DCS) is a program that enables national center staff to detect and correct inconsistencies in TIMSS background data files.
- The Trend-Scoring Reliability Software (TSRS) incorporates a database for each country containing a sample of student responses to constructed-response questions administered and scored as part of the TIMSS 1999 data collection. The TSRS software allowed NRCs to have their 2003 scoring staff rescore the 1999 student sample to document the reliability of the scoring process over time. This effort is described in Chapter 8.
- In a related effort, the Cross-Country Scoring Reliability Software (CCSRS) incorporates a database containing a sample of student responses to constructed-response items collected from English-speaking countries participating in TIMSS 2003. The CCSRS software enables every country with English-speaking scoring staff to score these common student responses in order to document the reliability of the scoring across countries participating in 2003. For more information, please refer to Chapter 8.

Each software package was supplied with a detailed manual describing how to install and use the software. In addition to the manuals, NRCs received hands-on training in the use of the WinW3S and WinDEM software from staff at the IEA Data Processing Center and Statistics Canada during a data entry seminar held before the field test.

6.2.3 Within-School Sampling Procedures

The study design anticipated relational analyses between student achievement and teacher-level data at the class level. For field operations, this meant that intact classes had to be sampled, and that for each sampled class the mathematics and science teachers had to be tracked and linked to their students. Although intact classes were the unit to be sampled in each school, the ultimate goal was a nationally representative sample of students. Consequently, in each country a classroom organization had to be chosen that ensured that every student in the school was in one class or another, and that no student was in more than one class. Such an organization is necessary for a random sample of classes to result in a representative sample of students. At the eighth grade in most countries, mathematics classes serve this purpose well, and so were chosen as the sampling units. In countries where students attended different classes for mathematics and science, classrooms were defined on the basis of mathematics instruction for sampling purposes.[2] At fourth grade, most schools use the same class for all subjects, including mathematics and science. Accordingly, the fourth-grade classroom was the sampling unit in these schools.

The TIMSS design required that for each student in each sampled class, all teachers teaching mathematics or science be identified and asked to complete a teacher questionnaire.

Although all students enrolled in the target grade were part of the target population and were eligible to be selected for testing, TIMSS recognized that some students in every school would be unable to take part in the 2003 assessment because of some physical or mental disability. Accordingly, the sampling procedures provide for the exclusion of students with any of several disabilities (see Chapter 5). Countries were required to track and account for all excluded students, and were cautioned that excluding an excessive proportion would lead to their results being annotated in the TIMSS 2003 international reports. It was important that the conditions under which students could be excluded be carefully delineated, because the definition of "disabled" students varied considerably from country to country.

2 For countries where a suitable configuration of classes for sampling purposes could not be identified, TIMSS also provided a procedure for sampling individual students directly from the eighth grade.

Exhibit 6.1 presents the major activities conducted by National Research Coordinators and school personnel while sampling classes within schools. These activities are incorporated in the WinW3S software, which automatically produces all necessary forms, lists, and labels, and assisted NRCs in keeping track of the field operations' status.

Exhibit 6.1 **Procedures for Sampling Classes in Partcipating Schools**

NRC activity	School activity
1. School Tracking	
• Contact schools participating schools	
• Prepare Class Listing Forms to be completed by schools.	
	2. **Complete the Class Listing Form listing all mathematics classes in the target grade (4 or 8) along with the names of their mathematics teachers.**
3. Class Tracking and Sampling	
• Sample a class or classes using the information on the Class Listing Form.	
• Prepare Student-Teacher Linkage Forms so that schools can list the students in the sampled class(es) and link them to their mathematics and science teachers.	
	4. **Complete Student-Teacher Linkage Forms by listing all of the students in the sampled class(es) (name, birth dates, sex) together with their mathematics and science teachers and course names.**
5. Student/ Teacher Tracking and Student-Teacher Linkage	
• Prepare a Student Tracking Form for each sampled class listing all students to be tested and their booklet assignments	
• Prepare a Teacher Tracking Form for each sampled class listing all mathemathematics and science teachers of the students in the class, their questionnaire assignments and their student-teacher link numbers	
• Send tracking forms, labels and test-instruments to schools.	
TEST ADMINISTRATION	
	6. **After the tests and questionnaires have been administered, record the participation status on Student and Teacher Tracking Forms; complete Test Administrator Forms.**
7. Record Participation Information and Test Administrator Information in Data Files.	

6.2.3.1 Survey Tracking Forms

As may be seen from Exhibit 6.1, TIMSS 2003 relied on a series of "tracking forms" to implement and record the sampling of classes, teachers, and students. It was essential that the tracking forms be completed accurately, since they determine which booklets and questionnaires should be given to which students and teachers, and record what happened as the assessment was administered in each school. In addition to facilitating the data collection, the tracking forms provided essential information for the computation of sampling weights and for evaluating the quality of the sampling procedures. All tracking forms were retained for review by staff of the TIMSS International Study Center and the IEA Data Processing Center.

Survey tracking forms were provided for sampling classes and students; for tracking schools, classes, teachers, and students; for linking students and teachers; and for recording information during test administration.

6.2.3.2 Linking Students, Teachers, and Classes

The Within-School Sampling Software (WinW3S) creates a hierarchical identification numbering system that uniquely identifies the sampled schools, teachers, classes, and students within each country. At the root of the system is a four-digit school identification number unique within each country that is assigned to each sampled school.

A class identification number is assigned to each class in the target grades listed on the class tracking form or entered in WinW3S. The six-digit class identification number consists of the four-digit school number followed by a two-digit number identifying the class within the school.

Each student listed on the student tracking form is assigned a student identification number. This eight-digit number consists of the six-digit class number followed by a two-digit number corresponding to the student's sequential position on the student tracking form. All students listed on the student tracking form, including those marked for exclusion, are assigned a student identification number.

Each mathematics and science teacher of the selected classes (i.e., those listed on the teacher tracking form) is assigned a teacher identification number consisting of the four-digit school number followed by a two-digit teacher number unique within the school. Since a teacher could be teaching both mathematics and science to some or all of the students in a class, it is necessary to have a unique identification number for each teacher/class and teacher/subject combination. This is achieved by adding a two-digit link number to the six digits of the teacher identification number, giving a unique

eight-digit teacher/class identification number. Careful implementation of these procedures is necessary so that during data analysis each class may be linked to a teacher, and student outcomes may be analyzed in relation to teacher-level variables.

6.2.4 Assigning Testing Materials to Students and Teachers

At both eighth and fourth grades, the mathematics and science assessment questions were packaged into 12 student test booklets. Each sampled student was required to complete one booklet, as well as the student questionnaire. Booklets were assigned to students by the WinW3S software using a random assignment procedure.

Each teacher listed on the teacher tracking form was assigned a teacher questionnaire. At eighth grade there were separate questionnaires for mathematics and science teachers. Where teachers taught both mathematics and science to the class, every effort was made to collect information about both subjects. However, NRCs had the final decision as to how much response burden to place on such teachers. Where a teacher taught both subjects to a class but completed only one questionnaire, the information from the general background part of the completed questionnaire was copied into the missing questionnaire.

6.2.5 Administering the Test Booklets and Questionnaires

The School Coordinator was the person in the school responsible for administrating the TIMSS 2003 assessment. The coordinator could be the principal, the principal's designee, or an outsider appointed by the NRC with the approval of the principal. The NRC was responsible for ensuring that the School Coordinators were familiar with their responsibilities.

The major responsibilities of the School Coordinators are detailed in the *TIMSS 2003 School Coordinator Manual* (TIMSS, 2002b). Prior to the test administration the tasks for the School Coordinator included:

- providing the NRC with all information necessary to complete the various tracking forms;
- checking the assessment materials when they arrived in the school to ensure that everything was in order;
- ensuring that the assessment materials were kept in a secure place before and after the administration;
- arranging the dates of the assessment administration with the national center;

- arranging for a Test Administrator and giving a briefing on the TIMSS 2003 study, the assessment materials, and the assessment sessions; and
- working with the school principal, the Test Administrator, and the teachers to plan the testing day – this involved arranging rooms, times, classes and materials.

The Test Administrator was responsible for administering the TIMSS tests and student questionnaires. Specific responsibilities were described in the *TIMSS 2003 Test Administrator Manual* (TIMSS, 2002c), and included:

- ensuring that each student received the correct testing materials which were specially prepared for him or her;
- administering the test in accordance with the instructions in the manual;
- ensuring the correct timing of the testing sessions by using a stopwatch and recording the time when the various sessions started and ended on the Test Administration Form; and
- recording student participation on the Student Tracking Form.

The responsibilities of the School Coordinator after the test administration included:

- ensuring that the Test Administrator returned all assessment materials, including the completed Student Tracking Form, the Test Administration Form, and any unused booklets;
- calculating the student response rate and arranging for makeup sessions if it was below 90 percent;
- distributing the teacher questionnaires to the teachers listed on the Teacher Tracking Form, ensuring that the questionnaires were returned completed, and recording teacher participation information on the Teacher Tracking Form;
- preparing a report for the NRC about the test administration in the school; and
- returning both completed and unused test materials and all tracking forms to the NRC.

The NRC prepared two packages for each sampled class. One contained the test booklets for all students listed on the Student Tracking Form and the other the student questionnaires. For each participating school, the test booklets and student questionnaires were bundled together with the Teacher Tracking Form and teacher questionnaires, the school questionnaire, and the materials prepared for briefing School Coordinators and Test Administrators, and were sent to the School Coordinator. A set of labels and prepaid envelopes addressed to the NRC was included to facilitate the return of testing materials.

6.2.6 National Quality Control Program

The International Study Center implemented an international quality control program whereby International Quality Control Monitors visited a sample of 15 schools in each country at each grade assessed and observed the test administration. In addition, NRCs were expected to organize a national quality control program, based upon the international model. This national program required Quality Control Observers to document data collection activities in their country. They visited a 10 percent sample of TIMSS 2003 schools, observed actual testing sessions, and recorded compliance of the test administration with prescribed procedures.

To assist NRCs in conducting their national quality control program, the TIMSS International Study Center prepared the *TIMSS 2003 National Quality Control Observer's Manual* (TIMSS, 2002g) which provided general information about TIMSS 2003 and detailed the role and responsibilities of the National Quality Control Observers.

6.3 Data Preparation

In the period immediately following the administration of the TIMSS 2003 assessment, the major tasks for the NRC included retrieving and collating the materials from the schools; recruiting and training scorers to score the constructed-response items; scoring these items, including double scoring a reliability sample of 1200 booklets; entering the data from the achievement tests and background questionnaires into computer files; checking and editing the data with the software provided by the IEA Data Processing Center; submitting the data files and materials to the IEA Data Processing Center; and preparing a report on survey activities.

When the testing materials were received back from the schools, NRCs had the following tasks:

- check that the appropriate testing materials were received for every student listed on the Student Tracking Form;
- verify all identification numbers on all instruments;
- check that the participation status recorded on the tracking forms matched the information on the test booklets and questionnaires; and
- follow up on schools that did not return the testing materials or for which forms were missing, incomplete, or inconsistent.

NRCs then organized the tests for scoring and data entry. The procedures involved were designed to maintain identification information that linked students to schools and teachers, minimize the time and effort spent handling the booklets, ensure reliability in the constructed-response coding, and document the reliability of the coding.

6.3.1 Scoring the TIMSS 2003 Constructed-Response Items

Reliable application of the scoring guides to the constructed-response questions, and empirical documentation of the reliability of the scoring process, were critical to the success of TIMSS 2003. The *TIMSS 2003 Survey Operations Manual* (TIMSS, 2002a) provided suggestions about arranging for staff and facilities for the constructed-response scoring effort required for the TIMSS 2003 main data collection; for effective training of the scorers; and for incorporating reliability scoring into the scheme for distributing booklets to scorers and monitoring the scoring. Countries were to double score 1200 booklets to document scoring reliability.

For all countries, the scope of the constructed-response scoring effort was substantial. The assessment contained 130 constructed-response questions at fourth grade and 146 constructed-response questions at eighth grade. These were distributed across 12 student booklets at each grade level.

6.3.1.1 Preparing to Train the Scorers

To ascertain the staff requirements for constructed-response scoring, it was necessary to estimate the amount of scoring to be done and the amount of time available to do it, and also to make provision for staff training and for clerical and quality control throughout the operation. The TIMSS International Study Center recommended at least one half-day of training on each of the 12 booklets, for a total of about a week for training activities.

In scoring the constructed-response items, it was vital that scoring staff apply the scoring rules consistently and in the same way in all participating countries. Hence, in selecting those who were to do the scoring, NRCs took care to arrange for persons who were conscientious and attentive to detail, knowledgeable in mathematics and science, and willing to apply the scoring guides as stated, even if they disagreed with a particular definition or category. Preference was given to individuals with educational backgrounds in the mathematics and science curriculum areas or who had taught at the middle school or primary level. Good candidates for scoring included teachers, retired teachers, college or graduate students, and staff of education agencies or ministries and research centers.

The success of assessments that, like TIMSS, include a large proportion of constructed-response questions is crucially dependent upon reliable scoring of student responses. In TIMSS 2003, scoring reliability was assured through the provision of detailed scoring guides (manuals), extensive training in their use, and continuous monitoring of the quality of the work. To support training in scoring, TIMSS 2003 provided training packets for training

in selected questions, and practice papers to help scorers achieve a consistent level of scoring.

At the international scoring training meetings, NRCs received training packets containing example responses and practice papers to help them achieve accuracy and consistency in scoring. For scoring guides that were difficult, example responses were selected to illustrate the scoring categories. The scores on these responses were explained and attached to the scoring guides. Practice sets were created for the more difficult guides. These papers illustrated a range of responses, beginning with several clear-cut examples. About 10 to 15 responses were enough for most guides, but sometimes more practice was necessary.

Each scorer received a copy of the *TIMSS 2003 Main Survey Scoring Guides for Mathematics and Science Constructed-Response Items* (TIMSS, 2002d; TIMSS, 2002e). These manuals explain the TIMSS scoring system, which was designed to produce a rich and varied profile of the range of students' competencies in mathematics and science, and provide detailed scoring guides and example student responses for each constructed-response question in the assessment.[3]

6.3.1.2 Conducting the Constructed-Response Scoring

TIMSS 2003 recommended that scorers be organized into teams of about six, headed by a team leader. The leader's primary responsibility was to monitor scoring reliability by continually checking and rechecking the scores that scorers had assigned. This process, known as back-reading, was essential for identifying scorers who did not understand particular guides or categories. Early detection of any misunderstandings permitted clarification and rectification of mistakes before too many responses had been scored. The back-reading systematically covered the daily work of each scorer. If a particular scorer appeared to have difficulty, however, then the percentage of back-reading for that scorer was increased. Any errors discovered were brought to the attention of the scorer responsible and corrected immediately. If a scorer was found to have been consistently making an error, then all of the booklets scored by that person were checked and any errors corrected.

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

6.3.1.3 Monitoring Scoring Reliability Within Each Country

To establish the reliability of the scoring within each country, NRCs were required to have a random sample of at least 100 booklets of each of the 12

3 See Chapter 2 for a description of the TIMSS constructed-response scoring system.

student test booklets scored independently by two different scorers. The reliability sample of booklets was selected randomly by the WinW3S software. The degree of agreement between the scores assigned by the two scorers is a measure of the reliability of the scoring process. Since the purpose of the double scoring was to document the consistency of the scoring procedure in each country, the procedure used for scoring the booklets in the reliability sample had to be as close as possible to that used for scoring the booklets in general. The procedure recommended by the TIMSS International Study Center was designed to blend the scoring of the reliability sample with the normal scoring activity, to take place at the same time, and to be systematically implemented across student responses and scorers.

In scoring the booklets for the main data set, scorers entered their scores directly into the student booklets. Therefore, in order that the reliability scoring be done "blind" (i.e., so that the two scorers did not know each other's scores), the reliability scoring had to be done before the scoring for the main data, and the reliability scores had to be recorded on a separate scoring sheet, and not in the booklets.

To implement the scoring plan effectively it was necessary that the scorers be divided between two equivalent teams (Team A and Team B), and that booklets be divided into two equivalent sets (Set A and Set B). The scorers in Team A scored around 600 of the booklets in Set B and all the booklets in Set A, while the scorers in Team B scored around 600 of the booklets in Set A and all of the booklets in Set B. Each team, therefore, handled both sets of booklets. For the set it handled first, the team did the reliablity scoring first and recorded the results on a separate answer sheet (this was the reliability sample). In the other set, the team scored all booklets and wrote the scores directly into the booklets.

Periodically during the day, the Team B scorers scored the reliability sample in the Set A batches, while the Team A scorers scored the reliability sample in the Set B batches. It was important that the reliability sample was scored as randomly drawn by the WinW3S software, and not just the top quarter in the set. When the reliability scoring was finished, Team B scorers marked it as completed and forwarded the batch to the Team A scorers. Similarly, the Team A scorers forwarded their scored reliability booklets from Set B to the Team B scorers. Once the booklets from Set A had been distributed to Team A scorers and the Set B booklets to the Team B scorers, all the constructed-response items were scored, and the scores were entered directly into the booklets.

6.3.1.4 Monitoring Scoring Reliability over Time (1999 to 2003)

The double scoring of a sample of the student test booklets provided a measure of the consistency within each country with which constructed-response questions were scored. To measure trends since 1999 and 1995, TIMSS 2003 included items from both of these assessments. TIMSS 2003 took steps to show that those constructed-response items used in 2003 that also had been used in 1999 were scored in the same way in both assessments. To make this possible, countries that participated in TIMSS 1999 sent samples of scored student booklets from the 1999 data collection to the IEA Data Processing Center, where they were digitally scanned and stored for later use. So that the student responses from 1999 could be rescored by 2003 scoring staff as a reliability check, the DPC developed software known as the Trend Scoring Reliability Software (TSRS) that presented the 1999 student responses without their 1999 scores. This enabled 2003 scoring staff to score these 1999 responses without seeing the scores awarded in 1999 and so provide a check on scoring consistency from 1999 to 2003. Those items from 1995 that were used in TIMSS 2003 all were in multiple-choice format, and therefore scoring reliability was not an issue.

6.3.1.5 Monitoring Scoring Reliability Across Countries

Because of the many different languages in use in TIMSS, establishing the reliability of constructed-response scoring across all countries was not feasible. However, TIMSS 2003 did conduct a cross-country study of scoring reliability among northern-hemisphere countries whose scorers were proficient in English. A sample of student responses to a subset of the mathematics and science constructed-response items was provided by the English-speaking southern hemisphere countries. These student responses were digitally scanned and incorporated into customized software known as the Cross-Country Scoring Reliability Software (CCSRS), developed by the DPC. English-speaking scorers in each of the northern-hemisphere countries used this software to independently score the student responses. The degree of agreement between scorers from the various countries may be taken as a measure of cross-country scoring reliability.

6.3.2 Data Entry

As described earlier in this chapter, the IEA Data Processing Center provided an integrated computer program for keyboard data entry and data verification known as DataEntryManager for Windows (WinDEM). This program works on all IBM-PC compatible personal computers running under MicroSoft's Windows operating system (Windows 95, 98, 2000, XP, and NT). WinDem

imports student and teacher tracking information directly from the W3S sampling software, facilitating keyboard data entry of responses to test booklets and questionnaires. WinDEM also offered data and file management capabilities, a convenient checking and editing mechanism, interactive error detection, and reporting and quality control procedures. Detailed information and operational instructions were provided in the WinDem manual. Since WinDEM incorporated the international codebooks describing all variables, use of the software ensured that the data files were produced according to the TIMSS 2003 rules and standards for data entry. Although use of WinDEM for all data entry tasks was strongly recommended, NRCs were permitted to use their own procedures and computer programs, as long as all data files conformed to the specifications of the international codebooks. DPC staff provided training to NRCs and national center personnel at various stages of the project, including prior to the field test and for six countries again prior to the main data collection.

NRCs who chose not to use WinDEM for data entry still had to ensure that all data files delivered to the DPC were in the international format and had passed all of the verification checks built into the WinDEM program. This can be accomplished by running WinDEM in data-checking mode on the data files. The WinDEM data-checking facility identifies a range of problems with identification numbers, out-of-range and otherwise invalid codes, and data file structure that can can be rectified before submitting the files to the DPC. In addition to the data-validation checks incorporated in WinDEM, NRCs were expected to use the WinLINK (or LinkCheck) program supplied by the DPC to verify the integrity of the links between the various student, teacher, and school files. Data files were acceptable at the DPC only if the reports generated by the WinDEM program and WinLINK programs indicated no errors.

During the TIMSS 2003 data collection, data were gathered from several sources, including students, teachers, and principals, as well as from a range of tracking forms. These data were recorded into data files as follows:

- The **school background file** contained information from the school background questionnaire.
- The **mathematics teacher background file** (eighth grade only) contained information from the eighth-grade mathematics teacher questionnaire.
- The **science teacher background file** (eighth grade only) contained information from the eighth-grade science teacher questionnaire.
- The teacher background file (fourth grade only) contained information from the fourth-grade classroom teacher questionnaire.

- The **student background file** contained data from the student background questionnaire.
- The student achievement file contained the achievement test booklet data.
- The **constructed-response scoring reliability file** contained the within-country scoring reliability data for the constructed-response items.

When all data files had passed the WinDEM and WinLINK/LinkCheck quality control checks, they were dispatched to the IEA Data Processing Center in Hamburg for further checking and processing.

6.3.3 Survey Activities Report

NRCs were requested to maintain a record of their experiences during the TIMSS 2003 data collection and to send a report to the TIMSS International Study Center when data-collection activities were completed. The report should describe any problems or unusual occurrences in selecting the sample or securing school participation, translating or preparing the data-collection instruments, administering the tests and questionnaires in the schools, scoring the constructed-response items, or creating and checking the data files.

6.3.4 Data Management Forms

NRCs were requested to document in a series of Data Management Forms any adaptations to the international instruments that they made while producing their national instruments. These forms were sent to the TIMSS International Study Center as well as to the IEA Data Processing Center. The information is used in the data editing and formatting process to recode data wherever possible to a form that allows for international comparisons. Additionally, the information provided in the Data Management Forms is included in a supplement to the *TIMSS 2003 User Guide for the International Database*.

6.4 Summary

This chapter has summarized the design and implementation of the TIMSS 2003 field operations from the point of first contact with the sampled schools to the submission of the ckecked and verified data files to the IEA Data Processing Center. Although the procedures were sometimes complex, each step was clearly documented in the TIMSS operations manuals and supported by training sessions at the NRC meetings. NRC Survey Activities Reports indicated that the field operations generally went well, and that the TIMSS 2003 data were of high quality.

References

TIMSS (2001), *TIMSS 2003 School Sampling Manual*, prepared by P. Foy & M. Joncas, Statistics Canada, Chestnut Hill, MA: Boston College.

TIMSS (2002a), *TIMSS 2003 Survey Operations Manual*, prepared by the International Study Center, Chestnut Hill, MA: Boston College.

TIMSS (2002b), *TIMSS 2003 School Coordinator Manual*, prepared by the International Study Center, Chestnut Hill, MA: Boston College.

TIMSS (2002c), *TIMSS 2003 Test Administrator Manual*, prepared by the International Study Center, Chestnut Hill, MA: Boston College.

TIMSS (2002d), *TIMSS 2003 Scoring Guides for Mathematics and Science Constructed-Response Items – Eighth Grade*, prepared by the International Study Center, Chestnut Hill, MA: Boston College.

TIMSS (2002e), *TIMSS 2003 Scoring Guides for Mathematics and Science Constructed-Response Items – Fourth Grade*, prepared by the International Study Center, Chestnut Hill, MA: Boston College.

TIMSS (2002f), *Manual for Entering the TIMSS 2003 Data*, prepared by the International Study Center in cooperation with the IEA Data Processing Center, Chestnut Hill, MA: Boston College.

TIMSS (2002g), *TIMSS 2003 National Quality Control Observer Manual*, prepared by the International Study Center, Chestnut Hill, MA: Boston College.