

Martin, M.O., Mullis, I.V.S., and Kelly, D.L. (1996) "Quality Assurance Procedures" in M.O. Martin and D.L. Kelly (eds.), *Third International Mathematics and Science Study (TIMSS) Technical Report, Volume I: Design and Development*. Chestnut Hill, MA: Boston College.

11. QUALITY ASSURANCE PROCEDURES.....11-1

Michael O. Martin, Ina V.S. Mullis, and Dana L. Kelly

11.1	OVERVIEW.....	11-1
11.2	STANDARDIZATION OF THE TIMSS PROCEDURES.....	11-2
11.3	PROCEDURES FOR TRANSLATION AND ASSEMBLY OF THE ASSESSMENT INSTRUMENTS....	11-4
11.4	SCORING THE OPEN-ENDED RESPONSES.....	11-5
11.5	NATIONAL QUALITY CONTROL PROGRAM.....	11-6
11.6	TIMSS QUALITY CONTROL MONITORS.....	11-6
11.7	THE QUALITY CONTROL MONITOR'S VISIT TO THE SCHOOLS.....	11-9

11. Quality Assurance Procedures

Michael O. Martin
Ina V.S. Mullis
Dana L. Kelly

11.1 OVERVIEW

A study as ambitious as TIMSS, and one that involves the collaboration of as many individuals, requires particular attention to all aspects of quality assurance to ensure that the design is properly implemented and that the data collected are comparable across all countries. As documented in previous chapters of this report, TIMSS has expended considerable effort in developing standardized materials and procedures so that the data collected in all countries are comparable to the greatest possible extent. In addition to setting high standards for data quality, the International Study Center has tried to ensure the overall quality of the study through a dual strategy of support to the national centers and quality control checks.

This chapter describes the procedures used to ensure high-quality data across all countries, and the support afforded to the national centers by the International Study Center in the form of standardized manuals, software aids, practical training, and technical assistance. The chapter describes also the development and implementation of an important aspect of the TIMSS quality assurance efforts—the program of site visits by trained Quality Control Monitors. The data collected during these visits are presented in Martin and Mullis (1996), together with extensive documentation of the quality of

translation activities, population sampling, free-response coding, and data checking and database construction.

11.2 STANDARDIZATION OF THE TIMSS PROCEDURES

One of the main ways in which TIMSS sought to achieve uniform project implementation was by providing clear and explicit documentation of all operational procedures. Such documentation was primarily in the form of operations manuals, supported where possible by computer software systems that implement the specified procedures. Forms accompanying some of the manuals serve to document the implementation of the procedures in each country. The manuals are described below. Bibliographic references can be found at the end of the chapter.

Sampling Manual: Defines the operational definitions of the school sample and details the procedures for selecting it for Populations 1 and 2. The forms provided in the *Sampling Manual* ensure that vital information at key stages is collected and recorded in a uniform manner. Target population definitions, choice of stratifying variables, definition of excluded populations, construction of school sampling frames, and selection of school samples are clearly documented.

Sampling Guide for Population 3: This outlines the school sampling procedures for Population 3.

Survey Operations Manual: The *Survey Operations Manuals* (one for Populations 1 and 2 and one for Population 3) was prepared by the IEA Data Processing Center for the National Research Coordinators (NRCs) and their colleagues who were responsible for implementing the TIMSS procedures. It describes the activities and responsibilities of NRCs from the moment the international testing materials arrive at the national center to the moment the cleaned data sets and accompanying documentation are submitted to the IEA Data Processing Center. The manual includes:

- Procedures for translating and assembling the test instruments and questionnaires
- Descriptions of the approved procedures for within-school sampling and guidelines for selecting the appropriate procedure
- Instructions for obtaining cooperation from the selected schools
- Instructions for installing and using the IEA software to prepare the sampling and tracking forms
- Explicit procedures for packing and sending materials to the schools
- Preparations for test administration and instructions for data entry and verification.
- An important feature of the *Survey Operations Manuals* is the detailed instructions for completing the various forms that are required to implement and document the within-school sampling procedure. The forms assist NRCs in their work by making each step

explicit, and also provide an audit trail that facilitates the International Study Center's evaluation of the implementation of the procedures.

School Coordinator Manual: Describes the steps the School Coordinator follows from the moment the testing materials arrive at the school to the moment they are returned to the NRC.

Test Administrator Manual: Covers the procedures from the beginning of testing to the return of the completed tests, questionnaires, and tracking forms to the School Coordinator. Included in this manual is an administration script to be read by the Test Administrator.

Guide to Checking, Coding and Entering the TIMSS Data: Provides further instructions on the procedures for coding, entering, and verifying the TIMSS data.

Performance Assessment Administration Manual: Provides instructions for selecting the sample of students, collecting the equipment for the tasks, and administering the TIMSS performance assessment.

Coding Guide for Performance Assessment: Together with the *Supplement to the Coding Guide for Performance Assessment*, contains the coding rubrics for the performance assessment items and exemplar coded student responses.

Coding Guides for Free-Response Items: Contain the coding rubrics for the free-response items and exemplar coded student responses.

Field Operations Software: This software was provided to assist the NRCs in selecting classes and students and preparing the booklet labels with the student identification. The sampling software targeted primarily within-school sampling activities, although it also provided for sampling of schools. It automatically generated all of the required documentation forms. Training in the use of the software was provided to NRCs, and Statistics Canada gave technical support.

International Codebooks: Contain the necessary information to code, enter, and verify the data from the tests and questionnaires. They are accompanied by data entry software (DATAENTRYMANAGER, DEM), which contains the codebooks.

Data Entry Software: Study participants received software specially developed to facilitate within-school sampling activities and data entry and management (DEM, see later section). Training in the use of these software products and technical support were also provided.

Throughout TIMSS, small-group training sessions during the semi-annual meetings of the National Research Coordinators (NRCs) dealt with desktop publishing, the use of the data entry software, and the use of the sampling software. Individual consultations between NRCs and staff members from the International Study Center, the IEA Data Processing

Center, and Statistics Canada provided further training in the TIMSS procedures. Presentations at NRC meetings, and progress reports disseminated via e-mail, fax, and mail, keep NRCs up to date on the status of TIMSS and their current and future tasks.

11.3 PROCEDURES FOR TRANSLATION AND ASSEMBLY OF THE ASSESSMENT INSTRUMENTS

In any comparative study of student achievement that takes place in more than one language it is crucial that procedures for ensuring comparable translations are followed. With the administration of the TIMSS survey instruments in languages, this was especially important. Furthermore, following translation, the NRCs had to assemble the test booklets according to a complicated booklet assembly plan. This step in the preparation of the instruments introduced another area of concern—uniformity of the test booklet and questionnaire layout. In order to ensure that all instruments administered in all languages and countries were equivalent, TIMSS established a series of procedures which were documented through manuals and supplementary material.

11.3.1 TRANSLATION OF THE TIMSS INSTRUMENTS

TIMSS participants were provided with a set of procedures to help them obtain reliable and high-quality translations. The *Survey Operations Manual* (TIMSS, 1994f, 1994g) contains the following guidelines for translators:

- Identify and minimize cultural differences
- Find equivalent words and phrases
- Make sure the reading level is the same in the target language as in the original English version
- Make sure the essential meaning does not change
- Make sure the difficulty level of achievement items does not change
- Be aware of changes in layout due to translation.

Also included were guidelines for decisions about vocabulary, meaning, and item and booklet layout, and guidelines for making cultural adaptations. Translators were also cautioned to ensure that another possible correct answer for a test item was not introduced. The translations were verified by an independent translation agency (this was coordinated by the International Coordinating Center in Vancouver). The independent translators prepared a translation verification report documenting the quality of the translations and corrections to be made to the booklets. A series of statistical checks were also conducted to identify problematic translations. The TIMSS translation procedures and verification process are further described by Beverley Maxwell in Chapter 8 of this report and in Mullis, Kelly, and Haley (1996).

11.3.2 ASSEMBLY OF THE TIMSS INSTRUMENTS

For the main survey, the International Study Center provided paper and electronic versions of the achievement test and questionnaire items, and paper versions of the completed test booklets and questionnaires for NRCs to use when assembling their national versions of the instruments. Instructions for the layout, printing, and assembly of the booklets also were provided in the *Survey Operations Manuals* and in the supplementary *Instructions for the Preparation of the Instruments at the National Centers*. These materials included directions for the layout of the item clusters, with special warnings related to editing and formatting; for verifying the translation; for printing the clusters from the electronic files; and for assembling the test booklets. In addition, the questionnaires were accompanied by notes on their adaptation by the national centers.

11.4 SCORING THE OPEN-ENDED RESPONSES

Because of the heavy reliance on the use of free-response questions, ensuring reliability of scoring was a major concern for TIMSS. As one step towards this goal, the International Study Center prepared coding guides for Populations 1 and 2 free-response items, for Population 3 mathematics and science literacy, physics, and advanced mathematics items, and for the performance assessment tasks. These contain the scoring rubrics for each item, and each of these is accompanied by exemplar coded student responses to illustrate how the codes are to be applied. In addition, the *Guide for Coding, Checking, and Entering the TIMSS Data* (TIMSS, 1995c) contained specific instructions related to coding. These instructions pertained to the following.

- Arranging for staff and facilities
- Distributing booklets to coders
- Procedures for coding the 10% reliability sample
- Procedures for monitoring the coding
- Preparing materials to train the coders
- Training the coders
- The roles and responsibilities of the coders
- The roles and responsibilities of the group leaders in coding.

Furthermore, an extensive training program was established in which representatives from each country were trained in the coding procedures (see Chapter 10).

In order to document the reliability of free-response coding (i.e., the degree of agreement between coders) in each country, two coders independently coded a random sample of 10% of the student responses (or, for main survey samples larger than 7,500 students, a random sample of 100 booklets from each booklet type). To help with this process, the International Study Center developed a procedure that separated the booklets into two equivalent samples as part of receipt control (by odd- and even-numbered school identifications). The scorers were also designated as belonging to one of two equivalent groups. First, scorers in

one group coded every tenth booklet on a separate coding sheet. These data constitute the 10% reliability sample. Then, scorers in the second group scored all the booklets and record codes in the booklets for data entry. This procedure ensures that the two coders do not know each other's codes, that each booklet is coded by two different scorers, and that the reliability scoring is distributed relatively equally among scorers.

In addition, the International Study Center conducted an international coding reliability study to obtain information about the degree of agreement among coders from different countries. A comprehensive study of inter-coder agreement across countries was beyond the resources of TIMSS. However, a limited study was designed and implemented in which 39 English-speaking coders from 21 countries coded a sample of booklets from 7 countries that tested in English. The results of the reliability studies are reported in Mullis and Smith (1996).

11.5 NATIONAL QUALITY CONTROL PROGRAM

As part of the national quality control efforts, NRCs were requested to arrange a program of unannounced visits by quality control observers to the schools on the day of testing. The main purpose of these visits was to ensure the proper implementation of the TIMSS policies and procedures in the schools and during test administration. The *Survey Operations Manuals* describe the steps to be taken to arrange for the quality control observation component and contains a list of the tasks of the quality control observer.

The International Study Center made available the manual and accompanying forms developed for the international quality assurance program. NRCs were encouraged to use the international materials to conduct their national quality control programs.

11.6 TIMSS QUALITY CONTROL MONITORS

As a major part of the TIMSS quality assurance efforts, a program of site by TIMSS Quality Control Monitors hired by the International Study Center was established. The purpose of this program was to observe the administration of the achievement tests in a sample of classrooms in participating classrooms, and document the degree of compliance with prescribed procedures.

In December 1994, the TIMSS International Study Center contracted Goodison Associates (United States) to help with the hiring, training, and overseeing of a team of Quality Control Monitors. In January 1995, NRCs were asked to nominate a person, such as a retired school teacher, to serve in that capacity for their country. The International Study Center reviewed the nominations and in almost all cases selected the NRC's first suggestion for a Quality Control Monitor. The monitors were trained centrally before returning to their countries to interview the NRC and to observe classroom testing sessions.

The TIMSS Quality Control Monitors (QC Monitors) were trained in a two-day session in which they were briefed on the design and purpose of TIMSS, the responsibilities of the NRC in conducting the study in each country, and their own roles and responsibilities. In

total, five training sessions were held for QC Monitors. Most of the monitors were trained during the three originally scheduled sessions: February 1995, London; March 1995, Enschede; April 1995, Paris. Two additional training sessions were held to train the remaining QC monitors, from Argentina (August 1995, Philadelphia) and Australia and New Zealand (July 1995, New Zealand).

The *Manual for the TIMSS Quality Control Monitors* (TIMSS, 1995e) was developed by the International Study Center with the assistance of Goodison Associates and was used as the basis for the training sessions. The manual included:

- An introduction to TIMSS, outlining the purpose of the study, study schedule, management arrangements, the major components of TIMSS (populations, sampling design, test and questionnaire design), and the purpose of the quality assurance program
- An overview of the roles and responsibilities of the TIMSS Quality Control Monitor
- An overview of the major tasks of the NRC
- Instructions for visiting the national center, interviewing the NRC, collecting the required materials from the NRC, and using the translation verification report to check the implementation of the suggestions made in the international review of the translations
- A report on the interview with the NRC
- Step-by-step procedures for selecting the schools for classroom observation
- Instructions for visiting these schools: arranging the visit, observing the testing sessions, completing the Classroom Observation Record, and interviewing the School Coordinator
- A copy of the Classroom Observation Record
- Instructions for returning materials to the International Study Center.

In addition to the *Manual for Quality Control Monitors*, each QC Monitor received copies of the *Survey Operations Manuals*, the *Test Administrator Manual*, the *School Coordinator Manuals*, and the *Guide to Checking, Coding and Entering the TIMSS Data*, which describe the procedures required for the implementation of TIMSS in each country. Although QC Monitors did not need to know every TIMSS policy and procedure in detail, they were encouraged to read through all the manuals in order to become familiar with the work of NRCs and the procedures to be followed in each country participating in TIMSS.

During each training session a staff member from the International Study Center explained the structure and major components of the study, emphasizing the NRC's tasks, especially as they related to the QC Monitor's duties. Goodison Associates reviewed the roles and responsibilities of the QC Monitor, and led QC Monitors through the Interview with the National Research Coordinator and the Classroom Observation Record. QC monitors also took part in an exercise to help them select the schools for classroom observation.

11.6.1 INTERVIEW WITH THE NATIONAL RESEARCH COORDINATOR

The QC Monitor's visit to the national center included an interview with the NRC and the selection of schools for classroom observation. The structured interview dealt with the NRCs' ten major responsibilities.

- Selecting the sample of students to be tested
- Working with the School Coordinators
- Translating the test instruments
- Assembling and printing the test booklets
- Packing and shipping the necessary materials to the designated School Coordinators
- Arranging for the return of materials from the school sites
- Arranging for coding the free-response and performance assessment questions
- Entering into data files the testing results and information from students, teachers, and principals
- Conducting on-site quality assurance observations for a 10% sample of schools
- Preparing the NRC report on survey activities.

The QC Monitor recorded the NRC's responses to questions about the implementation of these responsibilities, and any additional comments made regarding the TIMSS procedures. The interview questions were designed to ascertain the degree to which the procedures and policies described in the *Survey Operations Manuals*, the *Sampling Manual*, the *Guide to Coding, Checking, and Entering the TIMSS Data*, and other documents were followed. The results of the interviews with the NRCs are summarized in Martin, Hoyle, and Gregory (1996a).

Following the interview with the NRC, the QC Monitor and the NRC worked together to select ten schools for classroom observation, plus three extra schools as potential replacements. Using the School Tracking Form, the QC Monitor and NRC selected the schools by a random selection process (albeit one subject to a number of practical constraints). The schools selected for classroom observation had to be within easy traveling distance of the QC Monitor's home so that travel and observation could be done in one working day; the NRC or QC Monitor had to be able to contact the school to ascertain the date and time of testing and to arrange the visit; the school could not be taking part in the NRC's own national quality control observation program; and the testing could not yet have taken place in that school. After the schools, the classrooms for observation were selected. Where possible, the class chosen was the upper-grade class. The school name and classroom selected for observation were recorded on the Classroom Observation Tracking Form.

At the end of the visit to the national center, the QC Monitor collected the following materials from the NRC:

- *Test Administrator Manual*

-
- *School Coordinator Manual*
 - Test booklets (for each population assessed)
 - Performance assessment tasks (for each population assessed, if participating)
 - School questionnaires (for each population assessed)
 - Student questionnaires (for each population assessed)
 - Teacher questionnaires (for each population assessed)
 - Translation Verification Report (if this was not given to the QC Monitor at the training session)
 - Student Tracking Forms for each class selected for observation
 - Class Tracking Forms for each school selected for observation.

QC Monitors received the Translation Verification Report either from the International Study Center during training or from the NRC on their visit to the national center. The QC Monitor checked that any deviations in translation or booklet layout were corrected before test administration, recorded that information, and submitted it to the International Study Center together with the instruments and manuals collected from the NRC.

11.7 THE QUALITY CONTROL MONITOR'S VISIT TO THE SCHOOLS

The QC Monitor was given instructions for arranging the visits to the schools selected for observation, including guidelines for telephoning the School Coordinator and discussing the objectives of the QC Monitor.

To document the activities during the testing session in each school selected for a site observation, QC Monitor used the Classroom Observation Record, which documents the following:

- Activities preliminary to the testing session, including security of the test booklets, level of preparation of the Test Administrator, and adequacy of supplies and testing environment
- Activities during the testing session, including distribution of the test booklets to the appropriate students (using the Student Tracking Form), timing of the testing and breaks, and the Test Administrator's accuracy in reading the test administration script
- The QC Monitor's general impressions of the testing session, including the orderliness of the students, the Test Administrator's answering of students' questions, documentation of any cheating, handling of defective test booklets (if any), and handling of late students (if any).

11.7.1 INTERVIEW WITH THE SCHOOL COORDINATOR

Following the observation of the testing session, the QC Monitor met with the School Coordinator to conduct a brief interview covering the School Coordinator's evaluation of the TIMSS testing and suggestions for improvement, and any additional background

information. The QC Monitor documented responses to specific questions on the Classroom Observation Record. The questions focus on:

- The School Coordinator's impression of the success of the test session
- The attitude of school staff members toward the TIMSS testing
- The shipment of testing materials from the national center
- The level of communication with the national center
- The administration of the teacher questionnaires
- The security of the testing materials before the test date
- The accommodations for testing
- The use of make-up sessions
- The training of the Test Administrators
- Feedback on the sampling procedures used to select students in the school
- Any motivation talks, special instructions, or incentives provided to students to prepare them for the assessment
- Any use of practice questions to prepare the students for the assessment
- Suggestions for improving the *School Coordinator Manuals* (1994d, 1994e).

Finally, the QC Monitor checked the Class Tracking Form for that school with the School Coordinator to ensure that the information is accurate. The QC monitor verified with the School Coordinator:

- Whether the list of mathematics classes in the grade was complete
- Whether there were any students in the grade level who were not in any of the mathematics classes on the Class Tracking Form
- Whether there were any students in the grade level who were in more than one of the mathematics classes on the Class Tracking Form.

The information collected in the Classroom Observation Record is summarized in Martin, Hoyle, and Gregory (1996b).

REFERENCES

- Martin, M.O. and Mullis, I.V.S. (1996). *Third International Mathematics and Science Study: Quality Assurance in Data Collection*. Chestnut Hill, MA: Boston College.
- Martin, M.O., Hoyle, C., and Gregory, K. (1996a). "Monitoring the TIMSS Data Collection" in M.O. Martin and I.V.S. Mullis (eds.), *Third International Mathematics and Science Study: Quality Assurance in Data Collection*. Chestnut Hill, MA: Boston College.
- Martin, M.O., Hoyle, C., and Gregory, K. (1996b). "Observing the TIMSS Testing Sessions" in M.O. Martin and I.V.S. Mullis (eds.), *Third International Mathematics and Science Study: Quality Assurance in Data Collection*. Chestnut Hill, MA: Boston College.
- Mullis, I.V.S., Kelly, D.L., and Haley, K. (1996). "Translation Verification" in M.O. Martin and I.V.S. Mullis (eds.), *Third International Mathematics and Science Study: Quality Assurance in Data Collection*. Chestnut Hill, MA: Boston College.
- Mullis, I.V.S. and Smith, T.A. (1996). "Quality Control Steps for Free-Response Coding" in M.O. Martin and I.V.S. (eds.), *Third International Mathematics and Science Study: Quality Assurance in Data Collection*. Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1994a). *International Codebooks—Populations 1 and 2* (Doc. Ref.: ICC892-893/NRC428-428). Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1994b). *Performance Assessment Administration Manual for the Main Survey* (Doc. Ref.: ICC884/NRC421). Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1994c). *Sampling Manual—Version 4* (Doc. Ref.: ICC 439/NPC117). Prepared by Pierre Foy and Andreas Schleicher. Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1994d). *School Coordinator Manual—Populations 1 and 2* (Doc. Ref.: ICC891/NRC427). Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1994e). *School Coordinator Manual—Population 3* (Doc. Ref.: ICC907/NRC440). Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1994f). *Survey Operations Manual—Populations 1 and 2* (Doc. Ref.: ICC889/NRC425). Prepared by the IEA Data Processing Center. Chestnut Hill, MA: Boston College.

- Third International Mathematics and Science Study (TIMSS). (1994g).** *Survey Operations Manual–Population 3* (Doc. Ref.: ICC 906/NRC439). Prepared by the IEA Data Processing Center. Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1994h).** *Test Administrator Manual* (Doc. Ref.: ICC890/NRC426). Prepared by the IEA Data Processing Center. Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1995a).** *Coding Guide for Free-Response Items–Populations 1 and 2* (Doc. Ref.: ICC897/NRC433). Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1995b).** *Coding Guide for Free-Response Items–Population 3* (Doc. Ref.: ICC 913/NRC446). Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1995c).** *Guide to Checking, Coding, and Entering the TIMSS Data* (Doc Ref.: ICC918/NRC449). Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1995d).** *International Codebook–Population 3* (Doc. Ref.: ICC912/NRC445). Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1995e).** *Manual for the TIMSS Quality Control Monitors* (Doc. Ref.: ICC920/NRC450). Chestnut Hill, MA: Boston College.
- Third International Mathematics and Science Study (TIMSS). (1995f).** *Supplement to the Coding Guide for Performance Assessment* (Doc. Ref.: ICC933/NRC456). Chestnut Hill, MA: Boston College.