CHAPTER 6

Survey Operations Procedures in TIMSS 2015

Ieva Johansone

Overview

As data-based indicators of countries' student achievement profiles and learning contexts, TIMSS assessments are crucially dependent on the quality of the data collected by each participant. Whereas the development of the assessments is an intensely collaborative process involving all of the partners in the enterprise, the process of administering the assessments and collecting the data is uniquely the responsibility of each individual country or benchmarking participant.

To ensure the consistency and uniformity of approach necessary for high-quality, internationally comparable data, all participants are expected to follow a set of standardized operations procedures. These procedures have been developed over successive cycles of TIMSS through a partnership involving the TIMSS & PIRLS International Study Center, the IEA Data Processing and Research Center (IEA DPC), the IEA Secretariat, Statistics Canada, and National Research Coordinators (NRCs). With each new assessment cycle, the operations procedures are updated to enhance efficiency and accuracy and reduce burden, making use of developments in information technology to automate routine activities wherever possible.

In each country or benchmarking entity, the National Research Coordinator was responsible for the implementation of TIMSS 2015. Internationally, National Research Coordinators provided the country's perspective in all international discussions, represented the country at international meetings, and were the responsible contact persons for all project activities. Locally, National Research Coordinators were responsible for implementing all the internationally agreed-upon procedures and facilitating all of the national decisions regarding TIMSS, including any adaptations for the national context.

The daily tasks of the NRCs varied over the course of the TIMSS 2015 cycle. In the initial phases, National Research Coordinators participated in the TIMSS 2015 framework and assessment development process (see Developing the TIMSS 2015 Achievement Items), and collaborated with

IEA

TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

CHAPTER 6: SURVEY OPERATIONS PROCEDURES IN TIMSS 2015
METHODS AND PROCEDURES IN TIMSS 2015   6.1

Statistics Canada and the IEA DPC to develop a plan to implement the TIMSS 2015 sampling design within the country or benchmarking entity (see Sample Implementation).

Following the development of the draft achievement items and context questionnaires, all countries conducted a full-scale field test of all instruments and operational procedures in March-April 2014 in preparation for the TIMSS 2015 data collection, which took place in October-December 2014 in Southern Hemisphere countries and in March-May 2015 in Northern Hemisphere countries. The field test allowed the National Research Coordinators and their staff to become acquainted with the operational activities, and the feedback they provided was used to improve the procedures for the data collection. As expected, the field test resulted in some enhancements to survey operations procedures and most definitely contributed to ensuring the successful execution of TIMSS 2015.

As part of ongoing efforts to improve TIMSS operations, the National Research Coordinators were asked to complete a Survey Activities Questionnaire (SAQ), which sought feedback on all aspects of their experience conducting TIMSS 2015. The feedback solicited in the SAQ included an evaluation of the quality of the assessment materials and the effectiveness of the operations procedures and documentation. The results of the TIMSS 2015 Survey Activities Questionnaire are presented in the final section of this chapter.

TIMSS 2015 Survey Operations Units, Manuals, and Software

To support the National Research Coordinators in conducting TIMSS 2015, the TIMSS & PIRLS International Study Center provided step-by-step documentation of all operational activities. Organized into a series of units, the *Survey Operations Procedures* were made available at critical junctures of the project to ensure that NRCs had all the tools and information necessary to discharge their responsibilities.

The *Procedures Units* were accompanied by a series of manuals for use by School Coordinators and Test Administrators that National Research Coordinators could translate and adapt to their local situations. Consistent with the goal of automating and streamlining procedures wherever possible, the IEA DPC provided NRCs with a range of custom-built software products to support activities, including sampling and tracking classes and students, administering school, teacher, and home questionnaires, documenting scoring reliability, and creating and checking data files. The TIMSS & PIRLS International Study Center and the IEA DPC also provided NRCs and their staff with intensive training in constructed response item scoring and data management.

The *Survey Operations Procedures* units were crucial resources for the National Research Coordinators as the units described in detail the tasks the NRCs were responsible for conducting. In the event that some of these tasks were contracted out to other people or organizations, the units ensured that the NRCs had sufficient knowledge of these matters to supervise the activities of the people who helped conduct the assessment in their countries.

The following units, manuals, and software systems were provided for administering TIMSS 2015:

- *TIMSS 2015 Survey Operations Procedures Unit 1: Sampling Schools and Obtaining their Cooperation*

- *TIMSS 2015 Survey Operations Procedures Unit 2: Preparing for and Conducting the TIMSS 2015 Field Test*. Unit 2 consisted of the following four sections: Sampling Classes and Field Test Administration, Preparing Achievement Booklets and Background Questionnaires, Scoring the Constructed Response Items, and Creating the Databases. Unit 2 was accompanied by field test versions of the School Coordinator Manual, Test Administrator Manual, National Quality Control Monitor Manual, and three software systems (WinW3S, IEA DME, and IEA OSS – described below).

- *TIMSS 2015 Survey Operations Procedures Unit 3: Contacting Schools and Sampling Classes for the Data Collection*. Unit 3 was accompanied by the School Coordinator Manual and the Windows® Within-school Sampling Software (WinW3S) and its manual. The WinW3S software enabled TIMSS 2015 participants to randomly select classes in each sampled school and document in detail the class selection process. The software also was used to track school, teacher, student, and student-teacher linkage information; prepare the survey tracking forms (described later in this chapter); and assign test instruments to students, including printing labels for all the test booklets and questionnaires.

- *TIMSS 2015 Survey Operations Procedures Unit 4: Preparing Achievement Booklets and Context Questionnaires*. Unit 4 was accompanied by the IEA Online SurveySystem (OSS) and its manual. The IEA Online SurveySystem supported the online administration of the school, teacher, and home (Early Learning Survey for TIMSS at Grade 4) questionnaires.

- *TIMSS 2015 Survey Operations Procedures Unit 5: Conducting the Data Collection*. Unit 5 was accompanied by the Test Administrator Manual, National Quality Control Monitor Manual, and the International Quality Control Monitor Manual.

- *TIMSS 2015 Survey Operations Procedures Unit 6: Scoring the Constructed Response Items*. Unit 6 was accompanied by the TIMSS 2015 Scoring Guides, the IEA Coding Expert Software, the Trend Reliability Scoring Manual, and the Cross-country Reliability Scoring Manual. The IEA Coding Expert Software was used to facilitate the trend and cross-country reliability scoring tasks.

- *TIMSS 2015 Survey Operations Procedures Unit 7: Creating the Databases*. Unit 7 was accompanied by the IEA Data Management Expert (DME) software, its manual, and codebooks that specified information on the IEA DME data fields in each of the data files. The IEA DME software is used for data entry and data verification.

TIMSS 2015 Survey Tracking Forms

TIMSS uses a series of tracking forms to document class sampling procedures, assign assessment instruments, and track school, teacher, and student information, including the participation status of the respondents. The tracking forms also facilitate the data collection and data verification process. Five different tracking forms were used for TIMSS 2015:

- Class Listing Form: This form was completed for each sampled school, listing the eligible classes and providing details about the classes, such as the class stream (if applicable), the number of students, and the names of teachers.

- Student-Teacher Linkage Form: This form was completed for each class sampled, listing the names of the students and their teachers, student birth dates, gender, exclusion codes, and linking the students to their teachers.

- Student Listing Form (participants in TIMSS Numeracy only): This form was completed for each class sampled, listing the names of the students, student birth dates, gender, and exclusion codes.

- Student Tracking Form: This form was created for each class assessed and was completed by the Test Administrators during test administration. The Test Administrators used this form to verify the assignment of survey instruments to students and to indicate student participation.

- Teacher Tracking Form: This form was completed for each sampled school to indicate the completion of the Teacher Questionnaires.

Operations for Data Collection

The following sections describe the major operational activities coordinated by the National Research Coordinators.

- Contacting schools and sampling classes

- Overseeing translation and preparing assessment instruments

- Managing the administration of the TIMSS 2015 assessments

- Scoring of the constructed response items

- Creating the TIMSS 2015 data files

Three other major TIMSS 2015 operational activities—sampling schools, translation and translation verification of the assessment instruments, and layout verification of the assessment instruments—are described in separate sections of the *Methods and Procedures in TIMSS 2015* publication (see the Sample Design, Translation and Translation Verification, and Layout Verification chapters).

IEA

TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

CHAPTER 6: SURVEY OPERATIONS PROCEDURES IN TIMSS 2015
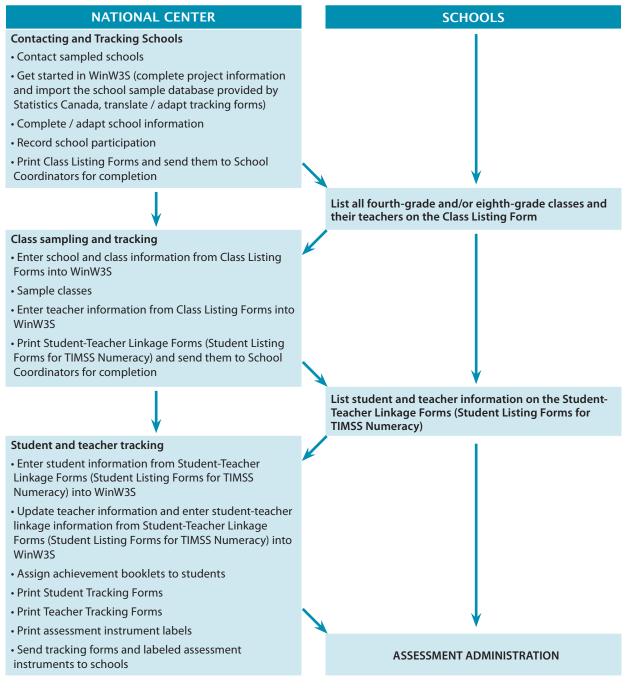METHODS AND PROCEDURES IN TIMSS 2015   6.4

Contacting Schools and Sampling Classes

Exhibit 6.1 illustrates the major steps of working with schools to sample classes and prepare for the TIMSS 2015 assessment administration. Once the school samples were drawn, National Research Coordinators were tasked with contacting schools and encouraging them to take part in the assessment. Depending on the national context, this could involve obtaining support from national or regional educational authorities. *Survey Operations Procedures Unit 1* outlines suggestions on ways to encourage schools to participate in the assessment.

Exhibit 6.1: **Diagram of the Sampling Procedures and Preparations for the Assessment Administration Implemented by National Centers and Schools**

NATIONAL CENTER	SCHOOLS
Contacting and Tracking Schools • Contact sampled schools • Get started in WinW3S (complete project information and import the school sample database provided by Statistics Canada, translate / adapt tracking forms) • Complete / adapt school information • Record school participation • Print Class Listing Forms and send them to School Coordinators for completion	
	List all fourth-grade and/or eighth-grade classes and their teachers on the Class Listing Form
Class sampling and tracking • Enter school and class information from Class Listing Forms into WinW3S • Sample classes • Enter teacher information from Class Listing Forms into WinW3S • Print Student-Teacher Linkage Forms (Student Listing Forms for TIMSS Numeracy) and send them to School Coordinators for completion	
	List student and teacher information on the Student-Teacher Linkage Forms (Student Listing Forms for TIMSS Numeracy)
Student and teacher tracking • Enter student information from Student-Teacher Linkage Forms (Student Listing Forms for TIMSS Numeracy) into WinW3S • Update teacher information and enter student-teacher linkage information from Student-Teacher Linkage Forms (Student Listing Forms for TIMSS Numeracy) into WinW3S • Assign achievement booklets to students • Print Student Tracking Forms • Print Teacher Tracking Forms • Print assessment instrument labels • Send tracking forms and labeled assessment instruments to schools	
	ASSESSMENT ADMINISTRATION

In cooperation with school principals, National Research Coordinators were responsible for identifying and training School Coordinators for all participating schools. A School Coordinator could be a teacher or guidance counselor in the school, or NRCs could appoint a member of the national center to fill this role. In some countries, a School Coordinator from the national center was responsible for several schools in an area. Each School Coordinator was provided with a School Coordinator Manual, which describes their responsibilities. The School Coordinator Manual was prepared by the TIMSS & PIRLS International Study Center and translated/adapted by National Research Coordinator staff, as necessary.

The responsibilities of the School Coordinator included providing the national center with information on the school; coordinating the date, time, and place for testing; identifying and training a Test Administrator to administer the assessment; coordinating the completion of the TIMSS 2015 tracking forms; distributing questionnaires; and obtaining parental permission (if necessary). School Coordinators also confirmed receipt of all assessment materials, oversaw the security of the assessment materials, and ensured the return of the assessment materials to the national center following the administration of the assessment.

School Coordinators also played a critical role in providing information for the sampling process, providing the national center with data on eligible classes in the school. With this information, the national centers used the Within-school Sampling Software (WinW3S) to sample class(es) within the school. WinW3S tracked school, teacher, and student information, and the software generated the necessary tracking forms and instrument labels facilitating the assessment administration process as well as data checking during the data cleaning process.

As TIMSS samples intact classes, one of the roles of the School Coordinator was to ensure that every student in the school was listed in one and only one class (course). This was necessary to ensure that the sample of classes results in a representative sample of students, and every student at the target grade has a chance of being selected. At fourth grade in most countries, students are taught mathematics and science in the same classroom, and therefore the fourth grade classroom was designated as the sampling unit. At the eighth grade, however, in many countries students are grouped differently for mathematics and science instruction. In other words, a student may take mathematics with one group of students and science with a different group of students. As the sampling required one set of students who could be considered a classroom, eighth grade classrooms usually were defined on the basis of mathematics instruction for the purposes of sampling.

Overseeing Translation and Preparing Assessment Instruments

National Research Coordinators also were responsible for preparing the assessment instruments (achievement booklets and context questionnaires) for their countries—a process that included overseeing the translation of the assessment instruments. The overarching goal of assessment instrument preparation is to create internationally comparable achievement booklets and context questionnaires that are appropriately adapted for the national context.

Each student was assigned one of 14 TIMSS achievement booklets (see the *TIMSS 2015 Assessment Frameworks* for more information on the matrix sampling design). The achievement booklets are composed of blocks of assessment items, with each block appearing in two booklets. From an operational perspective, each block needed to be translated only once, even though it was included in two different booklets. Adobe®InDesign® software is used by countries to link the translated and adapted assessment blocks to the appropriate booklets. Automating this process through Adobe®InDesign® decreased the chances of human error in the production process.

Twelve new assessment blocks at each grade level were developed for TIMSS 2015 (six mathematics and six science). The new assessment blocks replaced the ones released at the end of the previous assessment cycle. Also, eight new mathematics assessment blocks were developed for TIMSS Numeracy 2015. The new assessment items were tried out through the field test in order to investigate the psychometric characteristics of the achievement items and make well-informed decisions about the best items. Similarly, the context questionnaires were evaluated following the field test to gauge the validity and reliability of the various questionnaire scales.

TIMSS field tests around twice the number of items needed to fill the new assessment blocks. All participating countries and benchmarking entities translated and/or adapted the newly developed items into the test administration language(s) and did the same for the questionnaires. After the field test, the best assessment items were chosen for the main data collection and some edits were applied to both items and the questionnaires.

National Research Coordinators were responsible for applying these changes to the translated assessment items and questionnaires. Countries taking TIMSS at the fourth and/or eighth grade that did not participate in TIMSS 2011 or TIMSS 2007 had to translate and/or adapt the assessment blocks used in previous assessments (trend blocks) into their language(s) in preparation for the 2015 assessment administration. Countries that had participated in TIMSS 2007 and/or TIMSS 2011 were required to use the same translations they used in those cycles.

For both the field test and main data collection, the participating countries received the international version (English) of the achievement booklets and context questionnaires with all the necessary instrument production files, including fonts and graphics files. Instructions on how to use the materials to produce high-quality, standardized instruments, were included in the corresponding *Survey Operations Procedures* unit. The IEA Secretariat and the TIMSS & PIRLS International Study Center also provided a generic Arabic source version of the TIMSS 2015 assessment booklets and context questionnaires. Individual countries adapted the generic source version to local usage.

Once translated and/or adapted, first for the field test and then again for the main data collection, the achievement items and context questionnaires were submitted to the IEA Secretariat for translation verification. The IEA Secretariat worked with independent translators to evaluate each country's translations and when deemed necessary suggested changes to the text.

After the translations and adaptations had been verified by the IEA Secretariat, National Research Coordinators assembled the achievement booklets and context questionnaires using Adobe® InDesign® software, and print-ready copies of the instruments were sent to the TIMSS & PIRLS International Study Center for layout verification and a final review of national adaptations. This review checked that each booklet and questionnaire conformed to the international format and that any adaptations made to the instruments did not unduly influence their international comparability.

National Adaptations Forms (NAFs)

While preparing national achievement booklets and context questionnaires, countries sometimes by necessity made adaptations to the international versions. All national adaptations to the international assessment instruments, other than direct translation, were documented using the National Adaptations Forms. There is a separate set of NAFs for the achievement booklets and for the context questionnaires (per grade/assessment). During the translation verification and layout review, the verifiers checked whether the national adaptations were likely to influence the ability to produce internationally comparable data for the items involved. Any questions raised were directed to the NRC for consideration via the NAFs.

The NAFs were completed and reviewed at various stages of preparing national assessment instruments. Version I of the forms was completed during the internal translation and review process and sent along with the rest of the materials for international translation verification. After translation verification, the forms (Version II) were updated in response to the translation verifier's comments and reflecting any changes resulting from the verification, and sent along with the national assessment instruments for layout verification. Following layout verification, the national instruments and NAFs were finalized (Version III) and submitted to the IEA Secretariat, the TIMSS & PIRLS International Study Center, and the IEA DPC as the final documentation of the national adaptations.

Managing the Administration of the TIMSS 2015 Assessments

Printing assessment materials and distributing them to the participating schools required careful organization and planning on the part of the National Research Coordinator. Each student was assigned one of 14 achievement booklets according to a systematic distribution plan implemented by the WinW3S sampling software. This process is facilitated by the tracking forms and labels generated by WinW3S.

Each student also was assigned a Student Questionnaire, which was labeled so that it could be linked to the achievement booklet. For TIMSS at the fourth grade and for TIMSS Numeracy, the student's parents were assigned the Early Learning Survey, which also was linked to the achievement booklet. In addition, an individually labeled Teacher Questionnaire was sent to each teacher listed on the Teacher Tracking Form and a School Questionnaire was sent to the principal. These materials were packaged and sent to the School Coordinators prior to the testing date, giving

ample time for the School Coordinators to confirm the receipt and correctness of the materials. The School Questionnaire and Teacher Questionnaires were then distributed, while the other instruments were kept in a secure room until the testing date.

Each sampled class was assigned a Test Administrator who followed procedures described in the Test Administrator Manual to administer the achievement booklets and Student Questionnaire. This person was chosen and trained by the School Coordinator. In many cases, the School Coordinator doubled as the Test Administrator. The Test Administrator was responsible for distributing materials to the appropriate students, reading to the students the instructions provided in the Test Administrator's manual, and timing the sessions.

The Test Administrator documented the timing of the testing sessions on the Test Administration Form. The Test Administration Form also solicited information about anything out of the ordinary that took place during assessment administration.

The achievement booklets contained two sections, and the time allotted for each section of the assessment was standardized and strictly enforced by the Test Administrator. There was a required break in between the two sections of the assessment administration, and this break was not to exceed 30 minutes. To complete each part of the TIMSS achievement test, fourth grade students were allowed 36 minutes and eighth grade students were allowed 45 minutes. If a student completed part 1 or part 2 of the assessment before the allotted time, the student was not allowed to leave the testing room. Students completing the assessments early were asked to review their answers or read quietly, and some test administrators provided an activities sheet for the student.

To complete the Student Questionnaire, students were given at least 30 minutes, but extra time was given when necessary. Also, for fourth grade students, the Test Administrator was permitted to read the questionnaire items aloud together with the students.

The Test Administrator was required to use the Student Tracking Form and labels to distribute the booklets to the correct students and to document student participation. If the participation rate was below 90 percent in any class, it was the School Coordinator's responsibility to hold a makeup session for the absent students before returning all of the testing materials to the national center.

Linking Students to their Teachers and Classes

Exhibit 6.2 illustrates the hierarchical identification system codes that are used to link the data among schools, classes, students, and teachers. The school, class, and student IDs are strictly hierarchical, with classes nested within schools and students nested within classes.

Exhibit 6.2: Hierarchical Identification System Codes Used to Link Schools, Classes, Students, and Teachers

Participant	ID Components	ID Structure	Numeric Example
School	School	CCCC	0001
Class	School + Class within the school	CCCCKK	000101 000102
Student	School + Class within the school + Student within the class	CCCCKKSS	00010101 00010201
Teacher	School + Teacher within the school + Linkage number to the sampled class	CCCCTTLL	00010101 00010201

Each teacher is assigned a teacher identification number consisting of the four-digit school number followed by a two-digit teacher number. Since a teacher could be teaching mathematics and/or science to some or all of the students in a class, it is necessary to have a unique identification number for each teacher linked to a class and to certain students within the class. This is achieved by adding a two-digit link number to the six digits of the teacher identification number to create a unique eight-digit identification number.

Online Administration of the School, Teacher, and Home Questionnaires

Countries could choose to administer the school, teacher, and/or home questionnaires online. The benefits of administering the questionnaires online included saving money and time in printing, and improving the efficiency of questionnaire distribution, data entry, and data cleaning.

For the online administration of the questionnaires, the IEA DPC provided its IEA Online SurveySystem software that incorporates design, presentation, and monitoring components.

The design component, known as the Designer, supports the preparation of the online surveys, data management, and data output to the IEA DPC. Through the IEA Online SurveySystem Designer component, national centers could tailor the online questionnaires to their national language. To facilitate translation and adaptation, the Designer concurrently stored the original English question text and the translations and/or national adaptations. It also stored the variable names and data validation rules. If a national center decided not to administer a particular international question or option, it could be disabled in the Designer and would not be administered during the online questionnaire administration. The Designer also included an integrated preview function to allow for a visual side-by-side comparison of the paper/PDF and online versions of the questionnaires, facilitating the layout verification process.

For the online presentation, the Web Component presents the questionnaires to the respondents. The navigation capabilities of the Web Component are designed to allow respondents to pick and choose their order of response. Buttons marked "next" and "previous" facilitated navigation between adjacent pages, so users could browse through the questionnaire in the same way that they flip through the pages of the paper questionnaire. A hyperlinked interactive "table

of contents" allowed the respondents to fluidly navigate to specific questions. Overall, these two functions permitted the respondents to answer questions in the order of their choosing, and skip questions just as they could do if they were answering the paper questionnaire. Also, the online questionnaires could be accessed through any standard Internet browser on all standard operating systems without the user needing any additional software.

Finally, the Web-based Monitor component allows for monitoring the survey responses in real time. Many national centers made extensive use of the Web-based Monitor to follow-up with non-respondents.

The IEA Data Processing and Research Center followed a stringent set of procedures in order to safeguard the confidentiality of the respondents and maintain the integrity of the data. Each respondent received a statement of confidentiality, and information on how to access the online questionnaire. For most countries, the online questionnaire administration was hosted on the IEA DPC's customized high-performance server. The IEA DPC server allowed for the 24-hour availability of the questionnaires during the data-collection period, and it also ensured backup and recovery provisions for the data.

Scoring the Constructed Response Items

Constructed response items represent a substantial portion of the TIMSS assessments, and because reliable and valid scoring of these items is critical to the assessment results, the TIMSS & PIRLS International Study Center provided explicit scoring guides and extensive training in their use. Also, the *Survey Operations Procedures* units specified a procedure for efficiently organizing and implementing the scoring activity.

International scoring training sessions (one for the field test and two for the main data collection—one for Southern Hemisphere countries and another for Northern Hemisphere countries) were conducted where all National Research Coordinators (or country representatives appointed by the NRCs) were trained to score each of the constructed response items. At these training sessions, the scoring guide for each item was reviewed and applied to a sample set of example student responses that had already been scored. These example papers were actual student answers from pilot testing in several English-speaking countries and were chosen to represent a range of response types and to demonstrate the guides as clearly as possible. Following the example papers, the training participants applied the scoring guides to a different set of student responses that had not yet been scored. The scores to these practice papers were then shared with the group and any discrepancies were discussed.

Following the international scoring training, national centers trained their scoring staff on how to apply the scoring guides for the constructed response items. National Research Coordinators were encouraged to create additional example papers and practice papers from student responses collected in their country.

Documenting Scoring Reliability

Because reliable scoring of the constructed response items is essential for high quality TIMSS data, it is important to document the reliability of the scoring process. A high degree of scorer agreement is evidence that scorers have applied the scoring guides in the same way. The procedure for scoring the TIMSS 2015 constructed response items provided for documenting scoring reliability within each country (within-country reliability scoring), across countries (cross-country reliability scoring), and over time (trend reliability scoring).

The method for establishing the reliability of the scoring within each country was for two independent scorers to score a random sample of 200 responses for each constructed response item. The degree of agreement between the scores assigned by the two scorers is a measure of the reliability of the scoring process. In collecting the within-country reliability data, it was vital that the scorers independently scored the items assigned to them, and each scorer did not have prior knowledge of the scores assigned by the other scorer. The within-country reliability scoring was integrated within the main scoring procedure and ongoing throughout the scoring process.

The purpose of the trend reliability scoring was to measure the reliability of the scoring from one assessment cycle to the next (i.e., from TIMSS 2011 to TIMSS 2015). The trend reliability scoring required scorers of the current assessment to score student responses collected in the previous cycle. The scores of the current cycle were then compared with the scores awarded in the previous assessment cycle. Trend reliability scoring was conducted using the IEA Coding Expert Software provided by the IEA DPC.

Trend reliability scoring for TIMSS 2015 involved eight secured item blocks. Student responses included in the trend reliability scoring (150-200 responses per item) were actual student responses collected during the previous assessment cycle in each country and benchmarking entity. These responses were scanned and provided for each participating country and benchmarking entity along with the IEA Coding Expert Software. All scorers who scored the trend assessment blocks in 2015 were required to participate in the trend reliability scoring. If all scorers were trained to score all trend items, the software divided the student responses equally among the scorers. If scorers were trained to score specific item blocks, National Research Coordinators were able to specify within the software which scorers would score particular item blocks, and the software allocated the student responses accordingly. Similar to the within-country reliability scoring, the trend reliability scoring had to be integrated within the main scoring procedure.

Finally, cross-country reliability scoring gave an indication about how consistently the scoring guides were applied from one country to the next. The cross-country reliability scoring also was conducted using IEA Coding Expert Software. To begin the process, the IEA DPC compiled actual responses of students from English speaking countries participating in the previous TIMSS cycle. Cross-country reliability scoring included 21 items at the fourth grade and 26 items at the eighth grade. Two hundred student responses for each item were scanned by the IEA DPC and provided to

countries and benchmarking entities along with the IEA Coding Expert Software. All scorers who could score student responses written in English were required to participate in the cross-country reliability scoring, and the student responses were equally divided among the participating scorers in each country. The scoring exercise was completed immediately after all other scoring activities.

Creating the TIMSS 2015 Databases

The data entry process took place March-May 2014 for the field test, from December 2014–March 2015 following data collection in the Southern Hemisphere and June–September 2015 following data collection in the Northern Hemisphere. The procedure for creating the TIMSS 2015 databases included entering sampling and assessment administration information into the WinW3S database and adding responses from the context questionnaires and achievement booklets using the IEA Data Management Expert (DME) software.

The IEA DPC provided DME software to accommodate keyboard data entry and data verification. The DME software also offers data and file management capabilities, a convenient checking and editing mechanism, interactive error detection, and quality-control procedures. For the TIMSS 2015 context questionnaires administered online on the IEA DPC's server, the data were directly accessible by the IEA DPC and no further data entry was required.

Along with the DME software, the IEA DPC provided international codebooks describing all variables and their characteristics, thus ensuring that the data files met the internationally defined rules and standards for data entry. The files within the DME database for entering the TIMSS 2015 data were based on these codebooks. However, the codebooks had to match exactly the national assessment instruments so that the answers of the respondents could be entered properly. Therefore, any adaptations to the international instruments also required adaptations to the international codebooks. The adapted national codebooks then were used to create the TIMSS 2015 data files in each country, with the responses to the context questionnaires, achievement booklets, and Reliability Scoring Sheets keyed into the DME database.

Quality control throughout the data entry process was essential to maintain accurate data. Therefore, National Research Coordinators were responsible for performing periodic reliability checks during data entry and for applying a series of data verification checks provided by both WinW3S and DME software prior to submitting the databases to the IEA DPC. To ensure the reliability of the data entry process, the data-entry staff was required to double enter at least 5 percent of each instrument type. An error rate of 1 percent or less was acceptable for the background files. An error rate of 0.1 percent or less was required for the student achievement files and the reliability scoring files. If the required agreement was not reached, retraining of the key punchers was required.

Additionally, the data verification module of WinW3S and DME identified a range of problems, such as inconsistencies of identification codes and out-of-range or otherwise invalid

TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

IEA

codes. The data quality control procedures also verified the integrity of the linkage between the students, teachers, and schools entered into the DME database and tracking of information for those specified in WinW3S.

When all data files had passed the quality control checks, they were submitted to the IEA DPC, along with data documentation, for further checking and processing. For information on data processing at the IEA DPC, please refer to the [Creating the International Databases](#) chapter of this publication.

TIMSS 2015 Survey Activities Questionnaire

The Survey Activities Questionnaire was designed to elicit information about NRCs' experiences in preparing for and conducting the TIMSS 2015 data collection. The questionnaire was composed of six sections and focused on the following:

- Sampling schools and classes
- Preparing assessment instruments
- Administering the assessments
- Implementing the National Quality Control Program
- Preparing for and scoring the constructed response items
- Creating the databases

All items in the Survey Activities Questionnaire included accompanying comment fields, in which NRC respondents were encouraged to explain their responses, provide additional information, and suggest improvements in the process.

The *TIMSS 2015 Survey Activities Questionnaire* was administered online via the IEA's Online SurveySystem and was completed by a total of 59 NRCs. The following sections summarize information gathered from the Survey Activities Questionnaire, reflecting the quality of the TIMSS 2015 survey materials and procedures in the participating countries.

Sampling Schools and Classes

The first section of the Survey Activities Questionnaire asked NRCs about the *Survey Operations Procedures* for sampling both schools and classes within the sampled schools. As shown in Exhibit 6.3, all of the countries considered that *Survey Operations Procedures* Units 1 and 3 were clear and sufficient. Seven countries reported deviating from the basic TIMSS sampling design. Their reasons for these modifications to the sampling procedures included allowing for census participation, oversampling certain regions, and changing the target grade from previous cycles. One country reported selecting their TIMSS 2015 school sample at the national center in collaboration with Statistics Canada. Statistics Canada in cooperation with the IEA DPC selected the school samples for all other countries.

Exhibit 6.3: Survey Activities Questionnaire, Section One—Sampling (Numbers of NRC Responses)

Question	Yes	No	Not Answered
Was the information provided in the "TIMSS 2015 Survey Operations Procedures Unit 1 – Sampling Schools and Obtaining their Cooperation" clear and sufficient?	58	0	1
Were there any conditions or organizational constraints that necessitated deviations from the basic TIMSS sampling design described in the "Survey Operations Procedures Unit 1"?	7	51	1
Did you use the Within-school Sampling Software (WinW3S) to sample classes?	57	0	2
Did you experience any problems or inconveniences when using the WinW3S software?	16	40	2
Was the information provided in the "TIMSS 2015 Survey Operations Procedures Unit 3 – Contacting Schools and Sampling Classes for the Data Collection" clear and sufficient?	58	0	1
Did you follow the procedures outlined in "Survey Operations Procedures Unit 3" for working with the schools to sample classes (e.g., using the appropriate tracking forms in the proposed order to obtain information from School Coordinators)?	47	10	2

All countries selected classes within the sampled schools using the Windows® Within-school Sampling Software (WinW3S), provided by the IEA Data Processing and Research Center. Countries administering both the TIMSS fourth grade and TIMSS Numeracy achievement booklets encountered some organizational constraints in their systems that necessitated a modification to the sample design, and these countries also experienced some problems using the WinW3S software. Countries also noted that the WinW3S software was slow at times.

Ten NRCs applied some modifications to the procedures outlined in the *Survey Operations Procedures Unit 3*. For example, some NRCs did not use the Class Listing Forms because all classes at the target grade were tested or because a class level database was available at the ministry, and a number of countries did not use the fourth grade Teacher Tracking Forms because there was only one teacher per class. All modifications were reviewed and approved by the TIMSS & PIRLS International Study Center.

Translating, Adapting, and Producing Assessment Instruments

The second section of the Survey Activities Questionnaire asked NRCs about translating, adapting, assembling, and printing the test materials, as well as issues related to checking the materials and securely storing them. In the majority of cases, NRCs reported applying corrections to their survey instruments as suggested by the external translation verifier or the layout verifier.

As reported in Exhibit 6.4, all of the NRCs answered that they were able to assemble the test booklets and questionnaires according to the instructions provided. However, 10 countries reported

experiencing some problems using the survey instrument production materials. These problems mostly included the following: issues with fonts and special characters (e.g., for Cyrillic alphabet), difficulty fitting longer national text in the context questionnaires, and some problems with the layout style of tables. All of the identified problems were resolved either by specialists at the national center or with assistance from the TIMSS & PIRLS International Study Center.

Exhibit 6.4: Survey Activities Questionnaire, Section Two—Translating, Adapting, and Producing Assessment Instruments (Numbers of NRC Responses)

Question	Yes	No	Not Answered
Was the information provided in the "TIMSS 2015 Survey Operations Procedures Unit 4 – Preparing Achievement Booklets and Context Questionnaires" clear and sufficient?	55	3	1
Did you encounter any major problems using the assessment instrument production materials (e.g., instrument production files, fonts, support materials) provided by the TIMSS & PIRLS International Study Center?	10	48	1
After the translation verification, did you correct your translations/adaptations as suggested by the verifier in the majority of cases?			
TIMSS eighth-grade booklets	38	0	4 *(Not Answered)* 17 *(Not Applicable)*
TIMSS fourth-grade booklets	46	1	4 *(Not Answered)* 8 *(Not Applicable)*
TIMSS Numeracy booklets	8	0	4 *(Not Answered)* 47 *(Not Applicable)*
Eighth-grade context questionnaires	38	0	4 *(Not Answered)* 17 *(Not Applicable)*
Fourth-grade context questionnaires	49	0	4 *(Not Answered)* 6 *(Not Applicable)*
After the layout verification, did you correct your assessment instruments as noted by the verifier in the majority of cases?			
TIMSS eighth-grade booklets	39	0	3 *(Not Answered)* 17 *(Not Applicable)*
TIMSS fourth-grade booklets	47	1	3 *(Not Answered)* 8 *(Not Applicable)*
TIMSS Numeracy booklets	8	0	3 *(Not Answered)* 48 *(Not Applicable)*
Eighth-grade context questionnaires	39	0	3 *(Not Answered)* 17 *(Not Applicable)*
Fourth-grade context questionnaires	50	0	3 *(Not Answered)* 6 *(Not Applicable)*
Did you apply any quality control measures to check the achievement booklets and context questionnaires during the printing process (e.g., checking for missing pages, upside down pages, text too bright or too dark)?	54	2	3

Exhibit 6.4: Survey Activities Questionnaire, Section Two—Translating, Adapting, and Producing Assessment Instruments (Numbers of NRC Responses) (Continued)

Question	Yes	No	Not Answered
Did you take measures to protect the security of the assessment instruments during the translation, assembly, and printing process?	57	0	2
Did you detect any potential breaches in security of the assessment instruments?	1	56	2
Did you encounter any problems preparing the Online SurveySystem files for administering the school, teacher, and/or home (Early Learning Survey) questionnaires online?	5	14	3 *(Not Answered)* 37 *(Not Applicable)*

Nearly all of the countries conducted the recommended quality control checks during the process of printing the testing materials. The most common errors that countries detected and fixed during the printing process were pages that were missing or in the wrong order. One country expressed concerns about a breach of security, as the courier lost one package with the materials.

Five countries reported that they experienced problems with the Online SurveySystem. These problems were related to structural national adaptations, the national text being much longer than the original text in English, a very tight timeline for Southern Hemisphere countries, and some valid ranges (e.g., calendar dates) not being restricted.

Assessment Administration

The third section of the Survey Activities Questionnaire addressed the extent to which NRCs detected errors in the testing materials during packaging for shipment to schools. As shown in Exhibit 6.5, a small number of errors were found in the materials. About half of such errors were discovered before distributing materials to schools and fixed prior to their distribution. Errors found after distribution usually were very minor, and either were fixed by school coordinators or replacement materials were provided. The few cases where the errors could not be remedied were reported to the TIMSS & PIRLS International Study Center, where decisions were made about setting the problematic data to "Not Administered."

Exhibit 6.5: Survey Activities Questionnaire, Section Three—Assessment Administration (Numbers of NRC Responses)

Question	Yes	No	Not Answered
Was the information provided in the "TIMSS 2015 Survey Operations Procedures Unit 5 – Conducting the Data Collection" clear and sufficient?	57	0	2
Were any errors detected in any of the following assessment materials after they were sent to schools?			
Achievement booklets	18	38	3 (Not Answered) 0 (Not Applicable)
Achievement booklet ID labels	7	49	3 (Not Answered) 0 (Not Applicable)
Student Questionnaires	5	51	3 (Not Answered) 0 (Not Applicable)
Student Questionnaire ID labels	5	50	3 (Not Answered) 1 (Not Applicable)
Early Learning Surveys	1	48	3 (Not Answered) 7 (Not Applicable)
Early Learning Survey ID labels	4	45	3 (Not Answered) 7 (Not Applicable)
Student Tracking Forms	6	50	3 (Not Answered) 0 (Not Applicable)
Teacher Questionnaires	5	49	3 (Not Answered) 2 (Not Applicable)
Teacher Tracking Forms	2	51	3 (Not Answered) 3 (Not Applicable)
School Questionnaires	2	52	3 (Not Answered) 2 (Not Applicable)
School Coordinator Manuals	1	54	3 (Not Answered) 1 (Not Applicable)
Test Administrator Manuals	3	52	3 (Not Answered) 1 (Not Applicable)
If any errors were detected, did you correct the error(s) before the testing began?	15	24	2 (Not Answered) 18 (Not Applicable)
Does your country have a confidentiality policy that restricts putting student names on tracking forms and survey instrument covers?	16	41	2
Did you encounter any problems translating and/or adapting the School Coordinator Manual(s)?	2	55	2
Did you encounter any problems translating and/or adapting the Test Administrator Manual(s)?	1	55	3
Were School Coordinators appointed from within the participating schools?	50	7	2
Did you hold formal training session(s) for School Coordinators?	33	24	2

TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

IEA

CHAPTER 6: SURVEY OPERATIONS PROCEDURES IN TIMSS 2015
METHODS AND PROCEDURES IN TIMSS 2015   **6.18**

Exhibit 6.5: Survey Activities Questionnaire, Section Three—Assessment Administration (Numbers of NRC Responses) (Continued)

Question	Yes	No	Not Answered
Were Test Administrators trained by School Coordinators within the participating schools?	25	32	2
Did Test Administrators document any problems or special circumstances that occurred frequently during the assessment administration (please refer to the completed Test Administration Forms)?	20	36	3
If you administered school, teacher, and/or home (Early Learning Survey) questionnaires online, did any of the respondents in your country encounter any problems responding to the online questionnaires?	9	11	2 *(Not Answered)* 37 *(Not Applicable)*

Three NRCs reported difficulties translating the School Coordinator Manual and/or the Test Administrator Manual. Primarily, problems arose when the manual(s) had to be reorganized or adapted, and the standardized procedures were modified (e.g., no Class Listing Forms or Teacher Tracking Forms were used).

In 50 countries, School Coordinators were appointed within the participating schools, and in 25 of these countries, Test Administrators were trained by the School Coordinators. In the remaining countries, School Coordinators and/or Test Administrators either were from the national center or were contracted externally. In most countries, the National Research Coordinators organized training sessions for School Coordinators. In some, mostly larger countries, training was conducted either online or in a written form via extended manuals.

Among the problems documented by Test Administrators during assessment administration were the following: loud noises outside the classroom, many students asking questions, too much time, not enough time, and student complaints that the test was too difficult.

National Quality Control Program

The fourth section of the Survey Activities Questionnaire addressed the national quality control program that each country implemented during data collection. As part of the national quality assurance activities, NRCs were instructed to send National Quality Control Observers to 10 percent of the participating schools in order to observe test administration and document compliance with prescribed procedures. Due primarily to budgetary constraints, some countries sent national monitors to less than ten percent of participating schools, and three countries did not send monitors to any of the testing sessions.

As shown in Exhibit 6.6, when applicable, almost all of the national centers conducted their quality assurance program using the National Quality Control Monitor Manual provided by the TIMSS & PIRLS International Study Center. Among the few documented problems detected by the national monitors were students complaining about the length of the Student Questionnaire. In addition, one case was noted where the national monitor felt the Test Administrator was unprepared.

Exhibit 6.6: Survey Activities Questionnaire, Section Four—National Quality Control Program (Numbers of NRC Responses)

Question	Yes	No	Not Answered
Did you conduct a national quality control program that observed the data collection in the participating schools?	52	5	2
Did you use the National Quality Control Monitor (NQCM) Manual and the Classroom Observation Record provided by the TIMSS & PIRLS International Study Center to conduct your national quality control program?	49	4	3 *(Not Answered)* 3 *(Not Applicable)*
Did your national quality control monitors (NQCMs) document any major problems or special circumstances that occurred frequently during the assessment administration?	3	49	4 *(Not Answered)* 3 *(Not Applicable)*

Preparing for and Scoring the Constructed Response Items

Exhibit 6.7 provides data on responses to items asking NRCs about their experiences preparing for and scoring the constructed response items. Almost all NRCs found the scoring procedures as explained in the *Survey Operations Procedures Unit 6—Scoring the Constructed Response Items* to be clear and sufficient. Some countries reported that they would have liked to have scoring training practice materials for all items instead of select group of items. Countries reporting problems with the scoring training materials asked for more "borderline" examples including more detailed explanations within the scoring guides. All of the NRCs reported creating their own national examples and practice papers for training their scorers, as suggested by the TIMSS & PIRLS International Study Center.

Exhibit 6.7: Survey Activities Questionnaire, Section Five—Preparing for and Scoring the Constructed Response Items (Numbers of NRC Responses)

Question	Yes	No	Not Answered
Was the information provided in the "TIMSS 2015 Survey Operations Procedures Unit 6 – Scoring the Constructed Response Items" clear and sufficient?	55	2	2
Did you encounter any problems using the scoring training materials, provided by the TIMSS & PIRLS International Study Center?	11	46	2
Did you create national scoring training materials in addition to the international scoring training materials?	41	16	2
Did you scan the achievement booklets for electronic image scoring?			
TIMSS eighth-grade booklets	11	26	2 *(Not Answered)* 20 *(Not Applicable)*
TIMSS fourth-grade booklets	16	30	2 *(Not Answered)* 11 *(Not Applicable)*
TIMSS Numeracy booklets	0	9	2 *(Not Answered)* 48 *(Not Applicable)*

Exhibit 6.7: Survey Activities Questionnaire, Section Five—Preparing for and Scoring the Constructed Response Items (Numbers of NRC Responses) (Continued)

Question	Yes	No	Not Answered
Did you encounter any problems during the Trend Reliability Scoring?			
Procedural problems	1	51	2 *(Not Answered)* 5 *(Not Applicable)*
Technical, software related problems	10	42	2 *(Not Answered)* 5 *(Not Applicable)*
Did all your scorers participate in scoring student responses of the trend items?	30	21	3 *(Not Answered)* 5 *(Not Applicable)*
Did you encounter any problems during the Cross-country Reliability Scoring?			
Procedural problems	5	47	4 *(Not Answered)* 3 *(Not Applicable)*
Technical, software related problems	9	43	4 *(Not Answered)* 3 *(Not Applicable)*
Did all your scorers participate in the Cross-country Reliability Scoring?	17	36	3 *(Not Answered)* 3 *(Not Applicable)*

Eleven countries administering the eighth-grade assessment and 16 countries administering the fourth-grade assessment scanned their achievement booklets and scored student responses electronically. Some technical problems were encountered while using the Coding Expert Software and this feedback will be used by the IEA DPC to continue to improve the software. Because English was used for the cross-country reliability scoring task, three countries were unable to participate. For those countries that did not participate in the previous cycle of TIMSS, the question on the trend reliability scoring procedures did not apply.

Creating the Databases

The last section of the Survey Activities Questionnaire addressed data entry and quality control activities. As shown in Exhibit 6.8, all of the NRCs found the instructions in the *Survey Operations Procedures Unit 7* to be clear and sufficient. Some NRCs expressed a wish for a more automated data entry process in WinW3S (especially for entering the testing dates and time) and would like to have more detailed instructions on importing tables with information for multiple schools, teachers, and/or students. Most countries reported hiring temporary data entry staff to enter data manually, and a number of countries used optical scanning instead of manual data entry. A very positive finding of the TIMSS 2015 Survey Activities Questionnaire is that multiple countries reported exceeding the five percent requirement for double entry of each assessment instrument, with a couple of countries entering all of the instruments twice. All countries reported applying all required data quality checks.

Exhibit 6.8: Survey Activities Questionnaire, Section Six—Creating Databases (Numbers of NRC Responses)

Question	Yes	No	Not Answered
Was the information provided in the "TIMSS 2015 Survey Operations Procedures Unit 7 – Creating the Databases" clear and sufficient?	55	0	4
Did you encounter any problems entering test administration information and exporting your WinW3S database(s)?	19	36	4
Who primarily entered the data for your country?			
National center staff	13	-	0
Temporarily hired data entry staff	24	-	0
An external data entry firm	8	-	0
Combination of the above	8	-	0
Other	4	-	2
Did you use manual (key) data entry to create the data files for your country?			
TIMSS achievement booklets	43	12 *(Optical Scanning)*	3 *(Not Answered)* 1 *(Not Applicable)*
Context questionnaires	44	11 *(Optical Scanning)*	4 *(Not Answered)* 0 *(Not Applicable)*
Did you encounter any problems using the IEA's Data Manager Expert (DME) software?	10	45	4
If you entered data manually, did you enter 5% of each survey instrument twice as a quality control measure?	38	8	3 *(Not Answered)* 10 *(Not Applicable)*
Did you apply all the data quality checks described in the "TIMSS 2015 Survey Operations Procedures Unit 7 – Creating the Databases" before submitting your data to the IEA Data Processing and Research Center?	55	0	4
Have you stored all achievement booklets and context questionnaires in a secure storage area until the original documents can be discarded?	56	0	3