



# Chapter 6

## *PIRLS Survey Operations Procedures*

Juliane Barth, Ann M. Kennedy, and Kathleen L. Trong

### **6.1 Overview**

Conducting PIRLS 2006 was an ambitious enterprise in each country that required the careful coordination of schools, staff, and materials by the National Research Coordinator (NRC). In order to assist the NRCs and synchronize activities internationally, a standardized set of survey operations procedures was developed for each country to follow.

The design of the survey operations procedures was a collaborative effort between the TIMSS & PIRLS International Study Center, the IEA Secretariat, the IEA Data Processing and Research Center (DPC), and Statistics Canada. Procedures used successfully in PIRLS 2001, previous TIMSS studies, and other IEA studies, as well as feedback received from the countries that participated in the PIRLS 2006 field test, were used as a basis for developing these procedures.

Survey operations procedures included contacting schools and sampling classes, preparing materials for data collection, administering the assessment, scoring the assessment, and creating the data files. Procedures for quality control and attaining feedback on survey activities also were provided. Guidelines for each of these activities, outlined in subsequent sections of this chapter, were described in an international set of materials that was provided to each NRC.

## 6.2 Responsibilities of the National Research Coordinator

The NRC for each country was responsible for coordinating PIRLS survey activities at the national level. This included acting as the contact person for all those involved in PIRLS within the country, as well as being the representative of the country at the international level. With guidance from organizations that directed PIRLS and experts from within the country, the NRC ultimately made all of the national decisions regarding PIRLS, adapting procedures as necessary to make them appropriate for their national context.

## 6.3 Documentation and Software

Each NRC was provided with a comprehensive set of manuals and software to guide them through the survey operations procedures. Each of these is described below.

- The *School Sampling Manual* (TIMSS & PIRLS International Study Center, 2004) defines the PIRLS 2006 target population and sampling goals and describes the procedures for the sampling of schools.
- The *Survey Operations Procedures Units* are a series of documents that provided a framework for the survey operations. These were organized and distributed chronologically according to the activity and were meant to be used in conjunction with other more specialized manuals.
  - *Unit 1—Contacting Schools and Sampling Classes* (TIMSS & PIRLS International Study Center, 2005e)
  - *Unit 2—Preparing Materials for Data Collection* (TIMSS & PIRLS International Study Center, 2005f)
  - *Unit 3—Administering the PIRLS 2006 Assessment* (TIMSS & PIRLS International Study Center, 2005g)
  - *Unit 4—Scoring the PIRLS 2006 Assessment* (TIMSS & PIRLS International Study Center, 2005h)
  - *Unit 5—Creating the PIRLS 2006 Data Files* (TIMSS & PIRLS International Study Center, 2005i)
- The *School Coordinator Manual* (TIMSS & PIRLS International Study Center, 2005d) describes the steps to be taken by the School Coordinator, which included being responsible for all testing materials

and survey tracking forms, organizing the test administration, and returning the completed testing materials to the NRC.

- The *Windows Within-school Sampling Software and Manual* (IEA, 2005d) helps the NRC to randomly select the PIRLS classes in each sampled school, prepares the survey tracking forms, assigns test booklets to students, and prints labels for the test booklets and questionnaires.
- The *Test Administrator Manual* (TIMSS & PIRLS International Study Center, 2005j) describes the procedures for the Test Administrator to follow during testing, including the timing of and the script used to administer the test, as well as how materials should be returned to the School Coordinator.
- The *International and National Quality Control Manuals* (TIMSS & PIRLS International Study Center, 2005a, 2005b) describe the procedures that quality control monitors should follow when observing testing sessions, as well as the materials they should collect as part of quality control.
- The *Scoring Guides for Constructed-response Items* (TIMSS & PIRLS International Study Center, 2005c) provide detailed and explicit guides used to score each constructed-response item.
- The *Trend Scoring and Reliability Scoring Software and Manual* (TSRS) (IEA, 2005b) is used to ensure consistent scoring over time. This program incorporates a database for countries that participated in PIRLS 2001 and contains a sample of student responses from the PIRLS 2001 data collection. The software allows PIRLS 2006 scorers to rescore the 2001 student response sample, train scorers, and document the reliability of the scoring process over time.
- The *Cross-country Scoring and Reliability Software and Manual* (CCSRS) (IEA, 2005a) is used to document the reliability of scoring across countries. The program incorporates a database containing a sample of student responses to constructed-response questions, collected from English-speaking countries and enables every country to score a common set of student responses.
- The *Windows Data Entry Manager Software and Manual* (IEA, 2005c) captures all PIRLS 2006 responses using keyboard data entry and performs a number of validation checks on the data entered.

- The *Data Correction Software* (DCS) enables national center staff to detect and correct logical inconsistencies in the PIRLS background data.

## 6.4 Contacting Schools and Sampling Classes

One of the essential first steps in PIRLS 2006 was to establish good working relationships with the schools that have been sampled to participate in the study (for more information on sampling procedures, please refer to Chapter 4). NRCs were responsible for contacting these schools and encouraging participation in the assessment, which often involved obtaining support from national or regional education authorities, depending on the national context.

### 6.4.1 School Coordinators

Once a school agreed to participate, a School Coordinator was identified and trained by staff at the national center. This person was responsible for all PIRLS activities within that particular school and often was a teacher or staff member. In some cases, a School Coordinator was a member of the national center staff and was responsible for several schools in an area. School Coordinators were provided with the *School Coordinator Manual*, describing their responsibilities in detail and encouraging them to contact the NRC if they had any questions.

The responsibilities of the School Coordinator included providing information about their school; coordinating the date, time, and place for testing; distributing teacher and school questionnaires; obtaining parental permission (if necessary); and identifying and training a Test Administrator. They also ensured that all testing materials were received and kept secure until administration and returned the completed materials to the national center.

### 6.4.2 Survey Tracking Forms

A large part of the School Coordinator's activities involved providing information about the classes and students in their school. To do this in an organized manner, survey tracking forms were used. Most of these forms were generated by the *Windows Within-school Sampling Software* (WinW3S), completed by schools, and returned to the national centers. The forms were extremely useful in the facilitation of sampling and data collection and were retained for the purpose of data entry verification.

A Class Listing Form was provided to each School Coordinator who listed all of the eligible classes in the target population at that school and provided

details about these classes. From this information, a Class Sampling Form was produced by WinW3S for each school, indicating which classes in the school were selected as part of the sample. A Student Listing Form was created for these sampled classes so that the School Coordinator could list all of the students' names and their information (including exclusion codes, which are discussed in Chapter 9) and return this form to the national center. In addition, a Student Tracking Form was used to document the participation status of each student in the tested classes, and a Teacher Tracking Form was used to document the completion of the *Teacher Questionnaire*.

## 6.5 Administering the PIRLS 2006 Assessment

Distributing materials to the schools required careful organization and planning on the part of the NRC. Each sampled student was assigned 1 of 13 achievement booklets in a systematic rotation so that each achievement block within the booklets was given to an equal number of students in each country. Each student also was assigned a *Student Questionnaire* and a *Learning to Read Survey* for his or her parent to complete. These materials were packaged for each sampled class. In addition, a *Teacher Questionnaire* was sent for each teacher listed on the Teacher Tracking Form and a *School Questionnaire* for the principal. The packaged materials were sent to the School Coordinator, who confirmed receipt of all instruments, prior to the testing date. The *School Questionnaire* and *Teacher Questionnaire* then were distributed, while the other instruments were kept in a secure room until the testing date.

### 6.5.1 Test Administrators

The PIRLS 2006 assessment was conducted by the Test Administrator for each class. This person was chosen and trained by the School Coordinator, although in many cases, the School Coordinator also acted as the Test Administrator. Each Test Administrator was provided with the *Test Administrator Manual*, which outlined his or her responsibilities. The Test Administrator was responsible for distributing materials to the appropriate students, leading students through the assessment, and timing the sessions accurately. Following the assessment, they administered the *Student Questionnaire* and distributed the *Learning to Read Survey* for the students' parents.

### 6.5.2 Timing of the Testing Sessions

The administration of PIRLS 2006 consisted of two sessions, a test administration session and a student questionnaire session. The test administration session concerned the achievement booklets, which contained two parts. This was followed by the completion of the *Student Questionnaire*. The time allotted for each of these sections was standardized across countries, with 40 minutes allowed for each part of the achievement booklet. However, if all of the students finished after 30 minutes, the section could be ended sooner. Test Administrators were required to document the starting and ending time of each section on the Test Administration Form. The timing of the sessions was as follows:

- Preparation for Part 1: approximately 10 minutes, including instructions and booklet distribution
- Achievement booklet, Part 1: 40 minutes
- Break: approximately 15 minutes
- Preparation for Part 2: approximately 5 minutes
- Achievement booklet, Part 2: 40 minutes
- *Student Questionnaire*: at least 20 minutes
- Distribution of *Learning to Read Survey*: approximately 5 minutes

### 6.5.3 Documenting Participation

In addition to the information about the school and its students collected by the School Coordinator, the Test Administrator also used the *Student Tracking Form* during testing. This form was used to distribute the booklets to the correct students and to document student participation.

The School Coordinator used this information to calculate the participation rate. If this was below 90 percent in any class, it was the coordinator's responsibility to hold a makeup session for the absent students before returning all testing materials and survey tracking forms to the national center.

Once the materials had been returned to the national center, the NRC verified the materials, checking that all survey tracking forms had been completed. The national center verified that testing materials were returned for each student listed on the Student Tracking Form, and that the recorded participation status matched the information in the test instruments. Information recorded on the survey tracking forms was then recorded in

*Windows Within-school Sampling Software* (WinW3S). The software was used to check the data for missing and/or inconsistent information and for verification of the data entry process at a later stage, in conjunction with the data entry software, *Windows Data Entry Manager* (WinDEM).

#### 6.5.4 Quality Control

During the test administrations, 10 percent of schools were visited by an International Quality Control Monitor. These monitors were hired by the IEA to verify the quality of the materials and adherence to the test administration procedures in each country. During their school visits, they noted any changes made to the standardized administration script, timing, or procedures and interviewed the School Coordinator about his or her experiences with the PIRLS 2006 assessment. They also were responsible for a final verification of the translated achievement booklets. These were examined while reviewing the comments made by the international translation verifier, and the extent to which the verifier's suggested changes had been integrated was documented. These responsibilities were described in the *International Quality Control Monitor Manual*, and training was provided by staff from the IEA Secretariat and the TIMSS & PIRLS International Study Center.

Additionally, countries were asked to conduct their own quality control procedures in another 10 percent of sampled schools, based on the international program. To assist them, countries were provided with the *National Quality Control Observer Manual*, which was used to train their observers and modified to suit their national system.

### 6.6 Scoring the PIRLS 2006 Assessment

Scoring the PIRLS 2006 instruments in a reliable manner was critical to the quality of the results. To prepare for this task, NRCs were provided with suggestions on how to organize staff and materials. They also were given guidelines on how to select and train scorers to accurately and reliably score the constructed-response achievement items. NRCs were encouraged to employ scorers who were attentive to detail and familiar with education, particularly those with a background in reading instruction.

At international meetings, NRCs were trained to score each of the constructed-response items in the PIRLS 2006 assessment. At these training sessions (which were discussed in Chapter 2), each scoring guide was reviewed

together with examples of student responses that had already been scored according to the guide. The examples were chosen to represent a range of response types, intended to demonstrate the guides as clearly as possible. Following this, NRCs practiced applying the scoring guide to a different set of student responses that had not yet been scored. The scores NRCs gave to these practice papers were shared with the group and any discrepancies discussed. Following the training, NRCs were given a set of the correct scores for the practice papers along with rationales.

NRCs used this information to train their scoring staff on how to apply the *PIRLS 2006 Scoring Guides*. In some cases, NRCs created their own anchor and practice papers from student responses collected from the field test in their country.

In order to demonstrate the quality of the PIRLS 2006 data, it was important to document the reliability of the scoring process within countries, over time, and across countries.

To establish the reliability of the scoring *within each country*, NRCs were required to have a random sample of at least 200 student responses to each item scored independently by two different scorers. The double-scored booklets were selected randomly by *Windows Within-school Sampling Software*, indicated on the cover page of the test booklet. The degree of agreement between the scores assigned by the two independent scorers is a measure of the reliability of the scoring process. The scoring procedure, recommended by the TIMSS & PIRLS International Study Center, interspersed the scoring of the reliability sample with the normal scoring activity, with both taking place simultaneously in a systematic manner.

To measure the reliability of the scoring process *over time* (trend scoring), PIRLS 2006 took steps to document that the constructed-response questions that were carried over from PIRLS 2001 have been scored in the same way in both assessments. For this purpose, following the PIRLS 2001 data collection, countries that participated in this assessment sent samples of their administered and scored test booklets to the IEA DPC. These were digitally scanned and stored for later use in PIRLS 2006. Using this approach, the student responses from 2001 could be rescored by the 2006 scoring staff as a reliability check. The responses were made available to the scorers by the *Trend Scoring Reliability Software* (TSRS). This software allowed student responses to each of the items to be scored electronically. NRCs were asked to have at least two independent



scorers rescore all student responses presented by the software, totaling approximately 200 responses per item. Half of the items had to be scored before the normal scoring activity for PIRLS 2006 began. If the agreement of the scorers fell below 85 percent, retraining of the scorers was required and previously entered scores were disregarded and were scored again, as long as none of the items scored violated the agreement criteria. As soon as the 85 percent agreement criteria agreement was reached on the scored items, the second half of the TSRS student responses could be completed. Whereas the first half of the items were scored before the normal scoring activity for PIRLS 2006 took place, the second half of the items were scored at the same time as the PIRLS 2006 scoring.

In order to measure the reliability of the scoring process *across countries*, NRCs had to have at least two members of their PIRLS 2006 scoring staff score approximately 200 student responses to constructed-response items in English. Student responses to one half of the items were collected throughout the field test from English-speaking countries. The student responses to the second half of the items were taken from English-speaking countries' booklets from PIRLS 2001. Again, the *Cross-country Scoring Reliability Software* scanned the student responses and made them available to the scorers. The program allowed scorers to score the responses electronically by item. The cross-country scoring took place after the normal PIRLS 2006 scoring activity. The degree of agreement between scorers from the various countries may be taken as a measure of cross-country scoring reliability.

## 6.7 Creating the PIRLS 2006 Data Files

As described earlier in this chapter, the IEA DPC provided an integrated computer program for keyboard data entry and data verification known as WinDEM. The program worked in conjunction with WinW3S, so that it was not necessary to reenter tracking information that had been recorded in WinW3S. WinDEM was primarily used for the entry of data from test booklets and questionnaires. The software 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 manual that accompanied the software.

One of the benefits of using WinDEM was that it incorporated the international codebooks describing all variables and their characteristics, thus ensuring that the data files that were produced fulfilled the PIRLS 2006 rules and standards for data entry. Data entry training was provided to NRCs and their national center staff at various stages of the project, including an extensive 4-day training seminar before the field test and before the main data collection.

During the PIRLS 2006 assessment, data were gathered from students, parents, teachers, and school principals. These data were recorded into WinDEM data files as follows:

- *School background data file* contained principals' responses recorded from the *School Questionnaire*.
- *Teacher background data file* contained responses recorded from the *Teacher Questionnaire*.
- *Student background data file* contained responses recorded from the *Student Questionnaire*.
- *Student achievement data file* contained responses recorded from the test booklets.
- *Constructed-response scoring reliability data file* contained the within-country scoring reliability data for the constructed-response questions.

Quality control throughout the data entry process is essential in maintaining accurate data. Therefore, NRCs were responsible for performing periodic reliability checks on the data entry and for applying a series of data verification checks provided in WinDEM. NRCs had to ensure that all data files submitted to the IEA DPC followed the international format and had passed all verification checks. As part of this process, NRCs required their data entry staff to double enter at least 5 percent of each instrument type to ensure puncher reliability and retrain staff if agreement fell below 1 percent. Additionally, the data verification module of WinDEM identified any problems with identification codes and out-of-range and otherwise invalid codes. NRCs also were asked to verify the integrity of the linkage between the students, parents, teachers, and schools entered into the WinDEM data files and the tracking information for those specified in WinW3S. When all data files had passed the WinDEM quality control checks, they were submitted to the IEA DPC along with data documentation for further checking and processing.

## 6.8 Survey Activities Questionnaire

As a structured way to obtain feedback about the survey operations procedures from NRCs, the *Survey Activities Questionnaire* was administered. This consisted of a series of questions concerning each of the various survey activities, how the NRCs conducted them, and space for any comments or suggestions they had. This questionnaire was available online for the NRCs to complete as each of the survey activities was concluding. This format enabled the respondents to reflect on their experiences immediately and to more accurately provide information that can be used to improve survey operations in the future.

## 6.9 PIRLS 2006 Field Test

The PIRLS 2006 field test was a smaller administration of the PIRLS 2006 assessment, involving approximately 1,200 students from each country. It was conducted from March to April 2005 in each of the 40 participating countries and involved 12 newly developed blocks (6 for each reading purpose). One primary goal of the field test was to gather data on the newly developed items in order to analyze their statistical properties. These analyses were used to select six blocks to include (along with secure blocks from PIRLS 2001) in the PIRLS 2006 assessment. Another goal of the field test was to practice conducting the survey operations procedures. This allowed the NRCs and their staff members to become acquainted with the data collection activities and refine their national operations. The field test gave NRCs a basis from which to improve the procedures for the PIRLS 2006 data collection. The field test resulted in some modifications to survey operations procedures and contributed significantly to the successful administration of PIRLS 2006.

## References

---

- IEA. (2005a). *Cross-country scoring and reliability manual*. Hamburg: IEA Data Processing and Research Center.
- IEA. (2005b). *Trend scoring and reliability scoring manual*. Hamburg: IEA Data Processing and Research Center.
- IEA. (2005c). *WinDEM: Windows data entry manager manual*. Hamburg: IEA Data Processing and Research Center.
- IEA. (2005d). *WinW3S: Windows within-school sampling software manual*. Hamburg: IEA Data and Research Processing Center.

## References (continued)

---

- TIMSS & PIRLS International Study Center. (2004). *PIRLS 2006 school sampling manual*. Chestnut Hill, MA: Boston College.
- TIMSS & PIRLS International Study Center. (2005a). *PIRLS 2006 international quality control monitor manual*. Chestnut Hill, MA: Boston College.
- TIMSS & PIRLS International Study Center. (2005b). *PIRLS 2006 national quality control observer manual*. Chestnut Hill, MA: Boston College.
- TIMSS & PIRLS International Study Center. (2005c). *PIRLS 2006 scoring guides for constructed-response items*. Chestnut Hill, MA: Boston College.
- TIMSS & PIRLS International Study Center. (2005d). *PIRLS 2006 school coordinator manual*. Chestnut Hill, MA: Boston College.
- TIMSS & PIRLS International Study Center. (2005e). *PIRLS 2006 survey operations procedures unit 1: Contacting schools and sampling classes*. Chestnut Hill, MA: Boston College.
- TIMSS & PIRLS International Study Center. (2005f). *PIRLS 2006 survey operations procedures unit 2: Preparing materials for data collection*. Chestnut Hill, MA: Boston College.
- TIMSS & PIRLS International Study Center. (2005g). *PIRLS 2006 survey operations procedures unit 3: Administering the PIRLS assessment*. Chestnut Hill, MA: Boston College.
- TIMSS & PIRLS International Study Center. (2005h). *PIRLS 2006 survey operations procedures unit 4: Scoring the PIRLS assessment*. Chestnut Hill, MA: Boston College.
- TIMSS & PIRLS International Study Center. (2005i). *PIRLS 2006 survey operations procedures unit 5: Creating the PIRLS data files*. Chestnut Hill, MA: Boston College.
- TIMSS & PIRLS International Study Center. (2005j). *PIRLS 2006 test administrator manual*. Chestnut Hill, MA: Boston College.

