



for scoring the constructed-response items, creating and checking data files for achievement test and questionnaire responses, and dispatching the completed files to the IEA Data Processing Center in Hamburg.

In addition to administering the PIRLS 2001 instruments, the Trends in IEA's Reading Literacy Study was an option for PIRLS countries that also participated in the 1991 IEA Reading Literacy Study. A separate section in this chapter describes the survey operations for the trend study.

sampling and identification of classes, and ensuring that the administrative and testing materials were correctly distributed, completed, and collected. The School Coordinator also was responsible for identifying and training a Test Administrator to conduct the testing sessions. In some countries, the School Coordinator assumed the roles and responsibilities of the Test Administrator. Both the School Coordinator and Test Administrator received training materials that described their re

The NRC for each country had primary responsibility for carrying out the survey operations. The NRC was responsible for collecting and preparing the data for the PIRLS assessment according to the procedures specified internationally. Earlier chapters of this report describe the tasks of the NRC with regard to choosing a sample of schools, and translating the achievement tests and questionnaires. This chapter focuses on activities associated with the data collection itself.

An important responsibility of the NRC was to identify a School Coordinator for each of the sampled schools to act as a liaison between the national center and the school. The primary role of the School Coordinator was to assist the NRC in the assessment activities within the school, such as: the

- The *Sampling Manual* (PIRLS, 1999) gave an operational definition of the PIRLS main survey target population and sampling goals, and detailed the procedures for sampling schools.
- The *Sampling Manual* (PIRLS, 2001d) described the steps to be taken by the School Coordinator from the selection of the school for testing through the receipt of the survey tracking forms and testing materials at the school. It also specified procedures for returning the completed testing materials to the national center.
- The *Administration Manual* (PIRLS, 2001g) covered the procedures from the beginning of testing to the return of the testing materials and completed Student Tracking Forms to the School Coordinator, and contained the administration script for the testing sessions.
- The *Manual for Entering PIRLS Data* (PIRLS, 2001a) defined the variables and file formats in the data files, and provided instructions for coding, entering, and verifying the data. Codebooks containing detailed specifications for each variable were part of the documentation provided along with this manual.
- The *Scoring Guides* (PIRLS, 2001e) contained the guides developed for scoring each of the constructed-response items.
- The *Manual for Observations* (PIRLS, 2001c) provided instructions for conducting classroom observations in a sample of 10 percent of the participating schools. The observers who conducted the observations were hired by the national centers.
- The *Manual for International Quality Control Monitors* (PIRLS, 2001b) described the procedures employed by International Quality Control Monitors, who were hired by the International Study Center to visit each PIRLS country, interview the NRC, and observe the administration of the PIRLS achievement test in a sample of 15 of the PIRLS schools.

Additionally, three software packages and their corresponding manuals were supplied by the IEA Data Processing Center to assist NRCs in the main survey. These were:

1. The Within-School Sampling Software (W3S), a computer program designed to help NRCs select the within-school sample, prepare the survey tracking forms, assign test booklets to students, and print labels for test booklets and questionnaires (IEA, 2001a).
2. The DataEntryManager (WinDEM), a computer program for data entry and verification (IEA, 2001b).

b-1 -2. Inhe DaLinw dents.1(g01)]TJ4wu(Tw(by)Tj1.0691 NRCs in the main survele, pr- ification (IEssing Cebsmothe data[(ir)19.7(e)]TJ1e

The staff of the IEA Data Processing Center conducted hands-on training sessions in the installation and use of these software packages.

The goal of the PIRLS sampling procedures was to select a nationally representative sample of students in each country. Sampling intact classrooms within randomly selected schools offered the simplest solution from an operational perspective, while optimizing the information gathered about the students and their teachers. Although this was the standard procedure, it could only be implemented where classes in a school constituted an exhaustive and mutually exclusive partition of the students in the grade. In order for a random sample of classes to result in a representative sample of students, every student in the target grade in each country had to belong to one (and only one) of the classes in the school.

A key step in the PIRLS within-school sample selection was the correct identification of classes and teachers. Before classes could be sampled, all eligible classes within the sampled school had to be identified. To that end, the NRC asked the School Coordinator to list (on a Class Listing Form) all classes in the target grade, along with the names of the teacher or teachers responsible for teaching reading to the students in those classes. From the list, the NRC then prepared the Class Sampling Form and applied

a prescribed sampling algorithm to select at least one class at random. Within each school, a class identification number was assigned to each class in the target grades listed on the Class Sampling Form. The six-digit class ID consisted of the four-digit school ID plus the two-digit identification number for the class within the school. All students in this class were then selected to participate in the testing.

PIRLS 2001 relied on a set of tracking forms to implement and record the sampling of classes, teachers, and students. It was essential that these were used and completed accurately, since they indicated the particular test booklet that each student was to receive. They were also used to record participation in each school. In addition to facilitating the data collection, the tracking forms provided essential information for computing sampling weights, and in evaluating the quality of the sampling procedures. Although the tracking forms could be produced manually, NRCs were urged to use the W3S sampling software, which automated many of the sampling and tracking procedures. Once completed, all tracking forms were retained for review by staff at the ISC.

For each sampled class, the NRC asked the School Coordinator to provide a list of students, giving their names, dates of birth, and genders. The School Coordinator also confirmed the name(s) of the teacher(s) of the sampled class. Once the NRC received the list and verified that the requested

information had been supplied, he or she transcribed it onto a Student Tracking Form. Each student listed was assigned an eight-digit student identification number consisting of the six-digit class ID plus a two-digit number corresponding to the student's sequential position on the Student Tracking Form. Three extra student records were created at the end of the list to provide for additional students, or for assigning replacement materials in the event that a student received a damaged test booklet or questionnaire. These three extra records were also assigned a student ID.

The NRC produced a Teacher Tracking Form listing each teacher of the students in the sampled class. Each teacher of the selected class was assigned a teacher ID that consisted of the four-digit school ID, followed by a two-digit number for the teacher within the school, and a two-digit running number that sequentially numbered the entries on the Teacher Tracking Form. The two-digit running number was referred to as the Teacher Link Number, which identified each unique occurrence of a teacher in the Teacher Tracking Form.

During the test administration, the Test Administrator and School Coordinator used the tracking forms to record student, parent, and teacher participation, then returned them to the NRC after the test – together with the completed test booklets and questionnaires.

Although all students enrolled in the target grade were part of the target population and were eligible to be selected for testing, PIRLS recognized that some students in some schools would be unable to take part in the 2001 assessment because of a physical or mental disability. Accordingly, the sampling procedures provided for the exclusion of students with specified disabilities (see Chapter 5). Countries were required to track and account for all excluded students, and were cautioned that excluding an excessive proportion would lead to their results being annotated in international reports. The conditions under which students could be excluded were carefully delineated, because the definition of “disabled” students varied considerably from country to country.

The PIRLS reading assessment was packaged into nine student booklets and one magazine-like booklet known as the PIRLS Reader, which came with its own answer booklet (see Chapter 2). Each student was asked to complete just one of the 10 assessment booklets. Students recorded their answers directly in the booklet, except in the case of the Reader, where the accompanying answer booklet was used. All students completed the same student questionnaire. The assessment booklets were numbered sequentially 1 through 9, with R for the Reader answer booklet. Booklets were assigned to students following a systematic procedure that ensured an

used as a starting point to cycle through the list. For example, if 10 was chosen as the starting number in a particular class, the

even distribution of booklets throughout each class. Each student's booklet assignment was recorded in advance on the Student Tracking Form, and the Test Administrator was expected to ensure that the correct booklet was given to each student. To facilitate proper booklet distribution, each booklet was individually labeled. Exhibit 6.1 provides the rotation scheme whereby booklets were assigned to students. Within each class, a number from one to twelve was randomly selected and

The NRC prepared three packages for each sampled class. One package contained the test booklets for all students listed on the Student Tracking Form. The second and third packages contained the Student Questionnaires and Learning to Read Surveys, respectively. For each participating school, the packages for each sampled class were bundled together with the Teacher Tracking Form, the Teacher Questionnaire, the School Questionnaire, and any materials prepared for briefing school coordinators and test administrators, and were sent to the School Coordinator. Labels and prepaid envelopes addressed to the NRC were included – to facilitate the return of the testing materials.

The School Coordinator in each school was responsible for organizing the administration of the PIRLS 2001 test. The coordinator could be the principal, the principal's designee, or an outsider appointed by the NRC with the approval of the principal. The NRC was responsible for ensuring that the School Coordinator was familiar with his or her responsibilities.

The tasks of the School Coordinator were detailed in the *SCM*. Before the test administration, the School Coordinator, working with the NRC, had to arrange a testing date, select a Test

Administrator to conduct the testing sessions, and ensure that the Test Administrator was fully acquainted with the assessment procedures. In some countries, the School Coordinator assumed the roles and responsibilities of the Test Administrator as well as those of the School Coordinator.

In preparation for the testing day, the School Coordinator worked with the school principal, Test Administrator, and the teacher, to plan the testing sessions by arranging rooms, classes, and materials. In countries where obtaining parental permission for testing was required, the School Coordinator ensured that permission forms were signed and returned in time. Once the testing materials arrived from the national center, the School Coordinator checked that they were for the appropriate students and teachers, that there were enough copies, and that the materials would be kept in a secure place until the testing day.

The School Coordinator distributed Teacher Questionnaires to the teachers listed in the Teacher Tracking Form, and a School Questionnaire to the school principal, and ensured that they were completed and returned. Teacher participation was recorded on the Teacher Tracking Form when the questionnaires were returned. In some countries, it was also the responsibility of the School Coordinator to collect the completed Learning to Read Surveys from the schools, record parent participation on the Student Tracking Forms, and return the questionnaires to the NRC.





testing materials, including the completed Student Tracking Form, the Test Administration Form, and any unused materials. Any relevant information about the test administration in the school was communicated to the NRC by means of a Test Administration Form that was completed by the Test Administrator and School Coordinator.

The ISC implemented an international quality control program whereby international quality control monitors visited a sample of 15 schools in each country and observed the test administration. In addition, NRCs were also expected to organize an independent national quality control program based upon the international model. The latter program required Quality Control Observers to document data collection activities in their country. The Quality Control Observers visited a random sample of 10 percent of the schools (additional to those visited by the international Quality Control Monitors) and monitored the testing sessions, recording their observations for later analysis.

To assist NRCs in conducting their national quality control program, the International Study Center prepared the *Manual for National Quality Control Observers*, which contained information about PIRLS 2001, detailing the roles and responsibilities of the National Quality Control Observers.

In the period immediately following the administration of the PIRLS tests, the main tasks for the NRC included retrieving the materials from the schools and preparing the constructed-response items for scoring. This involved recruiting and training scorers; scoring the constructed-response items, including independent double-scoring of the reliability sample; entering the data from the achievement tests and background questionnaires; submitting the data files and materials to the IEA Data Processing Center; and preparing a report on survey activities.

When the testing materials were received back from the schools, NRCs were to do the following:

- Check that all survey tracking forms were returned from the schools
- Check that the appropriate testing materials were received for every student list-

NRCs then organized the test booklets and questionnaires for scoring and data entry. Procedures were provided to minimize the burden of sorting and handling the booklets, ensure reliability in the constructed-response coding, and document the reliability of the coding.

### Reliability of the Scoring Process

Reliable application of the scoring guides to the constructed-response questions and empirical documentation of the reliability of that process were essential to the success of PIRLS 2001. The *PIRLS Scoring Guide* contained information about arranging for staff and facilities for the constructed-response scoring effort required for the main survey. The manual outlined how to select and train the scorers, and specified the procedures for scoring the constructed-response items and double-scoring a random sample of at least 200 responses per item, to document scoring reliability.

In selecting those who were to do the scoring, NRCs took care to arrange for persons who were conscientious and attentive to detail, knowledgeable in reading, and willing to apply the scoring guides as stated – even if they disagreed with a particular definition or category. Good candidates for scoring included teachers, retired teachers, college or graduate students, and staff members from educational agencies, ministries, or research centers.

### Reliability of the Scoring Process

The success of assessments containing constructed-response questions depends upon the reliability of scoring student responses. In PIRLS 2001, reliability was assured through the provision of scoring guides (manuals), extensive training in their use, and monitoring of the quality of the work. In addition, PIRLS 2001 provided training packets for training in selected questions, along with practice papers to help scorers achieve a consistent level of scoring.

Each scorer received a copy of the *PIRLS Scoring Guide*. This document explained the PIRLS scoring system, which was designed to produce a rich and varied profile of the range of students' competencies in reading literacy. A description of the development and content of the scoring guides is provided in Chapter 2.

At international scoring training meetings for both the field test and the main survey, NRCs received training packets containing “anchor papers” (example student papers) and practice papers to help them achieve accuracy and consistency in scoring. About 10 to 15 responses were sufficient for most items, but the complexity of some scoring guides made additional practice papers necessary.

In order to demonstrate the quality of the PIRLS 2001 scoring, it was important to document the agreement between scorers. To establish the scoring reliability, NRCs were required to have a random sample of at least 200 responses for each constructed-response item independently scored by two scorers. This number is equal to approximately 25 percent of the responses based on the typical sample size. With the exception of the test booklet that accompanied the Reader, each item appeared in three booklets, meaning that the scorers needed to score a sample of 67 of each booklet. The items in the booklet accompanying the Reader do not appear in any other booklet; the scorers needed to score a sample of 200 of these booklets.

The *S O M* provides a procedure for interleaving the double scoring of the sample of reliability booklets with the regular booklet scoring, so that the reliability sample is scored in the same way and at the same time as the other booklets.

The activity described above provides evidence of the extent of agreement among scorers within each country, but does not address the question of scoring consistency across countries. Since PIRLS is administered in each country's own language, and since scorers generally do not know other countries' languages, cross-country scoring reliability is much more difficult to establish. As a partial solution to this problem, PIRLS took samples of student responses to a selection of constructed-response questions from each of the PIRLS countries that

tested in English, and had this common set of English responses scored by two scorers in each country where scorers could operate through English. The common set consisted of 200 student responses to 25 questions from four of the PIRLS reading passages, two of which were literary and two informational. The cross-country reliability scoring was conducted in each country after the scoring of the national PIRLS data had been completed.

The Trends in IEA's Reading Literacy Study was an option for countries that participated in the 1991 IEA Reading Literacy Study, allowing them to administer the 1991 test booklets and questionnaires again in 2001 – to compare reading achievement over time. A list of countries participating in the Trends in IEA's Reading Literacy Study is provided in Chapter 5.

Documentation for implementing the trend study was incorporated in the manuals for PIRLS 2001. The *S O M* provided instructions for preparing tracking forms, packaging materials for the schools, preparing, administering, and returning the test booklets and questionnaires, and for data entry. The *S C M* and *T A M* included sections specific to the 10-Year Trend Study wherever necessary.

The *S S g G 10-T S* (PIRLS, 2000) documented procedures for sampling schools for the Trends

in IEA's Reading Literacy Study. Essentially, the procedure was to sample an additional class from the target grade in half of the schools in the PIRLS sample. The survey tracking forms used in PIRLS were also used for tracking trend study schools, classes, and students.

Countries were expected to use test booklets identical to those used in 1991. Schools in which classes were selected for the trend study received two packages from the NRC. One contained the test booklets with the Student Tracking Forms, and the other the Student Questionnaires. Teachers of the classes selected for the trend study were not given a Teacher Questionnaire, nor was the School Questionnaire administered.

The data collection was conducted under the same conditions as in 1991. There were three data-collection sessions, with the same time limitations as in 1991. There was a short break between sessions:

- 1 minute and 30 seconds for answering questions in the Word Test (word recognition), followed by 35 minutes for answering questions in the first part of the Reading Test
- 40 minutes for answering the questions in the second part of the Reading Test
- At least 25 minutes for the completion of the Student Questionnaire.

The Test Administrator and School Coordinator followed the PIRLS procedures for collecting the test instruments, checking

for proper documentation on the survey tracking forms, and for calculating student response rates. Once any necessary makeup sessions had been held, the School Coordinator returned the materials to the NRC for data preparation. Since the trend study test booklets did not include constructed-response items for scoring, they were sorted separately for data entry – along with the trend study Student Questionnaires.

The IEA Data Processing Center provided each NRC with a copy of WinDEM (an integrated computer program for data entry and data verification) designed specifically for use with IEA studies. This program allowed data entry directly from the tracking forms and test instruments, and provided a convenient checking and editing mechanism. WinDEM also offered interactive error detection, error reporting, and quality control procedures. Detailed information and operational instructions were provided in the manual for the WinDEM software. WinDEM for PIRLS incorporates – for each PIRLS instrument – the international codebook, which describes the format and data characteristics of each variable in the instrument. Correct use of the WinDEM software ensured that the data files were produced according to the PIRLS 2001 standards for data entry.

Although WinDEM was strongly recommended for all data entry tasks, NRCs sometimes chose to use their own procedures and computer programs, which was accept-

able – provided all data files conformed to the specifications of the international codebooks. NRCs who did not to use WinDEM were responsible for ensuring that all data files were delivered to the Data Processing Center in the international format.

During the PIRLS 2001 main survey operations, data were gathered from several instruments, including: student assessment booklets, questionnaires from students, teachers, parents, and principals, as well as from a range of tracking forms. Before beginning data entry, the NRC needed to ensure that the instruments and corresponding tracking forms had been completed and sorted correctly. Data entry involved the following files:

- The school background file – information from the School Questionnaire and the School Tracking Form
- The teacher background file – information from the Teacher Questionnaire and the Teacher Tracking Form
- The student background file – data from the Student Questionnaire, the Test Administration Form, and Student and Teacher Tracking Forms
- The student achievement file – data from the student assessment booklets and the Data Entry Batch Header
- The home background file – data from the Learning to Read Survey
- The constructed-response scoring reliability file – data from the scoring sheets for the constructed-response items that were double-scored
- The Trends in IEA's Reading Literacy Study file consisted of the data from the Student Questionnaire, the achievement test booklet, the Test Administration Form, and the Student Tracking Form.

Quality control throughout the data entry process is essential in maintaining accurate data. Countries were responsible for performing periodic reliability checks on the data entry, and for applying a series of data verification options provided as part of WinDEM. NRCs not using WinDEM for data entry still had to apply the WinDEM data verification checks to their data before sending their files to the IEA Data Processing Center. The WinDEM data-checking facility could identify a range of problems that could then be fixed before submission to the Data Processing Center. Specifically, WinDEM checks for the following:

- Duplicate identification codes
- Inconsistencies in the hierarchical identification system
- Out-of-range values
- Mismatches between different student-level files.

In addition to the checks performed by WinDEM, the Linkcheck program detected mismatches between different files, as well as inconsistencies in important identification variables. Students whose IDs were not numbered sequentially are reported, and classrooms with unusually small numbers of students were detected. The booklet assignment was checked, and information about

Progress in International Reading Literacy Study  
(PIRLS). (2001e). *S g G*  
*C -R I* (PIRLS Ref. No. 01-  
0007). Prepared by the International Study Center  
at Boston College. Chestnut Hill, MA: Boston  
College.

Progress in International Reading Literacy Study  
(PIRLS). (2001f). *S O M* (PIRLS  
Ref. No. 01-0001). Prepared by the International  
Study Center at Boston College. Chestnut Hill,  
MA: Boston College.

Progress in International Reading Literacy Study  
(PIRLS). (2001g). *T A M*  
(PIRLS Ref. No. 01-0002). Prepared by the  
International Study Center at Boston College.  
Chestnut Hill, MA: Boston College.

