CHAPTER 4

Survey Operations Procedures for PIRLS 2021

Ieva Johansone

Overview

As data-based indicators of countries' student achievement profiles and learning contexts, the quality of the data collected by each participating country and benchmarking entity is critical. 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 a consistent and uniform approach necessary for high-quality, internationally comparable data, all participants are expected to follow a set of standardized operations procedures. These procedures were developed through a partnership involving the TIMSS & PIRLS International Study Center, IEA Amsterdam, IEA Hamburg, Statistics Canada, and National Research Coordinators (NRCs) from participating countries. The major steps of the operations and procedures are similar from one assessment cycle to the next. However, with each assessment cycle the operations procedures are updated to enhance efficiency and accuracy and reduce burden, making use of developments in technology to automate routine activities wherever possible.

Each new assessment cycle also brings something new and unique requiring the operations and procedures to be adapted. The 2021 cycle of PIRLS marked the transition to digital assessment, with about half of the participating countries switching from the previous paper-based version to the new digital format. Additionally, the ePIRLS tasks, introduced for the first time during the 2016 assessment cycle, were integrated and administered as part of digitalPIRLS 2021. Adapting operational procedures from the existing PIRLS operations and integrating the ePIRLS operations to create a workflow for this new assessment mode was a significant undertaking. In order to control for any assessment mode effects, in addition to the usual nationally representative sample, countries transitioning to digitalPIRLS were required to administer "bridge" paper instruments to an extra, equivalent sample of students. This bridge data collection also required integrating operations and procedures into the overall PIRLS 2021 assessment administration.

https://doi.org/10.6017/lse.tpisc.tr2103.kb8234

In each country or benchmarking entity, the National Research Coordinator was responsible for the implementation of PIRLS 2021. Internationally, National Research Coordinators provided the country's perspective in all international discussions, represented the country at international meetings, and served as 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 PIRLS, including any adaptations for the national context.

The daily tasks of the National Research Coordinators varied over the course of the PIRLS 2021 cycle. In the initial phases, National Research Coordinators participated in the PIRLS 2021 assessment frameworks and assessment development process (see Chapter 1 and Chapter 2) and collaborated with Statistics Canada and IEA Hamburg in developing a plan to implement the PIRLS 2021 sampling design within the country or benchmarking entity (see Chapter 3).

Following the development of the achievement test and context questionnaires, countries conducted a full-scale field test of all instruments and operational procedures in March through June 2020 in preparation for the main data collection. This was a longer field test administration period than usual due to the COVID-19 pandemic, but almost all countries managed to complete the field test activities. As well as providing crucial data to support the finalization of the assessment instruments (achievement items and questionnaires), the field test enabled the National Research Coordinators and their staff to become acquainted with the operational activities. The feedback they provided was used to improve the procedures for the main data collection. The field test, especially that of the digital administration mode, contributed significantly to ensuring the successful execution of PIRLS 2021.

The main data collection started in October 2020 for Southern Hemisphere countries, but due to the COVID-19 pandemic only two Southern Hemisphere countries managed the administration in the planned administration period through December 2020. From March through June 2021 most Northern Hemisphere countries managed to administer PIRLS 2021. The COVID-19 pandemic greatly disrupted the main data collection administration period, requiring that some countries delay the administration period anywhere from one month to one year. The national assessment administration periods stretched from the usual 4-6 weeks to several months, as countries did their best to administer PIRLS 2021 during the time periods their schools were open for instruction. These were unprecedented circumstances for the PIRLS study.

As part of ongoing efforts to improve operations, the National Research Coordinators were asked to complete a Survey Activities Questionnaire (SAQ) that sought feedback on all aspects of their experience conducting PIRLS 2021. 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 PIRLS 2021 Survey Activities Questionnaire are presented in the final section of this chapter.

PIRLS 2021 Survey Operations Units, Manuals, and Software

To support the National Research Coordinators in conducting PIRLS 2021, the TIMSS & PIRLS International Study Center provided step-by-step documentation for all operational activities. Organized into a series of units, the *PIRLS 2021 Survey Operations Procedures* were made available at critical junctures of the project to ensure that National Research Coordinators had all the tools and information necessary to perform their responsibilities. Also, the procedures units were accompanied by a series of manuals for use by School Coordinators and Test Administrators. National Research Coordinators could translate and adapt these to their local situations. Often, separate versions of the units and manuals were provided for paperPIRLS and for digitalPIRLS administrations. The TIMSS & PIRLS International Study Center and IEA Hamburg also provided National Research Coordinators and their staff with intensive training in constructed response item scoring and data management.

To digitally deliver the achievement test and student questionnaire to students, PIRLS 2021 used Assessment Master (AM), an assessment platform owned by RM (*www.rm.com*). For all other context questionnaires, IEA's Online SurveySystem (OSS) was used (see Chapter 5). Countries also were provided with the System Check program for checking the suitability of devices for PIRLS 2021. Monitoring the upload of student data was possible for both the RM's Assessment Master and IEA's Online SurveySystem.

In addition to the achievement test and context questionnaire delivery software and consistent with the goal of automating and streamlining procedures wherever possible, IEA Hamburg provided National Research Coordinators in both digitalPIRLS and paperPIRLS countries with a range of custom-built software products to support project activities. These included the Windows® Within-School Sampling Software (WinW3S) for sampling and tracking classes and students; the IEA CodingExpert Software for scoring constructed response items and documenting scoring reliability; and the IEA Data Management Expert (DME) software for creating and checking data files for all paper-based assessment instruments. These software products were accompanied by manuals to support the use of the software.

The *PIRLS 2021 Survey Operations Procedures* units were crucial resources for the National Research Coordinators, as they 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 contracted to conduct aspects of the assessment in their countries.

The following units, manuals, and software systems were provided for administering PIRLS 2021:

- *PIRLS 2021 Survey Operations Procedures Unit 1: Sampling Schools and Obtaining their Cooperation*
- *PIRLS 2021 Survey Operations Procedures Unit 2: Preparing for and Conducting the PIRLS 2021 Field Test*

  Unit 2 consisted of the following sections: Preparing the Field Test Instruments (paper or digital), Sampling Classes and Field Test Administration, Scoring the Field Test Constructed Response Items, and Creating and Submitting the Field Test Databases. A supplement describing online scoring of the digitalPIRLS constructed response items also was included.

  Unit 2 was accompanied by field test versions of the School Coordinator and Test Administrator Manuals for paperPIRLS and digitalPIRLS, instructions on "Preparing Computers for digitalPIRLS," and a National Quality Control Monitor Manual.

  In addition to the manuals, IEA Hamburg provided field test versions of the WinW3S within-school sampling software, the OSS software for questionnaire administration, CodingExpert software, and the DME data management software. RM provided the Assessment Master's Online Translation System and test delivery "Players" for the field test.
- *PIRLS 2021 Survey Operations Procedures Unit 3: Contacting Schools and Sampling Classes for the PIRLS 2021 Data Collection*

  Unit 3 was accompanied by the main data collection versions of the School Coordinator Manual and the WinW3S within-school sampling software and its manual. digitalPIRLS countries also received the System Check program and instructions on "Preparing Computers for digitalPIRLS," which provided the necessary information and tools for countries to test their devices for PIRLS compatibility and prepare them for the main data collection.
- *PIRLS 2021 Survey Operations Procedures Unit 4: Preparing the PIRLS 2021 Assessment Instruments*

  Separate versions of Unit 4 were provided for paperPIRLS and digitalPIRLS countries; the latter also received a manual on preparing the paper "bridge" booklets. The digitalPIRLS version provided access to the AM Online Translation System, which enabled National Research Coordinators to translate the PIRLS texts and items, as well as the Student Questionnaire, into their language(s) of instruction. The translated materials were available online for translation and layout verification by IEA Amsterdam and the TIMSS & PIRLS International Study Center (see Chapter 5).

Unit 4 also was accompanied by the main data collection version of the OSS online survey system for online administration of the school, teacher, and home (Early Learning Survey) questionnaires.

- *PIRLS 2021 Survey Operations Procedures Unit 5: Conducting the PIRLS 2021 Data Collection*

  Unit 5 was accompanied by the main data collection versions of the Test Administrator Manuals for paperPIRLS and digitalPIRLS, the National Quality Control Monitor Manual, and the International Quality Control Monitor Manual.

  digitalPIRLS countries also received the PIRLS Player for administering the digitalPIRLS assessment to students. Each country's PIRLS Player was customized to contain the country's translations of the assessment texts and items, as well as the national Student Questionnaire.

- *PIRLS 2021 Survey Operations Procedures Unit 6: Scoring the PIRLS 2021 Constructed Response Items*

  Unit 6 was accompanied by the main data collection versions of the scoring guides and IEA's CodingExpert Software and manuals. The CodingExpert Software was used to facilitate digitalPIRLS online scoring and the trend and cross-country reliability scoring tasks.

- *PIRLS 2021 Survey Operations Procedures Unit 7: Creating and Submitting the PIRLS 2021 Databases*

  Unit 7 was accompanied by the main data collection versions of the DME data management software, codebooks, and manual. The DME software was used for data entry and data verification for all paper instruments.

PIRLS 2021 Survey Tracking Forms

PIRLS 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. Four different tracking forms were used for PIRLS 2021:

- 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 Listing Form: This form was completed for each class sampled, listing the names of the students, student birth dates, genders, and exclusion codes.

- Student Tracking Form: This form was created for each class assessed and was completed by the Test Administrators during test administration. Separate Student Tracking Forms were provided for paperPIRLS and digitalPIRLS. The Test Administrators used this form to verify the assignment of survey instruments to students and to indicate participation status, including the return status of the Early Learning Surveys (home questionnaires).
- 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:

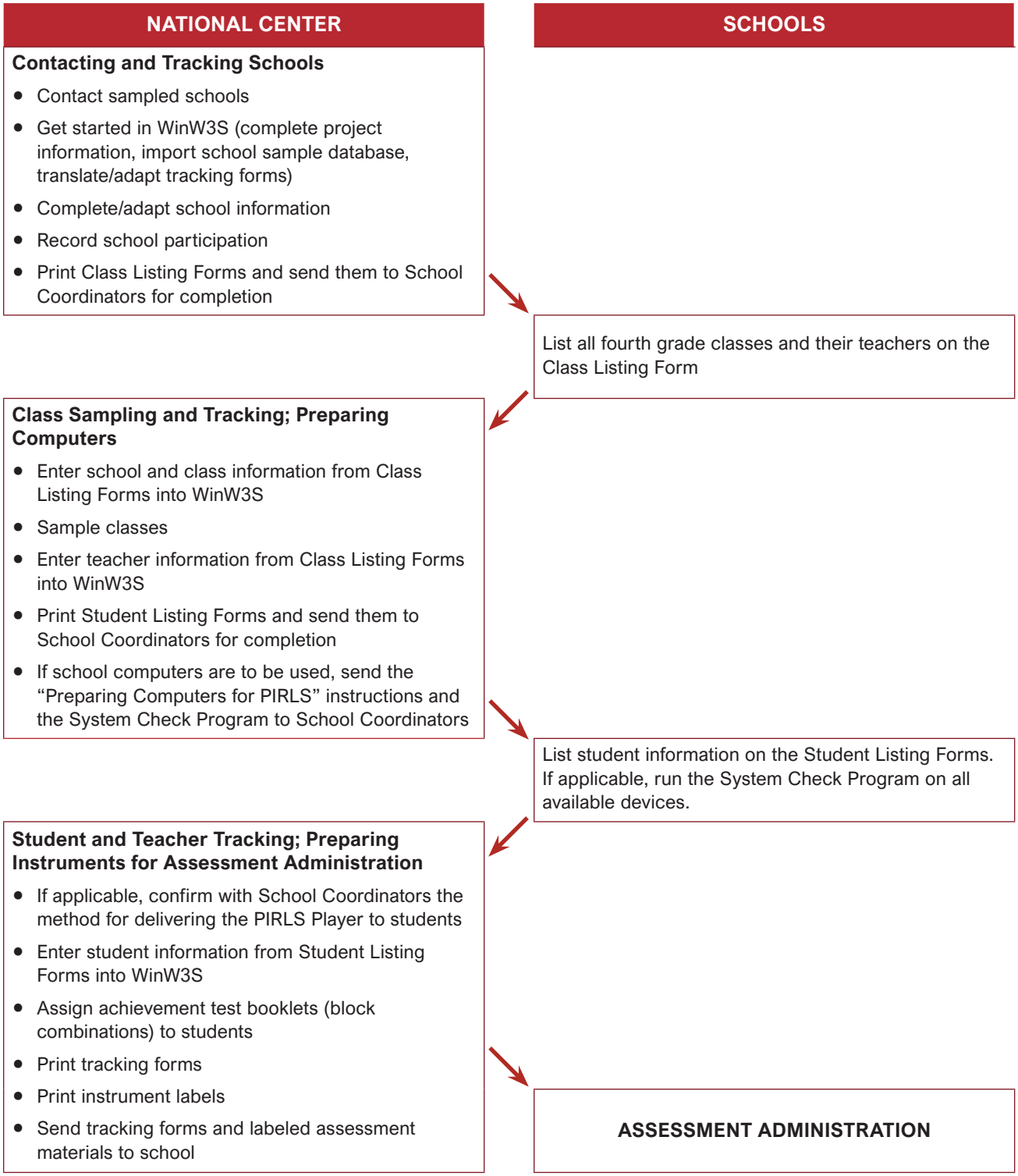
- Contacting schools and sampling classes
- Overseeing national assessment instrument preparation
- Managing the PIRLS 2021 assessment administration
- Scoring the constructed response items
- Creating the PIRLS 2021 database files

Two other major PIRLS 2021 operational activities are described in separate chapters of this publication: sampling schools (Chapter 3) and verification of the systems and the assessment instruments (Chapter 5).

Contacting Schools and Sampling Classes

Exhibit 4.1 illustrates the major steps in working with schools to sample classes and prepare for the PIRLS assessment administration. Once the school samples were drawn, National Research Coordinators were tasked with contacting schools and encouraging them to take part in PIRLS 2021. Depending on the national context, this could involve obtaining support from national or regional educational authorities. Survey Operations Procedures Unit 1 included suggestions on ways to encourage schools to participate in the assessment.

Exhibit 4.1: Diagram of Sampling Procedures and Preparations for the Assessment Administration Implemented by National Centers and Schools



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 National Research Coordinators 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. School Coordinators were provided with a School Coordinator Manual describing their responsibilities. The School Coordinator Manual was prepared by the TIMSS & PIRLS International Study Center and translated/adapted by national center staff in each country.

The responsibilities of the School Coordinators included providing the national center with information on the school; coordinating the dates, times, and places for testing; identifying and training Test Administrators to administer the assessment; coordinating the completion of the tracking forms; distributing questionnaires; and, when necessary, obtaining parental permission. If school devices were used for digitalPIRLS administration, School Coordinators were provided with the "Preparing Computers for PIRLS" instructions and the System Check Program in order to test the devices for PIRLS compatibility and prepare the compatible devices for testing. 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 assessment administration.

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

Overseeing National Assessment Instrument Preparation

National Research Coordinators also were responsible for preparing the assessment instruments (paperPIRLS achievement booklets or digitalPIRLS and ePIRLS texts and items, paper bridge booklets, if applicable, and context questionnaires) for their countries—a process that included overseeing the translation and/or adaptation of the assessment instruments. The overarching goal of assessment instrument preparation was to create internationally comparable instruments that were appropriately adapted for the national context of each participating country.

There were 18 assessment blocks—nine literary and nine informational (see Chapter 1). Each assessment block consisted of a reading passage and its corresponding achievement items. Also, there were five ePIRLS tasks, where students read through information presented online in a simulated internet environment and answered a series of comprehension questions about what they had read.

Twelve of the assessment blocks and three of the ePIRLS tasks were trend—used in 2016 and kept secure for the 2021 cycle. For these trend assessment blocks and tasks administered in 2016, countries were required to use the same translations as had been used previously. Countries that did not participate in the 2016 or 2011 cycles of PIRLS had to translate and/or adapt the trend blocks into their assessment language(s) in preparation for the 2021 assessment administration.

Six new assessment blocks were developed for PIRLS 2021, with the new blocks replacing the ones released at the end of the previous assessment cycle. Also, two new assessment tasks were developed for the ePIRLS 2021 assessment. The new assessment blocks and tasks were all tried out through the field test (12 new blocks were field tested) and the best assessment blocks were chosen, with some edits applied for the main data collection. Similarly, the context questionnaires were evaluated following the field test to gauge the validity and reliability of the various questionnaire scales. There were four context questionnaires for PIRLS 2021—School Questionnaire, Teacher Questionnaire, Student Questionnaire, and Home Questionnaire.

In addition to the main digitalPIRLS assessment, countries transitioning to the digital administration mode had to prepare eight bridge booklets. The bridge booklets were composed entirely of the trend assessment blocks that were previously used in PIRLS 2016 and kept secure for PIRLS 2021. For the bridge booklets, countries used Adobe® InDesign® software to link their translated and adapted assessment blocks from PIRLS 2016 to the appropriate bridge booklets.

In preparation for translation for both the field test and main data collection, the participating countries received the international version (English) of the instruments either as "Ready to Translate" in the translation system or as instrument production files. Instructions on how to use the materials to produce high-quality, standardized instruments were included in the corresponding Survey Operations Procedures units and manuals.

Once translated and/or adapted, first for the field test and then again for the main data collection, the national versions were submitted to IEA Amsterdam for translation verification (see Chapter 5). IEA Amsterdam worked with independent translators to evaluate each country's translations and, when deemed necessary, suggested changes to the text.

After the translations had been verified by IEA Amsterdam, the instruments were revised and sent to the TIMSS & PIRLS International Study Center for layout and adaptations verification (see Chapter 5). This review checked that the instruments conformed to the international format and that any adaptations made to the instruments did not unduly influence their international comparability.

Managing the PIRLS 2021 Assessment Administration

Preparing assessment materials and distributing them to the participating schools required careful organization and planning by National Research Coordinators. The assessment materials were provided to the School Coordinators prior to testing, giving ample time for the School Coordinators to confirm the receipt and correctness of the materials. The school and teacher questionnaires were then distributed, and 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 assessments and student questionnaire. Test Administrators were in most cases chosen and trained by School Coordinators, and in some cases, the School Coordinator doubled as the Test Administrator.

Test Administrators were responsible for distributing materials to the appropriate students, reading the instructions provided in the Test Administrator Manual to the students, and timing the sessions. WinW3S systematically assigned achievement booklets/block combinations and produced labels to facilitate the distribution of the assessment, and Test Administrators used the Student Tracking Form and these labels to distribute the assessment instruments to the correct students and to document student participation. When a class had a participation rate below 90 percent, 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. The COVID-19 pandemic disrupted this work significantly; many times the participation rate could not be met even with multiple makeup sessions. Using the Test Administration Form, the Test Administrators documented the timing of the testing sessions and information about any unusual incidents that took place during assessment administration.

The achievement test consisted of two parts, with each containing one assessment block. To complete each part of the achievement test, students were allowed 40 minutes. For digitalPIRLS countries, students were automatically logged out of the system once the time allowed had expired. For paperPIRLS, Test Administrators were responsible for reinforcing the timing of each part. There was a required break between the two parts of the achievement test administration. The break was not to exceed 30 minutes. Students who completed part 1 or part 2 of the achievement test before the allotted time were not allowed to leave the testing room and were asked to review their answers or read quietly. Some Test Administrators provided activity sheets for these students.

Following the administration of the achievement test, students were provided 30 minutes to complete the student questionnaire with extra time provided to students who needed it. Test Administrators were permitted to read the questionnaire items aloud together with the students.

Before students were dismissed, home questionnaires were distributed for parents/guardians to complete. If the home questionnaire was administered on paper, students were asked to bring it back to their teacher or the School Coordinator. If it was administered online, students brought

home printed information about the questionnaire and how to access it. Sometimes countries chose other means of communication with the parents about completing and returning the home questionnaire.

digitalPIRLS was administered online or via individual USB sticks on individual PIRLS compatible computers. In addition, the local PC server method was available using a Local Area Network (LAN) or Wi-Fi, where a single PIRLS compatible computer could be used as a local server and students using individual devices connected to the server computer. For the PC/USB administration offline, the Test Administrators and School Coordinators submitted/uploaded the data after each testing session. Due to computer shortages, multiple testing sessions were sometimes needed to assess all students in a class.

Linking Students to their Teachers and Classes

Exhibit 4.2 illustrates the hierarchical identification system codes that were used to link the data among schools, classes, students, and teachers. The school, class, and student IDs were strictly hierarchical, with classes nested within schools and students nested within classes.

Exhibit 4.2: 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 was assigned a teacher identification number consisting of the four-digit school number followed by a two-digit teacher number. Since the same teacher could be teaching more than one class within a school, it was necessary to have a unique identification number for each class linked to a teacher. This was achieved by adding a two-digit link number to the six digits of the teacher identification number to create a unique eight-digit number identifying the school, teacher, and class.

Scoring the Constructed Response Items

Reliable and valid scoring of constructed response items is critical to the assessment results. The TIMSS & PIRLS International Study Center provided explicit scoring guides for each individual item and extensive training in their use. Also, the Survey Operations Procedures units specified a

procedure for efficiently organizing and implementing the scoring activity. Scoring the digitalPIRLS constructed response items was done online via IEA's CodingExpert Software, which incorporated the IEA standards and reliability procedures.

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 National Research Coordinators) 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 set of example student responses that had already been scored. These example papers 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 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 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 PIRLS constructed response items provided for documenting scoring reliability within each country (within-country reliability scoring), over time (trend reliability scoring), and across countries (cross-country reliability scoring) (see results in Chapter 9).

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 that 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 PIRLS 2016 to PIRLS 2021). The trend reliability scoring required scorers of PIRLS 2021 to score student responses collected in 2016. The scores from 2021 were then compared with the scores awarded in 2016. Trend reliability scoring was conducted using IEA's CodingExpert Software provided by IEA Hamburg.

Student responses included in the trend reliability scoring (200 responses per item) were actual student responses to a total of 18-20 items from five passages. The responses were collected from the PIRLS trend assessment blocks during the PIRLS 2016 assessment administration in each country and benchmarking entity. These responses were scanned and provided to each respective participating country and benchmarking entity and then scored with IEA's CodingExpert Software. All scorers who scored the trend assessment blocks in 2021 were required to participate in the trend reliability scoring. If all scorers were trained to score all trend passages/items, the software divided the student responses equally among the scorers. If scorers were trained to score specific passages/items, National Research Coordinators were able to specify within the software which scorers would score particular passages/items, 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's CodingExpert Software. Student responses included in the cross-country reliability scoring (200 responses per item) were student responses to three passages/18 items. Student responses were collected from the English-speaking countries during the PIRLS 2016 assessment administration. 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 divided and then distributed equally to the participating scorers in each country. In most countries, the scoring exercise was completed immediately after all other scoring activities.

Creating the PIRLS 2021 Database Files

The procedure for creating the PIRLS 2021 database file included entering sampling and assessment administration information into WinW3S and adding responses from the paper context questionnaires and achievement booklets using IEA's Data Management Expert (DME) software. IEA Hamburg provided the DME software to accommodate keyboard data entry from the paper instruments. The DME software also offers data and file management capabilities, a convenient checking and editing mechanism, interactive error detection, and quality control procedures.

The digital achievement test and student questionnaire data were captured automatically and submitted through an online submission process after the assessment administration. The digital assessment constructed response scoring took place directly in the online database and thus did not require any manual data entry. For the teacher, school, and home questionnaires administered online through IEA's OSS via the IEA Hamburg server, the data were directly accessible to IEA Hamburg and no further data entry was required.

For manual data entry using the DME software, IEA Hamburg provided international codebooks describing all variables and their properties. These ensured that data files produced

with this system met the internationally defined rules and standards for data entry. Before being used, however, the international codebooks had to be updated to accommodate any national adaptations to the data collection instruments. These adapted national codebooks then were used to create the data files in each country, with the responses to the paper 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 systems prior to submitting the databases to IEA Hamburg. To ensure the reliability of the data entry process, data entry staff were required to independently reenter at least 5 percent of the records from each instrument type. An error rate of 1 percent or less was acceptable for the questionnaire 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.

Both WinW3S and DME systems offered a data verification module that checked for a range of problems, such as inconsistent identification codes, inconsistencies between participation status information and achievement and/or questionnaire data availability, and out-of-range or otherwise invalid codes. The data verification module also verified the integrity of the linkage between the students, teachers, and schools entered into the DME database and the tracking of information for those specified in WinW3S. For data captured online, it was possible to export data availability information and apply data verification to check for inconsistencies via the WinW3S and DME data verification modules.

When all data files had passed the quality control checks, they were submitted to IEA Hamburg along with data documentation for further checking and processing. For information on data processing at IEA Hamburg, please refer to Chapter 7 of this publication.

PIRLS 2021 Survey Activities Questionnaire

The Survey Activities Questionnaire was designed to elicit information about National Research Coordinators' experiences in preparing for and conducting the PIRLS 2021 data collection. The questionnaire was composed of six sections and focused on the following:

- Sampling schools and classes
- Translating, adapting, and producing the assessment instruments
- Administering the assessments
- Implementing the National Quality Control Program

- Preparing for and scoring the constructed response items
- Creating and submitting the databases and documentation

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 for the process.

The PIRLS 2021 Survey Activities Questionnaire was administered via IEA's Online SurveySystem and was completed by a total of 58 NRCs: 28 for digitalPIRLS and 30 for paperPIRLS. The following sections summarize information gathered from the Survey Activities Questionnaire.

Sampling Schools and Classes

The first section of the Survey Activities Questionnaire asked National Research Coordinators about the Survey Operations Procedures units for sampling schools (Unit 1) and sampling classes within the sampled schools (Unit 3). As shown in Exhibit 4.3, 55 National Research Coordinators considered Survey Operations Procedures Unit 1 to be clear and sufficient, and 56 considered Unit 3 to be clear and sufficient. Five countries reported deviating from the expected PIRLS sampling design. Their reasons for these modifications included a response to the effect of the COVID-19 pandemic on their samples. For example, countries chose larger school samples in anticipation of pandemic related non-participation. In addition, some countries adjusted their sampling procedures due to a change in the way a country identified schools; adjustments for classes based on gender, special needs, or mixed grade levels; and the need to oversample for enhanced reporting. Statistics Canada, in cooperation with IEA Hamburg, selected the school samples for all countries and benchmarking participants.

Exhibit 4.3: Survey Activities Questionnaire, Section One—Sampling (Numbers of NRC Responses)

Question	Yes	No	Not Answered
Was the information provided in the "PIRLS 2021 Survey Operations Procedures Unit 1: Sampling Schools and Obtaining their Cooperation" clear and sufficient?	55	1	2
Were there any conditions or organizational constraints that necessitated deviations from the basic PIRLS sampling design described in the "Survey Operations Procedures Unit 1"?	5	51	2
Did you use the Within-School Sampling Software (WinW3S) to sample classes?	55	1	2
If you answered "yes," did you experience any problems when using the WinW3S software?	6	48	4

Exhibit 4.3: Survey Activities Questionnaire, Section One—Sampling (Numbers of NRC Responses) (Continued)

Question	Yes	No	Not Answered
Was the information provided in the "PIRLS 2021 Survey Operations Procedures Unit 3: Contacting Schools and Sampling Classes" clear and sufficient?	56	1	1
Did you follow the procedures outlined in "PIRLS 2021 Survey Operations Procedures Unit 3: Contacting Schools and Sampling Classes" 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)?	46	11	1

All but one of the National Research Coordinators reported using Windows WinW3S provided by IEA Hamburg to select classes within the sampled schools. Six National Research Coordinators reported experiencing problems using the WinW3S software. Among the issues reported were difficulty in excluding classes, login issues, and technical issues that caused computers to run out of memory.

Eleven National Research Coordinators applied some modifications to the procedures outlined in the Survey Operations Procedures Unit 3. For example, one country sampled the schools and then included all the Grade 4 classes in the schools to help address any issues created by the pandemic. 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 National Research Coordinators about translating, adapting, assembling, and printing the test materials, as well as issues related to checking the materials and securely storing them. There were digitalPIRLS-specific questions asked in this section related to using RM's AM system, receiving the PIRLS Player, and, if applicable, preparing USBs in order to deliver digitalPIRLS to students.

As reported in Exhibit 4.4, almost all National Research Coordinators found the instructions on preparing assessment instruments to be clear and sufficient. However, three countries reported experiencing some problems using the paper-based instrument production materials. These problems included issues with fonts and special characters (e.g., for Cyrillic alphabet) and difficulties applying the translations in the InDesign instrument production files, as the CopyFlow Gold® plug-in for InDesign that supported the import of text was no longer available. The 11 National Research Coordinators who reported issues with the AM system noted the inability to

produce PDFs of their materials, the instability of the materials throughout the preparation process, font issues related to the use of accents in some languages, font options for certain languages, and the timeliness of the corrections that RM needed to make. The mentioned issues were dealt with by the TIMSS & PIRLS International Study Center and/or by RM.

Almost all National Research Coordinators reported applying corrections to their assessment instruments as suggested by the external translation verifier and/or the layout verifier. When suggestions were rejected, it was because the language suggested was not the most appropriate for the age group or was not consistent with styles used in trend items.

Exhibit 4.4: 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 "PIRLS 2021 Survey Operations Procedures Unit 4: Preparing the Assessment Instruments" (sections relevant for your country) clear and sufficient?	56	1	1
Did you encounter any major problems using the assessment instrument InDesign/RTF production/ translation materials (used for preparing the paper context questionnaires and achievement booklets, including the digitalPIRLS bridge booklets)?	3	54	1
Did you encounter any major problems using the digitalPIRLS Translation System for preparing the digitalPIRLS achievement test and Student Questionnaire?	11	16	1
After the translation verification (IEA Amsterdam), did you correct your translations/adaptations as suggested by the verifier in the majority of cases?			
paperPIRLS achievement booklets	41	1	1 *(Not Answered)* 15 *(Not Applicable)*
digitalPIRLS bridge booklets	24	2	1 *(Not Answered)* 31 *(Not Applicable)*
Context questionnaires	55	1	1 *(Not Answered)* 1 *(Not Applicable)*
digitalPIRLS achievement test	25	1	1 *(Not Answered)* 31 *(Not Applicable)*

Exhibit 4.4: Survey Activities Questionnaire, Section Two—Translating, Adapting, and Producing Assessment Instruments (Numbers of NRC Responses) (Continued)

Question	Yes	No	Not Answered
After the layout verification (TIMSS & PIRLS International Study Center), did you correct your assessment instruments as noted by the verifier in the majority of cases?			
paperPIRLS achievement booklets	41	1	1 *(Not Answered)* 15 *(Not Applicable)*
digitalPIRLS bridge booklets	25	1	1 *(Not Answered)* 31 *(Not Applicable)*
Context questionnaires	56	0	1 *(Not Answered)* 1 *(Not Applicable)*
digitalPIRLS achievement test	27	0	1 *(Not Answered)* 30 *(Not Applicable)*
Did you apply any quality control measures to check paper assessment instruments during the printing process (e.g., checking for missing pages, upside down pages, text too bright or too dark)?	56	1	1
Did you experience any problems with the digitalPIRLS Player(s) contents and, if applicable, preparing the digitalPIRLS USB sticks?	8	19	1
If applicable, did you apply quality control measures to check random digitalPIRLS USBs (e.g., number of files, size of the files, initiating the digitalPIRLS Player) before they were provided to schools?	20	1	1 *(Not Answered)* 6 *(Not Applicable)*
Did you take measures to protect the security of the assessment instruments during the preparing and duplicating process?	56	1	1
Did you detect any potential breaches in security of the assessment instruments?	2	55	1
Did you encounter any problems preparing the Online SurveySystem files for administering the school, teacher, and/or home (Early Learning Survey) questionnaires online?	8	29	1 *(Not Answered)* 20 *(Not Applicable)*

Nearly all of the countries conducted the recommended quality control checks during the process of printing the testing materials for paperPIRLS and preparing devices for digitalPIRLS. Samples of the printed material were checked for missing pages, pages in the wrong order, upside-down pages, and too dark or too light text. For digitalPIRLS, countries either randomly sampled USB sticks/tablets to ensure that the size of the files was correct and/or that they were operating properly or asked Test Administrators to check the USBs before test administration began.

Eight countries reported that they experienced problems with the IEA's OSS. They reported issues with adding national questions and noted that only one user at a time can use OSS to enter their translations. The NRCs noted the assistance and support from IEA Hamburg when using OSS.

Assessment Administration

The third section of the Survey Activities Questionnaire addressed the extent to which National Research Coordinators were notified about errors in the testing materials sent to schools. As shown in Exhibit 4.5, a small number of errors were found in the materials. More than half of such errors were corrected before distributing the materials to the respondents. Errors found after distribution were mostly minor and were either fixed by School Coordinators or were replaced with help from the national centers.

Exhibit 4.5: Survey Activities Questionnaire, Section Three—Assessment Administration (Numbers of NRC Responses)

Question	Yes	No	Not Answered
Was the information provided in the "PIRLS 2021 Survey Operations Procedures Unit 5: Conducting the Data Collection" clear and sufficient?	56	1	1
Were any errors detected in any of the following assessment materials after they were sent to schools?			
paperPIRLS achievement booklets	4	35	1 *(Not Answered)* 18 *(Not Applicable)*
paperPIRLS achievement booklet ID labels	5	34	1 *(Not Answered)* 18 *(Not Applicable)*
digitalPIRLS bridge booklets	4	23	1 *(Not Answered)* 30 *(Not Applicable)*
digitalPIRLS bridge booklet ID labels	2	27	1 *(Not Answered)* 28 *(Not Applicable)*
digitalPIRLS files on USB sticks	2	22	1 *(Not Answered)* 33 *(Not Applicable)*
Student Questionnaire (digital or paper)	1	56	1 *(Not Answered)* 0 *(Not Applicable)*
Student Questionnaire ID labels	1	54	1 *(Not Answered)* 2 *(Not Applicable)*
Early Learning Survey	0	53	1 *(Not Answered)* 4 *(Not Applicable)*
Early Learning Survey ID labels	1	52	1 *(Not Answered)* 4 *(Not Applicable)*

Exhibit 4.5: Survey Activities Questionnaire, Section Three—Assessment Administration (Numbers of NRC Responses) (Continued)

Question	Yes	No	Not Answered
Student Tracking Forms	5	52	1 *(Not Answered)* 0 *(Not Applicable)*
Teacher Questionnaires	0	56	1 *(Not Answered)* 1 *(Not Applicable)*
Teacher Tracking Forms	1	49	1 *(Not Answered)* 7 *(Not Applicable)*
School Questionnaire	1	56	1 *(Not Answered)* 0 *(Not Applicable)*
School Coordinator Manual	2	55	1 *(Not Answered)* 0 *(Not Applicable)*
Test Administrator Manual(s)	1	56	1 *(Not Answered)* 0 *(Not Applicable)*
If any errors were detected, did you correct the error(s) before the testing began?	15	10	1 *(Not Answered)* 32 *(Not Applicable)*
Did you provide access to the Data Protection Declaration (provided by IEA and/or prepared by your country) to respondents in your country?	29	10	1 *(Not Answered)* 18 *(Not Applicable)*
Does your country have a confidentiality policy that restricts putting respondents' names on tracking forms and assessment instrument covers?	15	42	1
Did you encounter any problems translating and/or adapting the School Coordinator Manual(s)?	1	56	1
Did you encounter any problems translating and/or adapting the Test Administrator Manual(s)?	0	57	1
Were most/all School Coordinators appointed from within the participating schools?	54	3	1
Did you hold formal training session(s) for School Coordinators?	37	20	1
Were most/all Test Administrators trained by School Coordinators within the participating schools?	31	26	1
Did the Test Administrators document any problems or special circumstances that occurred frequently during the assessment administration (please refer to the completed Test Administration Forms)?	22	35	1
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?	6	31	1 *(Not Answered)* 20 *(Not Applicable)*

Exhibit 4.5: Survey Activities Questionnaire, Section Three—Assessment Administration (Numbers of NRC Responses) (Continued)

Question	Yes	No	Not Answered
What kind of devices did you use for the digitalPIRLS testing?			
PCs	16	0	1
Apple computers	0	0	1
Apple iPads	0	0	1
Android tablets	1	0	1
Chromebooks	0	0	1
A mix of the above or other	10	0	1
Who did the devices used for digitalPIRLS testing belong to?			
Participating schools	9	0	1
Outsourced company	3	0	1
National center	2	0	1
A combination of above	13	0	1
Which delivery method did you use to administer digitalPIRLS in your country?			
Individual PCs/USB sticks	9	0	1
Online Administration	5	0	1
Local PC server method	0	0	1
A mix of methods was used	13	0	1
Did you require/suggest/provide an additional person to help the Test Administrators during the digitalPIRLS testing sessions?	22	5	1
Did you experience any software-specific problems with the digitalPIRLS Player(s)?	12	15	1
Did you have a sufficient number of devices available for all/most schools to test all of the selected students (the whole class) at the same time?	18	9	1
Did the COVID-19 pandemic significantly affect the assessment administration process in the participating schools?	40	17	1

In order for the PIRLS study to comply with the requirements of the General Data Protection Regulation (GDPR) law of the European Union, IEA provided countries with templates for a Data Protection Declaration for each of the PIRLS 2021 Context Questionnaires. The templates specifically reflected the content of the questionnaire, and were fully compliant with the GDPR of

Europe. All European countries prepared a Data Protection Declaration, complying with the GDPR and country-specific amendments to the law, and provided it to respondents along with each of the PIRLS 2021 national context questionnaires. Some non-European participating countries also adapted and adopted the declaration as required by law in those countries. Altogether, 29 National Research Coordinators responded that they prepared and provided Data Protection Declarations along with national context questionnaires.

One National Research Coordinator reported difficulties translating the School Coordinator Manual. Primarily, the reported problem arose when the manual(s) had to be reorganized or adapted. Countries administering both digitalPIRLS and bridge booklets also had two sets of manuals to prepare.

In 54 countries, School Coordinators were appointed from within the participating schools. In the remaining countries, School Coordinators came from the national center or were contracted externally. In many countries, the National Research Coordinators organized centralized training sessions for School Coordinators. In others, training was conducted through webinars and online and written materials, with phone support as needed. In 31 countries, Test Administrators were trained by the School Coordinators within the participating schools. In the remaining countries, Test Administrators were trained by members of the national center staff.

Although the PIRLS administration mostly went well, Test Administrators occasionally reported difficulties. Among the problems documented by Test Administrators were the following: technical issues (e.g., system not accessible for a moment, crashing player, internet access interruption), issues related to the COVID-19 pandemic such as a high absence rate and rules about number of students in a classroom for social distancing, loud noises outside the classroom, students having difficulty with the test being too long, students being unfamiliar with some digital skills such as scrolling, and students not having enough time to complete the test.

Six countries that administered the school, teacher, and/or home questionnaires online reported issues. These issues were mostly related to login information or technical issues. IEA Hamburg assisted countries in resolving the issues.

In most countries administering PIRLS 2021 digitally, at least one additional person helped the Test Administrators during the testing sessions. This was usually the classroom teacher, School Coordinator, or an information technology (IT) consultant/expert. Several countries had two people per classroom to help with computer setup as well as to troubleshoot any technical issues that occurred during the testing session. Several national centers offered telephone support for possible technical issues as well.

Twelve digitalPIRLS countries reported that some software-specific problems occurred. These issues were mostly related to logging in to the PIRLS Player or the system freezing or crashing during testing. In some rare cases, students experienced an item or some items in an incorrect

language for one or both parts of the achievement test. These cases were investigated by the TIMSS & PIRLS International Study Center and RM, and the issues were resolved right after such occurrences were reported. In most schools observed, PIRLS 2021 was successfully administered.

Eighteen of the 28 reporting countries administering PIRLS digitally had enough computers or tablets for each school to test all students in the selected class(es) at the same time. Where this was not the case, schools held multiple sessions with two to four sessions per school.

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 national quality assurance activities, National Research Coordinators were instructed to send National Quality Control Observers to ten percent of the participating schools to observe the PIRLS 2021 test administration and to document compliance with the prescribed procedures. The national program was in addition to the program of International Quality Control visits conducted by IEA and the TIMSS & PIRLS International Study Center. Some countries did not use national monitors mainly due to restrictions placed on traveling within their countries and on visitors to schools during the COVID-19 pandemic. In these cases, countries made additional efforts when training Test Administrators or used phone calls, surveys, and national center staff to gather information.

As shown in Exhibit 4.6, almost all of the national centers that conducted a quality assurance program did so using the National Quality Control Monitor Manual provided by the TIMSS & PIRLS International Study Center. Among the documented problems detected by the national monitors were some technical issues logging into the PIRLS Player, students having difficulty with scrolling during the test, and the need to plan testing around set schedules for school breaks.

Exhibit 4.6: 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?	42	15	1
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?	41	7	1 *(Not Answered)* 9 *(No NQC Program)*
Did your national quality control monitors (NQCMs) document any major problems or special circumstances that occurred frequently during the assessment administration?	6	41	1 (*Not Answered*) 10 *(No NQC Program)*

Preparing for and Scoring the Constructed Response Items

Exhibit 4.7 provides data on responses to items asking National Research Coordinators about their experiences preparing for and scoring the constructed response items. Almost all National Research Coordinators found the scoring procedures as explained in *Survey Operations Procedures Unit 6: Scoring the Constructed Response Items* to be clear and sufficient. Countries reporting problems with the scoring training materials asked for more "borderline" examples, more practice with the items in passages new to PIRLS 2021, and more detailed explanations within the scoring guides. More than half of National Research Coordinators reported creating their own national examples and practice papers for training their scorers, as suggested by the TIMSS & PIRLS International Study Center.

Exhibit 4.7: 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 "PIRLS 2021 Survey Operations Procedures Unit 6: Scoring the Constructed Response Items" clear and sufficient?	55	2	1
Did you encounter any major problems using the scoring training materials, provided by the TIMSS & PIRLS International Study Center?	8	48	2
Did you create national scoring training materials in addition to the international scoring training materials?	32	25	1
Did you scan any paper achievement booklets for electronic image scoring?	13	44	1
Did you encounter any major procedural problems during the PIRLS 2021 constructed response item scoring in your country?	7	48	3
Did you encounter any major problems with the Online Scoring System (IEA's CodingExpert Software)?	6	50	2
Did all your scorers participate in scoring student responses of the trend items, including the Trend Reliability Scoring?	30	15	3 *(Not Answered)* 10 *(Not Applicable - not a trend country)*
Did all your scorers participate in the Cross-country Reliability Scoring?	22	31	3 *(Not Answered)* 2 *(Not Applicable - no English-speaking scorers)*

Thirteen countries scanned their paperPIRLS achievement booklets and scored student responses electronically. A small number of countries reported some minor problems using IEA's CodingExpert Software, which was used for all digital student responses. The software also was

used for scanned paper responses for the trend and cross-country reliability scoring tasks in both paper and digitalPIRLS countries. The reported problems included issues "logging into" each item (which was in response to an intended behavior), some difficulties assigning items to scorers, and problems with some of the scanned images.

Because only English responses were used for the cross-country reliability scoring task, not all scorers were able to participate. For the countries that did not participate in the previous cycle of PIRLS, the question on the trend reliability scoring procedures did not apply.

Creating and Submitting the Databases and Documentation

The last section of the Survey Activities Questionnaire addressed data entry of the paper assessment instruments, test administration data entry, and data quality control activities. As shown in Exhibit 4.8, almost all of the National Research Coordinators found the instructions in *Survey Operations Procedures Unit 7: Creating and Submitting the Databases* to be clear and sufficient. Nine National Research Coordinators reported issues when using WinW3S, mainly related to import and export functions. IEA Hamburg was able to provide support to countries as needed and solve the issues.

Exhibit 4.8: Survey Activities Questionnaire, Section Six—Creating and Submitting the Databases and Documentation (Numbers of NRC Responses)

Question	Yes	No	Not Answered
Was the information provided in the "PIRLS 2021 Survey Operations Procedures Unit 7: Creating and Submitting the Databases" clear and sufficient?	54	1	3
Did you encounter any problems entering test administration information and exporting your WinW3S database(s)?	9	45	4 *(Not Answered)* 0 *(Not Applicable)*
Who primarily entered the data for your country?			
National center staff	23	0	3
Temporarily hired data entry staff	15	0	3
An external data entry firm	3	0	3
Combination of the above	12	0	3
Other	2	0	3

Exhibit 4.8: Survey Activities Questionnaire, Section Six—Creating and Submitting the Databases and Documentation (Numbers of NRC Responses) (Continued)

Question	Yes	No	Not Answered
Did you use manual (key) data entry to enter paper instrument data for your country?			
paperPIRLS achievement booklets	33	4 *(optical scanning)*	3 *(Not Answered))* 18 *(Not Applicable)*
digitalPIRLS bridge booklets	19	7 *(optical scanning)*	3 *(Not Answered)* 29 *(Not Applicable)*
Context questionnaires	45	8 *(optical scanning)*	3 *(Not Answered)* 2 *(Not Applicable)*
Did you encounter any major problems using the IEA's Data Management Expert (DME) software?	3	51	4
If you entered paper data manually, did you enter 5% of each assessment instrument twice as a quality control measure?	45	2	4 *(Not Answered)* 7 *(Not Applicable)*
Did you apply all the data quality checks described in the "PIRLS 2021 Survey Operations Procedures Unit 7: Creating and Submitting the Databases" before submitting your data and documentation to IEA Hamburg?	53	0	5
Have you stored all the hard copy assessment instruments in a secure storage area until the original documents can be destroyed?	54	0	4

In 23 countries, the national center staff entered data from the paper instruments, and the remaining 32 countries used a combination of national center staff, temporarily hired staff, and external data entry firms. Some countries used optical scanning instead of manual data entry. All countries that responded to the questionnaire applied all required data quality checks. All countries that responded reported having securely stored their original assessment instruments until all data were processed and reported.