

4

Assessment Design and Specifications

As in 2001, the PIRLS 2006 assessment includes a written test of reading comprehension to measure fourth-grade students' reading literacy achievement and a series of questionnaires focusing on contexts for reading literacy development to gather information about the contexts for developing reading literacy.

Reporting Reading Achievement

PIRLS 2006 will report reading literacy achievement of fourth-grade students in each country as well as achievement by reading purpose and comprehension process. To keep the assessment burden on any one student to a minimum, each student is presented with only part of the assessment, as described below. Following data collection, student responses are placed on a common scale using item response theory methods that provide an overall picture of the assessment results for each country. As well as a scale for overall reading literacy, there will be separate scales for the two purposes for reading:

- reading for literary experience
- reading to acquire and use information.

To provide information on the processes of reading comprehension, PIRLS 2006 also will provide separate scales for two processes of comprehension (Mullis, Martin, & González, 2004):

- retrieval and straightforward inferencing
- interpreting, integrating, and evaluating.*

Test Booklet Design

Given the broad coverage and reporting goals of the PIRLS framework and its emphasis on the use of authentic texts, it was inevitable that the specifications for the item pool would include extensive testing time. The PIRLS 2006 Reading Development Group found that a valid assessment of two purposes for reading, reading for literary experience and reading to acquire and use information, with reliable measures of two processes of comprehension required nearly seven hours of testing time. While the assessment material that can be presented in that time should provide good coverage of the reading material children meet in their everyday lives, it is not reasonable to expect to administer the entire set of reading passages and test items to any one child. Because of the difficulties of scheduling student assessments and because young children cannot

* Retrieval and straightforward inferencing will combine items from the Focus on and retrieve explicitly stated material and Make straightforward inferences comprehension processes. Similarly, Interpreting, integrating, and evaluating will be based on items from the Interpret and integrate ideas and information and Examine and evaluate content, language, and textual elements processes.

be subjected to long testing periods, the testing time is limited to 80 minutes per student, with an additional 15–30 minutes for a student questionnaire.

With a total testing time of nearly seven hours but just one hour and twenty minutes per student, the assessment material must be divided among students in some way. The PIRLS design uses a matrix sampling technique, whereby the passages and accompanying items are divided into groups or blocks, and individual student booklets are made up from these blocks according to a plan. In PIRLS 2006, the nearly seven hours of testing time is divided into ten 40-minute blocks of passages and items, labeled L1–L5 for the literary passages and I1–I5 for the informational texts (see Figure 3). Four of the ten blocks were retained from PIRLS 2001 to provide a foundation for measuring trends in reading achievement; the remaining six were developed specifically for the 2006 assessment.

Figure 3
PIRLS 2006 Matrix-Sampling Blocks

Purpose for Reading	Block				
Literary Experience	L1	L2	L3	L4	L5
Acquire and Use Information	I1	I2	I3	I4	I5

In the PIRLS 2006 design, the ten blocks are distributed across 13 booklets (see Figure 4). Each student booklet consists of two 40-minute blocks of passages and items. Each student responds to one assessment booklet and a student questionnaire. So as to present at least some passages in a more natural, authentic setting, two blocks (one literary and one informational) are presented in color in magazine-type format with the questions in a separate booklet. This booklet is referred to as the PIRLS “Reader.”

To enable linking among booklets, at least some blocks must be paired with others. Since the number of booklets can become very large if each block is to be paired with all other blocks, it was necessary to choose judiciously among possible block combinations. In the 13-booklet design used in PIRLS 2006, 12 test booklets are derived by combining four literary (L1, L2, L3, and L4) and four informational (I1, I2, I3, and I4) blocks. The 13th booklet, the Reader, accounts for the remaining literary block, L5, and informational block, I5.

Figure 4
PIRLS 2006 Student Booklet Design

Booklet	Part 1	Part 2
1	L1	L2
2	L2	L3
3	L3	L4
4	L4	I1
5	I1	I2
6	I2	I3
7	I3	I4
8	I4	L1
9	L1	I1
10	I2	L2
11	L3	I3
12	I4	L4
Reader	L5	I5

In this design, each of blocks L1 through L4 and I1 through I4 appear in three of the 12 booklets, each time paired with another, different, block. For example, literary block L1 appears with literary block L2 in booklet 1 and with informational blocks I4 and I1 in booklets 8 and 9. Similarly, literary block L2 appears not only with L1 in booklet 1 but also with literary block L3 in booklet 2 and with informational block I2 in booklet 10.

The pairing of blocks in booklets 1 through 12 ensures that there are good links both among the literary and among the informational passages and also between the two purposes for reading. The blocks in the Reader, L5 and I5, are not linked to any other blocks directly. However, because booklets are assigned to students using a randomized procedure, the group of students responding to the Reader is equivalent to those responding to the other booklets, within the margin of error of the sampling process. The Reader will be distributed so that the same proportion of students will respond to blocks L5 and I5 as to each of the other literary and informational blocks.

Selecting Reading Passages for the Assessment

To reach the goal of approximating an authentic reading experience in the assessment, the reading passages presented to students must be typical of those read by students in their everyday experiences. Texts that exist for students to read in and outside school have typically been written by successful authors who understand writing for a young audience. These are more likely than passages written specifically for a test to elicit the full range of comprehension processes. Furthermore, they are more likely to engage students' interests, and to yield assessment questions that will elicit a range of responses to text that are similar to those elicited in authentic reading experiences. In the context of an international study, attaining authenticity in the assessment reading experience may be somewhat constrained by the need to translate a text into numerous languages. Thus, care is taken to choose texts that can be translated without loss in meaning or in potential for student engagement.

In selecting texts for use in an international survey of reading literacy, the potential for cultural bias must be considered. The set of texts used must range as widely as possible across nations and cultures. No country or culture should be over-represented in the assessment texts. Text selection thus involves collecting potential stimulus texts from as many countries as possible. The final selection of texts is based, in part, on the national and cultural representation of the entire set of assessment texts. Texts that depend heavily on culture-specific knowledge are excluded.

The appropriateness and readability of texts for assessing fourth-grade students is determined through review by educators and curriculum specialists from countries participating in the assessment. Among the criteria used to select texts are topic and theme appropriateness for the grade level; fairness and sensitivity to gender, racial, ethnic, and religious considerations; nature and level of linguistic features; and density of information. The Fry Readability Index results are also provided. In addition, the time constraints of the test situation place some limits on the length of texts. Generally, texts selected will be no longer than 1,000 words so students have time to read the entire passage and answer the comprehension questions. However, length will vary somewhat because other text characteristics also affect rate of reading.

As a basis for measuring trends from 2001, PIRLS retained four passages and items from the 2001 assessment – two literary and two informational – to be included in the PIRLS 2006 assessment. To complete the design, six new passages and associated items were developed – three literary and three informational.

Question Types and Scoring Procedures

Students' ability to comprehend text through the four comprehension processes is assessed via comprehension questions that accompany each text. Two question formats are used in the PIRLS assessment – multiple-choice and constructed-response. Each multiple-choice question is worth one point. Constructed-response questions are worth one, two, or three points, depending on the depth of understanding required. Up to half of the total number of points

represented by all of the questions will come from multiple-choice questions. In the development of comprehension questions, the decision to use either a multiple-choice or a constructed-response format is based on the process being assessed, and on which format best enables test takers to demonstrate their reading comprehension.

Multiple-Choice Questions. Multiple-choice questions provide students with four response options, of which only one is correct. Multiple-choice questions can be used to assess any of the comprehension processes. However, because they do not allow for students' explanations or supporting statements, they may be less suitable for assessing students' ability to make more complex interpretations or evaluations.

In assessing fourth-grade students, it is important that linguistic features of the questions be developmentally appropriate. Therefore, the questions are written clearly and concisely. The response options are also written succinctly in order to minimize the reading load of the question. The options that are incorrect are written to be plausible, but not deceptive. For students who may be unfamiliar with this test question format, the instructions given at the beginning of the test include a sample multiple-choice item that illustrates how to select and mark an answer.

Constructed-Response Questions. For this type of test item students are required to construct a written response, rather than select a response from a set of options. The emphasis placed on constructed-response questions in the PIRLS assessment is consistent with the definition of literacy underlying the framework. It reflects the interactive, constructive view of reading – meaning is constructed through an interaction between the reader, the text, and the context of the reading task. This question type is used to assess any of the four comprehension processes. However, it is particularly well suited for assessing aspects of comprehension that require students to provide support or that result in interpretations depending upon students' background knowledge and experiences.

In the PIRLS assessment, constructed-response questions will be worth one, two, or three points, depending on the depth of understanding or the extent of textual support the question requires. In these questions, it is important to provide enough information to help students understand clearly the nature of the response expected.

Scoring guides for each constructed-response question describe the essential features of appropriate and complete responses. They focus on evidence of the type of comprehension the question assesses. They describe evidence of partial understanding and evidence of complete or extensive understanding. In addition, sample student responses at each level of understanding provide important guidance to raters.

In scoring students' responses to constructed-response questions, the focus is solely on students' understanding of the text, not on their ability to write well. Also, scoring takes into account the possibility of various interpretations that may be acceptable, given appropriate textual support. Consequently, a wide range of answers and writing ability may appear in the responses that receive full credit to any one question.

Score Points. In developing the assessment, the aim is to create blocks that each provide, on average, at least 15 score points – made up of approximately seven multiple-choice items (1 point each), two or three short-answer items (1 or 2 points each), and one extended-response item (3 points). The exact number of score points and the exact distribution of question types per block will vary somewhat, as different texts yield different types of questions.

Releasing Assessment Material to the Public

PIRLS 2006 is the second of a regular five-year cycle of studies that will provide data on trends in reading literacy. Administered for the first time in 2001, PIRLS will be administered again in 2011, 2016, and so on into the future. The design provides for the release of many of the passages and items into the public domain as the international reports are published, while safeguarding the trend data by not releasing a substantial proportion of the items. As passages and

items are released, new assessment material will be developed to take their place.

According to the PIRLS design, two blocks of literary passages and two blocks of informational passages from the 2001 assessment were published after the data collection, including the literary and informational passages in the PIRLS Reader. These will be replaced with new passages and items for the 2006 assessment. Following the 2006 data collection, another four blocks will be released, two from the original 2001 assessment (L2 and I2) and two from those developed for 2006 (L5 and I5).

Background Questionnaires

An important purpose of PIRLS is to study the home and school factors associated with children's reading literacy by the fourth grade. To that end, PIRLS will administer questionnaires to students, their parents, their teachers, and the principals of their schools. The questions are designed to measure key aspects of students' home and school environments.

Student Questionnaire. This questionnaire will be completed by each student who takes the PIRLS reading test. It asks about aspects of students' home and school lives, including classroom experiences and reading for homework, self-perception and attitudes toward reading, out-of-school reading habits, computer use, home literacy resources, and basic demographic information. The questionnaire requires 15-30 minutes to complete.

Learning to Read Survey. This short questionnaire is addressed to the parents or primary caregivers of each student taking part in the PIRLS data collection. It investigates child-parent literacy interactions, home literacy resources, parents' reading habits and attitudes, and home-school connections. Also, it collects basic demographic and socioeconomic information. Together with information collected from the students, parents' responses will provide a more complete picture of an important context for learning to read. This questionnaire is designed to take 10–15 minutes to complete.

Teacher Questionnaire. The reading teacher of each fourth-grade class in PIRLS will be asked to complete this questionnaire, which is designed to gather information about classroom contexts for developing reading literacy. The questionnaire asks teachers about characteristics of the class tested, such as size, reading level and language ability of the students; instructional time, materials, and activities for teaching reading and promoting the development of students' reading literacy; grouping of students for reading instruction; classroom resources; assessment practices; and home-school connections. It also asks teachers their views on their opportunities for collaboration with other teachers and for professional development, and for information about themselves and their education and training. This questionnaire requires about 30 minutes of the teacher's time.

School Questionnaire. The principal of each school in PIRLS will be asked to respond to this questionnaire. It asks about enrollment and school characteristics, such as location, resources available in the surrounding area, and indicators of the socioeconomic background of the student body; instructional time; emphasis and materials used in reading instruction for students in primary grades; school resources, such as the availability of instructional materials and staff; home-school connections; and school climate. It is designed to take about 30 minutes.

Curriculum Questionnaire. To provide information about the goals of reading instruction, the national research coordinator in each country will complete a questionnaire about the country's reading curriculum, including national policy on reading, goals and standards for reading instruction, time specified for reading, and provision of books and other literary resources.

PIRLS 2006 Encyclopedia

The *PIRLS 2006 Encyclopedia* will provide a profile of each country's education system, with a particular focus on reading education for primary-school children. The volume will provide general data on economic and educational indicators and describe how the education system is organized and how decisions about education are made.

The reading curriculum, including goals, materials, and instruction, will be discussed, along with information on assessment of reading achievement.

