CHAPTER 1

Developing the PIRLS 2016 Achievement Items

Ina V.S. Mullis
Caroline O. Prendergast

Unique Characteristics of the 2016 PIRLS Assessment

The general approach to developing the PIRLS achievement items is similar from assessment cycle to assessment cycle, but each assessment cycle tends to have some unique characteristics that influence instrument development. Besides providing measures on another cycle for the PIRLS trend lines monitoring changes in educational achievement, 2016 also was remarkable for two reasons.

- It was the inaugural year of the ePIRLS extension of PIRLS. ePIRLS was introduced in 2016 to assess online informational reading skills in a simulated Internet environment and was administered via computer (PCs). In ePIRLS, students are assessed on their ability to acquire and use information from webpages while investigating science and social studies topics through authentic, school-like assignments.
- The PIRLS Reading Achievement scale was extended to include PIRLS Literacy, which took the place of prePIRLS. PIRLS Literacy 2016 advanced prePIRLS by linking a less difficult version of the PIRLS assessment to the well-established PIRLS metric to enable assessing reading comprehension across a broader range of countries. Countries whose students were not yet prepared to take PIRLS were still able to participate in this important international project by administering PIRLS Literacy. Countries' results for the two different versions are both reported on the PIRLS scale.

ePIRLS 2016: Extending PIRLS to Assess Online Reading

Recognizing that the Internet has become the primary source for obtaining information at work, at home, and for school, PIRLS 2016 was extended to include ePIRLS on a voluntary basis for countries already participating in PIRLS and where students were familiar with using computers. ePIRLS used an engaging simulated Internet environment to measure fourth grade students' achievement

in reading for informational purposes. The assessment was administered via computer using a PC platform. Countries were responsible for using their own computers. In most cases, countries used the computers available in the schools or arranged for rental computers. ePIRLS presented students with authentic school-like assignments about science and social studies topics, which align with purposes for school reading. Led by a teacher avatar, students were asked to navigate through multiple, interconnected webpages containing both textual and visual information. ePIRLS allows for assessing reading comprehension skills beyond those used in "traditional" print material.

In addition to the data collected through the PIRLS 2016 Context Questionnaires (see Chapter 2), ePIRLS has its own short student questionnaire pertaining to students' familiarity with computers and online reading. Also, some process data will be analyzed to study students' navigation patterns.

PIRLS Literacy 2016

For a variety of reasons, there are some countries where most children in the fourth grade are still developing fundamental reading skills. Therefore, IEA offers options for matching the PIRLS reading assessment to the country's educational development. For some countries, the PIRLS Literacy version of PIRLS is a better match with students' learning. New for 2016, PIRLS Literacy has been placed on the same scale as PIRLS, with the two versions (PIRLS and PIRLS Literacy) having four passages in common with one another.

PIRLS Literacy reflects the same conception of reading as does PIRLS except the assessment is less difficult. The PIRLS Literacy assessment is consistent with the PIRLS framework for assessing reading comprehension. However, typically the passages are shorter with less complex syntax, and the questions include a different mix of items across the comprehension processes compared to PIRLS. PIRLS Literacy places somewhat greater emphasis on straightforward retrieval of information compared to PIRLS and less emphasis on straightforward inferencing, interpreting and integrating ideas and information, and evaluating and critiquing content and textual elements.

PIRLS Literacy was developed together with PIRLS. It uses the same context questionnaires, and the expert committees reviewed both the PIRLS and PIRLS Literacy passages, items, and scoring guides together. The challenge was identifying a range of passages with content suitable for fourth grade students that could be used in PIRLS, in both PIRLS and PIRLS Literacy, and only in PIRLS Literacy. Also, PIRLS Literacy passages contain questions placed throughout the passages to enable students to answer questions as they proceed through the text, rather than the PIRLS approach of presenting the entire passage followed by the set of questions.

The PIRLS Approach to Measuring Trends

Because PIRLS is designed to measure trends, the assessments cannot change dramatically from cycle to cycle. That is, PIRLS is based on a well-known premise for designing trend assessments (ascribed to John Tukey and Albert Beaton):

> "If you want to measure change, do not change the measure."

However, the achievement tests and questionnaires also need to be updated with each cycle to prevent the assessments from becoming dated and no longer relevant to current learning goals and policy issues. It is important that the content reflects the most recent discoveries in the field and is presented in ways consistent with students' instructional and everyday experiences.

To maintain continuity with past assessments while keeping up with current topics and technology, the PIRLS assessments evolve with each cycle. PIRLS has a specific design for rotating passages and items out of the assessment after each cycle and replacing them with newly developed passages and items for the following cycle. The remaining assessment items are kept secure to be readministered in subsequent cycles.

The design for passage/item replacement provides for each assessment to include passages and items from three cycles—essentially, one-third newly developed, one-third from the previous cycle, and one-third from two cycles before. With permission from IEA the replaced assessment passages and items are available on a restricted use basis for educational and research purposes (please see http://www.iea.nl/copyright-notice for permissions information).

Overview of the PIRLS 2016 Development Process

According to the PIRLS assessment design, it is necessary to replace a specific portion of the passages and achievement items for each upcoming cycle. Although the majority of the assessment items are carried forward from the previous assessment cycle to measure trends, the task of updating the instruments for each new cycle—every five years for PIRLS since 2001—is a substantial undertaking. All of the passages, and subsequently the items, must be reviewed by experts and agreed upon by the diverse participating countries.

The TIMSS & PIRLS International Study Center at Boston College uses a collaborative process to select the passages and develop the new items needed for each PIRLS cycle. A broad overview of the process includes:

- Updating the frameworks for the upcoming assessment
- Identifying and selecting appropriate reading passages
- Developing items and their scoring guides in accordance with the frameworks

- Conducting a full-scale field test
- Selecting the new assessment items based on the frameworks, field test results, and to complement existing passages and items from previous cycles
- Conducting training in how to reliably score responses to constructed response items (i.e., questions to which students provide a written response rather than choosing from a set of options)

The development process is directed and managed by the staff of the TIMSS & PIRLS International Study Center at Boston College, who collectively have considerable experience in the measurement and assessment of reading achievement. For PIRLS 2016, Executive Directors Ina Mullis and Michael Martin managed the assessment development process.

Also playing a key role in achievement item development were the National Research Coordinators (NRCs) designated by their countries to be responsible for the complex tasks involved in implementing PIRLS in their countries. The TIMSS & PIRLS International Study Center worked with the NRCs and experts from the countries throughout the development process to identify suitable PIRLS passages and develop new test items. To provide additional subject-matter expertise and support, staff consulted closely with external reading specialists. Continuing from PIRLS 2006 and 2011, the PIRLS 2016 Chief Reading Consultant was Marian Sainsbury, National Foundation for Educational Research (NFER), London, England. The Reading Development Group (RDG) provided additional advice and guidance in developing the PIRLS assessment through periodic reviews. The countries participating in PIRLS nominate RDG members for each PIRLS cycle.

Exhibit 1.1 lists the eight members of the PIRLS 2016 RDG.

Exhibit 1.1: PIRLS 2016 Reading Development Group (RDG)

Julian Fraillon
Australian Council for Educational Research
Australia

Jan Mejding
Aarhus University
Department of Education
Denmark

Galina Zuckerman
Russian Academy of Education
Russian Federation

Elizabeth Pang
Ministry of Education
Singapore

Jenny Wiksten Folkeryd
Uppsala University
Sweden

Ahlam Habeeb Msaiqer
Abu Dhabi Education Council
United Arab Emirates, Abu Dhabi

Donald Leu
University of Connecticut
United States

Karen Wixson
University of North Carolina, Greensboro
United States

RDG members met four times for PIRLS 2016. At the first RDG meeting in Copenhagen, Denmark (July 2013), the RDG reviewed the reading frameworks, potential passages, ePIRLS prototypes, and draft item writing guidelines. At the second meeting in London, England (April 2014), the RDG reviewed PIRLS field test passages and items and ePIRLS field test tasks and items. At the third meeting in Stockholm, Sweden (July 2015), the RDG reviewed field test results and made recommendations to the NRCs regarding which passages and items to include in the 2016 assessments. At the final meeting in Lübeck, Germany (May 2017), the RDG conducted the PIRLS 2016 scale anchoring process (see Chapter 13).

During busy periods in between RDG meetings, the Chief Reading Consultant and several RDG members served as a task force to assist in completing specific tasks, such as updating the framework (PIRLS Framework Task Force) or developing items (PIRLS Item Development Task Force).

The PIRLS 2016 Development Schedule

To accomplish the development work in timely fashion, the assessment was developed over three years of the five-year cycle according to a specific timeline. Essentially, one year or so was devoted to updating the framework and identifying appropriate passages, the second year was devoted to item development, and the third year to conducting the field test and selecting the materials for data collection. (The fourth year of the cycle was data collection and the fifth was analysis and reporting.)

Exhibit 1.2 shows the PIRLS 2016 development schedule from updating the frameworks to data collection.

Exhibit 1.2: PIRLS 2016 Development Schedule for Achievement Items

Date(s)		Group and Activity
July-December	2012	To begin work on updates to the Assessment Framework for PIRLS 2016, the TIMSS & PIRLS International Study Center summarized the curricular emphases in reading described in the *PIRLS 2011 Encyclopedia*
December	2012	Task Force of reading experts proposed updates for the 2016 Assessment Framework, incorporating information from the Encyclopedia (Boston, USA)
January	2013	TIMSS & PIRLS International Study Center sent proposed Assessment Framework updates to National Research Coordinators (NRCs) in preparation for the 1st NRC Meeting
February	2013	TIMSS & PIRLS International Study Center presented plans for ePIRLS, including a sample task, and NRCs reviewed proposed updates to Assessment Framework at 1st NRC meeting (Hamburg, Germany)
February-July	2013	TIMSS & PIRLS International Study Center incorporated feedback from 1st NRC meeting to further refine the *PIRLS 2016 Assessment Framework*
March-September	2013	NRCs submitted and reviewed proposed reading passages in preparation for the 2nd NRC meeting (Portorož, Slovenia)
March-July	2013	TIMSS & PIRLS International Study Center developed prototype ePIRLS tasks
May	2013	NRCs received a promotional ePIRLS video, which illustrated ePIRLS using the Polar Bears task
July	2013	Reading Development Group (RDG) reviewed proposed Assessment Framework, passages, ePIRLS prototypes, and draft *PIRLS 2016 Item Writing Guidelines* at the first RDG meeting (Copenhagen, Denmark)
August	2013	TIMSS & PIRLS International Study Center updated *PIRLS 2016 Item Writing Guidelines*
September	2013	TIMSS & PIRLS International Study Center prepared final drafts of *PIRLS 2016 Assessment Framework*, incorporating RDG and NRC comments
September	2013	NRCs performed final review of the *PIRLS 2016 Assessment Framework*, selected passages, reviewed storyboards for ePIRLS tasks, and developed draft field test items at the 2nd NRC meeting (Portorož, Slovenia)
October-January	2013-2014	TIMSS & PIRLS International Study Center further refined draft field test items and scoring guides and continued to develop ePIRLS tasks
November	2013	TIMSS & PIRLS International Study Center published *PIRLS 2016 Assessment Framework (1st edition)*
February	2014	PIRLS/PIRLS Literacy Item Development Task Force reviewed and edited draft field test items and scoring guides (Boston, USA)
March-April	2014	ACER and AIR conducted cognitive labs for two sample ePIRLS tasks
April	2014	RDG reviewed PIRLS field test passages and items for PIRLS as well as storyboards for six ePIRLS tasks and items at 2nd RDG meeting (London, England)
April-May	2014	TIMSS & PIRLS International Study Center revised draft field test passages and tasks, as well as their items and scoring guides, to address RDG comments
May	2014	NRCs reviewed and approved PIRLS/PIRLS Literacy field test passages and items and reviewed storyboards for five ePIRLS tasks at 3rd NRC meeting (Dublin, Ireland)
May-July	2014	TIMSS & PIRLS International Study Center assembled field test passages and items into assessment booklets

Exhibit 1.2: PIRLS 2016 Development Schedule for Achievement Items (Continued)

Date(s)		Group and Activity
July	2014	TIMSS & PIRLS International Study Center posted PIRLS field test achievement booklets for NRCs
August	2014	ePIRLS NRCs reviewed storyboards and items for Rivers
September	2014	TIMSS & PIRLS International Study Center posted PIRLS Literacy field test achievement booklets for NRCs
October	2014	NRCS received final storyboards for Mars, Rainforests, Blackwell, Migration, Troy, and Rivers as well as the ePIRLS student questionnaire
October-December	2014	TIMSS & PIRLS International Study Center worked with each of five English-speaking countries to administer PIRLS to several classes to collect student responses to constructed response items in order to develop scoring training materials
November	2014	TIMSS & PIRLS International Study Center administered PIRLS Literacy passages in a range of classrooms in the Boston area to collect student responses to constructed response items in order to develop scoring training materials
November	2014	TIMSS & PIRLS International Study Center posted ePIRLS tasks, software, system check, online translation system, and test administrator manual for the pilot test
November	2014	ePIRLS pilot test conducted in Australia, Ireland, and Canada (Ontario) to test the ePIRLS tasks and software in a classroom setting and inform scoring guides and training materials
November-February	2014-2015	TIMSS & PIRLS International Study Center posted systems and materials for the ePIRLS field test
December	2014	PIRLS/PIRLS Literacy Item Development Task Force modified scoring guides for constructed response items based on student responses and developed scoring training materials for 4th NRC meeting (Boston, USA)
February	2015	*PIRLS 2016 Assessment Framework (2nd edition)* published online, incorporating the introduction to PIRLS Literacy and the new integrated PIRLS/PIRLS Literacy assessment design
February	2015	ePIRLS Task Force reviewed students' typed responses from the pilot and developed scoring training materials for 4th NRC meeting (Boston, USA)
February	2015	NRCs received scoring training for PIRLS, PIRLS Literacy, and ePIRLS 2016 constructed response field test items at 4th NRC meeting (Floriana, Malta)
March-April	2015	Countries conducted PIRLS, PIRLS Literacy, and ePIRLS 2016 field tests
April-May	2015	Countries submitted field test achievement data for analysis and review
June	2015	PIRLS/PIRLS Literacy Item Development Task Force reviewed field test item statistics
June	2015	ePIRLS Task Force reviewed field test item statistics
June-July	2015	TIMSS & PIRLS International Study Center assembled proposed PIRLS/PIRLS Literacy passages and items in preparation for the 3rd RDG meeting
July	2015	RDG reviewed proposed PIRLS/PIRLS Literacy passages and items in conjunction with field test results and reviewed five proposed ePIRLS tasks via computer at the 3rd RDG meeting (Stockholm, Sweden)
July	2015	ePIRLS NRCs received "Preparing Computers for ePIRLS" instructions

Exhibit 1.2: PIRLS 2016 Development Schedule for Achievement Items (Continued)

Date(s)		Group and Activity
August	2015	NRCs reviewed and approved PIRLS/PIRLS Literacy passages and items and ePIRLS storyboards for PIRLS 2016 data collection at 5th NRC meeting (Jyväskylä, Finland)
August	2015	IEA Hamburg provided information to NRCs about ePIRLS software and operations at the 5th NRC meeting (Jyväskylä, Finland)
August	2015	TIMSS & PIRLS International Study Center distributed PIRLS/PIRLS Literacy and ePIRLS 2016 data collection achievement materials to NRCs
September	2015	ePIRLS NRCs received access to the Online Translation System for main data collection
October-December	2015	Southern Hemisphere countries conducted PIRLS 2016 data collection
October	2015	TIMSS & PIRLS International Study Center updated and prepared materials for PIRLS/PIRLS Literacy 2016 constructed response scoring training
November	2015	NRCs from Southern Hemisphere countries received scoring training for PIRLS/PIRLS Literacy constructed response items (Buenos Aires, Argentina)
November	2015	TIMSS & PIRLS International Study Center finalized scoring guides and training materials for PIRLS/PIRLS Literacy and ePIRLS constructed response items and distributed them to NRCs
February-March	2016	NRCs from Northern Hemisphere countries received scoring training for PIRLS/PIRLS Literacy and ePIRLS constructed response items at 6th NRC meeting (Hong Kong SAR)
March-June	2016	Northern Hemisphere countries conducted PIRLS/PIRLS Literacy and ePIRLS 2016 data collection

Updating the Assessment Framework for PIRLS 2016

Updating the PIRLS assessment for 2016 began with reviewing and modifying the assessment framework that describes the aspects of reading comprehension to be assessed.

The basic structure of the PIRLS assessment framework is based on two dimensions: purposes for reading and processes of comprehension. Reading for literary experience and reading to acquire and use information are the two major purposes assessed by PIRLS because they account for many of the reading experiences of young children.

The four comprehension processes assessed by PIRLS are:

- Focusing on and retrieving explicitly stated information
- Making straightforward inferences
- Interpreting and integrating ideas and information
- Evaluating and critiquing content and textual elements

For PIRLS 2016, the name of the fourth comprehension process was changed to "Evaluate and Critique Content and Textual Elements" from "Examine and Evaluate Content, Language, and Textual Elements" in 2011. This newer category name better describes the processes students use when answering items assigned to this category, clarifying for item writers the kinds of items to be developed. Also, a new section was added to the framework that described the components of online reading that should be addressed in ePIRLS.

The NRCs from the participating countries discussed the framework updates at their first meeting. Following the discussion at the 1st NRC meeting in Hamburg in February 2013, the NRCs consulted with their national experts about the PIRLS updates for 2016. Next, the RDG reviewed and revised the frameworks. Using an iterative process, the NRCs once again reviewed the RDG's revised version of the framework, which was updated a final time prior to publication of the 1st edition in November 2013.

Following that, however, further discussions with the NRCs revealed dissatisfaction with the 2011 design where prePIRLS was reported separately from PIRLS. Thus, the PIRLS 2016 design was updated to strengthen the assessment of reading for children still developing fundamental reading skills. PIRLS Literacy was developed to extend the PIRLS achievement scale to address the needs of a broader range of countries. PIRLS Literacy is equivalent in scope to PIRLS, and they are linked with four passages in common. This enables results for both assessments to be reported on the same PIRLS scale. However, the new design necessitated updating the PIRLS 2016 Assessment Framework, and a 2nd edition was published in February 2015. The first chapter of the *PIRLS 2016 Assessment Framework (2nd Edition)* describes the aspects of reading comprehension to be assessed by PIRLS 2016 in detail.

Identifying Reading Passages for PIRLS and PIRLS Literacy

In total, 18 new passages and item sets needed to be developed and field tested for PIRLS and PIRLS Literacy 2016. The PIRLS 2016 assessment required field testing 12 passages (8 of which were newly developed PIRLS passages and 4 of which were newly developed to be shared between PIRLS and PIRLS Literacy), which included a total of 203 new items. In addition to the four shared passages, the PIRLS Literacy component also required field testing 6 new passages, which included a total of 173 items.

Identifying appropriate passages for the PIRLS and PIRLS Literacy assessments was critical to their success, because readers make meaning from text in a variety of ways, depending not only on the purpose for reading but also on the difficulty of the text and the reader's prior knowledge. Examples of literary texts include contemporary short stories as well as traditional tales and fables. Informational texts can be from a variety of sources, such as informational books, textbooks, and journal articles and may include graphic support in the form of charts, tables, or diagrams.

At the beginning of the assessment cycle, the TIMSS & PIRLS International Study Center sent a call for passages to all NRCs. The criteria for suitable passages was discussed at the first NRC meeting in Hamburg in February 2013. In general, the PIRLS 2016 Chief Reading Consultant, Marian Sainsbury from NFER, explained that passages should:

- Be suitable for fourth grade students in content, interest, and reading ability
- Be well written in terms of depth and complexity to allow for a sufficient number of questions
- Avoid bias in that they are sensitive to cultural differences and are likely to be equally familiar or unfamiliar to all students

In March 2013, the TIMSS & PIRLS International Study Center created a discussion board so NRCs could review passages as they were submitted. At the same time, TIMSS & PIRLS International Study Center staff and the Chief Reading Consultant also began the search for suitable materials.

In conjunction with a qualitative evaluation of each text's characteristics and appropriateness for different languages and cultures, text length and readability guided passage selection. The TIMSS & PIRLS International Study Center computed the word count and readability for each passage as a quantitative check of the grade appropriateness of the recommended texts. The Flesch-Kincaid Grade Level Formula[1] was used as a measure of readability for this purpose because of its suitability for a wide range of texts and its extensive use in education. This quantitative information was provided alongside the texts to NRCs for their review.

The NRCs and the RDG conducted an iterative passage review process at meetings and online. During the year or so allocated to find texts, the NRCs and the RDG reviewed hundreds of passages in order to identify the approximately 18-20 passages for PIRLS and PIRLS Literacy that were needed to develop items for the field test. The TIMSS & PIRLS International Study Center relied on the professional judgment of the NRCs and their within-country experts to evaluate the grade appropriateness, translatability, and cultural suitability of the texts for their students.

The TIMSS & PIRLS International Study Center prepares an international version of all the PIRLS and ePIRLS assessment items in English. Subsequently, the items are translated by participating countries into their languages of instruction with the goal of creating high quality translations that are appropriately adapted for the national context and at the same time are internationally comparable. Therefore, a significant portion of the development and review effort by NRCs is dedicated to ensuring that the passages can be translated accurately.

1 See Kincaid, Fishburne, Rogers, and Chissom (1975).

Developing Website Texts and Items for ePIRLS

Reading for informational purposes on the Internet requires many of the same reading comprehension skills and strategies as does reading offline. However, reading online also requires some new skills and strategies and is done in a different environment containing a wider variety of texts. Developing ePIRLS involved creating six tasks that included simulated Internet webpages with multiple pages of text, and included 115 items in total.

Developing appropriate and engaging webpages for each ePIRLS assessment task involved creating a variety of texts that fit into an integrated website focused on a science or social studies topic. The texts included written descriptions and explanations, diagrams, interactive images and maps, and animated graphics. ePIRLS website text development followed the same guidelines as for PIRLS passages, taking into consideration suitability for fourth grade students regarding content, interest, reading ability, complexity, and cultural sensitivity.

Especially since it was for the first time, developing the ePIRLS tasks was extremely arduous and time consuming. The TIMSS & PIRLS International Study Center developed four ePIRLS tasks, all based on the TIMSS 2015 Science Framework for the fourth grade. The first task developed, called "Polar Bears," was about how the melting ice in the northern Polar Regions is affecting the habitat of the polar bears. The idea of a website about polar bears was part of the presentation on extending PIRLS 2016 to assess online reading that the TIMSS & PIRLS International Study Center made at the 1st PIRLS NRC meeting, where both ePIRLS and the topic of polar bears were well received by the NRCs.

There was considerable information about the polar bears topic on the Internet including a variety of texts and images. Developing the ePIRLS task proceeded slowly, involving sorting through choices and creating simulated websites that could be examined by the students. The staff at the TIMSS & PIRLS Study Center, including the Executive Directors, the Director of the Production Department, and the Communications Specialist, carefully selected the webpages for each of several websites (e.g., about polar bears, maps and data about polar ice melting, and opinions about the future of polar bears) and drafted the narrative for the teacher avatar, Mr./Ms. Webster. The TIMSS & PIRLS International Study Center owes a debt of gratitude to Dr. Don Leu who pioneered the idea of the teacher avatar and was a member of the PIRLS 2016 RDG.

The teacher avatar guided the students through the websites in the polar bear task, asking various questions about the information in the webpages. Answering the questions required students to navigate to the appropriate webpages and read various content. For some questions, students could choose their answers from multiple-choice questions or drop-down menus, and for other questions they were asked to type in their answers.

Once the Internet images were selected, the ideas for the websites created, and the script was drafted, the production staff at the TIMSS & PIRLS International Study Center prepared

storyboards covering the Polar Bear task from beginning to end. The Polar Bear storyboards provided the foundation for disseminating and reviewing the idea of ePIRLS with the NRCs and the RDG, and also allowed the TIMSS & PIRLS International Study Center to consult with Dr. Leu and his staff about how programming the tasks would work. Eventually, the Polar Bear task became the basis for a video the TIMSS & PIRLS International Study Center prepared to explain the characteristics of ePIRLS.

While the TIMSS & PIRLS International Study Center was working on task development, IEA Hamburg was working on the systems necessary to administer ePIRLS via PC. This included: an online translation system, a systems check, ePIRLS software so that USB sticks could be used to load the assessment tasks onto the countries' computers, provision to upload the student data to the IEA server in Hamburg, and a system in Hamburg to capture the data for scoring. The online translation system enables translators to adapt the international version of the ePIRLS tasks, including items and website text, into a target language directly in the online system. Additionally, the online translation system enables translators to review, revise, and verify translated text. The system check program allows test administrators to quickly check whether a given computer is able to support the ePIRLS software as delivered by the USB sticks or a local server. The data monitoring system allows NRCs to monitor collected student data through an online portal. The online scoring system streamlines the scoring process by providing scorers with student responses, scoring guides, and scoring capabilities for constructed response items.

Subsequent to the work with the Polar Bear task and the creation of the video, the TIMSS & PIRLS International Study Center developed three more tasks for the ePIRLS field test. The tasks were based on science topics and were developed using the same procedure of identifying websites and drafting a script. Then storyboards were developed, reviewed by the NRCs and RDG, and revised. Only then, were storyboards given to IEA Hamburg for programming the ePIRLS software.

Led by RDG member Dr. Julian Fraillon, the Australian Council for Educational Research (ACER) also developed three ePIRLS tasks. These tasks were in social science areas and followed a similar development path. The ideas and concepts were discussed with the TIMSS & PIRLS International Study Center and those selected for further development were then plotted out. The websites/webpages and scripts were reviewed by the TIMSS & PIRLS International Study Center before ACER drafted storyboards. The draft storyboards were thoroughly reviewed by the NRCs and RDG. The TIMSS & PIRLS International Study Center made the final revisions to the storyboards and forwarded them to IEA Hamburg for programming.

In the spring of 2014, the TIMSS & PIRLS International Study Center arranged for the American Institute for Research (AIR) to conduct cognitive labs in Washington, D.C. and ACER to conduct them in Camberwell, Victoria, Australia. Two ePIRLS tasks with 38 items in total were presented to approximately 21 students using an initial version of the test administration software. These students were observed and prompted to answer questions about the clarity, difficulty,

and familiarity of the item content and format, as well as questions about the simulated Internet environment and teacher avatar. As the students completed the tasks, their interactions with the software were monitored and recorded in order to collect information about the strengths and weaknesses of the software and the testing experience. The TIMSS & PIRLS International Study Center received the cognitive lab reports in the summer of 2014.

Based on the information from the cognitive labs, six ePIRLS tasks were developed, reviewed, and programmed for inclusion in the ePIRLS pilot. The pilot took place in October and November 2014 in Australia, Ireland, and Canada (Ontario). This process provided an additional opportunity to monitor the implementation of the ePIRLS software in a classroom setting while collecting student responses to the constructed response items. The typed responses gathered during the pilot test were then used to develop scoring guides for the constructed response items for the ePIRLS field test.

Writing and Reviewing the PIRLS/PIRLS Literacy 2016 Field Test Items and Scoring Guides

The TIMSS & PIRLS International Study Center uses a collaborative process involving the participating countries to develop test items and scoring guides for the field tests. Most of the 2nd PIRLS NRC meeting in Portorož, Slovenia in September 2013 was devoted to a workshop for developing the field test items. The NRCs, together with experienced item writers from participating countries and staff from the TIMSS & PIRLS International Study Center, created the newly developed items for the PIRLS and PIRLS Literacy passages.

Prior to the PIRLS item writing workshop, TIMSS & PIRLS International Study Center staff members identified the scope of the item writing task for the field test, examining the weight given to each purpose and comprehension process in the *PIRLS 2016 Assessment Framework*, as well as how many passages and items existed from previous assessments.

In preparation for the item writing workshop, the TIMSS & PIRLS International Study Center updated the Item Writing Guidelines, an item writing manual specifically developed for PIRLS assessments. The *PIRLS 2016 Item Writing Guidelines* contain general information about procedures for obtaining good measurement (for instance, items should be independent and not provide clues to the correct responses of other items) as well as specific information on how to deal with translation issues. The manual also includes the necessary steps for developing scoring guides, as well as checklists for reviewing the PIRLS 2016 items.

At the PIRLS/PIRLS Literacy item writing workshop, country representatives were divided into teams and given specific item writing assignments to ensure that enough field test items were developed in each of the purposes and processes of comprehension areas specified in the PIRLS 2016 framework. The TIMSS & PIRLS International Study Center staff and consultants

used the Item Writing Guidelines to provide training to the teams on item writing procedures for the PIRLS assessments. Once teams had completed their item writing assignments, each team reviewed the items drafted by other teams. In addition, some teams continued to send items to the TIMSS & PIRLS International Study Center for several weeks after the item writing workshop.

Exhibit 1.3 shows the number of participants in the PIRLS/PIRLS Literacy 2016 item writing workshop and the number of items written.

Exhibit 1.3: PIRLS/PIRLS Literacy 2016 Item Writing Workshop to Develop Field Test Items

Attendees	
Number of Countries and Benchmarking Entities	44
Number of Country Representatives	83
Approximate Number of Field Test Items Written at Item Writing Workshop	
PIRLS	394
PIRLS Literacy	134

Following the item writing workshop, the TIMSS & PIRLS International Study Center thoroughly reviewed the draft set of passages and field test items. Reviewers included the chief consultant and consultants experienced in developing assessment items such as those from NFER and ACER, as well as RDG members with particular item writing skills.

Finally, prior to field test instrument production, the PIRLS 2016 RDG members reviewed the proposed field test passages and items, followed by the NRCs at the 3rd NRC meeting in Dublin, Ireland in May 2014. The TIMSS & PIRLS International Study Center implemented the suggested revisions, produced the field test materials, and provided the final international version of the field test booklets to the NRCs so that they could begin translating the field test materials into their languages of instruction.

The PIRLS, PIRLS Literacy, and ePIRLS 2016 Field Tests

Because the TIMSS & PIRLS International Study Center generally field tests twice the number of passages and items actually required, the field test included the target number of new passages and items needed approximately multiplied by two. This included a total of 18 newly developed passages across PIRLS and PIRLS Literacy—8 passages for PIRLS, 4 passages to be shared in common between PIRLS and PIRLS Literacy, and 6 passages for PIRLS Literacy. Given that the field tests for PIRLS and PIRLS Literacy both included the passages in common, the PIRLS field test included 12 passages with 203 items and the PIRLS Literacy field test included 10 passages with 173 items.

The PIRLS and PIRLS Literacy field tests followed typical PIRLS procedures, where they served as full-scale "dress rehearsals" operationally for the assessments. That is, the data collection and

scoring procedures to be employed in the assessments were practiced in the field test. In addition, the field tests provided important information about how well each prospective item functioned and provided a basis for selecting items for the assessments. For the countries participating in ePIRLS, the PIRLS field test students were tested again via computer, typically on the day following the PIRLS field test. The ePIRLS field test involved schools using the ePIRLS software and systems as well as the students responding to the tasks.

All materials and operational procedures for PIRLS/PIRLS Literacy 2016 and ePIRLS were field tested with samples of students selected according to rigorous sampling procedures. The field tests were designed to be conducted in approximately 30 schools in each country. This yielded approximately 9,000 student responses to each PIRLS item, approximately 1,000 for each PIRLS Literacy item, and approximately 5,000 for ePIRLS. The school samples for the PIRLS 2016 field tests and assessments were drawn simultaneously, using the same random sampling procedures. This ensures that field test samples closely approximate assessment samples, and that a school is selected for either the field test or the assessment, but not both. For example, if a country needed 150 schools for the assessment and another 30 for the field test, then a larger sample of 180 schools was selected and a systematic sample of 30 schools was selected from the 180 schools.

Because ePIRLS was a brand new computer-based online reading assessment, preparation for the ePIRLS field test was quite complicated. It involved loading the ePIRLS software onto each computer and checking the compatibility of the computer with the software. The requirement that ePIRLS students also participated in PIRLS was part of the ePIRLS field test because ePIRLS is an extension of PIRLS. The countries participating in ePIRLS field tested ePIRLS with the same students that had already participated in PIRLS, typically on the day after the PIRLS field test.

The ePIRLS field test involved 13,701 students in 13 countries and 5 benchmarking entities. Implementing and monitoring the field test involved newly developed web based systems, including the online translation system, the online scoring system, and online data monitor. The ePIRLS tasks were delivered to the students' computers via USB sticks. Responses collected during the field test were used to evaluate the measurement properties of each item. Additionally, information about students' basic navigation behavior through the hyperlinks, tabs, and advertisements in the tasks was collected in order to analyze the ways students moved through and interacted with the test administration system. The item data and the navigation data were used to revise the ePIRLS tasks and items before the main data collection.

The PIRLS, PIRLS Literacy, and ePIRLS 2016 field tests were conducted in March–April 2015. Student responses were used to evaluate the measurement properties of each field test assessment item. Exhibits 1.4 through 1.6 provide a detailed summary of the field test effort, including the number of students, teachers, and schools that participated and the number of passages and items listed by format, purpose, and comprehension process.

Exhibit 1.4: Overview of the PIRLS 2016 Field Test

	PIRLS	PIRLS Literacy	ePIRLS
Passages/Tasks	12	10	6
Total Items	**203**	**173**	**115**
Responses per item (approx.)	9,000	1,000	5,000
Participants			
Countries	49	7	13
Benchmarking Entities	7	1	5
Students	58,078	6,795	13,701
Teachers	3,025	389	–
Schools	1,634	245	561

Note that four passages and their corresponding items are common to both the PIRLS and PIRLS Literacy assessments.

Exhibit 1.5: PIRLS 2016 Number of Field Test Items by Reading Purpose and Item Format

Reading Purpose	Number of Passages/ Tasks	Number of Multiple-Choice Items	Number of Constructed Response Items	Total Number of Items	Total Number of Score Points	Percentage of Score Points
			PIRLS			
Literary	6	45	56	101	130	49%
Informational	6	44	58	102	135	51%
Total	**12**	**89**	**114**	**203**	**265**	
			PIRLS Literacy			
Literary	5	41	46	87	104	51%
Informational	5	44	42	86	99	49%
Total	**10**	**85**	**88**	**173**	**203**	
			ePIRLS			
Informational	6	44	71	115	153	100%
Total	**6**	**44**	**71**	**115**	**153**	

Note that four passages and their corresponding items are common to both the PIRLS and PIRLS Literacy assessments.

Exhibit 1.6: PIRLS 2016 Number of Field Test Items by Comprehension Process and Item Format

Comprehension Process	Number of Multiple-Choice Items	Number of Constructed Response Items	Total Number of Items	Total Number of Score Points	Percentage of Score Points
		PIRLS			
Focus on and Retrieve Explicitly Stated Information	30	34	64	72	27%
Make Straightforward Inferences	37	23	60	70	26%
Interpret and Integrate Ideas and Information	9	39	48	85	32%
Evaluate and Critique Content and Textual Elements	13	18	31	38	14%
Total	**89**	**114**	**203**	**265**	
		PIRLS Literacy			
Focus on and Retrieve Explicitly Stated Information	33	52	85	92	45%
Make Straightforward Inferences	30	13	43	47	23%
Interpret and Integrate Ideas and Information	7	21	28	46	23%
Evaluate and Critique Content and Textual Elements	15	2	17	18	9%
Total	**85**	**88**	**173**	**203**	
		ePIRLS			
Focus on and Retrieve Explicitly Stated Information	11	14	25	25	16%
Make Straightforward Inferences	15	20	35	41	27%
Interpret and Integrate Ideas and Information	6	28	34	61	40%
Evaluate and Critique Content and Textual Elements	12	9	21	26	17%
Total	**44**	**71**	**115**	**153**	

Note that four passages and their corresponding items are common to both the PIRLS and PIRLS Literacy assessments.

Developing the Materials for PIRLS, PIRLS Literacy, and ePIRLS 2016 Field Test Scoring Training

In order for field test scoring to occur immediately upon completion of data collection, it was necessary to prepare scoring training materials for the newly developed constructed response items in advance of the field test.

For PIRLS, to provide "grist" for these scoring materials, Australia, Canada (Ontario), England, Ireland, and Singapore administered the newly developed constructed response field test items in a small selection of classrooms with English-speaking students. Approximately 100 sample responses to each newly developed constructed response field test item were collected in October–November 2014.

For PIRLS Literacy, the participating countries either were not English-speaking countries or on a Southern Hemisphere school schedule. Thus, the TIMSS & PIRLS International Study Center worked with the Boston College department responsible for working with local school districts to administer the newly developed PIRLS Literacy constructed response items to a range of third grade classrooms in the Boston area. Approximately 50–100 responses to each item were collected in October–November 2014.

For ePIRLS, about 50 responses to each constructed response item were collected in November 2014 as part of the ePIRLS pilot to test the systems in advance of the field test. Approximately 50 sample responses for each item were collected from students in Australia, Ireland, and Canada (Ontario).

Exhibit 1.7 provides the number of constructed response items included in the effort to collect student responses for developing scoring training materials and the number of student responses collected.

Exhibit 1.7: Collecting Student Responses for Developing Field Test Scoring Training Materials

	PIRLS	PIRLS Literacy	ePIRLS
Passages/Tasks	12	10	6
Items			
Total	**114**	**88**	**71**
Responses per item (approx.)	100	30	50
Participants			
Countries	Australia, Canada (Ontario), England, Ireland, Singapore	TIMSS & PIRLS International Study Center	Australia, Ireland, Canada (Ontario)

Note that four passages and their corresponding items are common to both the PIRLS and PIRLS Literacy assessments.

A working group consisting of Marian Sainsbury and Liz Twist from NFER, Prue Anderson from ACER, Karen Wixson from the RDG, and Ina Mullis from the TIMSS & PIRLS International Study Center created sets of example and practice responses for 41 fourth grade PIRLS and PIRLS Literacy items. The example and practice response sets for each item included a scoring guide, approximately 8–10 example responses illustrating the categories in the scoring guide, and approximately 8–10 practice responses so that country representatives could practice making distinctions among categories and reach agreement about how to make consistent scoring decisions across countries. For ePIRLS, Marian Sainsbury and the TIMSS & PIRLS International Study Center used computer produced Excel sheets of responses to develop scoring guides and example responses.

The PIRLS 2016 NRCs and their scoring supervisors received scoring training for the field test constructed response items in February 2015 in Floriana, Malta as part of the 4th PIRLS 2016 NRC meeting. This training was conducted by the scoring training team, which included Julian Fraillon and Prue Anderson of ACER and Marian Sainsbury of NFER. At the scoring training sessions, the trainers explained the purpose of each item and read it aloud. The trainer then described the scoring guide, explaining each category and the rationale for the score given to each example response. After the country representatives scored the practice responses, the NRCs and the scoring training team discussed any inconsistencies in scoring. When necessary, the field test guides were clarified and sometimes categories were revised.

Finalizing the PIRLS, PIRLS Literacy, and ePIRLS 2016 Achievement Items

Subsequent to the field test, the TIMSS & PIRLS International Study Center analyzed the field test data and prepared almanacs containing summary item statistics for each field test item. The data almanac for an item contained, row by row for each country: the sample size, the item difficulty and discrimination, the percentage of students answering each option (multiple-choice) or in each score category (constructed response), the point-biserial correlation for each multiple-choice option or constructed response category, and the degree of scoring agreement for constructed response items.

The field test data were used by the TIMSS & PIRLS International Study Center, the RDG, and NRCs to assess the quality of the field test items. The TIMSS & PIRLS International Study Center staff members, together with external consultants, first reviewed the field test data to make an initial judgment about the quality of each item based on its measurement properties (item statistics). Items were eliminated from further consideration if they had poor measurement properties, such as being too difficult or easy or having low discrimination. Particular attention was paid to unusual item statistics in individual countries because these could indicate errors in translation.

After the item-by-item review, the TIMSS & PIRLS International Study Center staff collaborated with consultants to assemble a set of recommended passages with their item sets and ePIRLS tasks for review by the RDG. RDG members scrutinized the recommendations for the newly developed assessment materials, reviewing each passage and item set as well as scoring guides for content accuracy, clarity, and adherence to the frameworks. In addition, the newly developed passages and items were considered in relation to the trend passages and item sets for overall coherence as a complete assessment. The ePIRLS tasks and items were reviewed via computer. Five of the six ePIRLS tasks that were field tested were recommended for inclusion in the ePIRLS assessment.

NRCs had the opportunity to review the recommended materials in light of the field test results and within the security of their own countries. Each country also could check any unusual national results that might indicate translation errors and correct the translation as necessary or recommend revisions to accommodate translation. The 5th NRC meeting held in Jyväskylä, Finland in August 2015 was devoted to reviewing all the recommended passages, tasks, and items for PIRLS, PIRLS Literacy, and ePIRLS. Following this meeting, the TIMSS & PIRLS International Study Center staff implemented revisions to the passages, tasks, and items as recommended by the NRCs. Final versions of the materials were distributed to the NRCs in August 2015.

Exhibit 1.8 includes descriptions of the PIRLS 2016 and PIRLS Literacy 2016 passages, including the newly developed passages for PIRLS 2016 and trend passages from PIRLS 2001, 2006, and 2011.

Exhibit 1.8: PIRLS 2016 Assessment Passages

Literary Passages	Informational Passages
PIRLS Passages	
Shiny Straw ■ – This animal story demonstrates heroism and the consequences of a reckless attitude.	**Leonardo Da Vinci** ◊ – This biographical text describes the inventions of Leonardo da Vinci and the ways that he was ahead of his time.
Macy and the Red Hen – This contemporary story portrays a complex character who meets a challenge when caring for a red hen.	**The Green Sea Turtle's Journey of a Lifetime** – This passage describes the life cycle of a female green sea turtle from the time she hatches from an egg to the time she lays her own eggs.
The Empty Pot * - This traditional tale set in China has a moral message about the importance of honesty.	**Where's the Honey?** * – This passage describes the relationship between the honeyguide bird and the Boran people in Africa using a combination of explanation, photographs, and graphic displays.
Oliver and the Griffin - In this fantasy story, a boy named Oliver meets an old griffin in a garden and decides to help him.	**Icelandic Horses** – This article describes the history and characteristics of Icelandic horses as they developed along with the people who lived near them.
Shared PIRLS/PIRLS Literacy Passages	
Flowers on the Roof ◊ – This contemporary story portrays friendship between the generations.	**Sharks** ■ – This article presents information about sharks in a variety of formats, using subheadings, a labeled diagram, and photographs.
Pemba Sherpa – This modern tale set in the Himalayan Mountains tells the story of a young girl determined to be a sherpa.	**How Did We Learn to Fly?** – This historical text explains how the modern airplane was developed.
PIRLS Literacy Passages	
Baghita's Perfect Orange * - This traditional tale set in Africa has a moral about greed and generosity.	**Training a Deaf Polar Bear** * – The passage describes how zookeepers worked with a polar bear that was found to be deaf.
The Pearl – This story about a young pearl merchant illustrates the power of home, friendship, and generosity above greed.	**African Rhinos & Oxpecker Birds** – This passage presents information about African rhinos and oxpecker birds and describes how the two animals depend on one another for food and survival.
The Summer My Father Was Ten * – In this thought-provoking story with a realistic contemporary setting, a boy is allowed to make amends for his thoughtless behavior.	**Ants** * – This article presents information about the lives of different types of ants, using subheadings, photographs, and diagrams.
Library Mouse – This story is about a mouse who lives in the library and inspires young children to be authors.	**Hungry Plant** – This scientific text describes the Venus Flytrap plant and explains how it captures insects for food.

◊ Passage from PIRLS 2001
■ Passage from PIRLS 2006
* Passage from PIRLS 2011

Exhibit 1.8: PIRLS 2016 Assessment Passages (Continued)

PIRLS 2016 Word Counts and Readability

Passage	Word Count	Flesch-Kincaid Grade Level
PIRLS Passages		
Shiny Straw	860	5.5
Macy and the Red Hen	913	4.4
The Empty Pot	767	4.9
Oliver and the Griffin	896	3.3
Leonardo Da Vinci	869	5.1
The Green Sea Turtle's Journey of a Lifetime	943	4.0
Where's the Honey?	870	3.2
Icelandic Horses	870	5.0
Shared PIRLS/PIRLS Literacy Passages		
Flowers on the Roof	811	2.8
Sharks	570	7.6
Pemba Sherpa	540	2.5
How Did We Learn to Fly?	514	6.3
PIRLS Literacy Passages		
Baghita's Perfect Orange	404	2.0
The Pearl	536	2.9
The Summer My Father Was Ten	484	4.0
Library Mouse	497	3.1
Training a Deaf Polar Bear	425	4.0
African Rhinos & Oxpecker Birds	449	4.7
Ants	415	2.9
Hungry Plant	509	3.5

The Flesch-Kincaid Grade Level Formula uses average syllables per word and average sentence length to produce a number that represents the US grade in which students can read the text.

Exhibit 1.9 includes descriptions of the ePIRLS tasks assessing online informational reading.

Exhibit 1.9: ePIRLS 2016 Assessment Tasks

Mars – In this science task, students learn what scientists know about Mars and investigate space exploration.
Dr. Elizabeth Blackwell – This biographical task is about the life and accomplishments of Elizabeth Blackwell, the first female doctor in both America and England.
Rainforests – This science task is about the plants and animals that live in the rainforest.
Zebra and Wildebeest Migration – Students learn about zebra and wildebeest migration through the Serengeti.
The Legend of Troy – This historical task is about the legend of Troy and archeological investigations of the ancient city.

Distribution of PIRLS 2016 Items by Reading Purpose and Comprehension Process

Exhibits 1.10 and 1.11 present the number of trend and newly developed items as well as the number of score points in the PIRLS 2016 assessments. The number of items represents the number of distinct questions in the assessment, while the number of score points represents the complexity and weight given to each item. Half the PIRLS and PIRLS Literacy items are based on literary passages and half are based on informational passages. ePIRLS assesses reading for information, but in an online environment.

Exhibit 1.10: PIRLS 2016 Achievement Items by Reading Purpose

Reading Purpose	Number of Passages/ Tasks	Number of Trend Items in PIRLS 2016	Percentage of Trend Score Points	Number of New Items in PIRLS 2016	Percentage of New Score Points	Total Items	Achieved Percentage of Score Points	Target Percentage of Score Points
				PIRLS				
Literary	6	44 (55)	49%	46 (58)	51%	90 (113)	51%	50%
Informational	6	37 (51)	46%	48 (59)	54%	85 (110)	49%	50%
Total	**12**	**81 (106)**		**94 (117)**		**175 (223)**		
				PIRLS Literacy				
Literary	6	43 (48)	45%	50 (59)	55%	93 (107)	50%	50%
Informational	6	40 (51)	49%	50 (54)	51%	90 (105)	50%	50%
Total	**12**	**83 (99)**		**100 (113)**		**183 (212)**		
				ePIRLS				
Informational	5	0 (0)	0%	91 (112)	100%	91 (112)	100%	100%
Total	**5**	**0 (0)**		**91 (112)**		**91 (112)**		

Score points are shown in parentheses.
Because percentages are rounded to the nearest whole number, some totals may appear inconsistent.
Note that four passages and their corresponding items are common to both the PIRLS and PIRLS Literacy assessments.

Exhibit 1.11: PIRLS 2016 Achievement Items by Comprehension Process

Comprehension Process	Number of Trend Items in PIRLS 2016	Percentage of Trend Score Points	Number of New Items in PIRLS 2016	Percentage of New Score Points	Total Items	Achieved Percentage of Score Points	Target Percentage of Score Points
			PIRLS				
Focus on & Retrieve Explicitly Stated Information	18 (21)	37%	32 (36)	63%	50 (57)	26	20
Make Straightforward Inferences	28 (30)	51%	25 (29)	49%	53 (59)	26	30
Interpret & Integrate Ideas and Information	24 (42)	53%	23 (37)	47%	47 (79)	35	30
Evaluate & Critique Content and Textual Elements	11 (13)	46%	14 (15)	54%	25 (28)	13	20
Total	**81 (106)**		**94 (117)**		**175 (223)**		
			PIRLS Literacy				
Focus on & Retrieve Explicitly Stated Information	31 (36)	40%	51 (55)	60%	82 (91)	43	50
Make Straightforward Inferences	27 (27)	50%	25 (27)	50%	52 (54)	25	25
Interpret & Integrate Ideas and Information	15 (26)	53%	16 (23)	47%	31 (49)	23	25
Evaluate & Critique Content and Textual Elements	10 (10)	56%	8 (8)	44%	18 (18)	8	
Total	**83 (99)**		**100 (113)**		**183 (212)**		
			ePIRLS				
Focus on & Retrieve Explicitly Stated Information	0 (0)	0%	22 (23)	100%	22 (23)	21	20
Make Straightforward Inferences	0 (0)	0%	27 (31)	100%	27 (31)	28	30
Interpret & Integrate Ideas and Information	0 (0)	0%	23 (38)	100%	23 (38)	34	30
Evaluate & Critique Content and Textual Elements	0 (0)	0%	19 (20)	100%	19 (20)	18	20
Total	**0 (0)**		**91 (112)**		**91 (112)**		

Score points are shown in parentheses.
Because percentages are rounded to the nearest whole number, some totals may appear inconsistent.
Note that four passages and their corresponding items are common to both the PIRLS and PIRLS Literacy assessments.

Distribution of PIRLS Item Formats within Reading Purposes and Comprehension Processes

As described in the *PIRLS 2016 Assessment Framework*, up to half of the total number of score points represented by all the questions come from multiple-choice items. Most PIRLS multiple-choice items are worth one score point, although some compound multiple-choice items are worth two score points. The 2-point compound multiple-choice items are scored as all parts answered correctly as fully correct (2 score points), and most parts answered correctly as partially correct (1 score point). Constructed response items generally are worth one, two, or three score points depending on the degree of complexity involved. The 1-point constructed response items are scored as correct (1 score point) or incorrect (0 score points), whereas 2-point constructed response items are scored as fully correct (2 score points), partially correct (1 score point), or incorrect (0 score points), and 3-point constructed response items are scored as fully correct (3 score points), partially correct (1 or 2 score points), or incorrect (0 score points). Fully correct responses show a complete or deeper understanding of a task while partially correct responses demonstrate only a partial understanding of the concepts embodied in the task.

Exhibits 1.12 and 1.13 display the number of passages or tasks and items (and score points) by item format for each purpose and comprehension process.

Exhibit 1.12: PIRLS 2016 Achievement Items by Reading Purpose and Item Format

Reading Purpose	Number of Passages/ Tasks	Multiple-Choice Items		Constructed Response Items			Total Items	Percentage of Score Points
		Four Response Options	Compound	1 Point	2 Points	3 Points		
PIRLS								
Literary	6	46 (46)	0 (0)	25 (25)	15 (30)	4 (12)	90 (113)	51%
Informational	6	40 (40)	0 (0)	24 (24)	17 (34)	4 (12)	85 (110)	49%
Total	**12**	**86 (86)**	**0 (0)**	**49 (49)**	**32 (64)**	**8 (24)**	**175 (223)**	
Achieved Percentage of Score Points		**39%**		**61%**				
Target Percentage of Score Points		**40%**		**60%**				
PIRLS Literacy								
Literary	6	47 (47)	0 (0)	33 (33)	12 (24)	1 (3)	93 (107)	50%
Informational	6	43 (43)	1 (2)	34 (34)	10 (20)	2 (6)	90 (105)	50%
Total	**12**	**90 (90)**	**1 (2)**	**67 (67)**	**22 (44)**	**3 (9)**	**183 (212)**	
Achieved Percentage of Score Points		**43%**		**57%**				
Target Percentage of Score Points		**40%**		**60%**				
ePIRLS								
Literary	0	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0 (0)	0%
Informational	5	36 (36)	4 (8)	37 (37)	11 (22)	3 (9)	91 (112)	100%
Total	**5**	**36 (36)**	**4 (8)**	**37 (37)**	**11 (22)**	**3 (9)**	**91 (112)**	
Achieved Percentage of Score Points		**39%**		**61%**				
Target Percentage of Score Points		**40%**		**60%**				

Score points are shown in parentheses.
Because percentages are rounded to the nearest whole number, some totals may appear inconsistent.
Note that four passages and their corresponding items are common to both the PIRLS and PIRLS Literacy assessments.

Exhibit 1.13: PIRLS 2016 Achievement Items by Comprehension Process and Item Format

Comprehension Process	Multiple-Choice Items		Constructed Response Items			Total Items	Percentage of Score Points
	Four Response Options	Compound	1 Point	2 Points	3 Points		
PIRLS							
Focus on and Retrieve Explicitly Stated Information	25 (25)	0 (0)	18 (18)	7 (14)	0 (0)	50 (57)	26%
Make Straightforward Inferences	35 (35)	0 (0)	12 (12)	6 (12)	0 (0)	53 (59)	26%
Interpret and Integrate Ideas and Information	11 (11)	0 (0)	12 (12)	16 (32)	8 (24)	47 (79)	35%
Evaluate and Critique Content and Textual Elements	15 (15)	0 (0)	7 (7)	3 (6)	0 (0)	25 (28)	13%
Total	**86 (86)**	**0 (0)**	**49 (49)**	**32 (64)**	**8 (24)**	**175 (223)**	
Achieved Percentage of Score Points	**39%**		**61%**				
Target Percentage of Score Points	**40%**		**60%**				
PIRLS Literacy							
Focus on and Retrieve Explicitly Stated Information	30 (30)	0 (0)	43 (43)	9 (18)	0 (0)	82 (91)	43%
Make Straightforward Inferences	35 (35)	0 (0)	15 (15)	2 (4)	0 (0)	52 (54)	25%
Interpret and Integrate Ideas and Information	8 (8)	1 (2)	8 (8)	11 (22)	3 (9)	31 (49)	23%
Evaluate and Critique Content and Textual Elements	17 (17)	0 (0)	1 (1)	0 (0)	0 (0)	18 (18)	8%
Total	**90 (90)**	**1 (2)**	**67 (67)**	**22 (44)**	**3 (9)**	**183 (212)**	
Achieved Percentage of Score Points	**43%**		**57%**				
Target Percentage of Score Points	**40%**		**60%**				
ePIRLS							
Focus on and Retrieve Explicitly Stated Information	10 (10)	0 (0)	11 (11)	1 (2)	0 (0)	22 (23)	21%
Make Straightforward Inferences	12 (12)	0 (0)	11 (11)	4 (8)	0 (0)	27 (31)	28%
Interpret and Integrate Ideas and Information	3 (3)	4 (8)	8 (8)	5 (10)	3 (9)	23 (38)	34%
Evaluate and Critique Content and Textual Elements	11 (11)	0 (0)	7 (7)	1 (2)	0 (0)	19 (20)	18%
Total	**36 (36)**	**4 (8)**	**37 (37)**	**11 (22)**	**3 (9)**	**91 (112)**	
Achieved Percentage of Score Points	**39%**		**61%**				
Target Percentage of Score Points	**40%**		**60%**				

Score points are shown in parentheses.
Because percentages are rounded to the nearest whole number, some totals may appear inconsistent.
Note that four passages and their corresponding items are common to both the PIRLS and PIRLS Literacy assessments.

PIRLS 2016 Constructed Response Scoring Training

In preparation for the main data collection scoring training, some PIRLS 2016 scoring guides were further refined or clarified based on the results of the field test. This included a thorough review of the field test scoring training materials to ensure that the student responses were still suitable for the updated scoring guides. In some cases, example and practice sets used in the field test were expanded to further illustrate particular aspects of a scoring guide. For PIRLS/PIRLS Literacy 2016 scoring training, the example and practice paper training sets included those used in PIRLS 2011 for the trend items and the updated training sets for the newly developed items selected for PIRLS 2016, resulting in 42 example and practice paper sets for PIRLS and 24 for PIRLS Literacy. Scoring training materials were developed for 8 ePIRLS items.

To provide scoring training for all the countries participating in PIRLS 2016, the TIMSS & PIRLS International Study Center conducted two training sessions. First, the NRCs for Southern Hemisphere countries and their scoring supervisors received PIRLS and PIRLS Literacy scoring training in November 2015 in Buenos Aires, Argentina. (No Southern Hemisphere countries participated in ePIRLS.) NRCs for Northern Hemisphere countries and their scoring supervisors received scoring training in March 2016 in Hong Kong SAR as part of the 6th PIRLS 2016 NRC meeting.

Exhibit 1.14 shows the number of participants in the two scoring training sessions.

Exhibit 1.14: PIRLS 2016 Scoring Training Participation

Participants	Southern Hemisphere	Northern Hemisphere
Number of Countries	6	49
Number of Benchmarking Entities	2	10
Number of Country Representatives	29	119

The Process Following Instrument Development

In general, after the participating countries received the international version of the assessment instruments, they began the process of translation and cultural adaptation (some adaptation to local usage typically is necessary even in English-speaking countries) and production of the materials for printing. At the same time, countries made final arrangements for data collection, including the host of activities necessary to obtain school participation, implement test administration, and score the responses to the tests and questionnaires (see following chapters).

Reference

Kincaid, J.P., Fishburne, R.P., Jr., Rogers, R.L., & Chissom, B.S. (1975). *Derivation of new readability formulas (Automated Readability Index, Fog Count and Flesch Reading Ease Formula) for Navy enlisted personnel.* (Research Branch Report 8-75). Millington, TN: Naval Technical Training Command Millington Tennessee Research Branch.