

categories) showed that private schools, high poverty, and 4th-grade enrollment were significant predictors of participation.

Results for the final sample of schools. In the analyses for the final sample of schools, all substitute schools were included with the original schools as responding schools, leaving nonresponding schools as those for which no assessment data were available.

The bivariate results for the final sample of 4th-grade schools indicated that three variables were statistically significant: school control, 4th-grade enrollment, and the percentage of Hispanic students. When all of these factors were considered simultaneously in a regression analysis, two variables remained significant predictors of participation: private schools and 4th-grade enrollment.

For the final sample of schools in grade 4 with school nonresponse adjustments applied to the weights,¹⁶ there were no statistically significant variables in the bivariate analysis. Note that the multivariate regression analysis cannot be conducted after the school nonresponse adjustments are applied to the weights.

These results suggest that there is some potential for nonresponse bias in the U.S. 4th-grade original sample based on the characteristics studied. It also suggests that, while there is little evidence that the use of substitute schools reduced the potential for bias, it has not added to it substantially. The application of school nonresponse adjustments substantially reduced the measurable potential for bias as no variables remained statistically significant.

Test development

PIRLS is a cooperative effort involving representatives from every education system participating in the study. For PIRLS 2011, the test development effort began with a review and revision of the frameworks that are used to guide the construction of the assessment (Mullis et al. 2009). The frameworks were updated to reflect changes in the curriculum and instruction of participating education systems. Extensive input from experts in reading education, assessment, and curriculum, and representatives from national educational centers around the world contributed to the final shape of the frameworks. Maintaining the ability to measure change over time was an important factor in revising the frameworks.

As part of the PIRLS dissemination strategy, approximately one-half of the 2006 assessment items were released for public use. To replace assessment items that had been released, education systems submitted items for review by subject-matter specialists, and additional items were written by the IEA Reading Review Committee in consultation with item-writing specialists in various countries to ensure that the content, as explicated in the frameworks, was covered adequately. Items were reviewed by an international Reading Item Review Committee and field-tested in most of the participating countries. Results from the field test were used to evaluate item difficulty, how well items discriminated between high- and low-performing students, the effectiveness of distracters in multiple-choice items, scoring suitability and reliability for constructed-response items, and evidence of bias toward or against individual countries or in favor of boys or girls. As a result of this review, 60 new items were selected for inclusion in the international assessment. In total, 135 reading items were included in the 2011 PIRLS assessment booklets. More detail on the distribution of new and trend items is included in table A-3.

¹⁶The international weighting procedures created a nonresponse adjustment class for each explicit stratum; see the *TIMSS and PIRLS Methods and Procedures* (Martin and Mullis 2011) for details. In the case of the U.S. 4th-grade sample, 12 explicit strata were formed by poverty level, school control, and Census region. The procedures could not be varied for individual education systems to account for any specific needs. Therefore, the U.S. nonresponse bias analyses could have no influence on the weighting procedures and were undertaken after the weighting process was complete.

Table A-3. Number and percentage distribution of reading items in the PIRLS assessment, by content domain and process: 2011

Content domain and process	All items		New items		Trend items	
	Number	Percent	Number	Percent	Number	Percent
Total items	135	100	60	100	75	100
Purposes of reading						
Literary experience	72	53	33	55	39	52
Acquire and use information	63	47	27	45	36	48
Processes of comprehension						
Focus on and retrieve explicitly stated information	33	24	14	23	19	25
Make straightforward inferences	46	34	20	33	26	35
Interpret and integrate ideas and information	38	28	18	30	20	27
Examine and evaluate content, language, and textual elements	18	13	8	13	10	13

NOTE: Detail may not sum to 100 percent due to rounding.

SOURCE: International Association for the Evaluation of Educational Achievement (IEA), Progress in International Reading Literacy Study (PIRLS), 2011.

Design of instruments

PIRLS 2011 included booklets containing assessment items as well as self-administered background questionnaires for principals, teachers, and students.

Assessment booklets

The assessment booklets were constructed such that not all of the students responded to all of the items. This is consistent with other large-scale assessments, such as NAEP.

The 2011 assessment consisted of 12 booklets and one reader (presented in a magazine-type format with the questions in a separate booklet). The assessment is given in 40-minute parts with a 5- to 10-minute break in between. The student questionnaire given after the second part of the assessment, while untimed, is allotted approximately 30 minutes of response time.

The booklets were rotated among students, with each participating student completing one booklet only. The reading items were each assembled separately into 10 blocks, or clusters, of items. Each of the 13 PIRLS 2011 booklets contained two blocks in total. Each booklet contained one block of literary experience items and one block of informational items only and each block occurred twice across the 13 booklets. Six of the ten blocks were included in previous PIRLS assessments. The remaining four blocks were new for PIRLS 2011.

The PIRLS booklets administered in the state sample were exactly the same as those administered in the national sample.

As part of the design process, it was necessary to ensure that the booklets showed an item distribution across the reading content domains as specified in the framework as well as a relatively equal distribution of items by item type. The number of reading items in the PIRLS 2011 assessment is shown in table A-4.

Table A-4. Number and percentage of reading items in the PIRLS assessment, by item format: 2011

Item Format	Number of items	Percent of items
Total	135	100
Multiple choice	74	55
Constructed response	61	45

SOURCE: International Association for the Evaluation of Educational Achievement (IEA), Progress in International Reading Literacy Study (PIRLS), 2011.

Background questionnaires

As in prior administrations, PIRLS 2011 included self-administered questionnaires for principals, teachers, and students. To create the questionnaires for 2011, the 2006 versions were reviewed extensively by the NRCs from the participating countries as well as a Questionnaire Item Review Committee (QIRC). The QIRC comprises 10–12 experienced NRCs from different participating countries who have analyzed PIRLS data and use it in their countries. The QIRC review resulted in items being deleted or revised, and the addition of several new ones. Like the assessment items, all questionnaire items were field tested, and the results reviewed carefully. As a result, some of the questionnaire items needed to be revised prior to their inclusion in the final questionnaires. The questionnaires requested information to help provide a context for the performance scores, focusing on such topics as students' attitudes and beliefs about learning, their habits and homework, and their lives both in and outside of school; teachers' attitudes and beliefs about teaching and learning, teaching assignments, class size and organization, instructional practices, and participation in professional development activities; and principals' viewpoints on policy and budget responsibilities, curriculum and instruction issues and student behavior, as well as descriptions of the organization of schools and courses. For 2011, online versions of the school and teacher questionnaires were offered to respondents as the primary mode of data collection. Detailed results from the student, teacher, and school surveys are not discussed in this report but are available in the international report, the *PIRLS 2011 International Report in Reading* (Mullis et al. 2012).

Translation

Source versions of all instruments (assessment booklets, questionnaires, and manuals) were prepared in English and translated into the primary language or languages of instruction in each education system. In addition, it was sometimes necessary to adapt the instrument for cultural purposes, even in countries that use English as the primary language of instruction. All adaptations were reviewed and approved by the International Study Center to ensure they did not change the substance or intent of the question or answer choices. For example, proper names were sometimes changed to names that would be more familiar to students (e.g., Marja-leena to Maria).

Each education system prepared translations of the instruments according to translation guidelines established by the International Study Center. Adaptations to the instruments were documented by each education system and submitted for review. The goal of the translation guidelines was to produce translated instruments of the highest quality that would provide comparable data across countries.

Translated instruments were verified by an independent, professional translation agency prior to final approval and printing of the instruments. Countries were required to submit copies of the final printed instruments to the International Study Center. Further details on the translation process can be found in the *PIRLS 2011 Technical Report* (Martin, Mullis, and Foy forthcoming).

Recruitment, test administration, and quality assurance

PIRLS 2011 emphasized the use of standardized procedures in all participating education systems, so that each collected its own data, based on comprehensive manuals and training materials provided by the international project team. These materials explained the survey's implementation, including precise instructions for the work of school coordinators and scripts for test administrators to use in testing sessions.

Recruitment of schools and students

With the exception of private schools, the recruitment of schools required several steps. Beginning with the sampled schools, the first step entailed obtaining permission from the school district to approach the sampled school(s) in that district. If a district refused permission, then the district of the first substitute school was approached and the procedure was repeated. With permission from the district, the school(s) was contacted in a second step. If a sampled school refused to participate, the district of the first substitute was approached and the permission procedure repeated. During most of the recruitment period sampled schools and substitute schools were being recruited concurrently. Each participating school was asked to nominate a school coordinator as the main point of contact for the study. The school coordinator worked with project staff to arrange logistics and liaise with staff, students, and parents as necessary.

On the advice of the school, parental permission for students to participate was sought with one of three approaches to parents: a simple notification; a notification with a refusal form; and a notification with a consent form for parents to sign. In each approach, parents were informed that their students could opt out of participating in the assessment.

Gifts to schools, school coordinators, and students

Schools, school coordinators, and students were provided with small gifts in appreciation for their willingness to participate. Schools were offered \$200, school coordinators received \$100, and students were given a clock-compass carabiner.