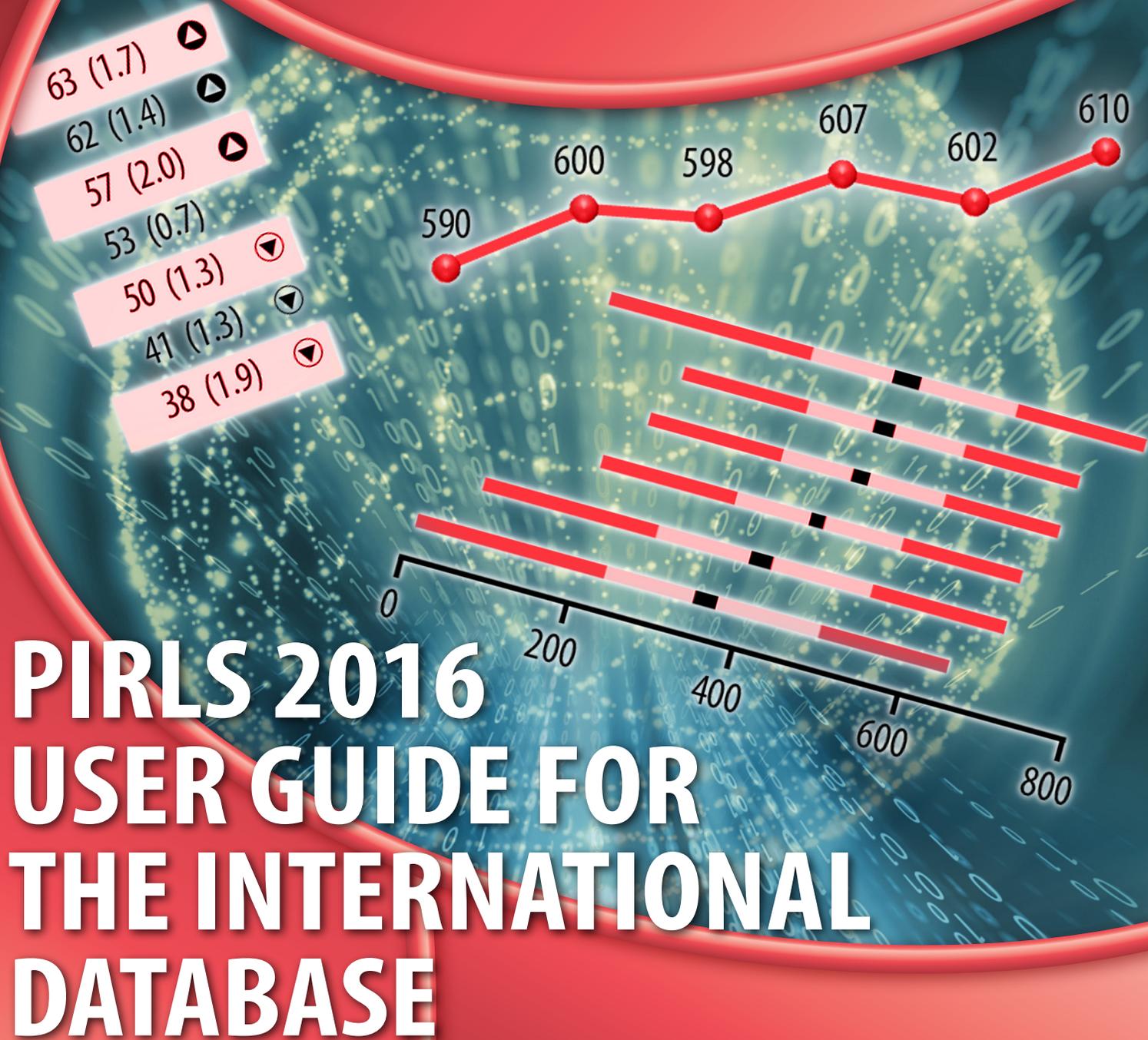


# PIRLS



# PIRLS 2016 USER GUIDE FOR THE INTERNATIONAL DATABASE

**PIRLS**  
**2016**

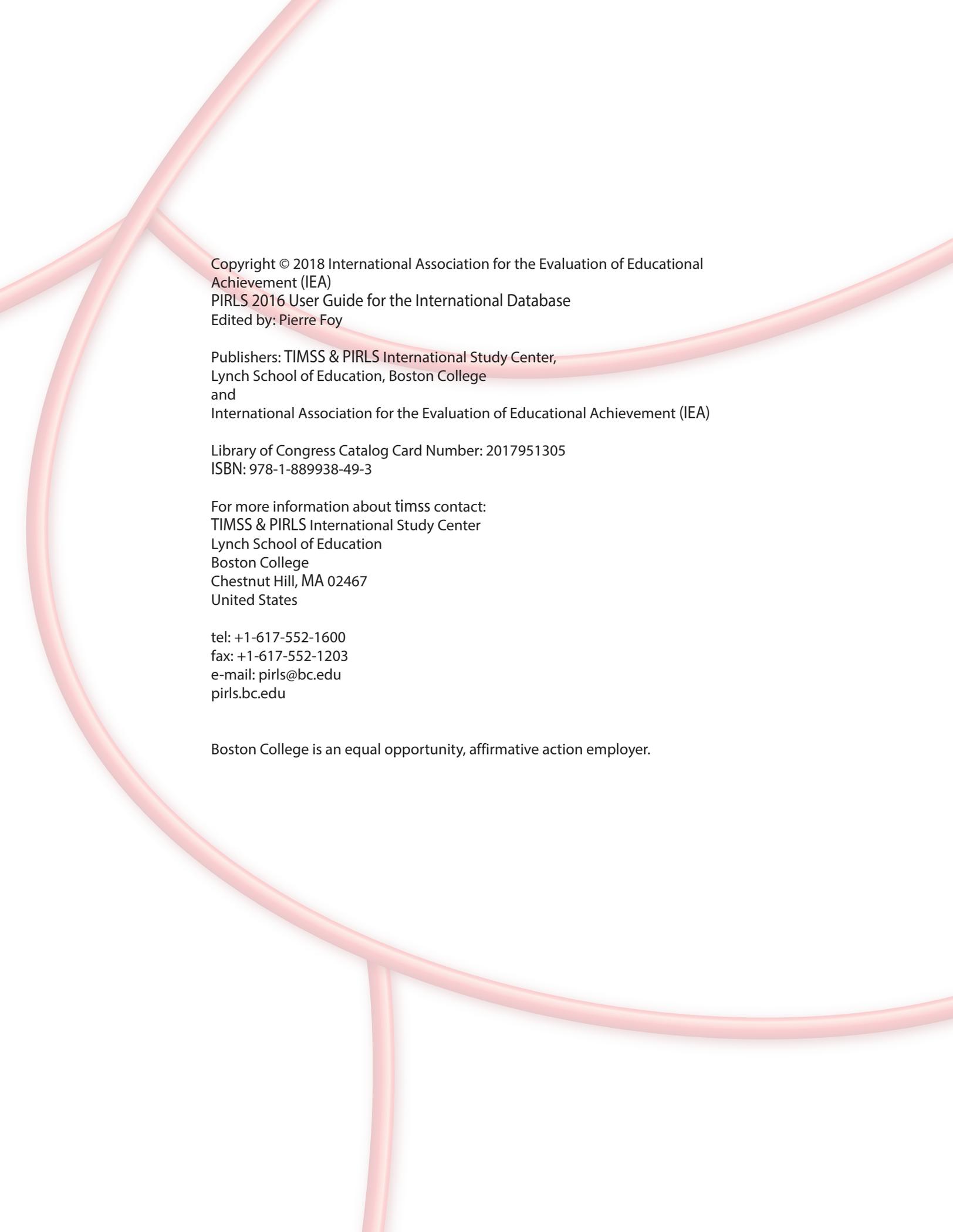
# PIRLS 2016 User Guide for the International Database

Pierre Foy

*Contributors: Charlotte E.A. Aldrich, Bethany G. Fishbein, Hannah Köhler, Kamil Kowolik, Jenny Liu, Sebastian Meyer, Caroline Prendergast, Erin Wry, and Liqun Yin*



**TIMSS & PIRLS**  
International Study Center  
Lynch School of Education  
BOSTON COLLEGE

The page features several thick, light pink circular lines that overlap and curve across the background, creating a modern, abstract design.

Copyright © 2018 International Association for the Evaluation of Educational Achievement (IEA)  
PIRLS 2016 User Guide for the International Database  
Edited by: Pierre Foy

Publishers: TIMSS & PIRLS International Study Center,  
Lynch School of Education, Boston College  
and  
International Association for the Evaluation of Educational Achievement (IEA)

Library of Congress Catalog Card Number: 2017951305  
ISBN: 978-1-889938-49-3

For more information about timss contact:  
TIMSS & PIRLS International Study Center  
Lynch School of Education  
Boston College  
Chestnut Hill, MA 02467  
United States

tel: +1-617-552-1600  
fax: +1-617-552-1203  
e-mail: [pirls@bc.edu](mailto:pirls@bc.edu)  
[pirls.bc.edu](http://pirls.bc.edu)

Boston College is an equal opportunity, affirmative action employer.

# Contents

---

## **Chapter 1**

*Introduction.* . . . . . 1

## **Chapter 2**

*Analyzing the PIRLS 2016 Data* . . . . . 6

## **Chapter 3**

*Special SPSS and SAS Programs* . . . . . 43

## **Chapter 4**

*The PIRLS 2016 International Database Files.* . . . . . 51

## **Appendix**

*Organizations and Individuals Responsible for PIRLS 2016* . . . . . 80

**PIRLS**  
**2016**

**PIRLS 2016 User Guide  
for the International Database**

# **Chapter 1**

## **Introduction**

# Chapter 1

## *Introduction*

### 1.1 Overview

IEA's PIRLS provides international comparative data about students' reading achievement after four years of schooling in participating countries around the world, while also measuring trends over time, monitoring curricular implementation, and identifying promising instructional practices. Conducted on a regular 5-year cycle, PIRLS has assessed reading in 2001, 2006, 2011, and 2016. PIRLS collects a rich array of background information to provide comparative perspectives on trends in achievement in the context of different educational systems, school organizational approaches, and instructional practices.

PIRLS Literacy 2016, introduced in 2011 as the prePIRLS assessment, is a less difficult reading assessment intended to assess the reading skills of populations whose students are earlier in the process of learning to read than is expected by the PIRLS assessment. The PIRLS Literacy assessment has shorter and easier texts than PIRLS, and places less emphasis on higher-order reading skills, while providing achievement results on the same scale as PIRLS. Five countries and one benchmarking participant participated in the PIRLS Literacy 2016 assessment, two of which also participated in PIRLS 2016.

As a new initiative in 2016, IEA developed ePIRLS, an online informational reading assessment designed to assess fourth grade students' online reading skills in a school context. The ePIRLS 2016 achievement data were scaled in order to make relative comparisons of performance between both ePIRLS and PIRLS assessments.

To support and promote secondary analyses aimed at improving reading education at the fourth grade, the PIRLS 2016 International Database makes available to researchers, analysts, and other users the data collected and processed by the PIRLS project, including the data from the PIRLS Literacy and ePIRLS assessments. This database comprises student achievement data as well as student, home, teacher, school, and curricular background data for 50 countries and 11 benchmarking participants. The database includes data from 346,852 students, 337,332 parents, 16,476 teachers, 12,124 school principals, and the National Research Coordinators of each country. All participating countries gave the IEA permission to release their national data.

For countries that participated in previous assessments, PIRLS 2016 (the fourth data collection in the PIRLS studies) provides trends for up to four cycles—2001, 2006, 2011, and 2016. In countries new to the study, the 2016 results can help policy makers and practitioners assess their comparative standing and gauge the rigor and effectiveness of their reading education

programs. Results of the assessments conducted in 2016 can be found in *PIRLS 2016 International Results in Reading* (Mullis, Martin, Foy, & Hooper, 2017).

PIRLS 2016 was an ambitious and demanding study, involving complex procedures for drawing student samples, assessing students' achievement, analyzing the data, and reporting the results. In order to work effectively with the PIRLS data, it is necessary to have an understanding of the characteristics of the study, which are described fully in *Methods and Procedures in PIRLS 2016* (Martin, Mullis, & Hooper, 2017). Therefore, this User Guide should be used in conjunction with the *Methods and Procedures* documentation. Whereas the User Guide describes the organization and content of the database, the *Methods and Procedures* documentation provides the rationale for the techniques used and for the variables created in the process of data collection and compilation.

## 1.2 The PIRLS 2016 User Guide

This User Guide serves as a primer for analyzing the data in the PIRLS 2016 International Database using the IEA International Database (IDB) Analyzer software (IEA, 2018). It also describes the content and format of the data in the PIRLS 2016 International Database. In addition to this introduction, the User Guide includes the following chapters:

- Chapter 2—This chapter introduces the IEA IDB Analyzer and presents examples of analyses of the PIRLS 2016 data using this software in conjunction with SPSS (IBM Corporation, 2016) and SAS (SAS Institute, 2012).
- Chapter 3—This chapter describes special SPSS and SAS programs needed to make full use of the PIRLS 2016 International Database.
- Chapter 4—This chapter serves as a reference for details about the structure and content of the PIRLS 2016 International Database.

The User Guide is accompanied by the following supplements:

- *Supplement 1*—This supplement contains the international version of all PIRLS 2016 background questionnaires.
- *Supplement 2*—This supplement describes all adaptations to the questions in the background questionnaires made by individual PIRLS 2016 participants.
- *Supplement 3*—This supplement describes how derived variables were constructed for reporting the PIRLS 2016 data.

The User Guide and its supplements are available on the PIRLS 2016 International Database and User Guide webpage: <http://timssandpirls.bc.edu/pirls2016/international-database.html>. The primary purpose of this User Guide is to introduce users to the PIRLS 2016 International Database and demonstrate the basic functionality of the IEA IDB Analyzer through simple examples of results published in *PIRLS 2016 International Results in Reading*. The IEA IDB Analyzer comes with its own manual, available through the Help Module, which describes the full functionality and features of the IEA IDB Analyzer.

### 1.3 The PIRLS 2016 International Database

The PIRLS 2016 International Database is available on the PIRLS 2016 International Database webpage: <http://timssandpirls.bc.edu/pirls2016/international-database.html>. The PIRLS 2016 International Database also is available for download at the IEA Study Data Repository website: <http://www.iea.nl/data.html>. The repository allows users to download data files and accompanying support materials from all recent IEA studies, including PIRLS 2016.

The International Database contains the PIRLS 2016 student achievement data files and student, home, teacher, and school background questionnaire data files, along with support materials. Exhibit 1.1 displays the general structure of the International Database and a brief description of the support materials available for download on the PIRLS 2016 International Database webpage.

**Exhibit 1.1 Contents of the PIRLS 2016 International Database**

<b>User Guide</b>	This User Guide with its supplements	
<b>Items</b>	The PIRLS 2016 item information files, IRT item parameters, and item percent correct statistics	
<b>International Database</b>	<b>SPSS Data</b>	PIRLS 2016 student, home, teacher, and school data files in SPSS format
	<b>SAS Data</b>	PIRLS 2016 student, home, teacher, and school data files in SAS export format <sup>1</sup>
	<b>Curriculum Data</b>	PIRLS 2016 curriculum questionnaire data files
	<b>Codebooks</b>	Codebook files describing all variables in the PIRLS 2016 International Database
	<b>Almanacs</b>	Data almanacs with summary statistics for all PIRLS 2016 items and background variables
	<b>SAS and SPSS programs</b>	

### 1.4 Two Versions of the PIRLS 2016 International Database

The PIRLS 2016 International Database is available in two versions: a public use version and a restricted use version. In the public use version, some variables are removed to minimize the risk of disclosing confidential information. Chapter 4 of this User Guide provides a list of variables removed from the public use version. The public use version is available for immediate access from the PIRLS 2016 website, as well as the IEA Study Data Repository, and users should be able to replicate all published PIRLS 2016 results with this version of the PIRLS 2016 International Database. Users who require any of the removed variables to conduct their analyses should contact the IEA (RandA@IEA-Hamburg.de) to obtain permission and access to the restricted use version of the PIRLS 2016 International Database.

<sup>1</sup> Because SAS data files are specific to their computing environment and may not be readable in all environments, the PIRLS 2016 International Database provides data files in SAS export (EXP) format. These SAS export files will generate SAS data files, suitable for analysis in the IEA IDB Analyzer, after conversion in the appropriate computing environment (see Chapter 3 of this User Guide).

## References

- IBM Corporation. (2016). *IBM SPSS Statistics for Windows*, version 24.0. Armonk, NY: IBM Corporation.
- International Association for the Evaluation of Educational Achievement. (2018). *International Database Analyzer*, version 4.0. Hamburg, Germany: IEA Hamburg.
- Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.). (2017). *Methods and Procedures in PIRLS 2016*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/publications/pirls/2016-methods.html>
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/pirls2016/international-results/>
- SAS Institute. (2012). *SAS System for Windows*, version 9.4. Cary, NC: SAS Institute.

**PIRLS**  
**2016**

**PIRLS 2016 User Guide  
for the International Database**

# **Chapter 2**

## **Analyzing the PIRLS 2016 Data**

# Chapter 2

## *Analyzing the PIRLS 2016 Data*

### 2.1 Overview

This chapter describes the general use of the IEA International Database Analyzer software (IEA, 2018) for analyzing the PIRLS 2016, PIRLS Literacy 2016, and ePIRLS 2016 data. Used in conjunction with either SPSS (IBM Corporation, 2016) or SAS (SAS Institute, 2012), the IEA IDB Analyzer provides a user-friendly interface to easily merge the various data file types of the PIRLS 2016 International Database and seamlessly takes into account the sampling information and the multiple imputed achievement scores to produce accurate statistical results.

Example analyses will illustrate some of the capabilities of the IEA IDB Analyzer (version 4.0) to compute a variety of statistics, including means and percentages of students in specified subgroups, mean student achievement in specified subgroups, regression coefficients, and percentages of students reaching achievement benchmark levels. The examples use student, home, teacher, and school background data to replicate some of the PIRLS 2016 results included in *PIRLS 2016 International Results in Reading* (Mullis, Martin, Foy, & Hooper, 2017).<sup>1</sup>

Users should be able to perform statistical analyses with the IEA IDB Analyzer with a basic knowledge of the PIRLS 2016 International Database. Chapter 4 gives a more detailed description of the data files contained in the International Database, including their structure and contents, along with a description of all the supporting documentation provided with the International Database.

### 2.2 The IEA IDB Analyzer

Developed by the IEA (IEA Hamburg), the IEA IDB Analyzer is an interface for SPSS and SAS, both well-known statistical analysis software. The IEA IDB Analyzer enables users to combine data files, either SPSS or SAS, from IEA's large-scale assessments and conduct analyses using either SPSS or SAS, without actually writing programming code. The IEA IDB Analyzer generates SPSS and SAS syntax that takes into account information from the sampling design in the computation of statistics and their standard errors. In addition, the generated syntax makes appropriate use of plausible values for calculating estimates of achievement scores and their standard errors, combining both sampling variance and imputation variance.

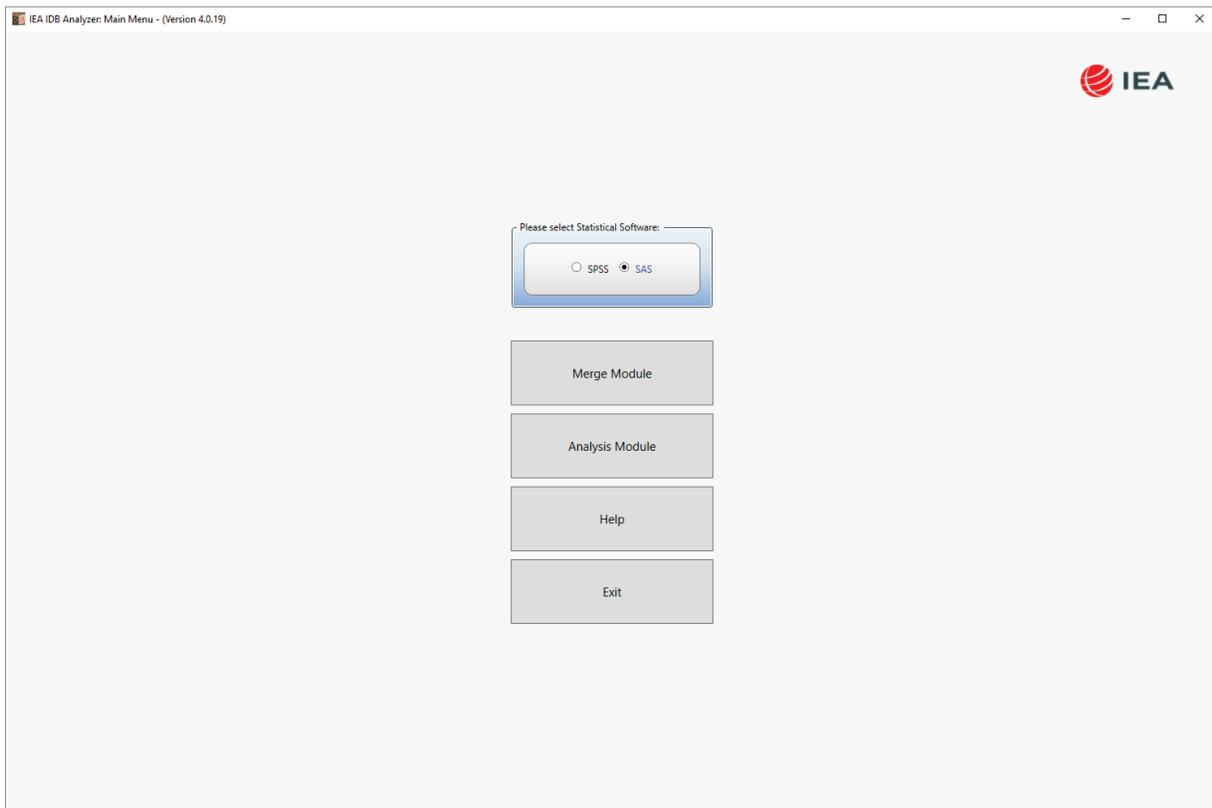
---

<sup>1</sup> Although this chapter does not offer examples of analyses with the ePIRLS 2016 data, users can use the same sets of instructions provided for the PIRLS 2016 examples to analyze the ePIRLS 2016 data.

The IEA IDB Analyzer consists of two modules—the merge module and the analysis module—which are integrated and executed in one common application. The merge module is used to create analysis datasets by combining data files of different types and from different countries, and selecting subsets of variables for analysis. The analysis module provides procedures for computing various statistics and their standard errors. The latest version of the IEA IDB Analyzer—version 4.0—is available for download from the IEA website: <http://www.iea.nl/data.html>. Once installed, the IEA IDB Analyzer application can be accessed by using the START menu in Windows or by clicking its icon, which is typically installed on the computer’s desktop window.

When the IEA IDB Analyzer application is launched, the main window will appear, as shown in Exhibit 2.1. Users are first directed to choose either SPSS or SAS as their statistical software of choice. For the examples in this chapter, we will use the SAS software. The main window also will direct users to the **Merge Module**, the **Analysis Module**, the **Help** manual, or simply **Exit** the application. The IEA IDB Analyzer has an extensive manual, accessible through the **Help** button, which users are encouraged to consult for full details on all the functionalities and features of the IEA IDB Analyzer.

**Exhibit 2.1** IEA IDB Analyzer Main Window



## 2.3 Merging Files with the IEA IDB Analyzer

The IEA IDB Analyzer uses the data files available from the PIRLS 2016 International Database and User Guide webpage (<http://timssandpirls.bc.edu/pirls2016/international-database/>) and from the IEA Study Data Repository (<http://www.iea.nl/data.html>). The PIRLS 2016, PIRLS Literacy 2016, and ePIRLS 2016 data files are disseminated separately by file type and for each country. In addition to allowing users to combine like datasets from more than one country for cross-country analyses, the merge module allows for the combination of data from different sources (e.g., student, home, teacher, and school) into one SPSS or SAS dataset for subsequent analyses.

Before doing any statistical analyses with the PIRLS 2016 International Database, users should download and copy the contents of the International Database either on their computer or on a server. For the purposes of this chapter, we will assume all data files are copied within the folder titled “C:\PIRLS2016\Data.” Additionally, users who plan to analyze PIRLS 2016 data with SAS will need to convert the SAS export files provided in the PIRLS 2016 International Database into SAS data files.<sup>2</sup> This process is described in Chapter 3 of this User Guide.

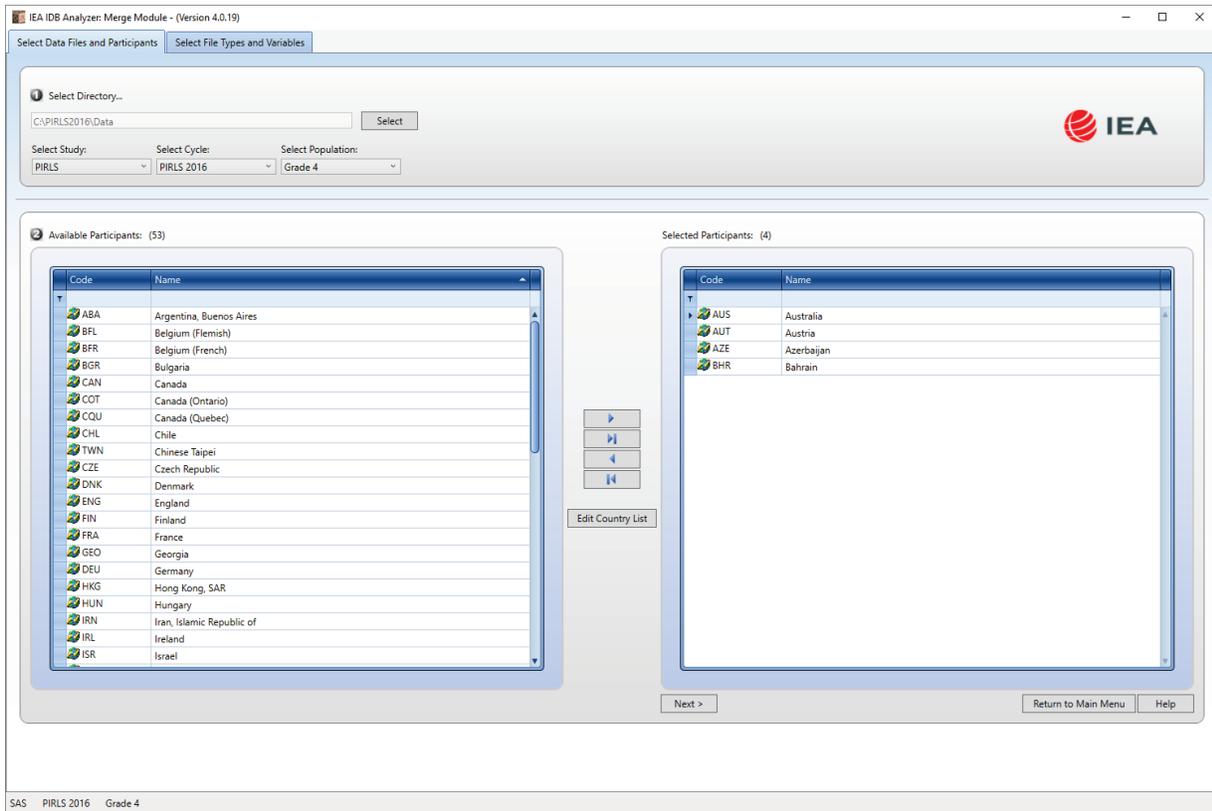
The following steps will create an SPSS or SAS data file with data from multiple countries and/or multiple file types:

1. Start the IEA IDB Analyzer from the START menu and click the **Merge Module** button.
2. Under the **Select Data Files and Participants** tab and in the **Select Directory** field, browse to the folder where all data files are located. For example, in Exhibit 2.2, all SAS data files are located in the folder “C:\PIRLS2016\Data.” The program will automatically recognize and complete the **Select Study**, **Select Cycle**, and **Select Population** fields and list all countries available in this folder as possible candidates for merging. If the folder contains data from more than one IEA study, or from more than one cycle or population, users should select the desired combination of study, cycle, and population for analyses. In Exhibit 2.2, PIRLS 2016 is selected.
3. Click a country of interest from the **Available Participants** list and click the **right arrow** button (▶) to move it to the **Selected Participants** panel. Individual countries can be moved directly to the **Selected Participants** panel by double-clicking on them. To select multiple countries, hold the CTRL key of the keyboard when clicking countries. Click the **tab-right arrow** button (▶) to move all countries to the **Selected Participants** panel. In Exhibit 2.2, Australia, Austria, Azerbaijan, and Bahrain are selected.
4. Click the **Next >** button to proceed to the next step. The software will open the **Select File Types and Variables** tab of the merge module, as shown in Exhibit 2.3, to select the file types and the variables to be included in the merged data file.
5. Select the files for merging by checking the appropriate boxes to the left of the window. For example, in Exhibit 2.3, the student background data files are selected.

---

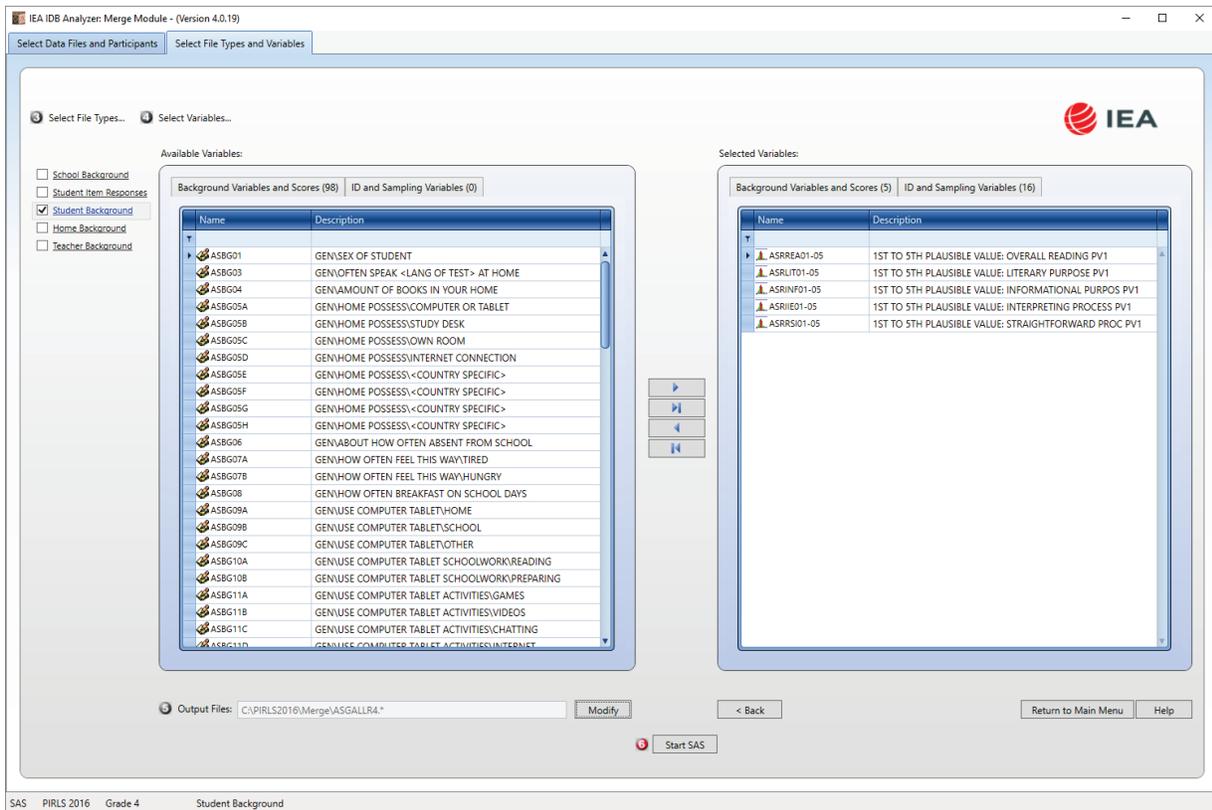
<sup>2</sup> Because SAS data files are specific to their computing environment and may not be readable in all environments, the PIRLS 2016 International Database provides data files in SAS export (EXP) format. These SAS export files will generate SAS data files, suitable for analysis in the IEA IDB Analyzer, after conversion in the appropriate computing environment (see Chapter 3).

**Exhibit 2.2 IEA IDB Analyzer Merge Module: Select Data Files and Participants**



6. Select the variables of interest from the **Available Variables** list in the left panel. The codebook files (described in Chapter 4 of this User Guide) provide the names of all variables available in the PIRLS 2016 International Database. Variables are selected by clicking on them and moved to the **Selected Variables** list by clicking the **right arrow** (▶) button. Clicking the **tab-right arrow** (▶) button selects and moves all variables to the **Selected Variables** list. Note that there are two tabs: “Background Variables and Scores” and “ID and Sampling Variables.” All achievement scores and all identification and sampling variables are selected automatically by the IEA IDB Analyzer.
7. Specify the desired name for the merged data file and the folder where it will be stored in the **Output Files** field by clicking the **Define/Modify** button. The IEA IDB Analyzer also will create a SAS syntax file (\*.SAS) of the same name and in the same folder with the code necessary to perform the merge. In the example shown in Exhibit 2.3, the syntax file ASGALLR4.SAS and the merged data file ASGALLR4.SAS7BDAT both will be created and stored in the folder “C:\PIRLS2016\Merge.” In SPSS, it will be the syntax file ASGALLR4.SPS and the merged data file ASGALLR4.SAV. The merged data file will contain all the variables listed in the **Selected Variables** panel on the right.
8. Click the **Start SAS** button to create the SAS syntax file and open it in a SAS editor window ready for execution. The syntax file can be executed by opening the **Run** menu of SAS and selecting the **Submit** menu option. In SPSS, it is the **All** option from the **Run** menu. The IEA IDB Analyzer will display a warning if it is about to overwrite an existing file in the specified folder.

**Exhibit 2.3 IEA IDB Analyzer Merge Module: Select File Types and Variables**



Once SPSS or SAS has completed its execution, it is important to check the SPSS output window or SAS log for possible warnings. If warnings appear, they should be examined carefully because they might indicate that the merge process was not performed properly and that the resulting merged data file might not be as expected.<sup>3</sup>

***Merging Student and Teacher Data Files***

The teachers in the PIRLS 2016 International Database do not constitute representative samples of teachers in the participating countries. Rather, they are the teachers of nationally representative samples of students. Therefore, analyses with teacher data should be made with students as the units of analysis and reported in terms of students who are taught by teachers with a particular attribute.

Teacher data are analyzed by linking the students to their teachers. The student-teacher linkage data files (AST) are used for this purpose and the IEA IDB Analyzer will make use of them automatically. Thus, to analyze teacher data, it is sufficient to select the **Teacher Background** file type in the **Select File Types and Variables** tab of the IEA IDB Analyzer merge module. To analyze student and teacher background data simultaneously, however, both the **Student Background** and **Teacher Background** file types must be selected in the

<sup>3</sup> For more information on how to use the IEA IDB Analyzer, and for troubleshooting, users should consult the Help manual.

**Select File Types and Variables** tab. The variables of interest need to be selected separately for both file types, as follows:

1. Click the **Student Background** file type so that it appears checked and highlighted. The **Background Variables and Scores** listed in the left panel will include all available variables from the student background data files.
2. Select the variables of interest from the left panel and click the **right arrow** (▶) button to move these variables to the **Selected Variables** panel on the right. Click the **tab-right arrow** (⏪) button to select all available variables.
3. Next, click the **Teacher Background** file type, selecting the variables of interest from the **Background Variables and Scores** panel on the left in the same manner shown in steps 1 and 2.
4. Specify the folder and merged data file name in the **Output Files** field, as described earlier.
5. Click the **Start SAS** button to create the SAS syntax file that will produce the required merged data file, which then can be run by opening the **Run** menu of SAS and selecting the **Submit** menu option.

Steps 1 and 2 above are required only if student background data and teacher background data are analyzed simultaneously. It is not recommended to combined both types of files and analyze only student data; the results may not be correct (see the *Sampling and Weighting Variables* section in Chapter 4 of this User Guide).

### ***Merging Student and School Data Files***

Because PIRLS 2016, PIRLS Literacy 2016, and ePIRLS 2016 have representative samples of schools, it is possible to compute reasonable statistics with schools as units of analysis. However, the school samples were designed to optimize the student samples and the student-level results. For this reason, it is preferable to analyze school-level variables as attributes of the students, rather than as elements in their own right. Therefore, analyzing school data should be done by linking the students to their schools.

To merge the student and school background data files, select both the **Student Background** and **School Background** file types in the **Select File Types and Variables** tab of the IEA IDB Analyzer merge module. The variables of interest to be included in the merged data file are selected separately by file type, as described above in “Merging Student and Teacher Data Files” and using the same set of instructions.

### ***Merging Student and Home Data Files***

The parents of the students participating in PIRLS 2016, PIRLS Literacy 2016, and ePIRLS 2016 responded to the Learning to Read Survey, usually called the Home Questionnaire. Their responses are included in the home background data files (ASH). Although home background variables are located in their own files, they are in essence attributes of the students and must be analyzed in the same manner as student background variables. This will require users to merge

the home background data files with the student background data files by selecting both the **Home Background** and **Student Background** file types in the merge module of the IEA IDB Analyzer. Variables of interest to be included in the merged data file are chosen separately by file type, as described above in “Merging Student and Teacher Data Files” and using the same set of instructions.

### *Combining Data from PIRLS and PIRLS Literacy*

With the inclusion of the PIRLS and PIRLS Literacy assessments in 2016 for reporting reading achievement on the same scale, proper analysis of reading requires combining PIRLS Literacy 2016 data with PIRLS 2016 data, especially for countries that participated in both assessments. This can be done using the **Merge Module** of the IEA IDB Analyzer to create a PIRLS 2016 data file (ASGALLR4, for example) and a PIRLS Literacy 2016 data file (ASGALLL2, for example) as described above, and then combining the two data files together. A short program, both in SPSS and SAS, is provided with the PIRLS 2016 International Database to perform this combination of files from both assessments and is described in Chapter 3 of this User Guide. The resulting combined file (ASGALLR4\_CMB, for example) can be used in the **Analysis Module** of the IEA IDB Analyzer to produce reading results that combine data from both assessments, as was done for all results reported in *PIRLS 2016 International Results in Reading*.

### *Merged Data Files for the User Guide Examples*

To conduct the analysis examples presented in this chapter, a number of merged data files were created following the instructions provided above. Because the examples presented in this User Guide report PIRLS reading achievement, we produced merged data files for both PIRLS 2016 and PIRLS Literacy 2016 and then used the special program to combine the files from both assessments. The following merged data files were created with all available background variables and achievement scores selected:

ASGALLR4_CMB	Merged student background data files for all countries
ASHALLR4_CMB	Merged home and student background data files for all countries
ATGALLR4_CMB	Merged teacher background data files for all countries
ACGALLR4_CMB	Merged school and student background data files for all countries

## **2.4 Performing Analyses with the IEA IDB Analyzer**

The IEA IDB Analyzer can perform statistical analyses on any files created using the **Merge Module**. The following statistical procedures are available in the **Analysis Module** of the IEA IDB Analyzer.

### Percentages and Means

Compute percentages, means, and standard deviations for selected analysis variables by subgroups defined by grouping variable(s). Plausible values can be included as analysis variables.

Percentages only

Compute percentages by subgroups defined by grouping variable(s).

Linear Regression

Compute linear regression coefficients for selected independent variables to predict a dependent variable by subgroups defined by grouping variable(s). Plausible values can be included as dependent or independent variables.

Logistic Regression

Compute logistic regression coefficients for selected independent variables to predict a dependent variable by subgroups defined by grouping variable(s). Plausible values can be included as dependent or independent variables. When used as a dependent variable, plausible values will be dichotomized using a specified cutpoint, such as one of the PIRLS International Benchmarks.

Correlations

Compute means, standard deviations, and correlation coefficients for selected analysis variables by subgroups defined by grouping variable(s). Plausible values can be included as analysis variables.

Benchmarks

Compute percentages of students meeting a set of user-specified achievement benchmarks, in particular the PIRLS International Benchmarks, by subgroups defined by grouping variable(s).

Percentiles

Compute the score points that separate a given proportion of the distribution of a continuous analysis variable by subgroups defined by the grouping variable(s). Plausible values can be included as analysis variables.

All available features of the IEA IDB Analyzer are described extensively in its Help manual. All statistical procedures offered in the analysis module of the IEA IDB Analyzer make appropriate use of sampling weights, and standard errors are computed using the jackknife repeated replication (JRR) method.<sup>4</sup> Percentages, means, linear regressions, correlations, and percentiles may be specified with or without achievement scores. When achievement scores are used, the analyses are performed five times (once for each plausible value) and the results are aggregated to produce accurate estimates of achievement and standard errors that incorporate both sampling and imputation errors. To conduct analyses using achievement scores, select the

---

<sup>4</sup> Starting with TIMSS 2015 and PIRLS 2016, the jackknife repeated replication method has been modified to include both replicates within each sampling zone, as described in [Chapter 4 of Methods and Procedures in PIRLS 2016](#).

**Use PVs** option from the **Plausible Value Option** drop-down menu. The various variables required to perform an analysis are input into specific variable fields according to their purpose.

#### Grouping Variables

This is a list of variables to define subgroups of interest. The list must consist of at least one grouping variable. By default, the IEA IDB Analyzer includes the variable IDCNTY used to distinguish the participating countries. Additional variables can be selected from the available list. If the **Exclude Missing from Analysis** option is checked, only cases that have non-missing values in the grouping variables will be used in the analysis. If it is not checked, missing values become reporting categories.

#### Analysis Variables

This is a list of variables for which means, percentages, correlations, or percentiles are to be computed. Usually, more than one analysis variable can be selected. To compute statistics based on achievement scores, it is necessary to select the **Use PVs** option in the **Plausible Value Option** drop-down menu, and select the achievement scores of interest in the **Plausible Values** field.

#### Plausible Values

This section is used to identify the set of plausible values to be used when achievement scores are the analysis variable for computing statistics. Select the **Use PVs** option in the **Plausible Value Option** drop-down menu before specifying the achievement scores in the **Plausible Values** field.

#### Independent Variables

This is a list of variables to be treated as independent variables for a linear or logistic regression analysis. More than one independent variable can be selected. Categorical variables and continuous variables can be specified as independent variables. When specifying categorical variables as independent variables, they can be treated either as “effect coding” or “dummy coding” using the **Contrast** drop-down menu.<sup>5</sup> Achievement scores also can be included as an independent variable. To specify achievement scores as an independent variable, it is necessary to select the **Use PVs** option in the **Plausible Value Option** drop-down menu and select the achievement scores of interest in the **Plausible Values** field.

#### Dependent Variable

This is the variable to be used as the dependent variable when a linear or logistic regression analysis is specified. Only one dependent variable can be listed and can be either a background variable or achievement scores. To use achievement scores as the dependent variable, select the

---

<sup>5</sup> Effect coding and dummy coding of categorical variables are described in the Help manual of the IEA IDB Analyzer.

Use **PVs** option in the **Plausible Value Option** drop-down menu, click on the **Plausible Values** radio button in the **Dependent Variable** section, and select the achievement scores of interest in the **Plausible Values** field.

#### Achievement Benchmarks

These are the values that will be used as cutpoints on an achievement scale, selected in the **Plausible Values** section, for computing the percentages of students meeting the specified benchmarks. Multiple cutpoints can be specified, each separated by a blank space. A drop-down menu is available to select the PIRLS International Benchmarks.

#### Percentiles

These are the percentiles that will be calculated from the distribution of a continuous analysis variable selected in the **Analysis Variables** section. Achievement scores can be selected as an analysis variable. Select the **Use PVs** option in the **Plausible Value Option** drop-down menu and select the achievement scores of interest in the **Plausible Values** field. Multiple percentiles can be specified, each separated by a blank space.

#### Weight Variable

This is the sampling weight variable that will be used in the analysis. The IEA IDB Analyzer automatically selects the appropriate weight variable for analysis based on the file types included in the merged data file. Generally, this will be TOTWGT, but SENWGT and HOUWGT also are available for student-level analyses with student, home, or school data. TCHWGT will be used when analyzing reading teacher data. Chapter 4 of this User Guide provides more information on the PIRLS 2016 sampling weights.

## 2.5 PIRLS Analyses with Student Background Data

Many analyses of the PIRLS 2016 International Database can be undertaken using the student background data. This section presents examples of actual analyses used to produce exhibits in *PIRLS 2016 International Results in Reading*. This section also includes examples of linear regression analyses and computing percentages of students reaching the PIRLS International Benchmarks.

A first example computes national average achievement, whereas a second example computes national average achievement by gender. In both cases, the IEA IDB Analyzer uses the sampling weights, implements the jackknife repeated replication method to compute appropriate sampling errors, effectively performs the computations five times (once for each plausible value), and aggregates the results to produce accurate estimates of average achievement and standard errors that incorporate both sampling and imputation errors. A third example expands on the second example by performing a test of significance on the gender difference using linear regression. A fourth example computes the percentages of students reaching the PIRLS International Benchmarks.

### *Student Background Data Analysis with Achievement*

In our first example, we want to replicate the analysis of the overall distribution of reading achievement. These results are presented in [Exhibit 1.1 of PIRLS 2016 International Results in Reading](#) and are repeated here in Exhibit 2.4. Because the results in this exhibit are based on plausible values, we need to make sure we include them when we create the file using the merge module, and also indicate that our analysis will make use of achievement scores.

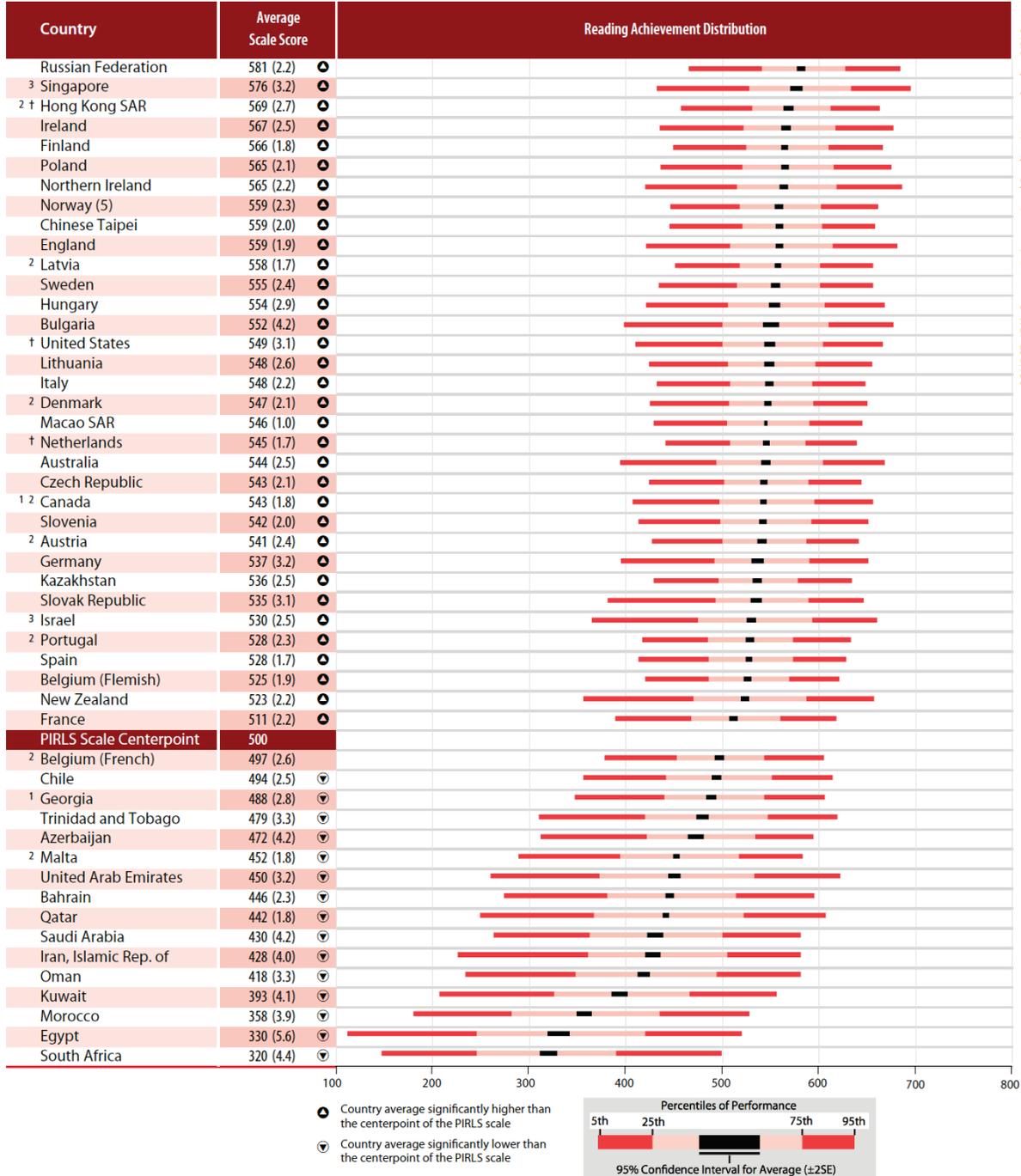
The **Percentages and Means** statistic type with the **Use PVs** option selected will compute percentages and average achievement scores based on plausible values and their respective standard errors. Because we will be analyzing reading achievement, we will create merged data files from both PIRLS 2016 (ASGALLR4) and PIRLS Literacy 2016 (ASGALLL2) and then combine them into a single merged data file (ASGALLR4\_CMB), as was explained earlier.

After creating the merged data file ASGALLR4\_CMB, the analysis module of the IEA IDB Analyzer is used to perform the analysis in the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.
2. Select the merged data file ASGALLR4\_CMB as the **Analysis File** by clicking the **Select** button.
3. Select **PIRLS (Using Student Weights)** as the **Analysis Type**.
4. Select **Percentages and Means** as the **Statistic Type**.
5. Select **Use PVs** as the **Plausible Value Option**.
6. The variable IDCNTRY is selected automatically as **Grouping Variables**. No additional grouping variables are needed for this analysis.
7. Specify the achievement scores to be used for the analysis by clicking the **Plausible Values** field to activate it. Select ASRREA01–05 from the list of available variables and move it to the **Plausible Values** field by clicking the **right arrow** (▶) button in this section.
8. The **Weight Variable** is selected automatically by the software; TOTWGT is selected by default because this example analysis uses student background data.
9. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button. In this example, the output files are labeled REA\_Overall and saved in the “C:\PIRLS2016\Analysis” folder.
10. Click the **Start SAS** button to create the SAS syntax file and open it in a SAS program editor window. The syntax file can be executed by opening the **Run** menu of SAS and selecting the **Submit** menu option. If necessary, the IEA IDB Analyzer will display a prompt to confirm the overwriting of existing files.

**Exhibit 2.4 Exhibit of Example Student Background Data Analysis with Achievement, Taken from PIRLS 2016 International Results in Reading (Exhibit 1.1)**

**Exhibit 1.1: Distribution of Reading Achievement**



SOURCE: IEA's Progress in International Reading Literacy Study – PIRLS 2016

Note: Five countries and one benchmarking entity participated in the PIRLS Literacy assessment: Egypt, Iran, Kuwait, Morocco, and South Africa as well as Denmark (3). Iran and Morocco also took part in the fourth grade assessment and their results are based on an average of both assessments.

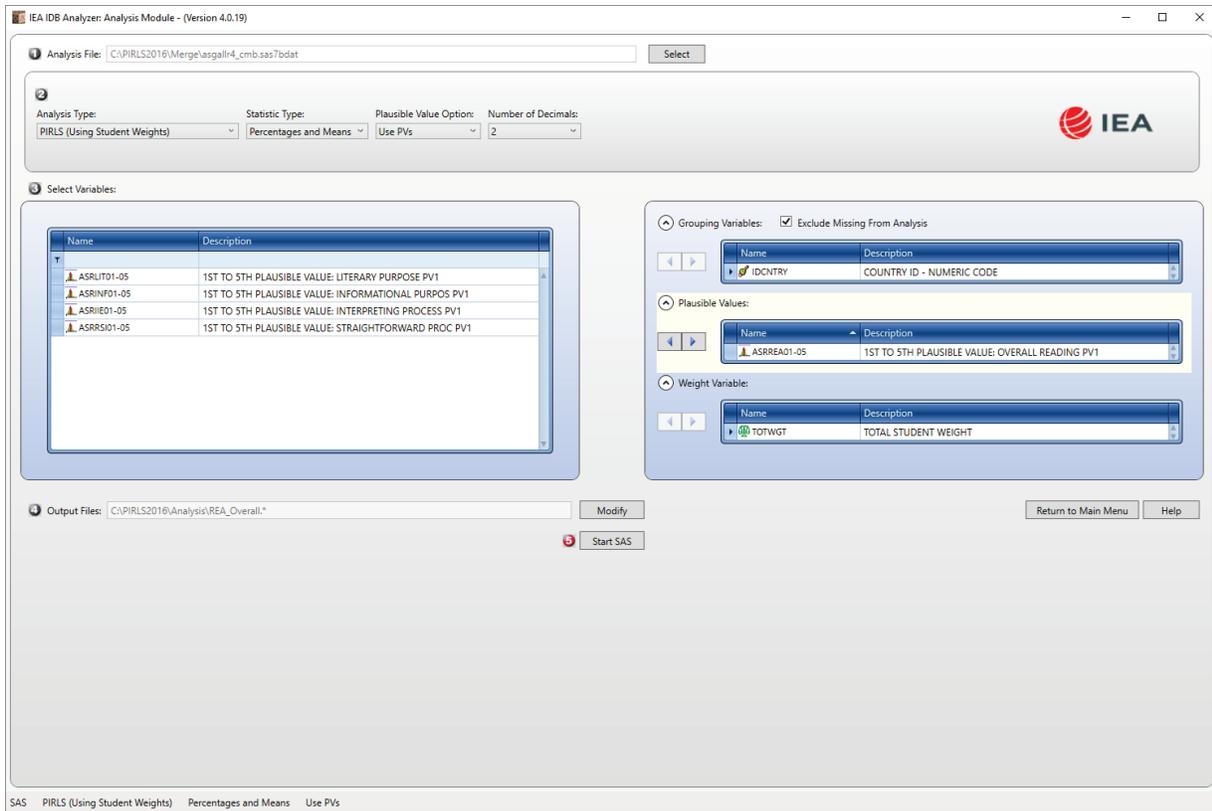
The PIRLS achievement scale was established in 2001 based on the combined achievement distribution of all countries that participated in PIRLS 2001. To provide a point of reference for country comparisons, the scale centerpoint of 500 was located at the mean of the combined achievement distribution. The units of the scale were chosen so that 100 scale score points corresponded to the standard deviation of the distribution.

See Appendix C.1 for target population coverage notes 1, 2, and 3. See Appendix C.4 for sampling guidelines and sampling participation notes †, ‡, and §.

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

Exhibit 2.5 shows the completed analysis module for this example analysis, and Exhibit 2.6 displays the results with four example countries: Australia, Bulgaria, Egypt, and Iran.

**Exhibit 2.5 IEA IDB Analyzer Setup for Example Student Background Data Analysis with Achievement**



Each country’s results are presented on a single line. The countries are identified in the first column and the second column reports the number of valid cases. The third column identifies the sum of weights of the sampled students. The fourth column is the standard error of the sum of weights. The next four columns report the percentage of students in each country and its standard error, followed by the estimated average reading achievement and its standard error. The standard deviation of the achievement scores and its standard error are reported in the next two columns and the last column reports the percentage of missing data.

**Exhibit 2.6 SAS Output for Example Student Background Data Analysis with Achievement**

**Average of ASRREA0 by (IDCNTRY)**

COUNTRY ID - NUMERIC CODE	n	TOTAL STUDENT WEIGHT	sumw_se	pct	pct_se	mnpv	mnpv_se	sdpv	sdpv_se	pctmiss
Australia	6341	280089.08	5090.07	1.46	0.03	544.36	2.53	83.60	1.58	0.00
Bulgaria	4281	58538.59	1050.12	0.31	0.01	551.54	4.24	85.20	2.70	0.00
Egypt	6957	1543298.62	35988.91	8.06	0.18	330.47	5.65	123.62	2.79	0.00
Iran, Islamic Republic of	8766	2401073.77	73621.89	12.54	0.35	427.90	3.95	107.50	2.85	0.00

As shown in the first line of Exhibit 2.6, Australia had valid data for 6,341 students and these sampled students represented a population of 280,089 eligible students. The average fourth grade reading achievement in Australia was 544.36 (standard error of 2.53) and its standard deviation was 83.60 (standard error of 1.58).

Because we used the file ASGALLR4\_CMB for this example, it is important to note that the results for Australia and Bulgaria are based on their participation in PIRLS 2016 and results for Egypt are based on its participation in PIRLS Literacy 2016. Also, results for Iran are based on the combined samples from PIRLS 2016 and PIRLS Literacy 2016 since Iran participated in both assessments. This will be true for all the examples we will present in this chapter.

### ***Student Background Data Analysis with Achievement by Gender***

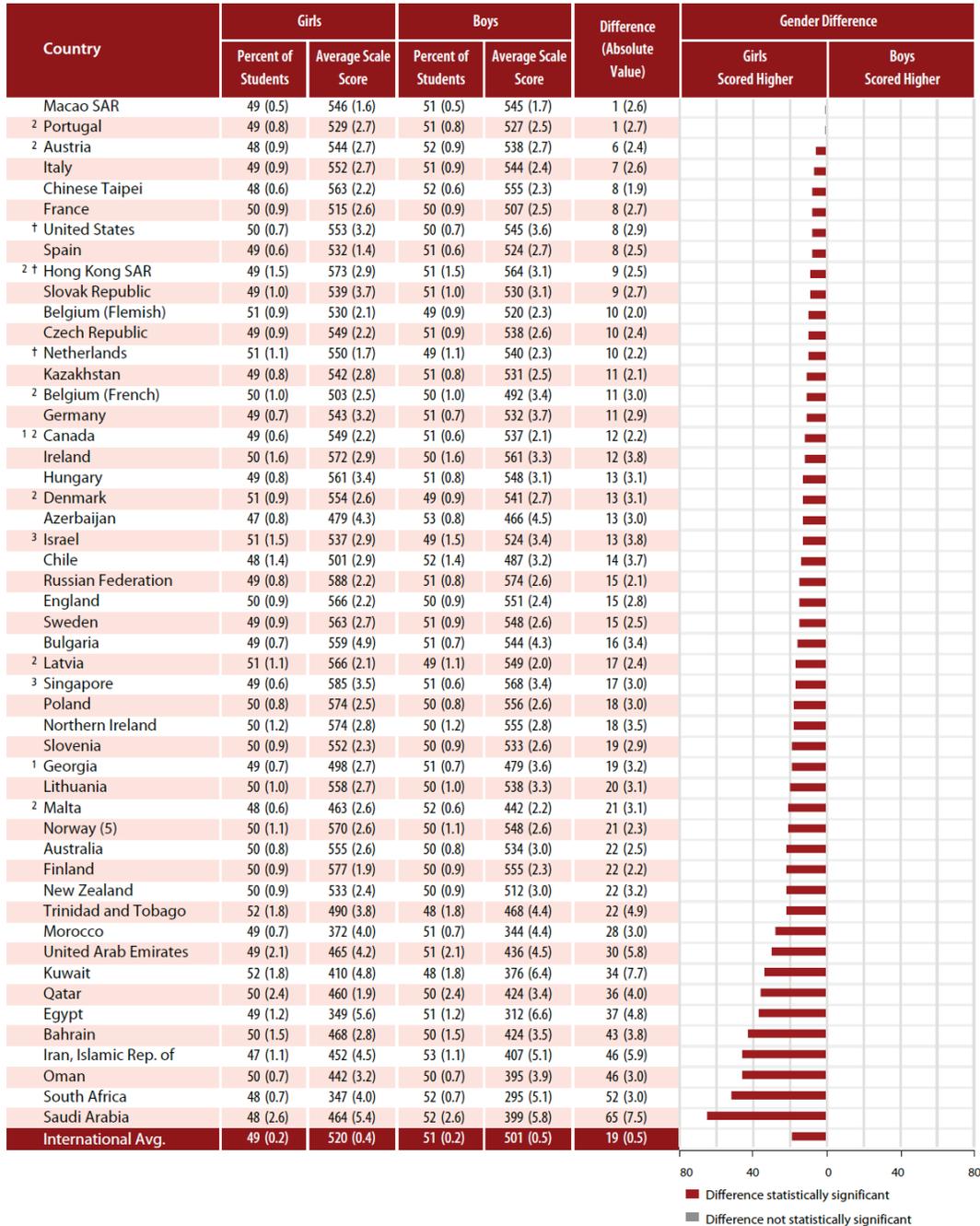
In our second example, we wish to replicate another set of results presented in [PIRLS 2016 International Results in Reading](#). We are interested in investigating the relationship between students' gender and reading achievement. These results, presented in [Exhibit 1.5 of PIRLS 2016 International Results in Reading](#), are repeated here in Exhibit 2.7. Because the results in this exhibit are based on plausible values, we must ensure that these values are included when creating the input file, and also indicate that this analysis will make use of achievement scores.

After reviewing the appropriate codebook, we observe that the variable ITSEX contains categorical information on the gender of students, and this variable is found in the student background data files. The **Percentages and Means** statistic type and the **Use PVs** plausible value option will allow us to compute the percentages and average achievement based on plausible values and their respective standard errors. The analysis module of the IEA IDB Analyzer is used to perform this analysis using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.
2. Select the merged data file ASGALLR4\_CMB as the **Analysis File** by clicking the **Select** button.
3. Select **PIRLS (Using Student Weights)** as the **Analysis Type**.
4. Select **Percentages and Means** as the **Statistic Type**.
5. Select **Use PVs** as the **Plausible Value Option**.
6. Specify the variable ITSEX as a second grouping variable by clicking the **Grouping Variables** field to activate it. Select ITSEX from the list of available variables and move it to the **Grouping Variables** field by clicking the **right arrow** (▶) button in this section.
7. Specify the achievement scores to be used for the analysis by clicking the **Plausible Values** field to activate it. Select ASRREA01–05 from the list of available variables and move it to the **Plausible Values** field by clicking the **right arrow** (▶) button in this section.

**Exhibit 2.7 Exhibit of Example Student Background Data Analysis with Achievement by Gender, Taken from PIRLS 2016 International Results in Reading (Exhibit 1.5)**

**Exhibit 1.5: Average Reading Achievement by Gender**

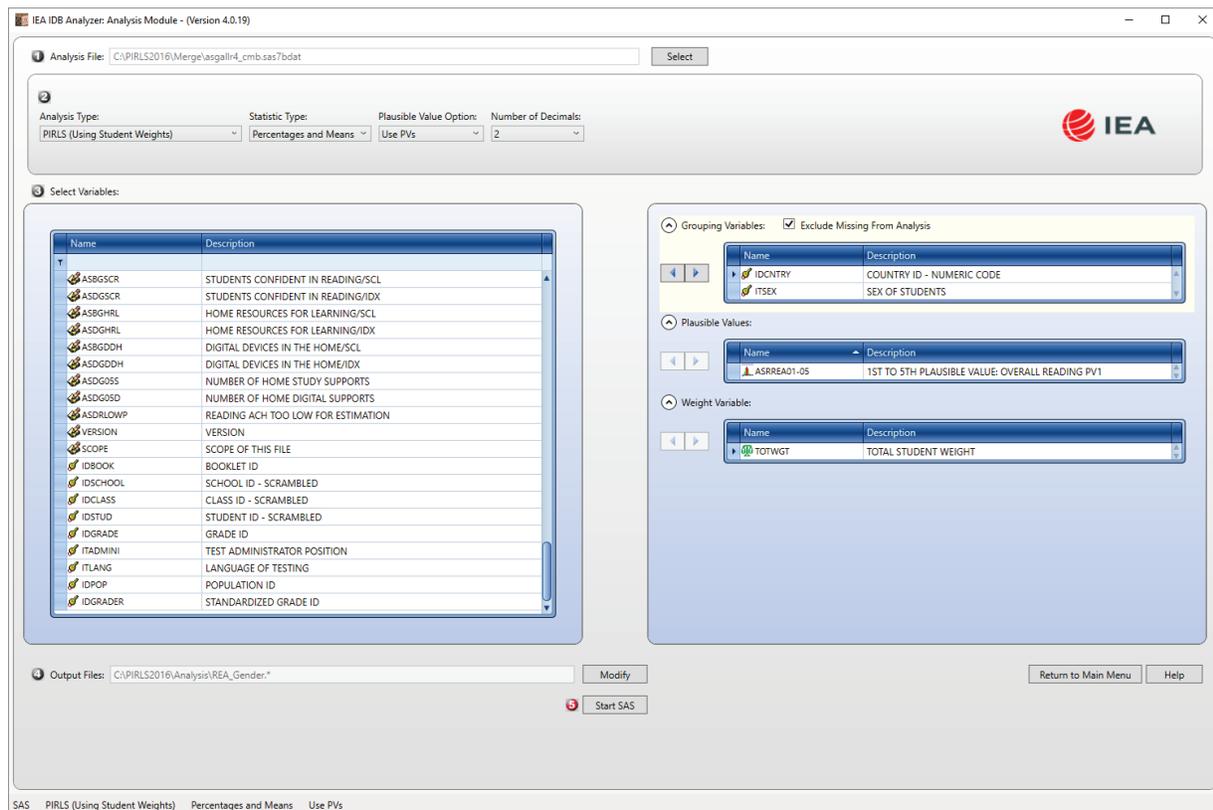


See Appendix C.1 for target population coverage notes 1, 2, and 3. See Appendix C.4 for sampling guidelines and sampling participation notes †, ‡, and ≡.  
( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

8. The **Weight Variable** is selected automatically by the software; TOTWGT is selected by default because this example analysis uses student background data.
9. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button. In this example, the output files are labeled REA\_Gender and saved in the "C:\PIRLS2016\Analysis" folder.
10. Click the **Start SAS** button to create the SAS syntax file and open it in a SAS editor window. The syntax file can be executed by opening the **Run** menu of SAS and selecting the **Submit** menu option. If necessary, the IEA IDB Analyzer will display a prompt to confirm the overwriting of existing files.

Exhibit 2.8 shows the completed analysis module for this example analysis and the results are presented in Exhibit 2.9.

**Exhibit 2.8 IEA IDB Analyzer Setup for Example Student Background Data Analysis with Achievement by Gender**



Each country's results are displayed on two lines, one for each value of the ITSEX variable. The countries are identified in the first column and the second column describes the category of ITSEX being reported (1 for girls and 2 for boys). The third column reports the number of valid cases and the fourth the sum of weights of the sampled students. The fifth column is the standard error of the sum of weights. The next four columns report the percentage of students in each category and its standard error, followed by the estimated average reading achievement

and its standard error. The standard deviation of the achievement scores and its standard error are reported in the next two columns.

**Exhibit 2.9 SAS Output for Example Student Background Data Analysis with Achievement by Gender**

**Average of ASRREA0 by (IDCNTY ITSEX)**

COUNTRY ID - NUMERIC CODE	SEX OF STUDENTS	n	TOTAL STUDENT WEIGHT	sumw_se	pct	pct_se	mnpv	mnpv_se	sdpv	sdpv_se	pctmiss
Australia	1	3153	140415.32	3448.19	50.13	0.78	555.10	2.62	79.50	1.79	0.00
	2	3188	139673.76	3249.52	49.87	0.78	533.56	2.98	86.19	1.91	0.00
Bulgaria	1	2119	28967.21	699.27	49.48	0.74	559.42	4.89	84.30	3.21	0.00
	2	2162	29571.38	657.84	50.52	0.74	543.82	4.27	85.37	2.86	0.00
Egypt	1	3410	763154.76	24127.97	49.45	1.16	349.36	5.55	115.29	2.83	0.00
	2	3547	780143.86	26516.48	50.55	1.16	311.99	6.60	128.59	3.41	0.00
Iran, Islamic Republic of	1	4283	1117514.16	38718.70	46.54	1.09	452.35	4.52	98.03	3.36	0.00
	2	4483	1283559.61	51492.80	53.46	1.09	406.61	5.09	110.80	3.44	0.00

From the first two lines of results, 50.13% of fourth grade students in Australia were girls and 49.87% were boys. The average reading achievement of girls was 555.10 (standard error of 2.62) and it was 533.56 for boys (standard error of 2.98).

**Linear Regression Analysis with Student Background Data**

This example is an extension of the previous example (Student Background Data Analysis with Achievement by Gender), where we will examine the difference in reading achievement between girls and boys and determine if it is statistically significant. The results of this example also are presented in *Exhibit 1.5 of PIRLS 2016 International Results in Reading* and shown in Exhibit 2.7, in the column labeled “Difference.”

It is worth noting that in our previous example the IEA IDB Analyzer performed the same gender difference significance test in the background. In the “C:\PIRLS2016\Analysis” folder, there is a CSV file (readable in Excel), REA\_Gender\_ASRREA0\_by\_ITSEX\_Sig, with the same results we will produce using the linear regression statistical method in our example. The IEA IDB Analyzer’s Help manual provides a description of this CSV file and its contents.

ITSEX has a value of one (1) for girls and two (2) for boys. By using ITSEX as a categorical variable in the IEA IDB Analyzer with dummy coding and category 1 as the reference category, the regression intercept will be the estimated average reading achievement of girls, and the regression slope will be the estimated change in average reading achievement for boys.

This example of a regression analysis is performed by the **Analysis Module** of the IEA IDB Analyzer using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.
2. Specify the data file ASGALLR4\_CMB as the **Analysis File** by clicking the **Select** button.
3. Select **PIRLS (Using Student Weights)** as the **Analysis Type**.
4. Select **Linear Regression** as the **Statistic Type**.

5. Select **Use PVs** as the **Plausible Value Option**.
6. The variable IDCNTY is selected automatically as **Grouping Variables**. No additional grouping variables are needed for this analysis.
7. Click the **Categorical Variables** field in the **Independent Variables** section to activate it and select the variable ITSEX as the independent variable. This is done by selecting ITSEX from the list of available variables and moving it to the **Categorical Variables** field by clicking the **right arrow** (▶) button in this section. By clicking the **Contrast** field of ITSEX, its drop-down menu will appear, from which **Dummy Coding** should be selected. By default, the IEA IDB Analyzer will recognize that ITSEX has two categories and it will select category 1 as the reference category. These settings should not be changed.
8. Click the **Plausible Values** radio button in the **Dependent Variable** section and select ASRREA01–05 as the analysis variable. This is done by selecting ASRREA01–05 from the list of available variables and moving it to the **Plausible Values** field by clicking the **right arrow** (▶) button in this section.
9. The **Weight Variable** is selected automatically by the software; TOTWGT is selected by default because this example analysis uses student background data.
10. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button. In this example, the output files are labeled REA\_Gender\_Diff and saved in the “C:\PIRLS2016\Analysis” folder.
11. Click the **Start SAS** button to create the SAS syntax file and open it in a SAS editor window. The syntax file will be executed by opening the **Run** menu of SAS and selecting the **Submit** menu option. If necessary, the IEA IDB Analyzer will display a prompt to confirm the overwriting of existing files.

Exhibit 2.10 shows the completed analysis module for this example analysis, and Exhibit 2.11 displays the results.

Each country’s results are displayed on two lines: the first for the intercept and the second for the ITSEX coefficient. Generally, there will be as many lines per country as there are regression coefficients, including the intercept. The first line of results in Exhibit 2.11 labeled “Intercept” is the estimated average reading achievement of fourth grade girls in Australia, which was 555.10 with a standard error of 2.62. This estimate concurs with the results obtained in the previous example (Exhibit 2.9). From the second line of results labeled “ITSEX\_D2,” the fourth grade boys in Australia had a negative average reading achievement difference of 21.55. With an estimated standard error of 2.47, this achievement difference is statistically significant at the 95% confidence level. Counting the two regression coefficients together (555.10 – 21.55) yields the estimated average reading achievement of fourth grade boys in Australia, which was 533.56 in Exhibit 2.9, within rounding. The IEA IDB Analyzer also computes standardized regression coefficients, whereby the dependent and independent variables are standardized to have a mean of zero (0) and standard deviation of one (1).

**Exhibit 2.10 IEA IDB Analyzer Setup for Example Student Background Data Linear Regression Analysis**

**Exhibit 2.11 SAS Output for Example Student Background Data Linear Regression Analysis**

**Equation Coefficients of (ITSEX\_D2) predicting (ASRREA0) by (IDCNTRY)**

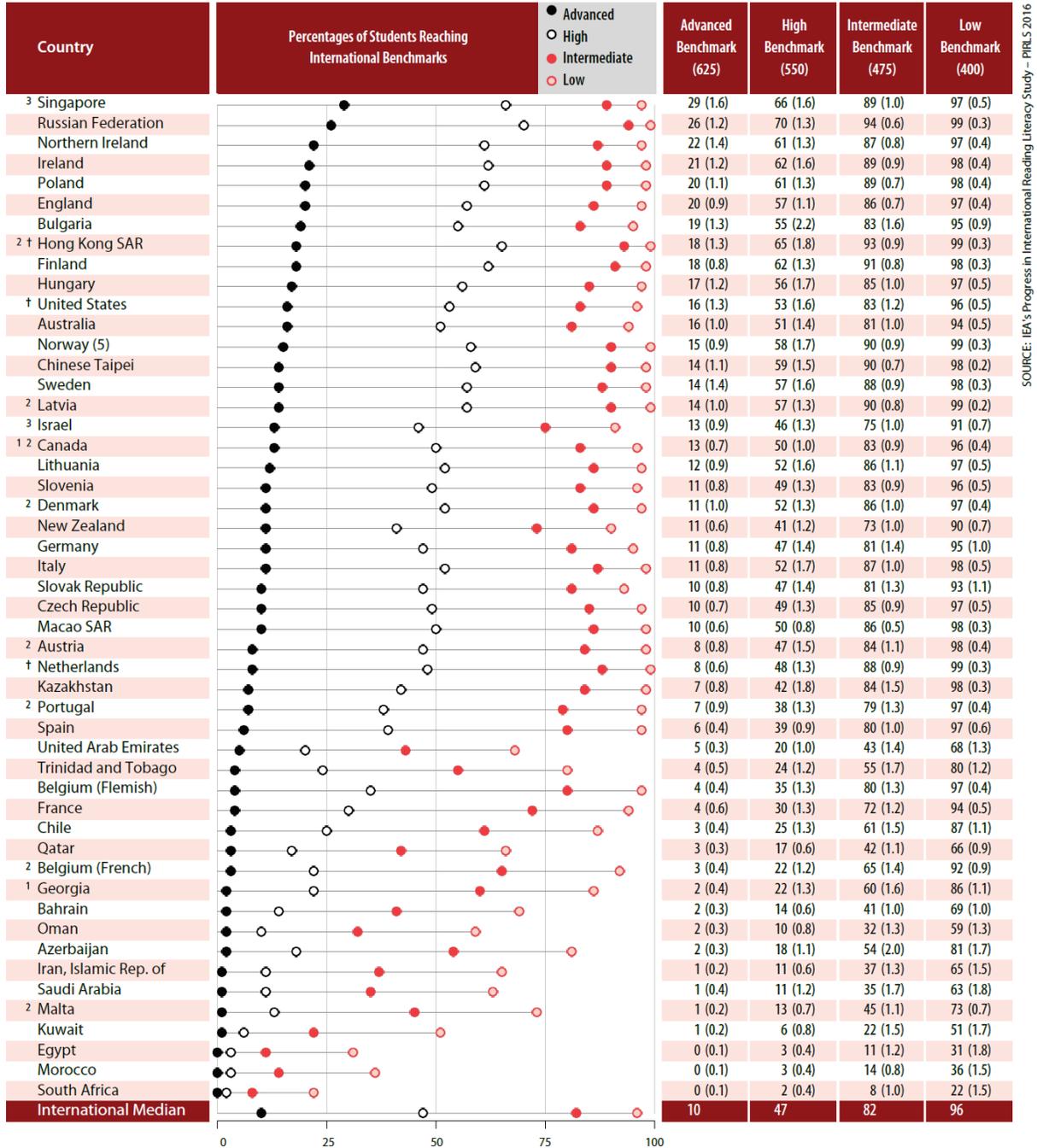
COUNTRY ID - NUMERIC CODE	variable	label	Parameter Estimate	Estimate_se	Estimate_t	Standardized Estimate	StandardizedEst_se
Australia	Intercept	Intercept	555.10	2.62	212.02	0.00	0.00
	ITSEX_D2		-21.55	2.47	-8.73	-0.13	0.01
Bulgaria	Intercept	Intercept	559.42	4.89	114.41	0.00	0.00
	ITSEX_D2		-15.60	3.45	-4.53	-0.09	0.02
Egypt	Intercept	Intercept	349.36	5.55	62.91	0.00	0.00
	ITSEX_D2		-37.37	4.77	-7.83	-0.15	0.02
Iran, Islamic Republic of	Intercept	Intercept	452.35	4.52	100.18	0.00	0.00
	ITSEX_D2		-45.74	5.85	-7.82	-0.21	0.03

### The PIRLS International Benchmarks

This section describes how to use the IEA IDB Analyzer to perform analyses of student achievement in relation to the PIRLS International Benchmarks. As an example, we will compute the percentages of students reaching the four PIRLS 2016 International Benchmarks of reading achievement (advanced, high, intermediate, and low) using the merged ASGALLR4\_CMB data file. These results, presented in [Exhibit 2.1 of PIRLS 2016 International Results in Reading](#), are shown here in Exhibit 2.12.

**Exhibit 2.12 Example Exhibit of PIRLS International Benchmarks Analysis, Taken from PIRLS 2016 International Results in Reading (Exhibit 2.1)**

**Exhibit 2.1: Performance at the International Benchmarks of Reading Achievement**



SOURCE: IEA's Progress in International Reading Literacy Study – PIRLS 2016

See Appendix C.1 for target population coverage notes 1, 2, and 3. See Appendix C.4 for sampling guidelines and sampling participation notes †, ‡, and ≡.  
( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

This example is performed by the analysis module of the IEA IDB Analyzer using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.
2. Specify the data file ASGALLR4\_CMB as the **Analysis File** by clicking the **Select** button.
3. Select **PIRLS (Using Student Weights)** as the **Analysis Type**.
4. Select **Benchmarks** as the **Statistic Type**.
5. Select the **Cumulative** option under the **Benchmark Option** drop-down menu to get cumulated percentages of students reaching the PIRLS International Benchmarks.
6. The variable IDCNTY is selected automatically as **Grouping Variables**. No additional grouping variables are needed for this analysis.
7. Specify the achievement scores to be used for the analysis by clicking the **Plausible Values** field. Select ASRREA01–05 from the list of available variables and move it to the **Plausible Values** field by clicking the **right arrow** (▶) button in this section.
8. Specify the PIRLS 2016 International Benchmarks—400, 475, 550, and 625 (low, intermediate, high, and advanced, respectively). These values can be entered manually in the **Achievement Benchmarks** field, each separated by a blank space, or they can be selected by clicking on the drop-down menu available for this field.
9. The **Weight Variable** is selected automatically by the software; TOTWGT is selected by default because this example analysis uses student background data.
10. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button. In this example, the output files are labeled REA\_Benchmarks and saved in the “C:\PIRLS2016\Analysis” folder.
11. Click the **Start SAS** button to create the SAS syntax file and open it in a SAS editor window. The syntax file will be executed by opening the **Run** menu of SAS and selecting the **Submit** menu option. If necessary, the IEA IDB Analyzer will display a prompt to confirm the overwriting of existing files.

Exhibit 2.13 shows the completed analysis module for this example analysis, and Exhibit 2.14 presents the results.

As shown in the first few lines of results, 94.49% of the fourth grade students in Australia performed at or above the Low International Benchmark of 400, with a standard error of 0.55; 80.85% of the students reached the Intermediate International Benchmark, with a standard error of 1.04; 51.23% of the students reached the High International Benchmark, with a standard error of 1.40; and 15.93% of the students reached the Advanced International Benchmark, with a standard error of 1.02.

**Exhibit 2.13 IEA IDB Analyzer Setup for Example PIRLS International Benchmarks Analysis**

**Exhibit 2.14 SAS Output for Example PIRLS International Benchmarks Analysis**

**Percent at or Above (400 475 550 625) of ASRREA0 by (IDCNTRY)**

COUNTRY ID - NUMERIC CODE	cutvar	n	TOTAL STUDENT WEIGHT	sumw_se	pct	pct_se
Australia	1.At or above 400	5917.0	264645.54	5133.37	94.49	0.55
	2.At or above 475	5039.4	226465.45	5155.69	80.85	1.04
	3.At or above 550	3187.2	143487.97	4906.93	51.23	1.40
	4.At or above 625	984.6	44610.64	3012.17	15.93	1.02
Bulgaria	1.At or above 400	4136.0	55486.81	1163.33	94.79	0.93
	2.At or above 475	3715.8	48369.04	1309.65	82.63	1.64
	3.At or above 550	2578.4	31946.95	1401.73	54.57	2.18
	4.At or above 625	948.8	11223.31	774.68	19.17	1.28
Egypt	1.At or above 400	2175.8	476028.79	30203.95	30.84	1.76
	2.At or above 475	822.8	177078.93	19963.41	11.47	1.24
	3.At or above 550	190.4	39104.13	6443.52	2.53	0.41
	4.At or above 625	32.4	6928.53	2302.29	0.45	0.15
Iran, Islamic Republic of	1.At or above 400	6228.6	1558378.85	57783.19	64.90	1.51
	2.At or above 475	3699.8	884747.89	41457.85	36.85	1.34
	3.At or above 550	1139.0	257034.13	16515.57	10.70	0.63
	4.At or above 625	146.6	34597.37	5969.74	1.44	0.25

### *Computing Correlations*

In addition to the analyses described above, the IEA IDB Analyzer is able to compute correlations among a set of variables. Thus it can compute correlations between background variables (such as the context questionnaire scales Students Like Reading and Home Resources for Learning), between achievement scores (such as reading for literary experience and reading to acquire and use information), and between a combination of both (such as students' age and their reading achievement). While these types of analyses will not be demonstrated here, the steps for conducting them are similar to those described previously: select the grouping variables, the analysis variables, the achievement scores (if necessary), and confirm the weight variable. The output will display, for each group defined by the grouping variables, the correlation coefficients for each possible pair of variables.

### *Calculating Percentiles of a Distribution*

The **Percentiles** statistic type is an additional tool provided by the IEA IDB Analyzer for analyzing the PIRLS 2016 data. This procedure will compute the percentiles of a distribution within any specified subgroups, along with appropriate standard errors. The distribution can be either a non-PV based analysis variable—such as a context questionnaire scale—or a specified set of plausible values.

### *Performing Logistic Regression*

The IEA IDB Analyzer can perform a logistic regression, with or without plausible values. Logistic regression is used to predict a binary response based on one or more predictor variables. Users can specify grouping variables, independent variables—with or without interactions—that can be categorical, continuous, or plausible values, and a dependent variable. Users will find useful information on performing logistic regression in the IEA IDB Analyzer's Help manual.

## **2.6 PIRLS Analyses with Home Background Data**

The PIRLS 2016 International Database includes data from a home context questionnaire completed by the parents of the students participating in either PIRLS 2016 or PIRLS Literacy 2016. Analyzing home background data from the PIRLS 2016 International Database requires that the home background data files (ASH) be merged with the student background data files (ASG) in order to retrieve the achievement scores and the required sample design variables. To illustrate how to perform analyses based on home background data, we will use an example that computes the average score on a context questionnaire scale, along with the percentages of students—with their average achievement—for the categories of the scale's corresponding index.

### *Analyses with a Context Questionnaire Scale*

PIRLS 2016 reports some context questionnaire data by creating context questionnaire scales based on Rasch modeling.<sup>6</sup> The context questionnaire scales are available in the International Database for analysis. Each context questionnaire scale variable is a Rasch score with an international centerpoint of 10 and an internationally set standard deviation of 2. From each context questionnaire scale, an index is derived that divides the range of scores on that scale into usually three categories: the most desirable scores (high values), the least desirable scores (low values), and the remaining scores in between.

These context questionnaire scales and their corresponding indices were reported in *PIRLS 2016 International Results in Reading*. Exhibit 2.15 shows one such example, *Exhibit 4.4 of PIRLS 2016 International Results in Reading*, reporting Parents Like Reading, based on parents' responses to a set of eight statements on this topic. Results on the Rasch scale are reported for each country as an "Average Scale Score" and its corresponding index is reported as the percentages of students in each category—Parents Very Much Like Reading, Parents Somewhat Like Reading, and Parents Do Not Like Reading—along with their average reading achievement in each category.

This example will replicate both the average scale score of the Parents Like Reading scale and the percentages of students and their average achievement in each category of its index. This will be done in two steps, both using the merged ASHALLR4\_CMB data file, where the student background data files were merged with the home background data files of both PIRLS 2016 and PIRLS Literacy 2016.

The first step of this example will compute the average scale score using the scale variable ASBHPLR. It is performed by the analysis module of the IEA IDB Analyzer using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.
2. Specify the data file ASHALLR4\_CMB as the **Analysis File**.
3. Select **PIRLS (Using Student Weights)** as the **Analysis Type**.
4. Select **Percentages and Means** as the **Statistic Type**.
5. Select **None Used** as the **Plausible Value Option**, because we will not use any achievement scores for this part of the analysis.
6. Specify the variable ASBHPLR as the analysis variable by clicking the **Analysis Variables** field to activate it. Select ASBHPLR from the list of available variables and move it to the **Analysis Variables** field by clicking the **right arrow (▶)** button in this section.
7. The **Weight Variable** is selected automatically by the software; TOTWGT is selected by default because this example analysis combines student background data with home background data.

<sup>6</sup> The context questionnaire scales are described in *Chapter 14 of Methods and Procedures in PIRLS 2016*.

**Exhibit 2.15 Example Exhibit of a Context Questionnaire Scale, Taken from PIRLS 2016 International Results in Reading (Exhibit 4.4)**

**Exhibit 4.4: Parents Like Reading**

Students Categorized by Parents' Reports

Country	Very Much Like		Somewhat Like		Do Not Like		Average Scale Score	Difference in Average Scale Score from 2011	
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement			
Ireland	47 (1.1)	588 (2.9)	40 (1.0)	560 (2.9)	13 (0.7)	544 (5.2)	10.3 (0.05)	-0.4 (0.07)	⬇️
Netherlands	s 46 (1.5)	566 (2.8)	39 (1.2)	548 (3.1)	16 (0.9)	525 (4.3)	10.0 (0.05)	s -0.4 (0.08)	⬇️
Malta	45 (0.8)	471 (2.3)	42 (0.9)	451 (2.6)	13 (0.5)	439 (3.9)	10.2 (0.03)	-0.4 (0.05)	⬇️
Denmark	44 (1.1)	564 (2.3)	38 (1.0)	542 (2.7)	17 (0.7)	530 (3.6)	10.1 (0.05)	-0.6 (0.06)	⬇️
Sweden	44 (1.1)	576 (2.9)	42 (1.1)	551 (3.0)	13 (0.8)	529 (4.4)	10.1 (0.05)	-0.8 (0.07)	⬇️
Azerbaijan	44 (1.3)	481 (5.1)	46 (1.1)	470 (4.4)	10 (0.9)	453 (6.8)	10.4 (0.06)	0.7 (0.08)	⬆️
Norway (5)	42 (1.0)	574 (2.6)	44 (0.9)	556 (2.5)	15 (0.9)	532 (3.9)	10.1 (0.05)	--	
Trinidad and Tobago	41 (1.0)	499 (3.6)	49 (0.9)	478 (4.3)	10 (0.6)	468 (6.1)	10.2 (0.04)	-0.5 (0.06)	⬇️
Finland	41 (1.0)	585 (2.2)	43 (0.9)	563 (2.5)	16 (0.7)	542 (3.4)	10.0 (0.05)	-0.5 (0.07)	⬇️
Spain	41 (0.8)	545 (2.1)	43 (0.7)	524 (2.3)	16 (0.7)	512 (2.5)	10.0 (0.04)	0.0 (0.05)	
Israel	40 (1.1)	554 (3.1)	47 (0.9)	522 (3.2)	13 (0.6)	519 (6.4)	10.0 (0.04)	r -0.5 (0.06)	⬇️
Austria	40 (1.1)	564 (2.3)	42 (0.8)	534 (2.7)	18 (0.8)	516 (3.7)	9.9 (0.05)	-0.4 (0.08)	⬇️
Canada	r 40 (0.8)	566 (2.1)	46 (0.7)	540 (2.5)	15 (0.5)	531 (3.3)	10.0 (0.03)	r -0.5 (0.05)	⬇️
Bulgaria	39 (1.4)	587 (3.1)	43 (1.4)	548 (4.2)	19 (1.7)	493 (9.1)	9.8 (0.09)	-0.2 (0.13)	⬇️
Hungary	38 (1.4)	583 (3.0)	45 (1.2)	547 (3.3)	17 (1.0)	516 (4.6)	9.8 (0.06)	-0.1 (0.08)	⬇️
Czech Republic	37 (1.0)	565 (2.4)	45 (0.7)	540 (2.2)	18 (0.8)	514 (4.2)	9.8 (0.05)	-0.2 (0.06)	⬇️
Italy	37 (1.0)	566 (2.5)	47 (0.9)	546 (2.7)	16 (0.8)	530 (3.8)	9.9 (0.04)	0.1 (0.06)	
Germany	r 36 (1.1)	578 (2.7)	44 (0.9)	540 (3.3)	20 (0.9)	509 (5.0)	9.7 (0.05)	r -0.4 (0.08)	⬇️
Poland	35 (0.8)	581 (2.9)	50 (0.8)	561 (2.3)	14 (0.8)	540 (4.3)	9.9 (0.04)	--	
Slovak Republic	35 (1.0)	566 (2.6)	46 (1.0)	532 (3.2)	19 (1.0)	489 (9.5)	9.6 (0.06)	-0.3 (0.08)	⬇️
Portugal	35 (0.9)	546 (3.0)	50 (0.9)	522 (2.4)	15 (0.7)	510 (5.4)	9.8 (0.04)	0.2 (0.05)	⬆️
Georgia	32 (1.0)	512 (2.9)	60 (1.1)	482 (2.9)	8 (0.8)	460 (11.3)	9.9 (0.04)	-0.2 (0.06)	⬇️
Belgium (French)	29 (0.9)	526 (2.8)	48 (0.9)	495 (3.2)	23 (0.8)	473 (3.6)	9.4 (0.04)	-0.4 (0.07)	⬇️
Belgium (Flemish)	28 (0.8)	546 (2.0)	48 (0.8)	527 (2.1)	24 (0.7)	509 (2.9)	9.3 (0.04)	--	
Lithuania	27 (1.0)	572 (3.2)	48 (1.3)	546 (3.5)	25 (1.1)	530 (3.9)	9.3 (0.04)	-0.3 (0.06)	⬇️
Slovenia	27 (0.9)	571 (3.0)	58 (1.1)	539 (2.2)	16 (0.7)	517 (3.6)	9.5 (0.03)	-0.3 (0.05)	⬇️
Bahrain	26 (0.7)	471 (3.2)	60 (0.9)	444 (2.9)	14 (0.6)	417 (4.3)	9.6 (0.02)	--	
Kazakhstan	26 (1.1)	545 (3.2)	67 (1.1)	533 (2.9)	7 (0.5)	531 (4.7)	9.9 (0.04)	--	
Latvia	26 (0.8)	579 (2.6)	52 (0.9)	557 (2.2)	22 (0.8)	541 (3.3)	9.4 (0.04)	--	
Kuwait	25 (1.0)	425 (6.1)	57 (1.0)	395 (4.5)	18 (0.8)	378 (7.5)	9.5 (0.04)	--	
Russian Federation	25 (0.8)	602 (2.5)	56 (0.8)	578 (2.3)	20 (0.7)	560 (3.6)	9.4 (0.03)	-0.2 (0.05)	⬇️
Singapore	25 (0.6)	603 (3.3)	57 (0.7)	572 (3.2)	18 (0.6)	561 (4.0)	9.4 (0.02)	-0.3 (0.03)	⬇️
Iran, Islamic Rep. of	25 (1.1)	459 (4.9)	62 (1.1)	429 (3.5)	14 (1.0)	374 (11.2)	9.5 (0.05)	-0.3 (0.06)	⬇️
South Africa	s 24 (0.7)	359 (7.0)	63 (1.1)	322 (5.3)	13 (0.9)	307 (7.0)	9.7 (0.03)	s -0.2 (0.05)	⬇️
Qatar	24 (0.7)	489 (3.3)	62 (0.8)	441 (2.0)	14 (0.4)	428 (4.3)	9.5 (0.02)	-0.2 (0.05)	⬇️
France	22 (0.8)	539 (3.2)	56 (0.8)	513 (2.2)	21 (0.9)	491 (4.0)	9.3 (0.04)	-0.2 (0.06)	⬇️
United Arab Emirates	22 (0.5)	496 (3.6)	65 (0.5)	445 (3.4)	13 (0.4)	436 (4.5)	9.5 (0.02)	-0.1 (0.03)	⬇️
Morocco	22 (0.9)	397 (4.3)	47 (1.5)	365 (4.4)	31 (1.8)	330 (6.0)	9.0 (0.06)	-0.3 (0.10)	⬇️
Chile	21 (0.8)	529 (3.3)	52 (0.9)	491 (2.8)	27 (0.9)	480 (3.7)	9.2 (0.04)	--	
Oman	21 (0.5)	450 (4.7)	67 (0.6)	418 (3.3)	12 (0.5)	380 (5.3)	9.5 (0.02)	0.0 (0.03)	
Saudi Arabia	21 (0.9)	454 (4.7)	62 (1.0)	428 (4.8)	17 (0.8)	414 (6.2)	9.4 (0.04)	-0.2 (0.06)	⬇️
Chinese Taipei	19 (0.8)	584 (2.9)	61 (0.7)	556 (2.5)	20 (0.8)	548 (2.8)	9.2 (0.03)	-0.2 (0.05)	⬇️
Hong Kong SAR	17 (0.8)	580 (3.1)	61 (0.9)	569 (3.2)	22 (0.8)	562 (3.5)	9.1 (0.03)	-0.2 (0.04)	⬇️
Macao SAR	17 (0.6)	563 (2.4)	62 (0.8)	544 (1.4)	22 (0.6)	538 (2.7)	9.1 (0.02)	--	
Egypt	13 (1.0)	385 (8.1)	54 (1.6)	345 (5.0)	33 (1.8)	286 (8.4)	8.7 (0.07)	--	
England	--	--	--	--	--	--	--	--	
United States	--	--	--	--	--	--	--	--	
Northern Ireland	x 49 (1.5)	602 (3.9)	36 (1.5)	577 (5.1)	15 (1.0)	568 (6.0)	10.3 (0.07)	x -0.4 (0.08)	⬇️
Australia	x 48 (1.4)	582 (3.5)	40 (1.3)	551 (3.6)	12 (0.9)	535 (5.3)	10.3 (0.06)	x -0.4 (0.09)	⬇️
New Zealand	x 47 (1.6)	567 (3.0)	40 (1.5)	524 (4.0)	13 (0.7)	511 (6.2)	10.3 (0.06)	x -0.6 (0.07)	⬇️
International Avg.	32 (0.1)	535 (0.5)	51 (0.1)	508 (0.5)	17 (0.1)	488 (0.8)			

SOURCE: IEA's Progress in International Reading Literacy Study – PIRLS 2016

This PIRLS questionnaire scale was established in 2011 based on the combined response distribution of all countries that participated in PIRLS 2011. To provide a point of reference for country comparisons, the scale centerpoint of 10 was located at the mean of the combined distribution. The units of the scale were chosen so that 2 scale score points corresponded to the standard deviation of the distribution.

Significantly higher than 2011 ⬆️  
Significantly lower than 2011 ⬇️

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

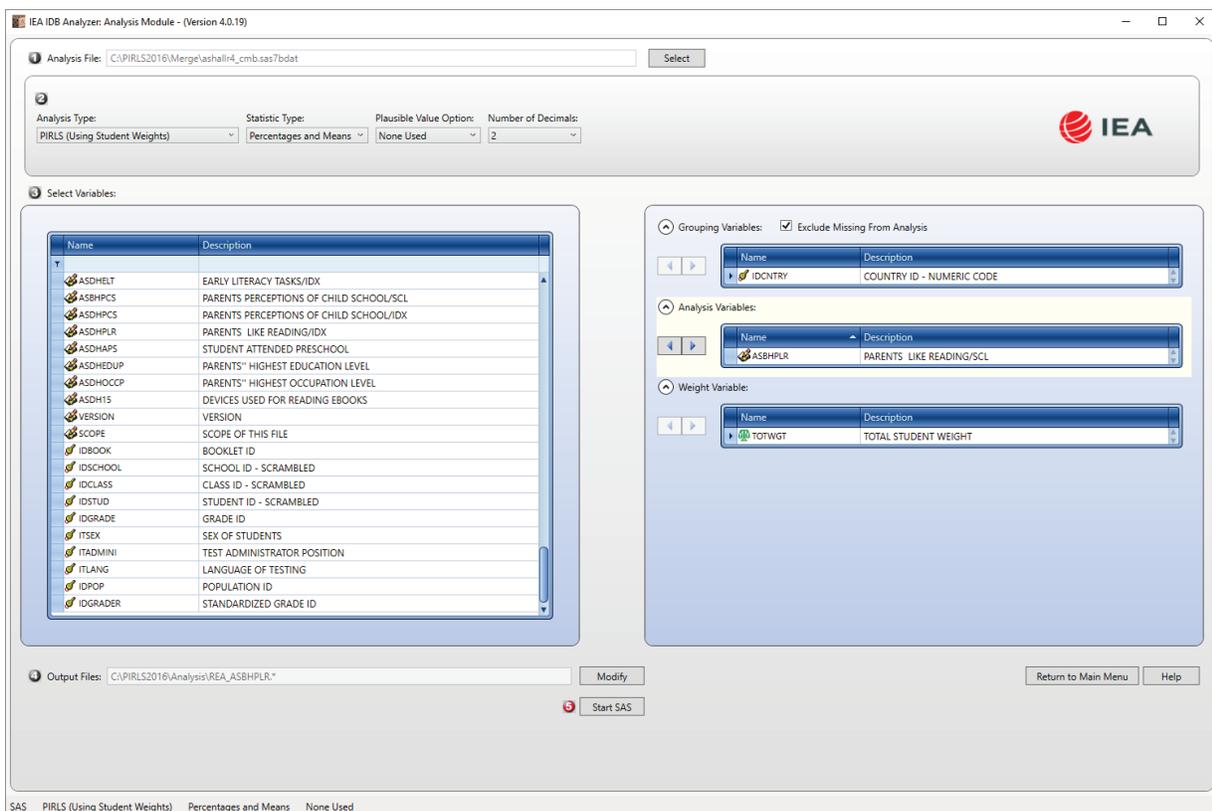
A dash (-) indicates comparable data not available. A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students. An "s" indicates data are available for at least 50% but less than 70% of the students. An "x" indicates data are available for less than 50% of the students—interpret with caution.

8. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button. In this example, the output files are labeled REA\_ASBHPLR and saved in the "C:\PIRLS2016\Analysis" folder.
9. Click the **Start SAS** button to create the SAS syntax file and open it in a SAS editor window. The syntax file can be executed by opening the **Run** menu of SAS and selecting the **Submit** menu option. If necessary, the IEA IDB Analyzer will display a prompt to confirm the overwriting of existing files.

Exhibit 2.16 shows the completed analysis module for this example analysis, and Exhibit 2.17 displays the results.

**Exhibit 2.16 IEA IDB Analyzer Setup for Example Context Questionnaire Scale Analysis (Step 1)**



As shown in the fourth line of the results for Iran, students in Iran scored 9.54 on the Parents Like Reading context questionnaire scale, with a standard error of 0.05. Note that this is markedly below the international centerpoint of 10. The IEA IDB Analyzer also computes the standard deviation of the context questionnaire scale and its standard error.

**Exhibit 2.17 SAS Output for Example Context Questionnaire Scale Analysis (Step 1)**

**Average of ASBHPLR by (IDCNTRY)**

COUNTRY ID - NUMERIC CODE	n	TOTAL STUDENT WEIGHT	sumw_se	pct	pct_se	mnx	mnx_se	sdx	sdx_se	pctmiss
Australia	2805	128059.98	4121.87	0.95	0.03	10.30	0.06	2.03	0.04	54.28
Bulgaria	4231	57638.53	1038.70	0.43	0.01	9.81	0.09	2.22	0.07	1.54
Egypt	6773	1505849.34	35729.74	11.21	0.25	8.75	0.07	1.63	0.04	2.43
Iran, Islamic Republic of	8606	2358178.41	68372.49	17.56	0.44	9.54	0.05	1.54	0.03	1.79

In the second step of this example, we will compute the percentages of students—with their average reading achievement—in each category of the corresponding index variable ASDHPLR. It is performed by the analysis module of the IEA IDB Analyzer using the following steps:

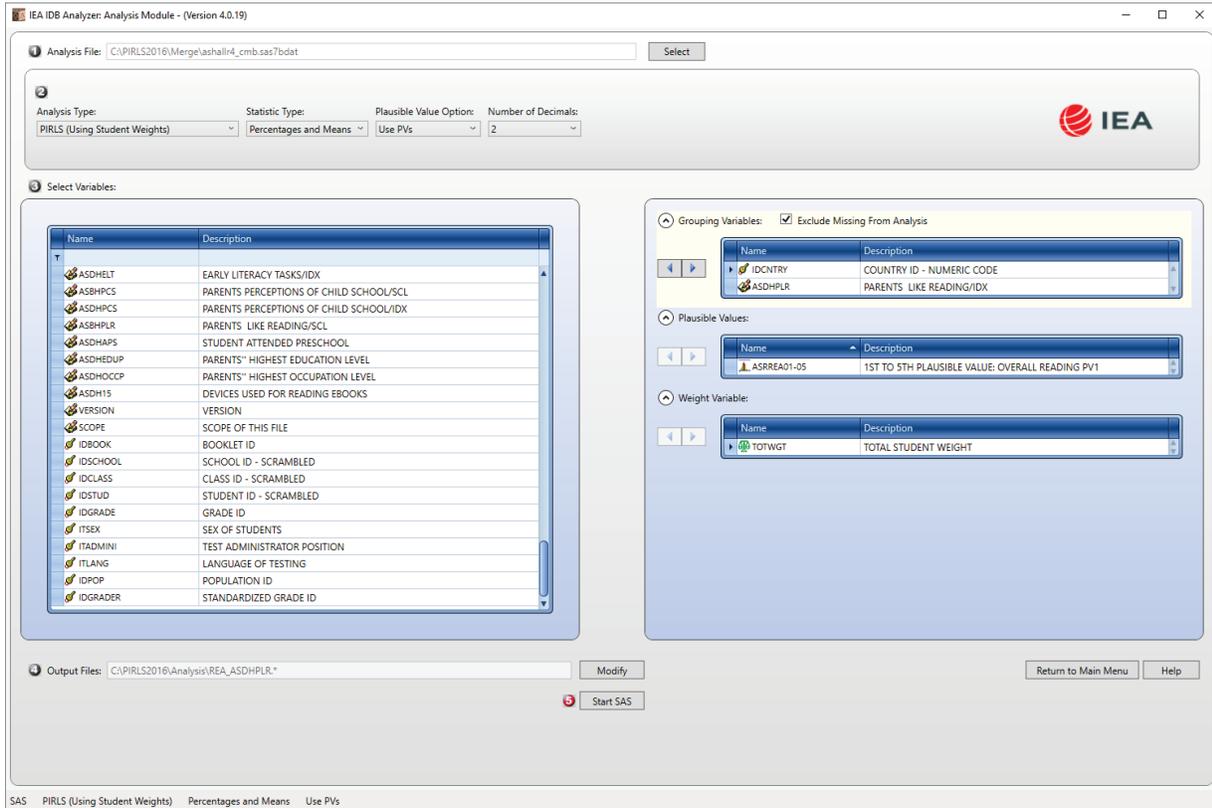
1. Open the **Analysis Module** of the IEA IDB Analyzer.
2. Specify the data file ASHALLR4\_CMB as the **Analysis File**.
3. Select **PIRLS (Using Student Weights)** as the **Analysis Type**.
4. Select **Percentages and Means** as the **Statistic Type**.
5. Select **Use PVs** as the **Plausible Value Option**, because we will be computing average achievement by the grouping variable ASDHPLR.
6. Specify the variable ASDHPLR as a second grouping variable by clicking the **Grouping Variables** field to activate it. Select ASDHPLR from the list of available variables and move it to the **Grouping Variables** field by clicking the **right arrow** (▶) button in this section.
7. Specify the achievement scores to be used for the analysis by clicking the **Plausible Values** field to activate it. Select ASRREA01–05 from the list of available variables and move it to the **Plausible Values** field by clicking the **right arrow** (▶) button in this section.
8. The **Weight Variable** is selected automatically by the software; TOTWGT is selected by default because this example analysis combines student background data with home background data.
9. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button. In this example, the output files are labeled REA\_ASDHPLR and saved in the “C:\PIRLS2016\Analysis” folder.
10. Click the **Start SAS** button to create the SAS syntax file and open it in a SAS editor window. The syntax file can be executed by opening the **Run** menu of SAS and selecting the **Submit** menu option. If necessary, the IEA IDB Analyzer will display a prompt to confirm the overwriting of existing files.

Exhibit 2.18 shows the completed analysis module for this example analysis and the results are presented in Exhibit 2.19.

As shown in the three lines of the results for Iran, 24.63% of fourth grade students in Iran had parents who very much like reading (standard error of 1.13) and their average reading

achievement was 458.52 (standard error of 4.94); 61.73% of students had parents who somewhat like reading (standard error of 1.13) and their average reading achievement was 429.26 (standard error of 3.54); and 13.64% of students had parents who do not like reading (standard error of 1.01) and their average reading achievement was 374.49 (standard error of 11.21).

**Exhibit 2.18 IEA IDB Analyzer Setup for Example Context Questionnaire Scale Analysis (Step 2)**



**Exhibit 2.19 SAS Output for Example Context Questionnaire Scale Analysis (Step 2)**

**Average of ASRREA0 by (IDCNTRY ASDHPLR)**

COUNTRY ID - NUMERIC CODE	PARENTS LIKE READING IDX	n	TOTAL STUDENT WEIGHT	sumw_se	pct	pct_se	mnpv	mnpv_se	sdpv	sdpv_se	pctmiss
Australia	1	1356	61209.57	3033.78	47.80	1.41	581.57	3.52	74.70	2.36	0.00
	2	1096	50999.37	1952.13	39.82	1.30	550.63	3.63	78.91	2.16	0.00
	3	353	15851.04	1261.74	12.38	0.91	535.29	5.34	74.35	5.72	0.00
Bulgaria	1	1823	22285.74	973.01	38.66	1.43	586.82	3.05	67.65	1.63	0.00
	2	1796	24689.09	976.21	42.83	1.36	547.84	4.20	80.15	2.61	0.00
	3	612	10663.71	958.43	18.50	1.73	492.85	9.09	89.84	6.00	0.00
Egypt	1	882	196733.34	16610.83	13.06	1.05	384.73	8.07	114.90	4.67	0.00
	2	3699	809324.01	29094.14	53.75	1.57	345.07	4.98	115.28	2.36	0.00
	3	2192	499791.99	29765.04	33.19	1.76	286.24	8.39	126.28	4.58	0.00
Iran, Islamic Republic of	1	2085	580798.91	27581.37	24.63	1.13	458.52	4.94	100.06	4.09	0.00
	2	5413	1455629.73	56626.78	61.73	1.13	429.26	3.54	100.40	2.23	0.00
	3	1108	321749.77	24987.00	13.64	1.01	374.49	11.21	127.01	6.25	0.00

## 2.7 PIRLS Analyses with Teacher Background Data

Analyses with teacher background data seek to make statements about students whose teachers have a given characteristic, rather than statements about teachers with a given characteristic. As our example of an analysis using teacher background data, we will investigate the percentage of fourth grade students and their reading achievement, according to their reading teachers' years of experience. The results of such an analysis are presented in [Exhibit 8.3 of PIRLS 2016 International Results in Reading](#) and are reproduced here in Exhibit 2.20.

We will use the **Percentages and Means** statistic type and select the **Use PVs** option to estimate the percentages of students with their average reading achievement by reporting categories of teachers' years of experience.

As with the previous examples, we first proceed to identify the variables relevant to the analysis in the appropriate files, and review the documentation for any specific national adaptations to the questions of interest. Because we are using a teacher-level variable, we need to look in the teacher background data files for the variable that contains the information on the years of experience of fourth grade reading teachers. The variable ATBG01 contains information on teachers' years of experience. That information was collapsed into four reporting categories and stored in the derived variable ATDG01 (see [Supplement 3 to this User Guide](#)).

The merged data file ATGALLR4\_CMB will be used for this example teacher background data analysis, which will be performed by the analysis module of the IEA IDB Analyzer using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.
2. Select the merged data file ATGALLR4\_CMB as the **Analysis File**.
3. Select **PIRLS (Using Teacher Weights)** as the **Analysis Type**.
4. Select **Percentages and Means** as the **Statistic Type**.
5. Select **Use PVs** as the **Plausible Value Option**.
6. Specify the variable ATDG01 as a second grouping variable by clicking the **Grouping Variables** field to activate it. Select ATDG01 from the list of available variables and move it to the **Grouping Variables** field by clicking the **right arrow** (▶) button in this section.
7. Specify the achievement scores to be used for the analysis by clicking the **Plausible Values** field to activate it. Select ASRREA01–05 from the list of available variables and move it to the **Plausible Values** field by clicking the **right arrow** (▶) button in this section.
8. The **Weight Variable** is selected automatically by the software; TCHWGT is selected by default because of the **Analysis Type** selected in step 3.
9. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button. In this example, the output files are labeled REA\_ATDG01 and saved in the "C:\PIRLS2016\Analysis" folder.

**Exhibit 2.20 Exhibit of Example Teacher Background Data Analysis, Taken from PIRLS 2016 International Results in Reading (Exhibit 8.3)**

**Exhibit 8.3: Teachers' Years of Experience**

Students Categorized by Teachers' Reports

Country	20 Years or More		At Least 10 but Less than 20 Years		At Least 5 but Less than 10 Years		Less than 5 Years		Average Years of Experience
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Australia	40 (3.3)	547 (4.4)	23 (2.9)	542 (4.2)	15 (2.7)	549 (7.4)	22 (2.8)	542 (7.6)	17 (0.8)
Austria	59 (3.7)	541 (3.1)	19 (2.9)	551 (5.3)	12 (2.2)	528 (6.8)	10 (2.0)	542 (6.8)	22 (0.9)
Azerbaijan	64 (3.3)	474 (5.7)	24 (2.9)	471 (9.6)	6 (1.8)	475 (17.7)	6 (1.6)	451 (10.0)	23 (0.7)
Bahrain	12 (2.5)	429 (9.9)	45 (4.8)	450 (4.8)	18 (3.9)	452 (9.0)	25 (2.6)	438 (7.8)	11 (0.5)
Belgium (Flemish)	37 (3.5)	529 (3.2)	37 (3.6)	529 (2.8)	10 (1.9)	516 (7.8)	17 (2.7)	514 (5.0)	17 (0.7)
Belgium (French)	47 (3.8)	503 (3.5)	31 (3.5)	494 (5.2)	15 (2.7)	494 (8.0)	6 (1.6)	479 (8.8)	19 (0.8)
Bulgaria	87 (2.9)	550 (4.7)	8 (2.3)	549 (13.1)	3 (1.2)	583 (8.8)	2 (0.8)	~ ~	28 (0.7)
Canada	28 (2.1)	545 (3.4)	37 (2.4)	545 (3.0)	20 (1.9)	540 (3.9)	16 (1.7)	540 (6.4)	14 (0.4)
Chile	28 (4.6)	503 (5.1)	27 (4.4)	498 (6.7)	28 (4.6)	487 (7.8)	17 (3.3)	501 (8.4)	15 (1.1)
Chinese Taipei	42 (4.0)	556 (2.6)	45 (4.1)	561 (3.0)	8 (2.3)	565 (5.8)	4 (1.6)	549 (9.9)	18 (0.6)
Czech Republic	54 (3.5)	545 (3.0)	24 (3.0)	545 (3.9)	11 (2.4)	537 (7.9)	11 (2.1)	538 (4.4)	20 (0.7)
Denmark	31 (3.8)	554 (3.6)	38 (4.3)	543 (3.7)	21 (3.1)	547 (4.3)	10 (2.2)	546 (6.6)	16 (0.8)
Egypt	43 (4.4)	315 (9.3)	32 (3.9)	353 (9.5)	18 (3.6)	332 (16.9)	7 (1.9)	318 (29.5)	17 (0.9)
England	20 (3.0)	563 (4.7)	26 (3.3)	554 (4.6)	18 (2.8)	563 (4.3)	35 (3.3)	557 (4.1)	11 (0.7)
Finland	38 (3.1)	567 (3.2)	25 (2.5)	565 (4.9)	18 (2.5)	563 (3.6)	18 (2.8)	571 (3.4)	16 (0.8)
France	30 (3.6)	521 (3.6)	41 (3.7)	510 (3.3)	18 (2.6)	509 (8.1)	12 (2.2)	495 (7.0)	16 (0.7)
Georgia	67 (3.5)	488 (3.2)	23 (3.1)	492 (6.4)	7 (1.9)	486 (11.4)	3 (1.2)	518 (16.0)	23 (0.8)
Germany	38 (3.8)	536 (6.7)	32 (3.7)	544 (4.0)	17 (2.8)	527 (8.7)	13 (2.5)	544 (8.0)	18 (0.9)
Hong Kong SAR	34 (3.9)	576 (5.0)	40 (4.3)	570 (4.8)	13 (3.0)	563 (8.6)	13 (2.4)	561 (6.6)	15 (0.6)
Hungary	61 (4.0)	558 (4.1)	22 (3.2)	560 (6.4)	13 (2.5)	535 (11.2)	4 (1.5)	532 (20.2)	24 (1.0)
Iran, Islamic Rep. of	49 (3.6)	444 (5.6)	23 (3.4)	415 (11.0)	15 (2.2)	404 (14.6)	13 (2.6)	415 (12.0)	17 (0.7)
Ireland	20 (3.3)	567 (5.0)	29 (3.4)	569 (4.0)	34 (4.1)	567 (4.8)	17 (2.6)	563 (5.5)	13 (0.9)
Israel	34 (3.5)	547 (6.3)	34 (4.0)	520 (8.1)	23 (3.8)	525 (9.6)	9 (2.4)	520 (16.6)	16 (0.8)
Italy	71 (3.0)	552 (2.7)	22 (2.9)	543 (5.4)	3 (0.8)	530 (12.9)	4 (1.7)	524 (13.8)	25 (0.8)
Kazakhstan	50 (4.0)	540 (3.5)	30 (3.7)	536 (5.2)	9 (2.2)	535 (8.5)	10 (2.2)	521 (7.5)	20 (0.8)
Kuwait	14 (3.6)	398 (22.8)	40 (5.5)	400 (8.8)	23 (4.0)	383 (8.0)	23 (4.6)	399 (15.6)	11 (0.7)
Latvia	81 (2.9)	557 (2.1)	12 (2.3)	565 (5.5)	4 (1.6)	556 (8.7)	3 (1.1)	548 (7.6)	27 (0.7)
Lithuania	86 (2.4)	548 (2.7)	11 (2.1)	549 (13.0)	2 (1.0)	~ ~	1 (0.6)	~ ~	27 (0.6)
Macao SAR	30 (0.1)	547 (1.7)	28 (0.1)	545 (2.4)	18 (0.1)	546 (2.3)	23 (0.1)	542 (2.8)	14 (0.0)
Malta	19 (0.1)	462 (3.8)	35 (0.2)	464 (2.6)	17 (0.1)	448 (3.8)	30 (0.1)	434 (2.5)	11 (0.0)
Morocco	56 (2.8)	346 (5.0)	21 (2.8)	364 (8.8)	13 (2.5)	369 (16.9)	10 (2.4)	383 (19.2)	21 (0.6)
Netherlands	39 (3.9)	548 (2.7)	29 (3.2)	545 (3.6)	15 (3.0)	545 (4.7)	17 (2.9)	540 (7.3)	17 (0.9)
New Zealand	21 (2.3)	529 (5.9)	40 (2.6)	532 (3.6)	18 (1.9)	521 (6.0)	22 (2.6)	520 (7.7)	13 (0.6)
Northern Ireland	45 (4.2)	558 (3.3)	29 (4.2)	570 (5.2)	15 (3.6)	577 (6.0)	11 (2.6)	554 (8.6)	18 (0.9)
Norway (5)	28 (3.5)	562 (4.0)	37 (4.1)	559 (2.8)	19 (3.3)	559 (4.5)	16 (2.9)	554 (5.2)	15 (0.8)
Oman	20 (2.9)	434 (8.0)	33 (3.1)	412 (4.4)	30 (3.2)	415 (7.2)	17 (2.2)	417 (5.4)	12 (0.5)
Poland	56 (3.8)	566 (2.9)	29 (3.3)	564 (4.0)	4 (1.3)	563 (8.8)	10 (2.6)	561 (8.2)	21 (0.9)
Portugal	46 (2.6)	529 (2.5)	48 (3.3)	527 (4.2)	4 (2.2)	538 (8.4)	2 (0.9)	~ ~	21 (0.4)
Qatar	14 (1.7)	452 (9.8)	27 (2.0)	440 (4.7)	33 (2.2)	440 (6.5)	27 (2.2)	442 (7.0)	10 (0.4)
Russian Federation	75 (3.3)	583 (2.9)	12 (2.5)	572 (9.1)	7 (1.8)	582 (10.6)	6 (1.7)	570 (10.2)	25 (0.9)
Saudi Arabia	33 (4.0)	424 (7.9)	36 (4.0)	420 (7.7)	23 (3.8)	435 (11.0)	8 (2.2)	458 (23.0)	15 (0.6)
Singapore	19 (2.3)	580 (9.1)	36 (2.9)	578 (5.3)	22 (2.0)	579 (6.4)	22 (2.4)	567 (6.3)	13 (0.5)
Slovak Republic	56 (3.7)	540 (3.9)	27 (3.1)	527 (6.6)	9 (1.8)	536 (9.8)	8 (1.9)	520 (13.3)	22 (0.8)
Slovenia	66 (3.7)	543 (2.6)	25 (3.3)	539 (3.3)	6 (1.8)	543 (6.9)	3 (1.2)	562 (7.2)	24 (0.7)
South Africa	40 (4.0)	315 (7.8)	24 (3.3)	325 (7.6)	13 (2.5)	313 (17.9)	23 (3.1)	322 (11.4)	15 (0.7)
Spain	47 (3.2)	527 (2.3)	26 (2.5)	531 (3.2)	18 (2.5)	526 (7.5)	8 (1.6)	525 (4.4)	19 (0.8)
Sweden	19 (3.2)	554 (4.9)	43 (4.7)	559 (3.1)	17 (3.8)	563 (6.4)	21 (3.5)	544 (4.3)	13 (0.7)
Trinidad and Tobago	44 (3.9)	478 (5.8)	36 (3.6)	472 (7.5)	9 (2.0)	510 (14.7)	11 (2.9)	482 (14.8)	18 (0.8)
United Arab Emirates	12 (1.4)	444 (11.7)	34 (2.6)	433 (5.0)	30 (2.2)	465 (7.6)	24 (2.4)	471 (6.5)	10 (0.3)
United States	30 (3.4)	557 (4.9)	37 (3.8)	549 (5.7)	16 (3.3)	553 (6.9)	17 (2.4)	535 (8.7)	15 (0.8)
International Avg.	42 (0.5)	513 (0.9)	30 (0.5)	511 (0.9)	15 (0.4)	510 (1.3)	13 (0.3)	505 (1.6)	17 (1.0)

SOURCE: IEA's Progress in International Reading Literacy Study - PIRLS 2016

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent. A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students. An "s" indicates data are available for at least 50% but less than 70% of the students.

- Click the **Start SAS** button to create the SAS syntax file and open it in a SAS editor window. The syntax file can be executed by opening the **Run** menu of SAS and selecting the **Submit** menu option. If necessary, the IEA IDB Analyzer will display a prompt to confirm the overwriting of existing files.

Exhibit 2.21 shows the completed analysis module for this example analysis, and Exhibit 2.22 displays the results.

**Exhibit 2.21 IEA IDB Analyzer Setup for Example Teacher Background Data Analysis**

The screenshot displays the IEA IDB Analyzer interface. The 'Analysis File' is set to 'C:\PIRLS2016\Merge\atgall4\_cmb.sas7bdat'. The 'Analysis Type' is 'PIRLS (Using Teacher Weights)', 'Statistic Type' is 'Percentages and Means', 'Plausible Value Option' is 'Use PVs', and 'Number of Decimals' is '2'. The 'Select Variables' section shows a list of variables with columns for Name and Description. The 'Grouping Variables' section includes 'IDCNTRY' (COUNTRY ID - NUMERIC CODE) and 'ATDG01' (TEACHERS YEARS OF EXPERIENCE). The 'Plausible Values' section includes 'ASRREA01-05' (1ST TO 5TH PLAUSIBLE VALUE: OVERALL READING PV1). The 'Weight Variable' section includes 'TCHWGT' (WEIGHT FOR RDG TEACHER DATA COMBINED). The 'Output Files' field is 'C:\PIRLS2016\Analysis\REA\_ATDG01.\*'. A 'Start SAS' button is visible at the bottom.

Each country’s results are displayed on up to four lines, one for each value of the ATDG01 variable. There will be fewer lines if any category does not have any observations. The results are presented in the same manner as in the previous examples, with countries identified in the first column and the second column describing the categories of the analysis variable ATDG01.

As shown in the four lines of results for Egypt, 43.37% (standard error of 4.40) of fourth grade students were taught by teachers with 20 years of experience or more (code 1 in column 2), 31.58% (standard error of 3.94) were taught by teachers with 10 to 19 years of experience (code 2 in column 2), 18.23% (standard error of 3.58) by teachers with 5 to 9 years of experience (code 3 in column 2), and 6.82% (standard error of 1.90) by teachers with less than 5 years of experience (code 4 in column 2).

Also, the estimated average reading achievement was 315.16 (standard error of 9.30) for fourth grade students taught by teachers with 20 years of experience or more, 353.33 (standard error of 9.55) for students taught by teachers with 10 to 19 years of experience, 331.73 (standard error of 16.95) for students taught by teachers with 5 to 9 years of experience, and 318.30 (standard error of 29.50) for students taught by teachers with less than 5 years of experience. The IEA IDB Analyzer also produces the standard deviations of achievement for all categories of ATDG01.

**Exhibit 2.22 SAS Output for Example Teacher Background Data Analysis**

**Average of ASRREA0 by (IDCNTY ATDG01)**

COUNTRY ID - NUMERIC CODE	TEACHERS YEARS OF EXPERIENCE	n	WEIGHT FOR TEACHER DATA	sumw_se	pct	pct_se	mnpv	mnpv_se	sdpv	sdpv_se	pctmiss
Australia	1	2705	109375.57	9184.43	39.99	3.26	546.90	4.37	79.88	1.98	0.00
	2	1622	63688.87	7613.70	23.29	2.86	541.56	4.24	84.12	2.41	0.00
	3	1059	41563.76	7498.13	15.20	2.69	549.39	7.37	83.81	3.51	0.00
	4	1330	58846.55	7864.51	21.52	2.81	542.14	7.58	88.30	3.82	0.00
Bulgaria	1	3520	48650.16	1958.87	87.01	2.85	550.06	4.67	85.90	3.01	0.00
	2	289	4437.46	1287.55	7.94	2.28	548.98	13.13	81.03	5.89	0.00
	3	156	1753.70	688.28	3.14	1.23	583.35	8.78	67.74	4.45	0.00
	4	111	1074.02	457.91	1.92	0.82	564.29	22.35	93.74	23.10	0.00
Egypt	1	2670	631853.31	70120.77	43.37	4.40	315.16	9.30	122.72	3.98	0.00
	2	2195	460177.76	57158.85	31.58	3.94	353.33	9.55	117.66	3.69	0.00
	3	1108	265546.06	53167.07	18.23	3.58	331.73	16.95	122.38	7.50	0.00
	4	571	99413.42	27443.60	6.82	1.90	318.30	29.50	145.90	10.55	0.00
Iran, Islamic Republic of	1	4353	1166863.55	86805.59	49.26	3.58	443.94	5.59	101.24	4.70	0.00
	2	2186	544866.44	87585.63	23.00	3.42	414.54	11.04	108.97	6.03	0.00
	3	1023	347178.71	53227.24	14.66	2.18	404.10	14.61	117.87	5.36	0.00
	4	1152	309915.87	60031.92	13.08	2.56	415.10	12.01	105.93	6.12	0.00

As a reminder, the results for Egypt are based on its PIRLS Literacy 2016 participation and the results for Iran include the data from its PIRLS 2016 and PIRLS Literacy 2016 assessments, since we used the combined merged data file ATGALLR4\_CMB.

**2.8 PIRLS Analyses with School Background Data**

When performing analyses with school background data, the data are analyzed to make statements about students attending schools with a given characteristic, rather than about schools with a given characteristic. Our example of an analysis using school background data will compute the percentages and average achievement of fourth grade students who attend schools composed of students with different levels of socioeconomic background. The results of this analysis are presented in *Exhibit 5.1 of PIRLS 2016 International Results in Reading* and are replicated here in Exhibit 2.23.

**Exhibit 2.23 Exhibit of Example School Background Data Analysis, Taken from PIRLS 2016 International Results in Reading (Exhibit 5.1)**

**Exhibit 5.1: School Composition by Socioeconomic Background of the Student Body**

*Students Categorized by Principals' Reports*

Country	More Affluent - Schools where more than 25% of the student body comes from economically affluent homes and not more than 25% from economically disadvantaged homes		Neither More Affluent Nor More Disadvantaged		More Disadvantaged - Schools where more than 25% of the student body comes from economically disadvantaged homes and not more than 25% from economically affluent homes	
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement
Australia	35 (3.2)	570 (4.6)	38 (3.3)	545 (3.2)	26 (2.9)	509 (4.3)
Austria	29 (4.0)	558 (3.5)	45 (3.9)	545 (2.3)	25 (3.6)	512 (5.7)
Azerbaijan	r 11 (2.1)	502 (13.6)	23 (4.0)	469 (8.5)	66 (4.1)	471 (4.9)
Bahrain	31 (1.6)	473 (5.5)	36 (3.0)	446 (3.4)	33 (3.0)	419 (5.4)
Belgium (Flemish)	67 (4.0)	535 (2.0)	23 (3.6)	515 (4.2)	10 (2.7)	487 (6.1)
Belgium (French)	56 (3.6)	515 (3.2)	24 (3.8)	493 (4.5)	21 (3.0)	461 (5.4)
Bulgaria	15 (3.5)	585 (7.9)	51 (4.3)	565 (5.4)	34 (3.6)	520 (7.9)
Canada	42 (2.7)	556 (2.2)	38 (2.3)	546 (2.9)	21 (2.1)	513 (4.5)
Chile	18 (2.3)	546 (3.1)	9 (2.7)	500 (9.3)	74 (3.1)	481 (3.5)
Chinese Taipei	23 (3.4)	575 (3.8)	62 (4.3)	558 (2.3)	15 (2.9)	536 (3.9)
Czech Republic	42 (3.7)	554 (2.6)	51 (4.1)	540 (2.5)	7 (1.9)	506 (15.8)
Denmark	63 (4.1)	554 (2.9)	29 (3.8)	538 (4.0)	7 (2.1)	538 (5.7)
Egypt	25 (3.7)	375 (14.6)	16 (3.6)	344 (14.9)	59 (4.7)	306 (7.4)
England	38 (3.6)	575 (3.5)	30 (3.9)	563 (3.3)	32 (3.7)	539 (3.5)
Finland	42 (3.7)	570 (3.0)	49 (4.1)	567 (2.9)	9 (1.8)	544 (8.0)
France	41 (4.5)	524 (3.4)	30 (4.1)	517 (4.4)	29 (3.3)	489 (3.9)
Georgia	23 (3.5)	507 (5.6)	31 (3.7)	484 (5.5)	46 (3.9)	486 (4.7)
Germany	22 (3.1)	560 (4.3)	48 (4.1)	547 (3.9)	30 (3.6)	504 (9.1)
Hong Kong SAR	32 (3.8)	570 (4.6)	28 (4.2)	573 (5.1)	39 (3.4)	571 (5.1)
Hungary	34 (3.6)	579 (4.9)	29 (4.2)	562 (4.1)	37 (3.5)	522 (5.2)
Iran, Islamic Rep. of	28 (3.9)	462 (6.8)	18 (3.0)	457 (8.1)	55 (4.2)	400 (7.0)
Ireland	43 (4.5)	582 (3.5)	32 (4.4)	570 (3.5)	25 (3.6)	534 (4.0)
Israel	30 (3.0)	561 (5.1)	33 (3.8)	549 (6.8)	36 (3.1)	489 (5.4)
Italy	32 (4.2)	555 (4.2)	50 (4.6)	547 (3.2)	18 (3.1)	539 (6.1)
Kazakhstan	70 (3.8)	542 (3.3)	24 (3.5)	517 (5.6)	6 (2.0)	524 (7.2)
Kuwait	r 39 (4.1)	408 (8.4)	27 (4.0)	382 (10.4)	34 (4.5)	385 (11.7)
Latvia	21 (3.7)	562 (4.7)	69 (4.4)	560 (2.3)	10 (2.4)	533 (7.8)
Lithuania	65 (3.4)	560 (2.8)	23 (3.3)	539 (5.3)	13 (2.6)	503 (8.7)
Macao SAR	37 (0.1)	541 (2.1)	26 (0.1)	548 (1.9)	36 (0.1)	549 (2.0)
Malta	39 (0.1)	441 (2.6)	58 (0.1)	461 (2.1)	3 (0.1)	415 (8.2)
Morocco	r 8 (1.9)	434 (14.1)	8 (1.7)	404 (13.2)	84 (1.9)	343 (5.0)
Netherlands	r 55 (4.9)	555 (2.4)	33 (4.8)	543 (3.3)	12 (3.1)	515 (7.2)
New Zealand	r 40 (3.4)	551 (2.9)	31 (3.6)	527 (6.0)	29 (3.0)	484 (7.5)
Northern Ireland	s 36 (5.2)	583 (4.3)	34 (5.8)	562 (5.1)	30 (3.9)	547 (5.0)
Norway (5)	54 (4.7)	563 (3.2)	38 (4.3)	555 (2.9)	8 (2.2)	546 (6.8)
Oman	44 (3.1)	430 (5.6)	39 (3.1)	413 (5.2)	18 (2.7)	402 (5.4)
Poland	19 (3.9)	572 (6.3)	60 (4.7)	566 (2.6)	21 (3.2)	545 (5.6)
Portugal	14 (2.4)	543 (5.3)	45 (4.7)	529 (4.2)	42 (4.8)	521 (3.2)
Qatar	r 70 (0.4)	449 (2.4)	22 (0.4)	450 (3.8)	8 (0.2)	422 (11.9)
Russian Federation	71 (3.3)	585 (2.9)	24 (3.2)	578 (5.7)	5 (1.4)	558 (9.4)
Saudi Arabia	r 43 (3.7)	437 (8.8)	29 (3.8)	424 (9.0)	28 (4.1)	424 (10.6)
Singapore	42 (0.0)	597 (4.8)	44 (0.0)	568 (4.9)	13 (0.0)	536 (10.4)
Slovak Republic	37 (3.6)	551 (3.7)	51 (4.1)	542 (3.6)	11 (2.7)	449 (16.7)
Slovenia	35 (4.6)	546 (3.8)	38 (4.2)	543 (3.4)	27 (4.8)	537 (4.5)
South Africa	r 9 (2.0)	428 (23.1)	16 (3.0)	310 (17.3)	75 (3.7)	309 (4.8)
Spain	54 (3.0)	535 (1.9)	31 (2.7)	520 (4.3)	15 (2.1)	511 (5.1)
Sweden	78 (3.9)	562 (2.8)	18 (3.5)	539 (4.0)	4 (1.4)	508 (18.7)
Trinidad and Tobago	r 25 (3.7)	526 (6.8)	21 (3.8)	493 (8.6)	54 (4.5)	463 (5.7)
United Arab Emirates	r 50 (2.2)	473 (6.2)	25 (1.9)	453 (6.5)	25 (2.0)	409 (6.5)
United States	15 (3.0)	588 (4.3)	21 (3.7)	566 (5.4)	63 (4.0)	532 (3.7)
International Avg.	38 (0.5)	530 (0.9)	33 (0.5)	513 (0.9)	29 (0.4)	487 (1.1)

SOURCE: IEA's Progress in International Reading Literacy Study – PIRLS 2016

( ) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

A tilde (~) indicates insufficient data to report achievement.

An "r" indicates data are available for at least 70% but less than 85% of the students. An "s" indicates data are available for at least 50% but less than 70% of the students.

We will use the **Percentages and Means** statistic type and select the **Use PVs** option to estimate the percentages of students with their average reading achievement by reporting categories of students' economic background as reported by school principals.

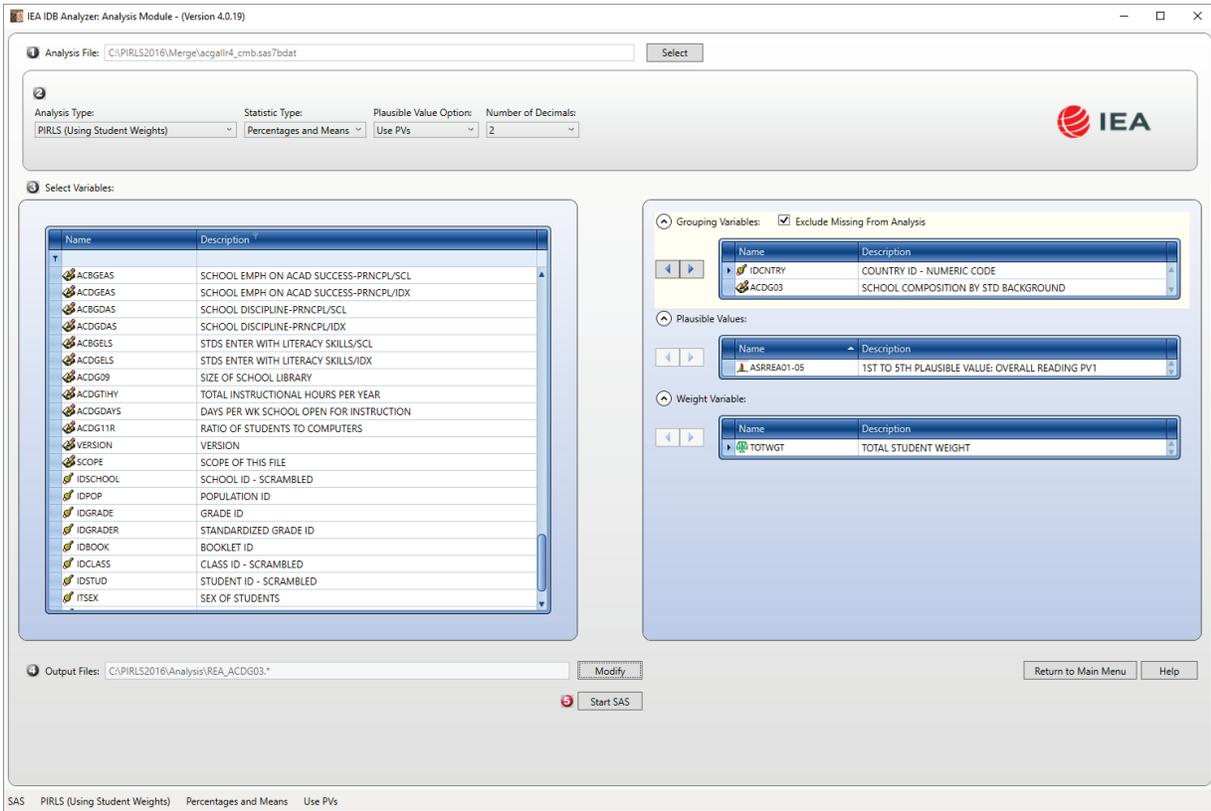
The information for this analysis is found in the school-level derived variable ACDG03 (see [Supplement 3 to this User Guide](#)), where schools are characterized as being composed of more affluent students than disadvantaged students, more disadvantaged students than affluent students, or neither more affluent nor more disadvantaged students.

The merged data file ACGALLR4\_CMB will be used for this example analysis and it is performed by the analysis module of the IEA IDB Analyzer using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.
2. Select the merged data file ACGALLR4\_CMB as the **Analysis File**.
3. Select **PIRLS (Using Student Weights)** as the **Analysis Type** because we want to analyze school background data as student attributes.
4. Select **Percentages and Means** as the **Statistic Type**.
5. Select **Use PVs** as the **Plausible Value Option**.
6. Specify the variable ACDG03 as a second grouping variable by clicking the **Grouping Variables** field to activate it. Select ACDG03 from the list of available variables and move it to the **Grouping Variables** field by clicking the **right arrow** (▶) button in this section.
7. Specify the achievement scores to be used for the analysis by clicking the **Plausible Values** field to activate it. Select ASRREA01–05 from the list of available variables and move it to the **Plausible Values** field by clicking the **right arrow** (▶) button in this section.
8. The **Weight Variable** is selected automatically by the software; TOTWGT is selected by default because this example analysis uses school background data linked to student background data.
9. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button. In this example, the output files are labeled REA\_ACDG03 and saved in the "C:\PIRLS2016\Analysis" folder.
10. Click the **Start SAS** button to create the SAS syntax file and open it in a SAS editor window. The syntax file can be executed by opening the **Run** menu of SAS and selecting the **Submit** menu option. If necessary, the IEA IDB Analyzer will display a prompt to confirm the overwriting of existing files.

Exhibit 2.24 shows the completed analysis module for this example analysis and the results are presented in Exhibit 2.25. In this example, each country's results are presented on three lines, one for each value of the ACDG03 variable. The results are presented in the same manner as in previous examples, with countries identified in the first column and the second column describing the categories of ACDG03.

**Exhibit 2.24 IEA IDB Analyzer Set-Up for Example School Background Data Analysis**



**Exhibit 2.25 SAS Output for Example School Background Data Analysis**

**Average of ASRREA0 by (IDCNTRY ACDG03)**

Country ID - Numeric Code	School Composition by Std Background	n	TOTAL STUDENT WEIGHT	sumw_se	pct	pct_se	mnpv	mnpv_se	sdpv	sdpv_se	pctmiss
Australia	1	2292	92822.22	8583.07	35.49	3.17	569.68	4.65	76.41	2.47	0.00
	2	2126	99457.64	9392.99	38.03	3.34	544.72	3.21	80.63	2.50	0.00
	3	1480	69272.17	7514.97	26.49	2.87	508.97	4.34	86.02	2.76	0.00
Bulgaria	1	723	8259.56	1928.03	15.15	3.46	584.58	7.88	70.50	5.53	0.00
	2	2302	27600.70	2371.38	50.62	4.31	564.52	5.37	80.29	4.95	0.00
	3	951	18666.02	2075.81	34.23	3.58	520.21	7.92	87.03	3.56	0.00
Egypt	1	1517	341738.59	53043.08	25.37	3.68	375.36	14.61	123.65	8.14	0.00
	2	1006	211006.63	48748.23	15.67	3.57	344.47	14.91	118.37	6.18	0.00
	3	3617	794197.68	68283.17	58.96	4.75	306.21	7.43	119.04	3.23	0.00
Iran, Islamic Republic of	1	1821	568933.56	89721.51	27.54	3.92	462.37	6.76	93.28	3.85	0.00
	2	1666	367398.59	61132.88	17.78	2.97	456.89	8.10	97.27	5.16	0.00
	3	4150	1129575.47	95662.16	54.68	4.15	399.82	6.99	109.06	4.22	0.00

As shown in the three lines of results for Bulgaria, 15.15% (standard error of 3.46) of fourth grade students attended schools with more affluent than disadvantaged students (code 1 in column 2), 50.62% (standard error of 4.31) attended schools with neither more affluent nor more disadvantaged students (code 2 in column 2), and 34.23% (standard error of 3.58) attended schools with more disadvantaged than affluent students (code 3 in column 2).

Also, the estimated average reading achievement was 584.58 (standard error of 7.88) for students in schools with more affluent students, 564.52 (standard error of 5.37) for students in schools with neither more affluent nor more disadvantaged students, and 520.21 (standard error of 7.92) for students in schools with more disadvantaged students. The IEA IDB Analyzer also produces the standard deviations of achievement for all categories of ACDG03.

## References

- IBM Corporation. (2016). *IBM SPSS Statistics for Windows*, version 24.0. Armonk, NY: IBM Corporation.
- International Association for the Evaluation of Educational Achievement. (2018). *International Database Analyzer*, version 4.0. Hamburg, Germany: IEA Hamburg.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/pirls2016/international-results/>
- SAS Institute. (2012). *SAS System for Windows*, version 9.4. Cary, NC: SAS Institute.

**PIRLS**  
**2016**

**PIRLS 2016 User Guide  
for the International Database**

# **Chapter 3**

## **Special SPSS and SAS Programs**

# Chapter 3

## *Special SPSS and SAS Programs*

### 3.1 Overview

This chapter presents special SPSS (IBM Corporation, 2016) and SAS (SAS Institute, 2012) programs essential to make full use of the PIRLS 2016 International Database. A first set of SAS programs converts the SAS export files into SAS data files, needed as input to the IEA IDB Analyzer (IEA, 2018). There is a second set of SPSS and SAS programs to recode the responses to the individual PIRLS 2016, PIRLS Literacy 2016, and ePIRLS 2016 items from the achievement data files to their appropriate score levels. There is a third set of SPSS and SAS programs to combine PIRLS 2016 and PIRLS Literacy 2016 data files in order to include reading results of all countries participating in either assessment and to estimate properly results for countries that participated in both assessments.

Users who intend to use the IEA IDB Analyzer with SAS should execute the SAS programs described in this chapter, in particular the conversion programs. Users who intend to use the IEA IDB Analyzer with SPSS should execute the SPSS programs.

### 3.2 Converting the SAS Export Files

The PIRLS 2016 International Database provides three SAS programs—P16\_CONVERT.SAS, PL16\_CONVERT.SAS, and eP16\_CONVERT.SAS—to convert the SAS export files provided in the International Database into SAS data files. This conversion is necessary because the IEA IDB Analyzer requires the use of SAS data files; it will not function with the SAS export files provided in the PIRLS 2016 International Database. The P16\_CONVERT.SAS program will convert the PIRLS 2016 SAS export files. The PL16\_CONVERT.SAS program will convert the PIRLS Literacy 2016 SAS export files; eP16\_CONVERT.SAS will convert the ePIRLS 2016 SAS export files. To convert SAS export files into SAS data files, users should perform the following steps:

1. Open P16\_CONVERT.SAS, PL16\_CONVERT.SAS, or eP16\_CONVERT.SAS with the SAS software.
2. At the end of the program, specify the data file types of interest, each separated by a blank space, in the parameter "TYPE." By default, the CONVERT programs list all available file types.
3. Specify the folder where the SAS export files are located in the parameter "INDIR."
4. Specify the folder where the converted SAS data files will be located in the parameter "OUTDIR."

5. List all the countries of interest in the parameter "COUNTRY." By default, the program lists all available countries for any given assessment.
6. Submit the edited code for processing by SAS.

Exhibit 3.1 shows an example of the P16\_CONVERT.SAS program. The PL16\_CONVERT.SAS and eP16\_CONVERT.SAS programs have similar structures. This example converts the SAS export files of all file types and for all countries. For this example, all SAS export files are located in the folder "C:\PIRLS2016\Data," where the converted SAS data files also will be located.

### Exhibit 3.1 The P16\_CONVERT SAS Program

```
%MACRO DOIT (TYPE = ,
             INDIR = ,
             OUTDIR = ) ;

. . .

%LET COUNTRY = <List of PIRLS 2016 countries> ;

. . .

PROC CIMPORT FILE="%INDIR%\&FTYPE&CTRY.R4.EXP"
             DATA=OUTDIR.&FTYPE&CTRY.R4 ;
RUN ;

. . .

%MEND DOIT ;

%DOIT (TYPE = ACG ASA ASG ASH ASR AST ATG ,
      INDIR = C:\PIRLS2016\Data ,
      OUTDIR = C:\PIRLS2016\Data ) ;
```

SAS users are advised to run the P16\_CONVERT.SAS, PL16\_CONVERT.SAS, and eP16\_CONVERT.SAS programs for all countries and all file types. The available file types for all three assessments are ACG, ASA, ASG, ASH, ASR, AST, and ATG. These file types are described in Chapter 4 of this User Guide. In principle, this program needs to be run only once for all file types and countries and should be one of the first things SAS users do with the PIRLS 2016 International Database before undertaking any data analyses with SAS and the IEA IDB Analyzer, in particular the data analysis examples presented in this User Guide.

### 3.3 Scoring the PIRLS 2016 Items

Student achievement in PIRLS 2016 is represented by sets of five plausible values for reading, the two purposes for reading, and the two processes of comprehension. They are the preferred scores for any analysis of student achievement. However, analyzing performance on individual items may be of interest to some users. Carrying out such analyses may require that the individual items in the PIRLS 2016 International Database be assigned their correctness score levels, rather than the actual response options selected by students for multiple-choice items. The International Database provides SPSS and SAS programs to perform this task.

For multiple-choice items, the numbers 1 through 4 (1 through 6 for some ePIRLS items) are used to represent response options A through D (A through F for some ePIRLS items),

respectively, in the PIRLS 2016 achievement data files. These responses must be converted to their appropriate score level (“1” for correct and “0” for incorrect) based on each multiple-choice item’s correct response key. The constructed-response items, worth either a total of one, two, or three points, already are coded as the appropriate point values of the students’ responses—either zero, one, two, or three points.

For both types of items, special codes are set aside to represent missing data as either “Not Administered,” “Omitted,” or “Not Reached.” These special missing codes also must be converted to their appropriate score level. By default, the “Not Administered” response code is left as missing and the “Omitted” and “Not Reached” response codes are recoded as incorrect (zero). These default settings can be modified within the programs, depending on the requirements of the item-level analyses. For example, “Not Reached” responses were treated as missing for the purpose of calibrating the PIRLS 2016 items, whereas they were treated as incorrect when deriving achievement scores for students.

The PIRLS 2016 International Database includes three SPSS programs—ASASCRR4.SPS for the PIRLS items, ASASCRL2.SPS for the PIRLS Literacy items, and ASASCRE1.SPS for the ePIRLS items—and three SAS programs—ASASCRR4.SAS, ASASCRL2.SAS, and ASASCRE1.SAS—which will recode the responses to individual items from the achievement data files (ASA) to their appropriate score levels.

Before executing a scoring program, users first must create a merged data file of all required student achievement data files (ASA) using the Merge Module of the IEA IDB Analyzer, as described in Chapter 2 of this User Guide. The scoring programs will then create a merged data file with scored achievement items that can be used by the Analysis Module of the IEA IDB Analyzer. To score each individual PIRLS 2016 item, the program code in the SAS program must be adapted by completing the following steps:

1. Open ASASCRR4.SAS, ASASCRL2.SAS, or ASASCRE1.SAS with the SAS software.
2. At the end of the program, specify the folder where the merged SAS data file of student achievement data files is located in the parameter “INDIR.”
3. Specify the folder where the merged SAS data file of scored achievement items will be located in the parameter “OUTDIR.”
4. Specify the name of the merged SAS data file of student achievement data files in the parameter “INFILE.”
5. Submit the edited code for processing by SAS.

Exhibit 3.2 shows a condensed version of the SAS program to score the individual PIRLS 2016 items. The PIRLS Literacy 2016 and ePIRLS 2016 SAS score programs have similar structures.

In this example, the merged SAS data file of student achievement data files is called ASAALLR4, located in the folder “C:\PIRLS2016\Merge”, and contains the PIRLS achievement items. The resulting merged data file of scored achievement items will be called ASAALLR4\_SCR and located in the same folder. The merged data file of scored achievement items will have the

same data structure as the student achievement data files (ASA), but with the score levels stored in the item variables instead of the student responses. The Analysis Module of the IEA IDB analyzer will be able to perform analyses with this merged data file.

**Exhibit 3.2 The ASASCRR4 SAS Program**

```
%MACRO SCOREIT (ITEM, TYPE, RIGHT, NR, NA, OM, OTHER) ;
. . .
%MEND SCOREIT ;

%MACRO DOIT (INDIR = ,
            OUTDIR = ,
            INFILE = ) ;
. . .

%LET ARIGHT = <List of multiple choice items where A is correct> ;
DO OVER ARIGHT ; %SCOREIT (ARIGHT, "MC", 1, .R, .A, ., .I) ; END ;
%LET BRIGHT = <List of multiple choice items where B is correct> ;
DO OVER BRIGHT ; %SCOREIT (BRIGHT, "MC", 2, .R, .A, ., .I) ; END ;
%LET CRIGHT = <List of multiple choice items where C is correct> ;
DO OVER CRIGHT ; %SCOREIT (CRIGHT, "MC", 3, .R, .A, ., .I) ; END ;
%LET DRIGHT = <List of multiple choice items where D is correct> ;
DO OVER DRIGHT ; %SCOREIT (DRIGHT, "MC", 4, .R, .A, ., .I) ; END ;
%LET CONSTR = <List of constructed response items> ;
DO OVER CONSTR ; %SCOREIT (CONSTR, "CR", , .R, .A, ., .I) ; END ;

. . .

%MEND DOIT ;

%DOIT (INDIR = C:\PIRLS2016\Merge ,
      OUTDIR = C:\PIRLS2016\Merge ,
      INFILE = ASAALLR4 ) ;
```

If Not Reached responses are to be treated as missing rather than incorrect, users should replace the following SAS statement (which appears twice in the program):

```
IF &ITEM = &NR THEN SCORE = 0 ; * Not Reached ;
```

with this statement:

```
IF &ITEM = &NR THEN SCORE = . ; * Not Reached ;
```

Users should be careful to use the score program appropriate for the student achievement data files under consideration; the ASASCRR4 program for the PIRLS achievement items, the ASASCRL2 program for the PIRLS Literacy achievement items, and the ASASCRES1 program for the ePIRLS achievement items.

Executing the equivalent SPSS program requires the same steps as the SAS program. Exhibit 3.3 shows a condensed version of the SPSS program to score the individual PIRLS 2016 items.

### Exhibit 3.3 The ASASCRR4 SPSS Program

```

DEFINE SCOREIT <List of macro parameters> .
. . .
!ENDDDEFINE .
DEFINE DOIT (INDIR = !CHAREND('/') /
            OUTDIR = !CHAREND('/') /
            INFILE = !CHAREND('/') ) .
. . .
SCOREIT TYPE = MC / ITEM = <List of multiple choice items where A is correct> /
        RIGHT = 1 / NR = 6 / NA = SYSMIS / OM = 9 / OTHER = 7 .
SCOREIT TYPE = MC / ITEM = <List of multiple choice items where B is correct> /
        RIGHT = 2 / NR = 6 / NA = SYSMIS / OM = 9 / OTHER = 7 .
SCOREIT TYPE = MC / ITEM = <List of multiple choice items where C is correct> /
        RIGHT = 3 / NR = 6 / NA = SYSMIS / OM = 9 / OTHER = 7 .
SCOREIT TYPE = MC / ITEM = <List of multiple choice items where D is correct> /
        RIGHT = 4 / NR = 6 / NA = SYSMIS / OM = 9 / OTHER = 7 .
SCOREIT TYPE = CR / ITEM = <List of constructed response items> /
        RIGHT = 0 / NR = 6 / NA = SYSMIS / OM = 99 / OTHER = 90 .
. . .
!ENDDDEFINE .
DOIT INDIR = C:\PIRLS2016\Merge /
     OUTDIR = C:\PIRLS2016\Merge /
     INFILE = ASAALLR4 .

```

If Not Reached responses are to be treated as missing rather than incorrect, users should replace the following SPSS statement (which appears twice in the program):

```
(!NR = 0) /* Not Reached */
```

with this statement:

```
(!NR = SYSMIS) /* Not Reached */
```

### 3.4 Combining PIRLS and PIRLS Literacy Data

In *PIRLS 2016 International Results in Reading* (Mullis, Martin, Foy, & Hooper, 2017), all exhibits report results from countries that participated in the PIRLS 2016 assessment and the PIRLS Literacy 2016 assessment. Thus, Denmark (at the third grade), Egypt, Kuwait, and South Africa—which participated in PIRLS Literacy 2016—appear in all PIRLS 2016 exhibits. As well, results for Iran and Morocco—which participated in both assessments—include data from both assessments.

In order to replicate the published results, merged data files produced by the IEA IDB Analyzer must combine data from both assessments. The PIRLS 2016 International Database includes a short program, in both SPSS and SAS, to perform this function. Exhibit 3.4 shows the SAS version of this program.

#### Exhibit 3.4 The P16\_PL16\_COMBINE SAS Program

```
%MACRO CMB_P16_PL16(P16FILE = ,
                    PL16FILE = ,
                    INDIR = ,
                    OUTDIR = ) ;

LIBNAME INDIR "&INDIR" ;
LIBNAME OUTDIR "&OUTDIR" ;

OPTIONS NONOTES FMTSEARCH=(INDIR,OUTDIR) MPRINT ;

DATA OUTDIR.&P16FILE._CMB ;
  SET INDIR.&P16FILE
      INDIR.&PL16FILE ;
RUN ;

%MEND CMB_P16_PL16 ;

%CMB_P16_PL16 (P16FILE = ASGALLR4 ,
              PL16FILE = ASGALLL2 ,
              INDIR = C:\PIRLS2016\Merge ,
              OUTDIR = C:\PIRLS2016\Merge ) ;
```

Before this program can be executed, users must first create merged data files for both PIRLS 2016 and PIRLS Literacy 2016. To combine merged data files from both assessments, the program code in the SAS program must be adapted by completing the following steps:

1. Open P16\_PL16\_COMBINE.SAS with the SAS software.
2. At the end of the program, specify the name of the PIRLS 2016 merged data file in the parameter "P16FILE."
3. Specify the name of the PIRLS Literacy 2016 merged data file in the parameter "PL16FILE."
4. Specify the folder where the merged data files are located in the parameter "INDIR." Both merged data files must be in the same folder. In this example, they are located in the folder "C:\PIRLS2016\Merge".
5. Specify the folder where the combined merged data file will be located in the parameter "OUTDIR." In this example, it will be located in the same folder as the input merged data files.
6. Submit the edited code for processing by SAS.

The same set of instructions can be used to edit the SPSS program and execute it. Exhibit 3.5 shows the SPSS version of this program. It is important to note that both merged files must be of the same type and located in the same folder. In the example given here, the ASG files of both assessments are combined.

**Exhibit 3.5 The P16\_PL16\_COMBINE SPSS Program**

```

DEFINE CMB_P16_PL16 (P16FILE = !CHAREND('/') !DEFAULT("") /
                    PL16FILE = !CHAREND('/') !DEFAULT("") /
                    INDIR = !CHAREND('/') !DEFAULT("") /
                    OUTDIR = !CHAREND('/') !DEFAULT("") ) .

ADD FILES
  / FILE = !QUOTE(!CONCAT(!UNQUOTE(!INDIR), "\", !P16FILE, ".SAV"))
  / FILE = !QUOTE(!CONCAT(!UNQUOTE(!INDIR), "\", !PL16FILE, ".SAV")) .

SAVE OUTFILE = !QUOTE(!CONCAT(!UNQUOTE(!OUTDIR), "\", !P16FILE, "_CMB.SAV")) .

!ENDDEFINE .

CMB_P16_PL16 P16FILE = ASGALLR4 /
             PL16FILE = ASGALLL2 /
             INDIR = C:\PIRLS2016\Merge /
             OUTDIR = C:\PIRLS2016\Merge .

```

**References**

- IBM Corporation. (2016). *IBM SPSS Statistics for Windows*, version 24.0. Armonk, NY: IBM Corporation.
- International Association for the Evaluation of Educational Achievement. (2018). *International Database Analyzer*, version 4.0. Hamburg, Germany: IEA Hamburg.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/pirls2016/international-results/>
- SAS Institute. (2012). *SAS System for Windows*, version 9.4. Cary, NC: SAS Institute.

**PIRLS**  
**2016**

**PIRLS 2016 User Guide  
for the International Database**

# **Chapter 4**

**The PIRLS 2016  
International Database  
Files**

# Chapter 4

## *The PIRLS 2016 International Database Files*

### 4.1 Overview

The PIRLS 2016 International Database contains student achievement data and student, home, teacher, and school background data collected in the 50 countries and 11 benchmarking participants that took part in PIRLS 2016, including the PIRLS Literacy and ePIRLS assessments.<sup>1</sup> Exhibit 4.1 lists all of the PIRLS 2016 countries and benchmarking participants included in the International Database, along with the identification codes used in the International Database. The International Database also contains materials that provide additional information on its structure and contents.

This chapter describes the contents of the International Database with a special emphasis on the actual data files used in Chapter 2 of this User Guide to analyze the PIRLS 2016 assessment results. The next section introduces the Supplements to this User Guide. The following section describes the item-related materials included in the international database. The next five sections present the different data files and related materials included in the database, as described in Exhibit 1.1 of Chapter 1.

**Exhibit 4.1 Countries Included in the PIRLS 2016 International Database**

Countries	Identification Codes		PIRLS	PIRLS Literacy	ePIRLS
	Alpha	Numeric			
Australia	AUS	036	●		
Austria	AUT	040	●		
Azerbaijan	AZE	031	●		
Bahrain	BHR	048	●		
Belgium (Flemish)	BFL	956	●		
Belgium (French)	BFR	957	●		
Bulgaria	BGR	100	●		
Canada	CAN	124	●		●
Chile	CHL	152	●		
Chinese Taipei	TWN	158	●		●

<sup>1</sup> Many aspects of the PIRLS Literacy 2016 and ePIRLS 2016 assessments are identical to the PIRLS 2016 assessment in terms of database structure and contents. All references to PIRLS 2016 in general also will apply to the PIRLS Literacy 2016 and ePIRLS 2016 assessments, unless explicitly stated otherwise.

**Exhibit 4.1 Countries Included in the PIRLS 2016 International Database (Continued)**

Countries	Identification Codes		PIRLS	PIRLS Literacy	ePIRLS
	Alpha	Numeric			
Czech Republic	CZE	203	●		
Denmark	DNK	208	●		●
Egypt	EGY	818		●	
England	ENG	926	●		
Finland	FIN	246	●		
France	FRA	250	●		
Georgia	GEO	268	●		●
Germany	DEU	276	●		
Hong Kong SAR	HKG	344	●		
Hungary	HUN	348	●		
Iran, Islamic Rep. of	IRN	364	●	●	
Ireland	IRL	372	●		●
Israel	ISR	376	●		●
Italy	ITA	380	●		●
Kazakhstan	KAZ	398	●		
Kuwait	KWT	414		●	
Latvia	LVA	428	●		
Lithuania	LTU	440	●		
Macao SAR	MAC	446	●		
Malta	MLT	470	●		
Morocco	MAR	504	●	●	
Netherlands	NLD	528	●		
New Zealand	NZL	554	●		
Northern Ireland	NIR	928	●		
Norway (5 <sup>th</sup> grade)	NOR	578	●		●
Oman	OMN	512	●		
Poland	POL	616	●		
Portugal	PRT	620	●		●
Qatar	QAT	634	●		
Russian Federation	RUS	643	●		
Saudi Arabia	SAU	682	●		
Singapore	SGP	702	●		●
Slovak Republic	SVK	703	●		
Slovenia	SVN	705	●		●
South Africa	ZAF	710		●	
Spain	ESP	724	●		
Sweden	SWE	752	●		●
Trinidad and Tobago	TTO	780	●		
United Arab Emirates	ARE	784	●		●
United States	USA	840	●		●

**Exhibit 4.1 Countries Included in the PIRLS 2016 International Database (Continued)**

Countries	Identification Codes		PIRLS	PIRLS Literacy	ePIRLS
	Alpha	Numeric			
<b>Benchmarking Participants</b>					
Buenos Aires, Argentina	ABA	32001	●		
Ontario, Canada	COT	9132	●		
Quebec, Canada	CQU	9133	●		
Denmark (3 <sup>rd</sup> grade)	DN3	208001		●	
Norway (4 <sup>th</sup> grade)	NO4	5784	●		
Moscow City, Russian Fed.	RMO	643001	●		
Eng/Afr/Zulu – RSA (5)	ZA5	7105	●		
Andalusia, Spain	EAN	72401	●		
Madrid, Spain	EMA	724005	●		
Abu Dhabi, UAE	AAD	7842	●		●
Dubai, UAE	ADU	7841	●		●

## 4.2 The Supplements to the PIRLS 2016 User Guide

The PIRLS 2016 User Guide is an important resource to the PIRLS 2016 International Database. It provides a description of the structure and contents of the International Database, as well as instructions to make the best use of its data. The User Guide includes a series of supplements in printable PDF format:

- **Supplement 1:** International Version of the PIRLS 2016 Context Questionnaires—Supplement 1 includes the international version of all context questionnaires administered in PIRLS 2016, including the Curriculum Questionnaire. It is a good reference guide to know what questions were asked and the variable names under which the responses are recorded in the International Database.
- **Supplement 2:** National Adaptations of International Context Questionnaires—Supplement 2 provides details on all national adaptations that were applied to the national version of all PIRLS 2016 context questionnaires. Users should refer to this supplement for any special national adaptations to background variables that could potentially affect the interpretation of analysis results.
- **Supplement 3:** Variables Derived from the Student, Home, Teacher, and School Questionnaire Data—Supplement 3 describes how the derived background variables were computed to produce exhibits in *PIRLS 2016 International Results in Reading* (Mullis, Martin, Foy, & Hooper, 2017).

## 4.3 The Restricted Use Version of the PIRLS 2016 International Database

There are two versions of the PIRLS 2016 International Database. The public use version is available for immediate access from the PIRLS 2016 International Database webpage (<http://timssandpirls.bc.edu/pirls2016/international-database.html>) and the IEA Study Data Repository (<http://www.iea.nl/data.html>). A number of variables have been removed from the

public use version in order to minimize the risk of disclosing confidential information. Exhibit 4.2 lists the variables removed from the public use version of the PIRLS 2016 International Database. More details for all of these variables are available in the appropriate codebook files, described later in this chapter. Users who require any of the removed variables to conduct their analyses should contact the IEA (RandA@IEA-Hamburg.de) to obtain permission and access to the restricted use version of the PIRLS 2016 International Database. It is important to note that the students' age remains available in the public use version of the International Database.

**Exhibit 4.2 Variables Removed from Public Use Version of PIRLS 2016 International Database**

<b>Variables</b>	<b>Description</b>
ITBIRTHY ITBIRTHM	Students' year and month of birth, respectively, from the tracking forms
ASBG02A ASBG02B	Students' year and month of birth, respectively, from the student questionnaire
ITDATE	PIRLS 2016 test administration date
ACBG01	Total school enrollment
ACBG02	School enrollment in the target grade
ITMODE_C ITMODE_H ITMODE_T	Administration mode for school, home, and teacher context questionnaires, respectively

#### 4.4 PIRLS 2016 Achievement Items Documentation

A number of documents related to the PIRLS 2016, PIRLS Literacy 2016, and ePIRLS 2016 achievement items are available for download as part of the PIRLS 2016 International Database. They include summary information on the PIRLS 2016 items, the IRT item parameters estimated from the PIRLS 2016 item calibrations and their associated linear transformation constants, and percent correct statistics for all PIRLS 2016, PIRLS Literacy 2016, and ePIRLS 2016 items.

Achievement item information files are provided to enable users of the PIRLS 2016 International Database to readily produce summaries of item characteristics. There are separate achievement item information files for PIRLS, PIRLS Literacy, and ePIRLS. These files are available in Excel and PDF formats and include the following information for each item in the PIRLS 2016 assessments:

- The item's unique identifier;
- The item's passage and its sequential location within the passage;
- The item's label;
- The item's purpose for reading and process of comprehension;
- The item's type, either multiple choice or constructed response;
- The number of options for a multiple choice item;

- The correct response key for a multiple choice item;
- The item’s point value;
- An indicator showing if the item was included in the IRT scaling; and
- An indicator showing if the item was secured after the 2016 assessment.

The International Database includes Excel and PDF files with the IRT item parameters estimated for all PIRLS 2011 and PIRLS 2016 items from the concurrent item calibrations. There are also IRT item parameter files for the PIRLS Literacy 2016 items and the ePIRLS 2016 items. These same item parameters are presented in the appendices for [\*Chapter 12 of Methods and Procedures in PIRLS 2016\*](#) (Martin, Mullis, & Hooper, 2017). The PIRLS 2016 International Database also includes files with the linear transformation constants that were used to set the PIRLS 2016 achievement scores on the PIRLS trend scale. Users should note that the PIRLS 2016 transformation constants also were used for the PIRLS Literacy 2016 and ePIRLS 2016 achievement scores.

The International Database also includes percent correct statistics for all PIRLS 2016, PIRLS Literacy 2016, and ePIRLS 2016 items in Excel and PDF format.

#### **4.5 The PIRLS 2016 Data Files**

The PIRLS 2016 International Database includes the actual data from all instruments administered to the students, their parents, their teachers, and their school principals. This includes the student responses to the achievement items and the responses to the student, home, teacher, and school context questionnaires. These data files also include the achievement scores estimated for participating students, as well as background variables derived for reporting in [\*PIRLS 2016 International Results in Reading\*](#). National Research Coordinators’ responses to the curriculum questionnaire also are part of the International Database and are described later in this chapter.

This section describes the format and contents of the PIRLS 2016 data files. With the exception of the curriculum data file, the PIRLS data files are provided in SPSS (IBM Corporation, 2016) format (.SAV) and SAS (SAS Institute, 2012) export format (.EXP). Data files are provided for each country that participated in PIRLS 2016 and for which internationally comparable data are available. The file names given to the various data file types in PIRLS 2016, PIRLS Literacy 2016, and ePIRLS 2016 are shown in Exhibit 4.3. For example, ASGCHLR4.SAV is an SPSS file that contains Chile’s PIRLS 2016 student background data. For each file type, a separate data file is provided for each participating country. All data files and the variables they contain are described in the following sections.

**Exhibit 4.3 PIRLS 2016 Data File Names**

<b>Assessment</b>	<b>File Name</b>	<b>Description</b>
PIRLS 2016	ACG●●●R4	PIRLS school background data files
	ASA●●●R4	PIRLS student achievement data files
	ASG●●●R4	PIRLS student background data files
	ASH●●●R4	PIRLS home background data files
	ASR●●●R4	PIRLS within-country scoring reliability data files
	AST●●●R4	PIRLS student-teacher linkage files
	ATG●●●R4	PIRLS teacher background data files
PIRLS Literacy 2016	ACG●●●L2	PIRLS Literacy school background data files
	ASA●●●L2	PIRLS Literacy student achievement data files
	ASG●●●L2	PIRLS Literacy student background data files
	ASH●●●L2	PIRLS Literacy home background data files
	ASR●●●L2	PIRLS Literacy within-country scoring reliability data files
	AST●●●L2	PIRLS Literacy student-teacher linkage files
	ATG●●●L2	PIRLS Literacy teacher background data files
ePIRLS 2016	ACG●●●E1	ePIRLS school background data files
	ASA●●●E1	ePIRLS student achievement data files
	ASG●●●E1	ePIRLS student background data files
	ASH●●●E1	ePIRLS home background data files
	ASR●●●E1	ePIRLS within-country scoring reliability data files
	AST●●●E1	ePIRLS student-teacher linkage files
	ATG●●●E1	ePIRLS teacher background data files

●●● = 3-character country abbreviation based on the ISO 3166 coding scheme (the Identification Codes in Exhibit 4.1).

#### 4.5.1 PIRLS 2016 Student Achievement Data Files (ASA)

The PIRLS 2016 student achievement data files contain the student responses to the individual achievement items in the PIRLS 2016 assessments. The student achievement data files are best suited for performing item-level analyses. Achievement scores (plausible values) for all of the PIRLS 2016 achievement scales are available in the student achievement data files, as well as in the student background data files and student-teacher linkage data files (described later in this chapter). Users should note that all PIRLS 2016 achievement scores also are available in the ePIRLS 2016 data files.

Students who participated in PIRLS 2016 were administered one of 16 assessment booklets, each with a literary passage and an informational passage and items associated with these passages.<sup>2</sup> Students who participated in PIRLS Literacy 2016 also were administered one of 16 assessment booklets, consisting of a literary passage and an informational passage and their items. Students who participated in ePIRLS 2016 were administered two of five available simulated online reading tasks and their items. Some of these items were multiple choice items and some were constructed response items. The student achievement data files contain the actual responses to the multiple choice questions and the scores assigned to the constructed response items through the PIRLS 2016 scoring guides.<sup>3</sup>

The ePIRLS 2016 student achievement files also include a short set of background questions related to the students' experience and familiarity with computers. They also include a number of navigation variables that describe some of the events and actions taken in the ePIRLS simulated online environment. These variables are described in the ePIRLS 2016 codebook for the ASA files.

### *Item Variable Naming Convention*

The achievement item variable names are based on an 8-character alphanumeric code, which observes the following rules:

- The first character is either "R" for PIRLS items, "L" for PIRLS Literacy items, or "E" for ePIRLS items.
- The second character indicates the assessment cycle when the item was first presented in PIRLS. The items in the PIRLS 2016 assessment have either "1" for items introduced in 2001, "2" for items introduced in 2006, "3" for items introduced in 2011, or "4" for new items in 2016. The PIRLS Literacy 2016 items have either "1" for items introduced in the prePIRLS 2011 assessment or "2" for new items in 2016. All ePIRLS 2016 items have code "1."
- The third character is always "1" for all PIRLS 2016, PIRLS Literacy 2016, and ePIRLS 2016 items.
- The fourth character indicates the PIRLS or PIRLS Literacy passage, or ePIRLS task, to which the item belongs. The item information spreadsheets show the characters used in 2016.
- The fifth and sixth characters are a two-digit sequential number that identifies the sequential position of the item within the passage.
- The seventh character indicates the item type. The letter "C" is used for constructed response items and the letter "M" for multiple choice items.
- The eighth character indicates the item part, and appears only when required. It is generally a letter from "A" to "F," depending how many parts there are to a particular item. The letter "Z" is used to represent a derived item from past PIRLS assessments (derived items are described later in this chapter).

<sup>2</sup> The PIRLS 2016 assessment designs are described in Chapter 3 of *PIRLS 2016 Assessment Framework* (Mullis & Martin, 2015).

<sup>3</sup> The PIRLS 2016 scoring guides are described in *PIRLS 2016 Item Writing Guidelines* (Mullis & Martin, 2013).

For example, R31M17CA is the first part of a constructed response PIRLS reading item introduced in 2011 and that is the 17<sup>th</sup> item in the passage “The Empty Pot.”

### *Item Response Code Values*

A series of conventions were adopted to code the data included in the PIRLS 2016 data files. This section describes these conventions for the achievement items.

The values assigned to each of the achievement item variables depend on the item format. For multiple choice items, numerical values from 1 through 4 (1 through 6 for some ePIRLS items) were used to correspond to the response options A through D (A through F for some ePIRLS items), respectively. For these items, the correct response key is included in the item information files (described earlier) and as part of the variable label in the achievement codebook files (described later in this chapter). SPSS and SAS programs are included as part of the PIRLS 2016 International Database to derive correctness response codes for these items (see Chapter 3 of this User Guide). For the constructed response items, numerical values ranging from 0 to 3 are used that correspond to the score points obtained from the students’ responses.

### *Derived Items Code Values*

For some items, students were asked to provide more than one answer, each one being scored separately. The pattern of responses across these item parts determined the score on the item as a whole. Generally, students provided answers to either two or three parts and the score on the item was the number of correct responses.<sup>4</sup> For all of these items, each item part was worth one point, and the derived items were given the code values shown in Exhibit 4.4. The total score for a multi-part item is contained in a derived variable, identified by word “DERIVED” in the item label. For example, derived variable R31M17CZ contains the combined score for its three parts labeled R31M17CA, R31M17CB, and R31M17CC.

PIRLS Literacy and ePIRLS included special multi-part derived items, worth either one point or two points. For a derived one-point item (e.g., item E11T18C in ePIRLS), a student was given one point (full credit) if all parts were answered correctly, or zero points otherwise. For a derived two-point item (e.g., item L11A13CZ in PIRLS Literacy), a student was given two points (full credit) if all parts were answered correctly, only one point (partial credit) if fewer parts were answered correctly, and zero points otherwise.

All derived items in the PIRLS 2016 assessments are described in [\*Appendix E of Chapter 10 of Methods and Procedures in PIRLS 2016.\*](#)

---

<sup>4</sup> Although this section describes 2-point and 3-point derived items, the PIRLS 2016 assessments included only 3-point derived items of this type.

**Exhibit 4.4 PIRLS 2016 Derived Item Variable Codes**

<b>Codes</b>	<b>Descriptions</b>
For 3-point Derived Items:	
3	Full credit (all three parts answered correctly)
2	Partial credit (only two parts answered correctly)
1	Partial credit (only one part answered correctly)
0	No credit (none of the parts answered correctly)
For 2-point Derived Items:	
2	Full credit (both parts answered correctly)
1	Partial credit (only one part answered correctly)
0	No credit (neither part answered correctly)

***Codes for Missing Values***

A subset of values was reserved for specific codes related to different categories of missing data. We recommend that the user read the following section with particular care because the way in which these missing codes are used may have implications for analyses.

Not Administered Response Codes (SAS: .A ; SPSS: sysmis)

Special codes were given to items that were “Not Administered” to distinguish these cases from data that were missing due to non-response. In general, the not administered code was used when an item was not administered, either by design arising from the rotation of items across the assessment booklets and tasks, or unintentionally when an item was misprinted or otherwise unavailable for a student to respond to. The not administered code was used in the following cases:

- Achievement item not assigned to the student: All students participating in PIRLS 2016 received one of 16 available test booklets, students participating in PIRLS Literacy 2016 received one of 16 test booklets, and ePIRLS 2016 students were administered two randomly assigned simulated online reading tasks. All variables corresponding to items that were not present in a student’s assigned booklet, or ePIRLS tasks, were coded as “Not Administered.”
- Student absent from session: When a student was not present for a particular testing session (either part 1 or part 2 of an assessment booklet, or pair of tasks) all items relevant to that session were coded as “Not Administered.”
- Item left out or misprinted: When a particular item (or a whole page) was misprinted or otherwise not available to the student, the corresponding variable was coded as “Not Administered.”

- Item mistranslated or deleted: An item identified during translation verification or item review as having a translation error such that the nature of the item was altered, or as having poor psychometric properties, was coded as “Not Administered.”

Omitted Response Codes (SAS: . ; SPSS: 9, 99)

“Omitted” response codes were used for items that a student should have answered but did not. An omitted response code was given when an item was left blank or when two or more response options were checked for a multiple choice item. Omitted responses generally are treated as incorrect responses.

Not Reached Response Codes (SAS: .R ; SPSS: 6, 96)

An item was considered “Not Reached” when, within part 1 or part 2 of a booklet, or pair of ePIRLS tasks, the item itself and the item immediately preceding it were not answered, and there were no other items completed in the remainder of that part of the booklet or task. For most purposes, PIRLS 2016 treated the not reached items as incorrect responses, except during the item calibration step of the IRT scaling, when not reached items were considered to have not been administered (see [Chapter 12 of Methods and Procedures in PIRLS 2016](#)).

### ***PIRLS Achievement Scores***

Achievement scales were produced for overall reading achievement, its two purposes for reading, and its two processes of comprehension, for both PIRLS 2016 and PIRLS Literacy 2016, as shown in Exhibit 4.5. For ePIRLS 2016, achievement scales were produced for overall online informational reading and its two processes of comprehension. A total of five achievement scales were produced in PIRLS 2016 and PIRLS Literacy 2016; an additional three scales were produced in ePIRLS 2016. Users should note that all PIRLS 2016 achievement scores also are available in the ePIRLS 2016 data files. A detailed description of the PIRLS, PIRLS Literacy, and ePIRLS scaling methods and how these achievement scales were created is available in [Chapter 12 of Methods and Procedures in PIRLS 2016](#).

For each achievement scale, the PIRLS 2016 International Database provides five separate estimates of each student’s score on that scale. The five estimated scores are known as “plausible values,” and the variability between them encapsulates the uncertainty inherent in the scale estimation process. The plausible values for any given scale are the best available measures of student achievement on that scale in the PIRLS 2016 International Database, and should be used as the outcome measure in any study of student achievement. Plausible values can be analyzed readily using the IEA IDB Analyzer (IEA, 2018), as described in Chapter 2 of this User Guide.

**Exhibit 4.5 PIRLS 2016 Achievement Scales**

<b>PIRLS 2016 Achievement Scales</b>			
PIRLS and PIRLS Literacy	Reading	REA	Overall Reading
	Purposes for Reading	LIT	Reading for Literary Experience
		INF	Reading to Acquire and Use Information
	Processes of Comprehension	RSI	Retrieving and Straightforward Inferencing
		IIE	Interpreting, Integrating, and Evaluating
ePIRLS	Reading	REA	Overall Reading
	Processes of Comprehension	RSI	Retrieving and Straightforward Inferencing
		IIE	Interpreting, Integrating, and Evaluating

The achievement score variable names are based on an 8-character alphanumeric code, which adheres to the following rules:

- The first character is always “A,” corresponding to a fourth grade score;
- The second character is always “S,” to indicate it is a student level variable;
- The third character is “R” for PIRLS and PIRLS Literacy, to indicate a reading score, whether it is an overall score, a purpose score, or a process score; it is “E” for ePIRLS, to indicate an online reading score;
- The fourth through sixth characters are a three-character code describing the achievement scale, as shown in Exhibit 4.5; and
- The seventh and eighth characters are a two-digit number indicating the plausible value—“01,” “02,” “03,” “04,” or “05.”

For example, ASRINF01 is the first plausible value on the PIRLS and PIRLS Literacy informational purpose achievement scale. ASEREA01 is the first plausible value on the ePIRLS online informational reading achievement scale.

### ***PIRLS International Benchmarks of Reading Achievement***

To help users of the PIRLS 2016 International Database understand what performance on the overall reading achievement scales signifies in terms of the reading skills and strategies demonstrated by fourth grade students, PIRLS identified four points on the overall reading scale to serve as International Benchmarks. As shown in Exhibit 4.6, the PIRLS International Benchmark scores are 625, 550, 475, and 400, which correspond to the Advanced International Benchmark, the High International Benchmark, the Intermediate International Benchmark, and the Low International Benchmark, respectively. PIRLS used a technique known as scale anchoring to summarize and describe student achievement at these four points on the scale.<sup>5</sup> The

<sup>5</sup> The scale anchoring procedure is described in *Chapter 13 of Methods and Procedures in PIRLS 2016*.

*PIRLS 2016 International Results in Reading* presents the results of this scale anchoring, and reports the percentage of students in each country reaching each of the PIRLS International Benchmarks.

**Exhibit 4.6 PIRLS 2016 International Benchmarks of Reading Achievement**

Scale Score	International Benchmark
625	Advanced International Benchmark
550	High International Benchmark
475	Intermediate International Benchmark
400	Low International Benchmark

The PIRLS 2016 International Database contains a set of variables indicating which International Benchmark the students have reached. There is a benchmark variable for each plausible value of the PIRLS overall reading scale and of the ePIRLS online informational reading scale. The International Benchmark variables follow the achievement score variable naming convention where the fourth through sixth positions have the letters “IBM.” Thus, ASRIBM01–05 are the five benchmark variables for overall reading in PIRLS and PIRLS Literacy; ASEIBM01–05 are the five benchmark variables for overall online informational reading in ePIRLS 2016. The codes defined for all the benchmark variables are described in Exhibit 4.7.

**Exhibit 4.7 PIRLS 2016 International Benchmark Variable Codes**

Code	Description
1	Student performed below the Low International Benchmark
2	Student performed at or above the Low International Benchmark, but below the Intermediate International Benchmark
3	Student performed at or above the Intermediate International Benchmark, but below the High International Benchmark
4	Student performed at or above the High International Benchmark, but below the Advanced International Benchmark
5	Student performed at or above the Advanced International Benchmark

#### 4.5.2 PIRLS 2016 Within-Country Scoring Reliability Data Files (ASR)

The PIRLS 2016, PIRLS Literacy 2016, and ePIRLS 2016 within-country scoring reliability data files contain data that can be used to investigate the reliability of the constructed response item scoring. The scoring reliability data files contain one record for each student whose responses to constructed response items, in whole or in part, were double scored during the within-country scoring reliability exercise (see *Chapter 10 of Methods and Procedures in PIRLS 2016*). For each constructed response item in the assessment, the following three variables are included in the scoring reliability data files:

- Original Score: the score assigned by the first scorer and also present in the student achievement files;
- Second Score: score assigned by the second scorer and present only in the scoring reliability files; and
- Score Agreement: dichotomous variable indicating agreement between the two scorers.

It should be noted that the second score data were used only to evaluate within-country scoring reliability and were not used in computing the achievement scores included in the International Database and presented in [\*PIRLS 2016 International Results in Reading\*](#).

### ***Scoring Reliability Variable Naming Convention***

The variable names for the Original Score, Second Score, and Score Agreement variables are based on the same naming convention as that for the achievement item variables discussed earlier. The third character in the variable name differentiates between the three reliability variables:

- The Original Score variable has the number “1” as the third character, as in the achievement item naming convention (e.g., R31M17CA);
- The Second Score variable has the letter “R” as the third character (e.g., R3RM17CA); and
- The Score Agreement variable has the letter “I” as the third character (e.g., R3IM17CA).

### ***Scoring Reliability Variable Score Values***

The values contained in both the Original Score and Second Score variables are the scores assigned using the PIRLS 2016 scoring guides. The Score Agreement variable may have one of two values, denoting either agreement or disagreement between the two scorers, as described in Exhibit 4.8.

**Exhibit 4.8 PIRLS 2016 Score Agreement Variable Codes**

<b>Code</b>	<b>Description</b>
1	Agreement—Identical scores from both scorers
0	Disagreement—Different scores from both scorers

### **4.5.3 PIRLS 2016 Context Questionnaire Data Files**

There are six types of PIRLS 2016 context questionnaire data files: four (student, parent, teacher, and school) data files correspond to the four types of context questionnaires administered in PIRLS 2016, the fifth corresponds to the curriculum questionnaire administered to the PIRLS National Research Coordinators, and the sixth serves to link the student and teacher background data files. The first four types of data files contain the responses to the questions asked in their respective context questionnaires.

Because the ePIRLS 2016 samples are in effect subsamples of the PIRLS 2016 samples, all ePIRLS 2016 context questionnaire data files contain a subset of the records from their corresponding PIRLS 2016 context questionnaire data files. In other words, users will find the same set of responses for the same students, whether it be in the associated student background data files, home background data files, school background data files, or teacher background data files.

### ***PIRLS 2016 Student Background Data Files (ASG)***

Students who participated in PIRLS 2016 were administered a context questionnaire with questions related to their home background, school experiences, and attitudes towards reading. The student background data files contain students' responses to these questions. They also contain students' reading achievement scores (plausible values) to facilitate analyses of relationships between student background characteristics and achievement.

The student background data files also contain a number of identification variables, tracking variables, sampling and weighting variables, and derived variables that were used for producing exhibits in *PIRLS 2016 International Results in Reading*. These variables are described later in this chapter.

### ***PIRLS 2016 Home Background Data Files (ASH)***

Countries participating in PIRLS 2016 and PIRLS Literacy 2016 administered the Learning to Read Survey questionnaire, which was completed by the students' parents or guardians. This questionnaire is usually called the Home Questionnaire. It asked questions about preparations for primary schooling, including attendance in preschool and literacy-centered activities in the home before the child began school, such as reading books, singing songs, or writing letters or words. Parents answered questions about home resources in addition to information about their highest level of education and their employment situation. Analyzing data from the home questionnaires requires that the home background data files be merged with the student background data files using the country and student identification variables. Details of the merging procedure with the IEA IDB Analyzer are described in Chapter 2 of this User Guide.

### ***PIRLS 2016 School Background Data Files (ACG)***

The school background data files contain school principals' responses to the questions in the PIRLS 2016 school context questionnaires. Although school-level analyses where the schools are the units of analysis can be performed, it is preferable to analyze school-level variables as attributes of students. To perform student-level analyses with school data, the school background data files must be merged with the student background data files using the country and school identification variables. Details of the merging procedure with the IEA IDB Analyzer are described in Chapter 2 of this User Guide.

### ***PIRLS 2016 Teacher Background Data Files (ATG)***

The reading teachers of the students that were sampled in PIRLS 2016 were administered a questionnaire with questions pertaining to their background and their teaching practices in the classes of the sampled students. Each teacher was asked to respond to a questionnaire for each class taught that contained sampled students. The teacher background data files contain one record for each of the sampled classes taught by a reading teacher. If teachers taught more than one class, they were expected to complete only one part A (general background questions) and a separate part B (class-specific questions) for each class they taught. In some cases, although the teacher was to respond to more than one questionnaire, responses to only one were obtained. In these cases, there were as many records entered in the teacher background data file as sampled classes were taught by the teacher, and the background information in part A from the completed questionnaire was entered into these teacher records.

In the teacher background data files, each teacher has a unique identification number (IDTEACH) and a link number (IDLINK) that is specific to the sampled class taught by the teacher and to which the information in the data record corresponds. The IDTEACH and IDLINK combination uniquely identifies, within a country, a teacher teaching a specific class. Thus, students linked to teachers identified by the same IDTEACH but different IDLINK are taught by the same teacher but in different classes. The teacher background data files cannot be merged directly with the student data files and they do not contain sampling and weighting information, nor achievement scores. The student-teacher linkage data files serve that purpose.

It is important to note that the teachers in the teacher background data files do not constitute a representative sample of teachers in a country, but rather are the teachers who taught a representative sample of students. The teacher data, therefore, should be thought of as attributes of the students to which they are linked, and should be analyzed only in conjunction with the student-teacher linkage data files. Chapter 2 of this User Guide describes student-level analyses combining the teacher data and the student-teacher linkage data files with the IEA IDB Analyzer software.

### ***PIRLS 2016 Student-Teacher Linkage Data Files (AST)***

The PIRLS 2016 student-teacher linkage data files contain information required to link the student and teacher data files. The student-teacher linkage data files contain one entry per student-teacher linkage combination in the data. For instance, if three teachers are linked to a student, there are three entries in the file corresponding to that student. The sole purpose of the student-teacher linkage data files is to link teacher-level data with student-level data in order to perform appropriate student-level analyses where teacher characteristics are considered as attributes of the students. The student-teacher linkage data files also include achievement scores and sampling and weighting information.

### ***PIRLS 2016 Curriculum Data File***

The PIRLS 2016 curriculum questionnaire data file contains the responses provided by the PIRLS National Research Coordinators of the participating countries to the PIRLS 2016 curriculum questionnaire. There is a single PIRLS 2016 curriculum questionnaire data file combining PIRLS and PIRLS Literacy participants. This file is available in Excel in the PIRLS 2016 International Database.

### ***Background Variable Naming Convention***

The background variable naming convention for the variables in the student, home, teacher, and school background data files is based on a 7- or 8-character string. The following rules are applied in naming the background variables:

- The first character is always “A”, representing fourth grade data.
- The second character indicates the type of respondent. The letter “C” identifies data from the school principals, the letter “T” for teacher data, and the letter “S” for student and parent data.
- The third character is used to indicate the source of the data. The letter “B” is used for all background variables. The letter “D” is used for variables derived from responses in the context questionnaires. In addition, the letter “B” is used for the Rasch scores associated with context questionnaire scales derived from questionnaire data and the letter “D” is used for the index variables constructed from these context questionnaire scale Rasch scores.<sup>6</sup>
- The fourth character is used to indicate the topic to which a background question refers. The following letters are used:
  - G—General questions;
  - H—Home questionnaire questions;
  - R—Questions related to reading; and
  - E—ePIRLS questions related to the students’ familiarity with computers.
- The fifth through eighth characters of all background variables represent the sequential numbering of the questions as presented in their respective questionnaires.<sup>7</sup>

The curriculum data files follow their own variable naming convention. The first three characters of a variable name are either “GEN” for general questions, or “READ” for questions related to reading instruction. The remaining characters in the curriculum data variable naming convention refer to the question location, as shown in [\*Supplement 1 to this User Guide\*](#).

### ***Background Variable Location Convention***

The background variable naming convention shows explicitly the ordering of questions in the context questionnaires. Each question also was assigned a unique location code as shown in

<sup>6</sup> The context questionnaire scales are described in [\*Chapter 14 of Methods and Procedures in PIRLS 2016\*](#).

<sup>7</sup> Because of the modular structure of the student and teacher questionnaires, the sequential numbering is specific to each module.

Exhibit 4.9. This unique code includes the sequence number of the question within the questionnaire, the same sequence number found in the question’s variable name. For example, if the location variable is given as SQG-08, it refers to general question 8 in the student context questionnaire. This convention is followed in the codebooks, the data almanacs, and in the description of the variables included in the Supplements to this User Guide.

**Exhibit 4.9 PIRLS 2016 Context Questionnaire Variable Location Convention**

Questionnaire	Location Variable	
Student Questionnaire	SQG-●●●	General questions
	SQR-●●●	Reading questions
Home Questionnaire	HQ-●●●	All questions
Teacher Questionnaire	TQG-●●●	General questions
	TQR-●●●	Reading questions
School Questionnaire	SCQ-●●●	All questions
Curriculum Questionnaire	CQG-●●●	General questions
	CQR-●●●	Reading questions

●●● = sequential numbering of the question location in the questionnaire

### *Data Coding Conventions*

A series of conventions was adopted to code the data included in the background data files. This section describes these conventions.

#### Background Question Response Code Values

The values assigned to each of the background variables depend on the item format and the number of options available. For categorical questions, sequential numerical values were used to correspond to the response options available. The numbers correspond to the sequence of appearance of the response options. For example, the first response option was represented with a 1, the second response option with a 2, etc. Open-ended questions—such as “How many students are in this class?”—were coded with the actual number given as a response.

#### Codes for Missing Values

A subset of values is reserved for specific codes related to various categories of missing data.

Not Administered Response Codes (SAS: .A ; SPSS: sysmis)

Special codes were given to items that were “Not Administered” to distinguish these cases from data that were missing due to non-response. In general, the “Not Administered” code was used when an entire questionnaire was not completed or a specific question was not administered,

such as when a question was left out of the instrument or misprinted. The “Not Administered” code was used in the following cases:

- Background questions removed: Variables corresponding to questions in the student, home, teacher, or school context questionnaires that were considered not appropriate in some countries were not included in the national versions of the questionnaires. These questions were coded as “Not Administered.”
- Background question left out or misprinted: When a particular question (or a whole page) was misprinted, or otherwise not available to the respondent, the corresponding variables were coded as “Not Administered.”
- Background questions mistranslated or not internationally comparable: In some cases, questions in the international version of the questionnaires were mistranslated or modified to fit the national context. Whenever possible, modified questions were recoded to match as closely as possible the international version. When this was not possible, modified questions were coded as “Not Administered.”

Omitted Response Codes (SAS: . ; SPSS: 9, 99, 999, ...)

“Omitted” response codes were used for questions that a student, parent, teacher, or school principal should have answered but did not. The length of the omitted response code given to a variable in the SPSS data files depends on the number of characters needed to represent the variable. In all cases, the space necessary to represent the variable is filled with 9’s. For questionnaire data, no distinction was made between items left blank and items with invalid answers, such as checking two or more response options in a categorical question, or unreadable or uninterpretable responses to open-ended questions. In a few cases, data received from a country in an invalid or inconsistent manner also were coded as “Omitted.”

Not Applicable Response Codes (SAS: .B ; SPSS: 6, 96, 996, ...)

“Not Applicable” response codes were used for the context questionnaire items for which responses were dependent upon a filter question. Generally, a “No” response to a filter question leads to any follow-up questions being coded as “Not Applicable” because there were no appropriate responses to these follow-up questions.

### ***Context Questionnaire Scales and Derived Variables***

In the PIRLS 2016 context questionnaires, often several questions were asked about various aspects of a single construct. In these cases, responses to the individual items were combined to create a score, using Rasch scaling, which provided a more comprehensive picture of the construct of interest than the individual variables could on their own. These context questionnaire scales also were categorized, usually into three groups, to create an index. The context questionnaire scales and their indices are included in the PIRLS 2016 International Database and described in [\*Chapter 14 of Methods and Procedures in PIRLS 2016\*](#).

Additional variables were derived from responses to multiple questions to provide more pertinent information for analysis and reporting. Parents’ education is an example where responses from both parents, or guardians, were combined into a single variable to report a single educational level. [Supplement 3 to this User Guide](#) provides a description of the derived variables included in the PIRLS 2016 International Database.

***Sampling and Weighting Variables***

Several sampling and weighting variables are included in the PIRLS 2016 data files; they are listed and described in Exhibit 4.10. Exhibit 4.11 illustrates the location of the various sampling and weighting variables among the different types of data files. It is important to note that the teacher background data files do not have any sampling and weighting variables, nor do the home background data files.

**Exhibit 4.10 PIRLS 2016 Sampling and Weighting Variables**

<b>Variable Names</b>	<b>Descriptions</b>
TOTWGT	Total student weight—sums to the national population
SENWGT	Student senate weight—sums to 500 in each country
HOUWGT	Student house weight—sums to the national student sample size
SCHWGT	School weight
TCHWGT	Teacher weight
JKZONE	The sampling zone, or stratum, to which the student’s school is assigned
JKREP	The sampling replicate, or primary sampling unit, to which the student’s school is assigned
JKCZONE	The sampling zone, or stratum, to which the school is assigned
JKCREP	The sampling replicate, or primary sampling unit, to which the school is assigned
WGTFAC1	School weighting factor
WGTADJ1	School weighting adjustment
WGTFAC2	Class weighting factor
WGTADJ2	Class weighting adjustment
WGTFAC3	Student weighting factor
WGTADJ3	Student weighting adjustment

As a general rule, TOTWGT—the overall student sampling weight—is the preferred sampling weight to use when analyzing student-level data. Although TOTWGT has desirable properties, it may have drawbacks for some analyses. Because TOTWGT sums to the student population size in each country, analyses using TOTWGT that combine countries will have proportionately more students from larger countries and fewer from smaller countries, which may not be desirable for

some purposes. For cross-country analyses in which countries should be treated equally, PIRLS provides SENWGT, a transformation of TOTWGT, that results in a weighted sample size of 500 in each country. Additionally, because TOTWGT inflates sample sizes to estimate the population size, software systems that use the actual sample size to perform significance tests may give misleading results for analyses weighted by TOTWGT. HOUWGT, another transformation of TOTWGT, ensures that the weighted sample corresponds to the actual sample size in each country.

**Exhibit 4.11 Location of Sampling and Weighting Variables in the PIRLS 2016 Data Files**

Sampling and Weighting Variables	Data File Types			
	ASG	ASA	AST	ACG
JKREP	●	●	●	
JKZONE	●	●	●	
JKCREP				●
JKCZONE				●
TOTWGT	●	●		
SENWGT	●	●		
HOUWGT	●	●		
TCHWGT			●	
SCHWGT				●
WGTFAC1	●			●
WGTADJ1	●			●
WGTFAC2	●			
WGTADJ2	●			
WGTFAC3	●			
WGTADJ3	●			

The weight variables TOTWGT, SENWGT, and HOUWGT are designed for use in student-level analyses from all student-level and school-level files, including the home background data files. The weight variable SCHWGT is designed for use in school-level analyses where the schools are the units of analysis.

The weight variable TCHWGT is specifically designed for using teacher background data in student-level analyses and is based on TOTWGT. The teacher weight is located in the student-teacher linkage data files (AST), not in the actual teacher background data files (ATG). Analyses

with teacher data will be weighted properly by merging the teacher files with the student-teacher linkage files, which is done automatically when using the IEA IDB Analyzer.

The sampling variables beginning with the letters “JK” are used to compute standard errors based on the jackknife repeated replication methodology.<sup>8</sup> All weighting variables beginning with the letters “WGT” provide insight into the multi-stage sampling and weighting methodology applied to the PIRLS data. All weighting variables are described in [Chapter 3 of \*Methods and Procedures in PIRLS 2016\*](#).

A word of caution is in order to prospective users of the PIRLS 2016 International Database. As tempting as it seems to merge all file types into one all-encompassing merged data file, analyses of student, home, and school background data may not produce correct results if the weight variable TOTWGT is used. This problem arises from the student-teacher linkage and how it apportions student weights (TOTWGT) to the student-teacher linkage records in the teacher weights (TCHWGT). Thus only the use of the weight variable TCHWGT will produce correct results in this context, even for analyses that do not include teacher background data.

### ***Structure and Design Variables***

Besides the variables used to store responses to the context questionnaires and achievement booklets, the PIRLS 2016 data files also contain variables meant to store information that identifies and describes the respondents and design information required to properly analyze the data.

#### Identification Variables

In all PIRLS 2016 data files, several identification variables are included that provide information to identify countries, students, teachers, or schools. These variables also are used to link cases between the different data file types. The identification variables have the prefix “ID” and are described below.

#### IDCNTY

IDCNTY is a six-digit country identification code based on the ISO 3166 classification as shown in Exhibit 4.1. This variable should always be used as the first linking variable whenever files are linked within and across countries.

#### IDSCHOOL

IDSCHOOL is a four-digit identification code that uniquely identifies the participating schools within each country. The school codes are generated and assigned specifically for PIRLS and are not meant to represent actual national school identifiers in the participating countries. They are

---

<sup>8</sup> Starting with TIMSS 2015 and PIRLS 2016, the jackknife repeated replication method has been modified to include both replicates within each sampling zone, as described in [Chapter 4 of \*Methods and Procedures in PIRLS 2016\*](#).

not unique across countries. Schools across countries can be identified uniquely only with the IDCNTRY and IDSCHOOL combination of linking variables.

#### IDCLASS

IDCLASS is a six-digit identification code that uniquely identifies the sampled classrooms within a country. The variable IDCLASS has a hierarchical structure and is formed by concatenating the IDSCHOOL variable and a two-digit sequential number identifying the sampled classrooms within a school. Classrooms can be identified uniquely in the database by the combination of IDCNTRY and IDCLASS as linking variables.

#### IDSTUD

IDSTUD is an eight-digit identification code that uniquely identifies each sampled student in a country. The variable IDSTUD also has a hierarchical structure and is formed by concatenating the IDCLASS variable and a two-digit sequential number identifying all students within each sampled classroom. Students can be identified uniquely in the database by the combination of IDCNTRY and IDSTUD as linking variables.

#### IDTEACH

IDTEACH is a six-digit identification code that uniquely identifies a teacher within a school. It has a hierarchical structure and is formed by the concatenation of IDSCHOOL and a two-digit sequential number within each school.

#### IDLINK

IDLINK uniquely identifies the class for which a teacher answered a questionnaire. The combination of linking variables IDCNTRY, IDTEACH, and IDLINK uniquely identifies all teacher-class combinations in the database.

#### IDGRADE

IDGRADE identifies the target grade of the participating students. In PIRLS 2016, the usual value is "4" for most countries. Some countries and benchmarking participants can have the values "3" or "5."

#### IDBOOK

IDBOOK identifies the specific PIRLS 2016 or PIRLS Literacy 2016 assessment booklet that was administered to each student. The booklets are given a numerical value from "1" through "16" in PIRLS 2016, and from "17" through "32" in PIRLS Literacy 2016.

ROTATION

ROTATION identifies the specific pair of ePIRLS 2016 tasks that was administered to each student. It is a two-digit number consisting of the values 1 through 5, representing the two assigned tasks (e.g., code 41 is given to students who were assigned task 4 first, and task 1 second).

Exhibit 4.12 shows in which data files the various identification variables are located. It also highlights the combinations of variables used to identify uniquely the records contained in the different data file types. In the student background, home background, and achievement data files, the variables IDCNTRY and IDSTUD provide a unique identification number to identify all students in the database. Since teachers may teach more than one class, the combination of the IDCNTRY, IDTEACH, and IDLINK variables in the teacher background data files is needed to identify uniquely all teachers and the classes they teach. Teacher background variables are linked to the appropriate students using the student-teacher linkage data files. The variable IDSCHOOL, contained in all files, is a unique identification number for each school within a country. Combined with IDCNTRY, it can be used to link school background data to corresponding students or teachers.

**Exhibit 4.12 Location of Identification Variables in the PIRLS 2016 Data Files**

Identification Variables	Data File Types					
	ASA	ASG	AST	ATG	ACG	ASH
IDCNTRY	●	●	●	●	●	●
IDSCHOOL	●	●	●	●	●	●
IDCLASS	●	●	●	●		●
IDSTUD	●	●	●			●
IDTEACH			●	●		
IDLINK			●	●		
IDGRADE	●	●	●	●	●	●
IDBOOK	●	●				
ROTATION	●	●				

Tracking Variables

Information about students, teachers, and schools provided by the survey tracking forms is stored in the tracking variables.<sup>9</sup> These variables have the prefix "IT." All tracking variables are included in the student background data files. ITLANG is included in the student achievement data files, student background data files, and home background data files.

<sup>9</sup> Survey tracking forms are lists of students, teachers, and schools used for sampling and administrative purposes.

ITSEX

Gender of each student as stated in the Student Tracking Forms.

ITBIRTHM and ITBIRTHY

Month and year of birth of each student as stated in the Student Tracking Forms. Both variables are available only in the restricted use version of the PIRLS 2016 International Database.

ITDATE

Testing date for each student. This variable is available only in the restricted use version of the PIRLS 2016 International Database.

ITLANG

Language of testing for each student. The valid codes for ITLANG are specified in the codebook files.

#### **4.6 PIRLS 2016 Codebook Files**

All information related to the structure of the PIRLS 2016 data files, as well as the source, format, descriptive labels, and response option codes for all variables, is contained in codebook files.

There are three codebook files in Excel format:

P16_Codebook	Codebook for the PIRLS 2016 data files
PL16_Codebook	Codebook for the PIRLS Literacy 2016 data files
eP16_Codebook	Codebook for the ePIRLS 2016 data files

In the codebook Excel files, there is a tab for each appropriate data file type in the PIRLS 2016 International Database. These tabs describe the contents and structure of the PIRLS 2016 data files. Important codebook fields include LABEL, which contains extended textual information for all variables, QUESTION LOCATION, which provides the location of questions and achievement items within their respective survey instruments, and VALUE SCHEME DETAILED, which lists the acceptable responses allowed for each variable.

#### **4.7 PIRLS 2016 Data Almanac Files**

Data almanacs provide weighted summary statistics for all variables in the PIRLS 2016 data files. There are two basic types of data almanacs: achievement data almanacs for the achievement items and background data almanacs for the background variables. There are separate sets of data almanac files for PIRLS 2016, PIRLS Literacy 2016, and ePIRLS 2016. All data almanac files are provided in Word format and printable PDF format.

### *Achievement Data Almanacs*

The achievement data almanacs provide weighted summary statistics for each participating country on each individual achievement item included in the PIRLS 2016 assessments. The achievement data almanacs display for each item its classification in the purposes for reading and processes of comprehension, the passage or task to which it belongs, a brief description of the item, its variable name, whether it is a multiple choice or constructed response item, the correct response key if it is a multiple choice item, and its point value if it is a constructed-response item. The achievement data almanacs also display the international averages for each item, with each country weighted equally. The benchmark participants, listed below the international averages, are not included in the calculation of international averages.

The achievement data almanac files available in the International Database are listed in Exhibit 4.13. The trend item almanacs provide summary statistics for achievement items used in both the 2011 and 2016 assessments. Users should note that the PIRLS Literacy 2016 trend item almanac includes only South Africa. Also, there is no trend item almanac for ePIRLS 2016 since it is the first assessment cycle.

**Exhibit 4.13 PIRLS 2016 Achievement Data Almanacs**

<b>Achievement Data Almanacs</b>	<b>Contents</b>
P16_INF_ItemAlmanac	Almanac for PIRLS 2016 informational items
P16_LIT_ItemAlmanac	Almanac for PIRLS 2016 literary items
P16_INF_TrendItemAlmanac	Trend almanac for PIRLS 2016 informational items
P16_LIT_TrendItemAlmanac	Trend almanac for PIRLS 2016 literary items
PL16_INF_ItemAlmanac	Almanac for PIRLS Literacy 2016 informational items
PL16_LIT_ItemAlmanac	Almanac for PIRLS Literacy 2016 literary items
PL16_INF_TrendItemAlmanac	Trend almanac for PIRLS Literacy 2016 informational items
PL16_LIT_TrendItemAlmanac	Trend almanac for PIRLS Literacy 2016 literary items
eP16_INF_ItemAlmanac	Almanac for ePIRLS 2016 items

There are two types of displays in the achievement data almanacs, depending on whether an item is a multiple choice item or a constructed response item. The statistics displayed in these almanacs are as follows:

- N—The number of students to whom the item was administered.
- DIFF—The percentage of students that responded correctly to a multiple choice item.
- A, B, C, and D—The percentage of students choosing each one of the response options for a multiple choice item (options E and F are possible in ePIRLS).
- Scoring Guide Codes (e.g., 0, 1, 2, 3)—The percentage of student responses assigned each of the codes in the scoring guide for a constructed response item.

- **OMITTED**—The percentage of students that omitted a response to the item.
- **NOT REACHED**—The percentage of students that did not reach the item.
- **V1, V2, and V3**—The percentage of students that scored 1 point or better on the item (V1), 2 points or better on the item (V2), or 3 points (V3), for constructed-response items.
- **GIRL PCT RIGHT and BOY PCT RIGHT**—The percentage of girls and boys that either got a multiple choice item right, or obtained the maximum score on a constructed response item.

### **Background Data Almanacs**

Background data almanac files contain weighted summary statistics for each participating country on each variable in the student, home, teacher, and school context questionnaires, including the context questionnaire scales and their indices and the derived variables based on these background variables. There is also a computer navigation almanac specific to ePIRLS 2016.

Among the statistics reported are mean reading achievement by response category. The background data almanacs also display for each variable the question as it was asked, its location in the corresponding questionnaire, and its variable name in the data files. The background data almanac files available in the PIRLS 2016 International Database are listed in Exhibit 4.14.

**Exhibit 4.14 PIRLS 2016 Background Data Almanacs**

<b>Background Data Almanacs</b>	<b>Contents</b>
P16_StudentAlmanac	PIRLS student background almanac with reading achievement
P16_HomeAlmanac	PIRLS home background almanac with reading achievement
P16_TeacherAlmanac	PIRLS teacher background almanac with reading achievement
P16_SchoolAlmanac	PIRLS school background almanac with reading achievement
PL16_StudentAlmanac	PIRLS Literacy student background almanac with reading achievement
PL16_HomeAlmanac	PIRLS Literacy home background almanac with reading achievement
PL16_TeacherAlmanac	PIRLS Literacy teacher background almanac with reading achievement
PL16_SchoolAlmanac	PIRLS Literacy school background almanac with reading achievement
eP16_StudentAlmanac	ePIRLS student background almanac with ePIRLS reading achievement
eP16_HomeAlmanac	ePIRLS home background almanac with ePIRLS reading achievement
eP16_TeacherAlmanac	ePIRLS teacher background almanac with ePIRLS reading achievement
eP16_SchoolAlmanac	ePIRLS school background almanac with ePIRLS reading achievement
eP16_ComputerNavAlmanac	ePIRLS computer navigation almanac with ePIRLS reading achievement

The background data almanacs also display the international averages for each variable, with each country weighted equally. The benchmark participants, listed below the international averages, are not included in the calculation of international averages.

There are two types of displays in the background data almanacs, depending on whether the data are categorical (i.e., have a small number of discrete values) or continuous. The almanac display for categorical variables includes the following:

- The sample size (number of students, parents, teachers, or schools included in the sample);
- The number of valid cases (number of students, parents, teachers, or schools for whom valid data were obtained);
- The weighted percentages of students corresponding to each valid response option (percentages based only on the students with valid data, as well as “Not Applicable” codes when used);
- The weighted percentages of students for whom none of the valid response options were selected, coded as “Not Administered” or “Omitted” (percentages based on the sample size);
- The weighted mean achievement values of students corresponding to each valid response option, as well as the “Not Administered” and “Omitted” codes; and
- In cases where a variable can be coded as “Not Applicable” because of an earlier filter question, the weighted percentage of students for whom the variable is coded as “Not Applicable” also is displayed, along with the corresponding weighted mean achievement.

The almanac display for continuous variables includes the following:

- The sample size (number of students, parents, teachers, or schools included in the sample);
- The number of valid cases (number of students, parents, teachers, or schools for whom valid data were obtained);
- The weighted percentages of students for whom the variable is coded as “Not Administered” or “Omitted” (percentages based on the sample size);
- The weighted mean, mode, minimum, maximum, and the 5<sup>th</sup>, 10<sup>th</sup>, 25<sup>th</sup>, 50<sup>th</sup>, 75<sup>th</sup>, 90<sup>th</sup>, and 95<sup>th</sup> percentiles across students (based on the students with valid data); and
- In cases where a variable can be coded as “Not Applicable” because of an earlier filter question, the weighted percentage of students for whom the variable is coded as “Not Applicable” also is displayed.

It is important to note that all statistics reported in the background data almanacs, with the exception of the sample sizes and the number of valid cases, always are based on student-level calculations—for example, the percentage of students whose teachers or schools gave a particular response to a question, because teacher data and school data usually are analyzed as student attributes.

## 4.8 PIRLS 2016 Program Files

The PIRLS 2016 International Database includes a number of SPSS and SAS programs designed to facilitate the manipulation of the PIRLS 2016 data files. The International Database contains SAS programs that convert the SAS export files into SAS data files. It includes SPSS and SAS syntax

files that will recode the responses to individual items from the achievement data files to their appropriate score levels. It also includes SPSS and SAS programs designed to combine PIRLS 2016 data files and PIRLS Literacy 2016 data files, which may be necessary when analyzing PIRLS reading results. All programs are described in Chapter 3 of this User Guide.

## References

- IBM Corporation. (2016). *IBM SPSS Statistics for Windows*, version 24.0. Armonk, NY: IBM Corporation.
- International Association for the Evaluation of Educational Achievement. (2018). *International Database Analyzer*, version 4.0. Hamburg, Germany: IEA Hamburg.
- Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.). (2017). *Methods and Procedures in PIRLS 2016*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <https://timssandpirls.bc.edu/publications/pirls/2016-methods.html>
- Mullis, I. V. S., & Martin, M. O. (Eds.). (2015). *PIRLS 2016 Assessment Framework* (2nd ed.). Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/pirls2016/framework.html>
- Mullis, I. V. S., & Martin, M. O. (2013). *PIRLS 2016 Item Writing Guidelines*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: [https://timssandpirls.bc.edu/publications/pirls/2016-methods/pdf/P16\\_ItemWritingGuidelines.pdf](https://timssandpirls.bc.edu/publications/pirls/2016-methods/pdf/P16_ItemWritingGuidelines.pdf)
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017). *PIRLS 2016 International Results in Reading*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/pirls2016/international-results/>
- SAS Institute. (2012). *SAS System for Windows*, version 9.4. Cary, NC: SAS Institute.

**PIRLS**  
**2016**

**PIRLS 2016 User Guide  
for the International Database**

# **Appendix**

**Organizations and  
Individuals Responsible  
for PIRLS 2016**

# Organizations and Individuals Responsible for PIRLS 2016

## Introduction

PIRLS (Progress in International Reading Literacy Study) is a collaborative effort involving hundreds of individuals around the world. This appendix acknowledges the individuals and organizations who contributed to the assessment. Given that work on PIRLS 2016 has spanned approximately five years and has involved so many people and organizations, this list may not include all who contributed. Any omission is inadvertent. PIRLS 2016 also acknowledges the students, parents, teachers, and school principals who contributed their time and effort to the study. It would not be possible without them.

## Management and Coordination

PIRLS is a major undertaking of IEA, and together with TIMSS (Trends in International Math and Science Study), comprises the core of IEA's regular cycles of studies. The PIRLS assessment at the fourth grade complements TIMSS, which regularly assesses science and math achievement at the fourth and eighth grades.

PIRLS was conducted by IEA's TIMSS & PIRLS International Study Center at Boston College, which has responsibility for the overall direction and management of the TIMSS and PIRLS projects, including design, development, and implementation. Headed by Executive Directors Drs. Ina V.S. Mullis and Michael O. Martin, the study center is located in the Lynch School of Education. In carrying out the project, the TIMSS & PIRLS International Study Center worked closely with IEA Amsterdam, which managed country participation, was responsible for verification of all translations produced by the participating countries, and coordinated the school visits by International Quality Control Monitors. Staff at IEA Hamburg worked closely with participating countries to organize sampling and data collection operations and to check all data for accuracy and consistency within and across countries; Statistics Canada in Ottawa was responsible for school and student sampling activities; The Australian Council for Educational Research (ACER) participated in developing the ePIRLS tasks and items, and ACER and the National Foundation for Educational Research in England (NFER) participated in developing

the PIRLS 2016 passages and items; and Educational Testing Service in Princeton, New Jersey consulted on psychometric methodology, provided software for scaling the achievement data, and replicated the achievement scaling for quality assurance.

The Project Management Team, comprising the study directors and representatives from the TIMSS & PIRLS International Study Center, IEA Amsterdam and IEA Hamburg, Statistics Canada, and ETS met twice a year throughout the study to discuss the study's progress, procedures, and schedule. In addition, the study directors met with members of IEA's Technical Executive Group twice yearly to review technical issues.

To work with the international team and coordinate within-country activities, each participating country designates an individual to be the PIRLS National Research Coordinator (NRC). The NRCs have the challenging task of implementing PIRLS in their countries in accordance with the PIRLS guidelines and procedures. In addition, the NRCs provide feedback and contributions throughout the development of the PIRLS assessment. The quality of the PIRLS assessment and data depends on the work of the NRCs and their colleagues in carrying out the complex sampling, data collection, and scoring tasks involved. Continuing the tradition of exemplary work established in previous cycles of PIRLS, the PIRLS 2016 NRCs performed their many tasks with dedication, competence, energy, and goodwill.

## Funding

Funding for PIRLS 2016 was provided primarily by the participating countries. The National Center for Education Statistics of the U.S. Department of Education was a major funding partner, providing funding under contract number ED08C00117. The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. Boston College also is gratefully acknowledged for its generous financial support and stimulating educational environment.

## TIMSS & PIRLS International Study Center at Boston College

Ina V.S. Mullis, *Executive Director*

Michael O. Martin, *Executive Director*

Pierre Foy, *Director of Sampling, Psychometrics, and Data Analysis*

Paul Connolly, *Director, Graphic Design and Publications*

Marcie Bligh, *Manager, Events and Administration*

Katie Trong Drucker, *PIRLS Coordinator (through 2013)*

Susan Farrell, *Lead Web and Database Designer*

Bethany Fishbein, *Research Specialist, Instrument Development and Reporting*

Elena Forzani, *Assistant Research Director, PIRLS (from 2015)*

Joseph Galia, *Lead Statistician/Programmer*

Shirley Goh, *Assistant Director, Communications and Media Relations*

Christine Hoage, *Manager of Finance*

Kathleen Holland, *Administrative Coordinator*

Martin Hooper, *Assistant Research Director, TIMSS and PIRLS Questionnaire Development and Policy Studies*

Ieva Johansone, *Associate Research Director, Operations and Quality Control*

Cristián Leiva, *Front-End Web Developer*

Lauren Palazzo, *Research Associate, TIMSS and PIRLS Questionnaire and Technical Reporting*

Yenileis Pardini, *Lead Designer/Developer for eAssessments*

Mario Pita, *Lead Graphic Designer*

Jyothsna Pothana, *Statistician/Programmer*

Betty Poulos, *Administrative Coordinator (through 2016)*

Caroline Prendergast, *Research Associate, PIRLS*

Ruthanne Ryan, *Senior Graphic Designer*

Jennifer Moher Sepulveda, *Data Graphics Specialist (through 2015)*

Steven A. Simpson, *Senior Graphic Designer*

Erin Wry, *Research Associate, TIMSS and PIRLS Operations and Quality Control*

Liqun Yin, *Research Psychometrician*

## IEA Amsterdam

Dirk Hastedt, *Executive Director*

Hans Wagemaker, *Executive Director (through 2014)*

Paulína Koršňáková, *Director of the IEA Secretariat (through 2016)*

Andrea Netten, *Director of the IEA Secretariat*

Barbara Malak, *Manager, Member Relations (through 2013)*

Roel Burgers, *Financial Director*

Juriaan Hartenberg, *Financial Manager (through 2013)*

Gabriela Nausica Noveanu, *Senior Research Advisor*

David Ebbs, *Senior Research Officer*

Michelle Djekić, *Research and Liaison Officer*

Isabelle Braun-Gémin, *Financial Officer*

Dana Vizkova, *Financial Officer*

Gillian Wilson, *Publications Officer*

Manuel Butty, *Public Outreach Officer (through 2016)*

## IEA Hamburg

Heiko Sibberns, *IEA Hamburg Director*

Oliver Neuschmidt, *Senior Research Analyst, Unit Head, International Studies*

Milena Taneva, *Senior Research Analyst, Project Co-Manager, PIRLS Data Processing*

Juliane Hencke, *Senior Research Analyst, Project Co-Manager, PIRLS Data Processing*

Sebastian Mayer, *Research Analyst, Deputy Project Manager, PIRLS Data Processing*

Mark Cockle, *Research Analyst, Deputy Project Manager, PIRLS Data Processing*

Yasin Afana, *Research Analyst*

Alena Becker, *Research Analyst*

Clara Beyer, *Research Analyst*

Christine Busch, *Research Analyst*

Tim Daniel, *Research Analyst*

Limiao Duan, *Programmer*

Eugenio Gonzalez, *Senior Research Analyst*

Michael Jung, *Research Analyst*

Deepti Kalamadi, *Programmer*

Hannah Köhler, *Research Analyst*

Kamil Kowolik, *Research Analyst*

Sabine Meinck, *Senior Research Analyst, Head of Research, Analysis & Sampling Unit*

Ekaterina Mickheeva, *Research Analyst*

Dirk Oehler, *Research Analyst*

Duygu Savaşçı, *Research Analyst, Sampling Team*

Sabine Tieck, *Research Analyst, Sampling Team*

Meng Xue, *Unit Head, Software*

## Statistics Canada

Sylvie LaRoche, *Senior Methodologist*  
Marc Joncas, *Senior Methodologist (through 2015)*  
Ahmed Almaskut, *Methodologist*  
Shou Xiang Chen, *Methodologist (through 2016)*

## Educational Testing Service

Edward Kulick, *Research Director*  
Jonathan Weeks, *Associate Research Scientist*  
Timothy C. Davey, *Research Director Special Projects*  
Sandip Sinharay, *Principal Research Scientist*  
Scott Davis, *Senior Research Data Analysis Consultant*

## Sampling Referee

Keith Rust, *Vice President and Associate Director of the Statistical Group Westat, Inc.*

## PIRLS 2016 Reading Development Group

Julian Fraillon  
Australian Council for Educational Research  
**Australia**

Jan Mejdning  
Department of Education  
Aarhus University  
**Denmark**

Galina Zuckerman  
Psychological Institute  
Russian Academy of Education  
**Russian Federation**

Elizabeth Pang  
Curriculum Planning and Development  
Division  
Ministry of Education  
**Singapore**

Jenny Wiksten Folkeryd  
Uppsala University  
**Sweden**

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

Marian Sainsbury, *Chief Reading Consultant*  
National Foundation for Educational Research  
**United Kingdom**

Donald Leu  
University of Connecticut  
**United States**

Karen Wixson  
University of North Carolina, Greensboro  
**United States**

### PIRLS 2016 Item Development Task Force

Prue Anderson  
Australian Council for Educational Research  
**Australia**

Liz Twist  
National Foundation for Educational Research  
**United Kingdom**

Marian Sainsbury, Chief Reading Consultant  
National Foundation for Educational Research  
**United Kingdom**

Karen Wixson  
University of North Carolina, Greensboro  
**United States**

### PIRLS 2016 Questionnaire Development Group

Joanne Latourelle  
Coordonnatrice aux études pancanadiennes  
et internationales, Sanction des  
études, Ministère de l'Éducation, du Loisir  
et du Sport  
**Canada**

Maryam A. Al-Ostad  
National Centre for Education Development  
**Kuwait**

Hwa Wei Ko  
Graduate Institute of Learning and  
Instruction  
National Central University  
**Chinese Taipei**

Megan Chamberlain  
Comparative Education Research Unit  
Ministry of Education  
**New Zealand**

Marc Colmant  
Ministère de l'éducation nationale  
Direction de l'évaluation, de la prospective et  
de la performance  
**France**

João Maroco  
Instituto de Avaliação Educativa, I.P.  
**Portugal**

Sarah Howie  
Centre for Evaluation and Assessment  
University of Pretoria  
**South Africa**

### PIRLS 2016 National Research Coordinators

**Australia**  
Sue Thomson  
Australian Council for Educational Research

**Bahrain**  
Huda Al-Awadi  
Ministry of Education

**Austria**  
Christina Wallner-Paschon  
Birgit Suchan (through 2014)  
Bundesinstitut fuer Bildungsforschung,  
Innovation und Entwicklung des  
Oesterreichischen Schulwesens (BIFIE)

**Belgium (Flemish)**  
Kim Bellens  
Centrum voor Onderwijseffectiviteit en  
evaluatie

**Azerbaijan**  
Narmina Aliyeva  
Ministry of Education

**Belgium (French)**  
Anne Matoul  
Geneviève Hindryckx (through 2014)  
Université de Liège

**Bulgaria**

Marina Vasileva Mavrodieva  
Center for Assessment in Pre-School and  
School Education (CAPSE)

**Canada**

Kathryn O'Grady  
Tanya Scerbina  
Pierre Brochu (through 2016)  
Mélanie Labrecque (through 2015)  
Council of Ministers of Education

**Chile**

Carolina Leyton  
Maria Victoria Martinez Muñoz (through  
2016)  
Departamento de Estudios Internacionales  
División de Estudios  
Agencia de Calidad de la Educación

**Chinese Taipei**

Hwa Wei Ko  
Graduate Institute of Learning and  
Instruction  
National Central University

**Czech Republic**

Zuzana Janotová  
Czech School Inspectorate

**Denmark**

Jan Mejding  
Aarhus University Department of Education  
(DPU)

**Egypt**

Shokry Sayed Ahmed  
National Center of Examinations and  
Educational Evaluation

**England**

Grace Grima  
Kath Thomas (through 2015)  
Pearson

**Finland**

Kaisa Leino  
Inga Carita Arffman (through 2016)  
Finnish Institute for Educational Research  
University of Jyväskylä

**France**

Marc Colmant  
Ministère de l'éducation nationale  
Direction de l'évaluation, de la prospective et  
de la performance (DEPP)

**Georgia**

Natia Andguladze  
Nutsa (Magda) Kobakhidze (through 2013)  
National Assessment and Examination Center

**Germany**

Wilfried Bos  
Heike Wendt  
Institut für Schulentwicklungsforschung  
TU Dortmund University

**Hong Kong SAR**

Tse Shek Kam  
The University of Hong Kong  
Faculty of Education

**Hungary**

Ildikó Balázs  
Péter Balkányi  
Educational Authority  
Department of Assessment and Evaluation

**Iran, Islamic Republic of**

Abdol'azim Karimi  
Research Institute for Education (RIE)  
Ministry of Education

**Ireland**

Eemer Eivers  
Educational Research Centre

**Israel**

Inbal Ron-Kaplan  
National Authority for Measurement and  
Evaluation in Education (RAMA)

**Italy**

Laura Palmerio  
Elisa Caponera (through 2013)  
INVALSI – Istituto Nazionale per la Valutazione  
del Sistema Educativo di Istruzione e di  
Formazione

**Kazakhstan**

Aigul Baigulova  
Gulmira Berdibayeva (through 2014)  
Ministry of Education and Science JSC  
Information-Analytic Center

**Kuwait**

Maryam A. Al-Ostad  
National Centre for Education Development

**Latvia**

Antra Ozola  
University of Latvia  
Faculty of Education, Psychology, and Art

**Lithuania**

Ramute Skripkiene  
Irina Mackeviciene (through 2016)  
National Examination Centre  
Ministry of Education and Science

**Macao SAR**

Wong ILin, Irene  
Wai Ki Yiao  
Man Hio Fai (through 2016)  
Education and Youth Affairs Bureau (DSEJ)

**Malta**

Charles Mifsud  
Centre for Literacy  
University of Malta

**Morocco**

Mohammed Sassi  
Centre National de L'Evaluation et des  
Examens et de l'Orientalion  
Ministere de l'Education Nationale et de la  
Formation Professionnelle

**The Netherlands**

Joyce Gubbels  
Andrea Netten (through 2016)  
National Language Education Center,  
Expertisecentrum Nederlands

**New Zealand**

Megan Chamberlain  
Comparative Education Research Unit, EDK  
Ministry of Education

**Northern Ireland**

Juliet Sizmur  
National Foundation for Educational Research  
(NFER)

**Norway**

Egil Gabrielsen  
National Centre for Reading Education and  
Research  
University of Stavanger

**Oman**

Zuwaina Saleh AlMaskari  
Ministry of Education

**Poland**

Michal Federowicz  
Krzysztof Konarzewski  
Educational Research Institute

**Portugal**

João Maroco  
Ana Sousa Ferreira (through 2013)  
Instituto de Avaliação Educativa, I.P.

**Qatar**

Badriya Salman Al-Mohannadi  
Ameena Abdulla Alobaidli (through 2015)  
Saada Hassan Alobaidli (through 2014)  
Student Assessment Office  
Ministry of Education and Higher Education

**Russian Federation**

Galina Kovaleva  
Institute of the Strategies of Educational  
Development  
Russian Academy of Education  
Moscow Center Quality of Education

**Saudi Arabia**

Mohammed Majre Al-Sobeiy  
Saleh Alshaya (through 2013)  
General Directorate of Evaluation  
Ministry of Education

**Singapore**

Elizabeth Pang  
Chan Lee Shan  
Ng Huey Bian (through 2015)  
Curriculum Planning and Development  
Division  
Ministry of Education

**Slovak Republic**

Kristína Čevorová  
Eva Ladányiová (through 2016)  
Soňa Gallová (through 2013)  
National Institute for Certified Educational  
Measurements (NÚCEM)

**Slovenia**

Marjeta Doupona  
Educational Research Institute

**South Africa**

Sarah Howie  
Surette van Staden (through 2014)  
Centre for Evaluation and Assessment (CEA)  
University of Pretoria

**Spain**

Verónica Díez Girado  
Beatriz Sanz Sumelzo (through 2016)  
Jose Maria Sanchez-Echave (through 2015)  
Ministerio de Educación, Cultura y Deporte  
Instituto Nacional de Evaluación Educativa

**Sweden**

Agnes Tongur  
Elina Ekberg  
Tomas Matti (through 2014)  
Swedish National Agency for Education  
(SKOLVERKET)

**Trinidad & Tobago**

Peter Smith  
Division of Educational research and  
Evaluation  
Ministry of Education

**United Arab Emirates**

Moza Rashid AlGhufli  
Ayesha Ghanim Khalfan Almerri (through  
2016)  
Nada Abu Baker Husain Ruban (through  
2015)  
Maryam Mohammed Sulaiman (through  
2015)  
Aljawhara Ali AlSebaiei (through 2014)  
Assessment Department  
Ministry of Education

**United States**

Sheila D. Thompson  
National Center for Education Statistics  
U.S. Department of Education

**Benchmarking Participants**

**Buenos Aires, Argentina**

Tamara Viancur  
Ines Cruzalegui (through 2015)  
Silvia Montoya (through 2015)  
Unit for the Comprehensive Evaluation of  
Educational Quality and Equity

**Ontario, Canada**

Richard Jones  
Angela Hinton (through 2015)  
Education Quality and Accountability Office

**Quebec, Canada**

Joanne Latourelle  
Sanction des études, Ministère de l'Éducation,  
et de l'Enseignement Supérieur

**Moscow City, Russian Federation**

Zozulya Elena Stanislavovna  
Moscow Center for Quality of Education

**Andalusia, Spain**

Sebastián Cárdenas Zabala  
Agencia Andaluza de Evaluación Educativa

**Madrid, Spain**

Luis Pires Jimenez  
Regional Ministry of Education, Youth, and  
Sports

**Abu Dhabi, United Arab Emirates**

Shaikha Ali Al Zaabi  
Nasreen Hussain Al Marzooqi (through 2013)  
Ahlam Habeeb Msaiger  
Assessment Department  
Abu Dhabi Education Council

**Dubai, United Arab Emirates**

Mariam Al Ali  
Rabaa AlSumaiti (through 2014)  
Knowledge and Human Development  
Authority



BOSTON  
COLLEGE

[pirls.bc.edu](http://pirls.bc.edu)

ISBN: 978-1-889938-49-3



**TIMSS & PIRLS**  
International Study Center  
Lynch School of Education  
BOSTON COLLEGE

© IEA, 2018  
International Association  
for the Evaluation of  
Educational Achievement