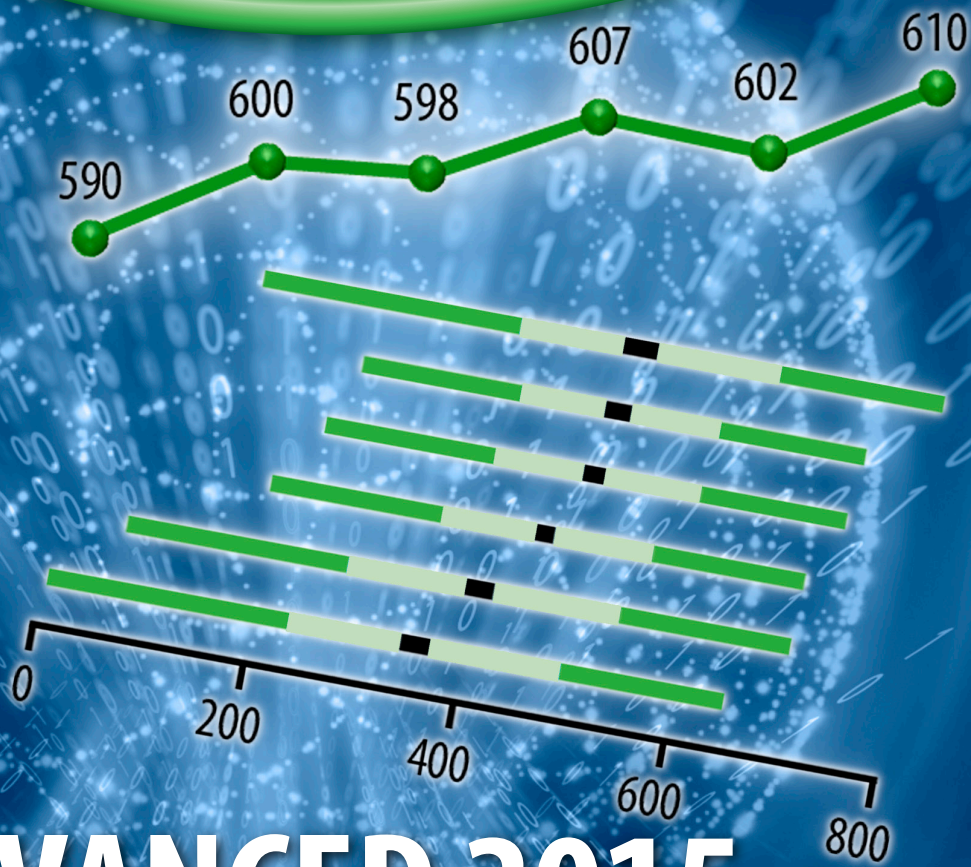


TIMSS



TIMSS ADVANCED 2015 USER GUIDE FOR THE INTERNATIONAL DATABASE

Pierre Foy



TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

TRENDS IN INTERNATIONAL MATHEMATICS AND SCIENCE STUDY

TIMSS

TIMSS Advanced 2015 User Guide for the International Database

Pierre Foy

*Contributors: Victoria A.S. Centurino , Kerry E. Cotter, Bethany G. Fishbein,
Martin Hooper, Hannah Köhler, Kamil Kowolik, Lauren Palazzo, & Erin Wry*



IEA

TIMSS & PIRLS
International Study Center
Lynch School of Education, Boston College

Copyright © 2017 International Association for the Evaluation of Educational
Achievement (IEA)
TIMSS ADVANCED 2015 USER GUIDE FOR THE INTERNATIONAL DATABASE
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: 2017930180
ISBN: 978-1-889938-39-4

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: timss@bc.edu
timss.bc.edu

Boston College is an equal opportunity, affirmative action employer.

Contents

Chapter 1

Introduction 1

Chapter 2

*Using the IEA IDB Analyzer to Analyze the
TIMSS Advanced 2015 International Database* 5

Chapter 3

Special SPSS and SAS Programs 37

Chapter 4

The TIMSS Advanced 2015 International Database Files 43

Appendix

Organizations and Individuals Responsible for TIMSS Advanced 2015. . . 68

Chapter 1

Introduction

1.1 Overview of the TIMSS Advanced 2015 User Guide and International Database

IEA's TIMSS Advanced measures trends in advanced mathematics and physics achievement at the final year of secondary schooling in participating countries around the world, while also monitoring curricular implementation and identifying promising instructional practices. TIMSS Advanced has assessed advanced mathematics and physics in 1995, 2008, and 2015. TIMSS Advanced 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.

To support and promote secondary analyses aimed at improving advanced mathematics and physics education at the end of secondary schooling, the TIMSS Advanced 2015 International Database makes available to researchers, analysts, and other users the data collected and processed by the TIMSS Advanced project. This database comprises student achievement data as well as student, teacher, school, and curricular background data for 9 participating countries. The database includes data from 56,802 students, 4,650 teachers, 2,982 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, TIMSS Advanced 2015 provides trends for up to three cycles—1995, 2008, and 2015. In countries new to the study, the 2015 results can help policy makers and practitioners assess their comparative standing and gauge the rigor and effectiveness of their advanced mathematics and physics programs. Results of the assessments conducted in 2015 can be found in *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics* (Mullis, Martin, Foy, & Hooper, 2016).

TIMSS Advanced 2015 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 TIMSS Advanced data, it is necessary to have an understanding of the characteristics of the study, which are described fully in *Methods and Procedures in TIMSS Advanced 2015* (Martin, Mullis, & Hooper, 2016). It is intended, therefore, that this User Guide 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 TIMSS Advanced 2015 User Guide

This User Guide describes the content and format of the data in the TIMSS Advanced 2015 International Database. In addition to this introduction, the User Guide includes the following chapters:

- Chapter 2—This chapter introduces the IEA International Database (IDB) Analyzer software (IEA, 2017) and presents examples of analyses with the TIMSS Advanced 2015 data using this software in conjunction with SPSS (IBM Corporation, 2013) and SAS (SAS Institute, 2012).
- Chapter 3—This chapter describes special SPSS and SAS programs needed to make full use of the TIMSS Advanced 2015 International Database.
- Chapter 4—This chapter serves as a reference for important details about the structure and content of the TIMSS Advanced 2015 International Database.

The User Guide is accompanied by the following supplements:

- Supplement 1—This supplement contains the international version of all TIMSS Advanced 2015 context questionnaires.
- Supplement 2—This supplement describes all adaptations to the questions in the context questionnaires made by individual TIMSS Advanced 2015 participants.
- Supplement 3—This supplement describes how derived variables were constructed for reporting the TIMSS Advanced 2015 data.

The User Guide and its supplements are available from the TIMSS Advanced 2015 International Database and User Guide webpage: <http://timssandpirls.bc.edu/timss2015/advanced-international-database/>. The primary purpose of this user guide is to introduce users to the TIMSS Advanced 2015 International Database and demonstrate the basic functionality of the IEA IDB Analyzer through simple examples of results published in *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*. The IEA IDB Analyzer comes with its own manual, available through its Help Module, which describes the full functionality and features of the IEA IDB Analyzer.

1.3 The TIMSS Advanced 2015 International Database

The TIMSS Advanced 2015 International Database is available from the TIMSS Advanced 2015 International Database and User Guide webpage: <http://timssandpirls.bc.edu/timss2015/advanced-international-database/>. The TIMSS Advanced 2015 International Database also is available for download from the IEA Study Data Repository website: <http://www.iea.nl/data.html>. The repository allows users to download subsets of files and the corresponding support material through customizable queries from all recent IEA studies, including TIMSS Advanced 2015.

The Database contains the TIMSS Advanced 2015 student achievement data files and student, teacher, and school context 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 TIMSS Advanced 2015 International Database and User Guide webpage.

Exhibit 1.1 Contents of the TIMSS Advanced 2015 International Database

User Guide	This User Guide with its supplements	
Items	The TIMSS Advanced 2015 item information files, IRT item parameters, and item percent correct statistics	
International Database	SPSS Data	TIMSS Advanced 2015 student, teacher, and school data files in SPSS format
	SAS Data	TIMSS Advanced 2015 student, teacher, and school data files in SAS format
	Curriculum Data	TIMSS Advanced 2015 curriculum questionnaires data files
	Codebooks	Codebook files describing all variables in the TIMSS Advanced 2015 International Database
	Almanacs	Data almanacs with summary statistics for all TIMSS Advanced 2015 items and background variables
	TCMA	National item selection data for the TIMSS Advanced 2015 Text-Curriculum Matching Analysis
	SPSS and SAS programs	

1.4 Two Versions of the TIMSS Advanced 2015 International Database

The TIMSS Advanced 2015 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. The list of variables removed from the public use version is given in Chapter 4. The public use version is available for immediate access from the TIMSS Advanced 2015 International Database website, as well as from the IEA Study Data Repository, and users should be able to replicate all published TIMSS Advanced 2015 results with this version of the TIMSS Advanced 2015 International Database. Users who require any of the removed variables to conduct their analyses should contact the IEA through its Study Data Repository (<http://www.iea.nl/data.html>) to obtain permission and access to the restricted use version of the TIMSS Advanced 2015 International Database.

References

IBM Corporation. (2013). *IBM SPSS Statistics* (version 22.0). Somers, NY: IBM Corporation.

International Association for the Evaluation of Educational Achievement. (2017). *IDB Analyzer* (version 4.0). Hamburg, Germany: IEA Hamburg. Available online at <http://www.iea.nl/data.html>.

Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.). (2016). *Methods and Procedures in TIMSS Advanced 2015*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timss.bc.edu/publications/timss/2015-a-methods.html>.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/timss2015/international-results/advanced/>.

SAS Institute. (2012). *SAS System for Windows* (version 9.4). Cary, NC: SAS Institute.

Chapter 2

Using the IEA IDB Analyzer to Analyze the TIMSS Advanced 2015 International Database

2.1 Overview

This chapter describes the general use of the IEA International Database Analyzer software (IEA, 2017) for analyzing the TIMSS Advanced 2015 data. Used in conjunction with either SPSS (IBM Corporation, 2013) or SAS (SAS Institute, 2012), the IEA IDB Analyzer provides a user-friendly interface to easily merge the various data file types of the TIMSS Advanced 2015 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 benchmark levels. The examples use student, teacher, and school background data to replicate some of the TIMSS Advanced 2015 results included in *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics* (Mullis, Martin, Foy, & Hooper, 2016).

Users should be able to perform statistical analyses with the IEA IDB Analyzer with a basic knowledge of the TIMSS Advanced 2015 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 Data Processing and Research Center (IEA DPC), 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.

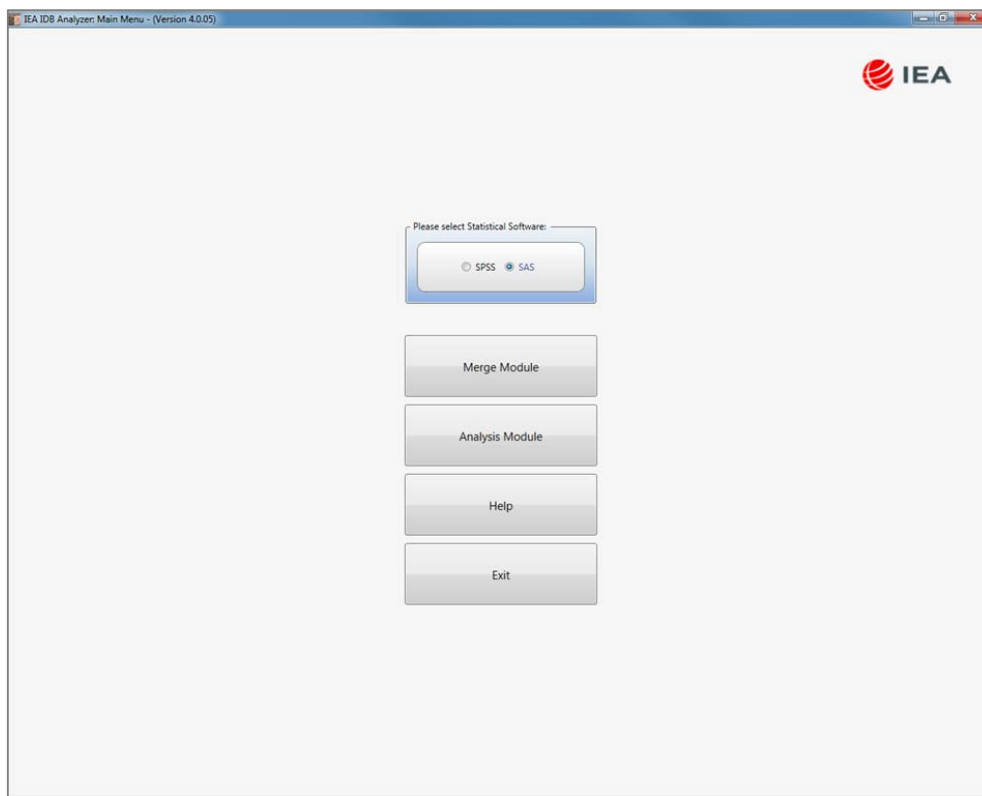
The IEA IDB Analyzer consists of two modules—the merge module and the analysis module—which are integrated and executed in one common application. The merge module is used to create

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 can be accessed by using the START menu in Windows:

Start ⇒ All Programs ⇒ IEA ⇒ IDB AnalyzerV4 ⇒ IEA IDBAnalyzer

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 TIMSS Advanced 2015 International Database and User Guide webpage (<http://timssandpirls.bc.edu/timss2015/advanced-international-database/>) and from the IEA Study Data Repository (<http://www.iea.nl/data.html>). The

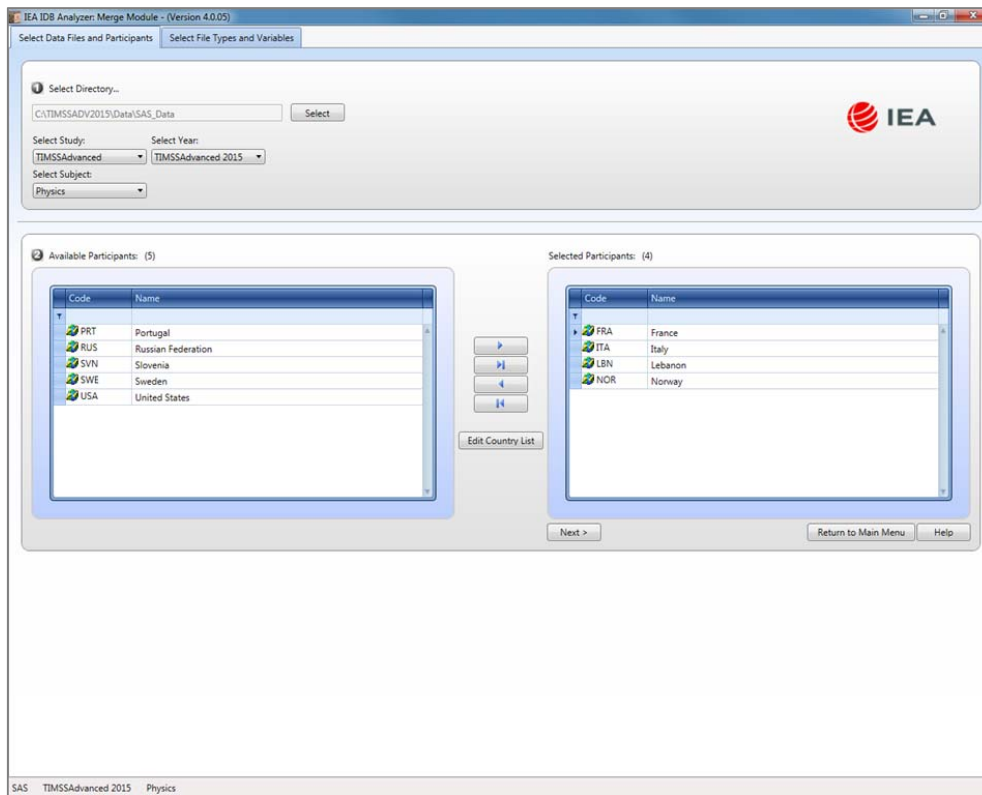
TIMSS Advanced 2015 data files are disseminated separately by file type and for each country. In addition to allowing users to combine like datasets from more than one country for cross-country analyses, the merge module allows for the combination of data from different sources (e.g., student, teacher, and school) into one SPSS or SAS dataset for subsequent analyses.

Before doing any statistical analyses with the TIMSS Advanced 2015 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 have been copied within the folder titled “C:\TIMSSADV2015\.” Additionally, users who plan to analyze TIMSS Advanced 2015 data with SAS will need to convert the SAS export files provided in the TIMSS Advanced 2015 International Database into SAS data files. 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 titled “C:\TIMSSADV2015\Data\SAS_Data.” The program will automatically recognize and complete the **Select Study**, **Select Year**, and **Select Subject** 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 subject (or grade), the IEA IDB Analyzer will prompt users to select files from the desired combination of study and subject (or grade) for analyses. In Exhibit 2.2, physics is the subject 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, France, Italy, Lebanon, and Norway 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.
6. Select the variables of interest from the **Available Variables** list in the left panel. Supplement 1 to this User Guide provides the variable names for all questions in the context questionnaires. Variables are selected by clicking on them and then clicking the **right arrow** (▶) button. Clicking the **tab-right arrow** (⌘) button selects all variables. 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.

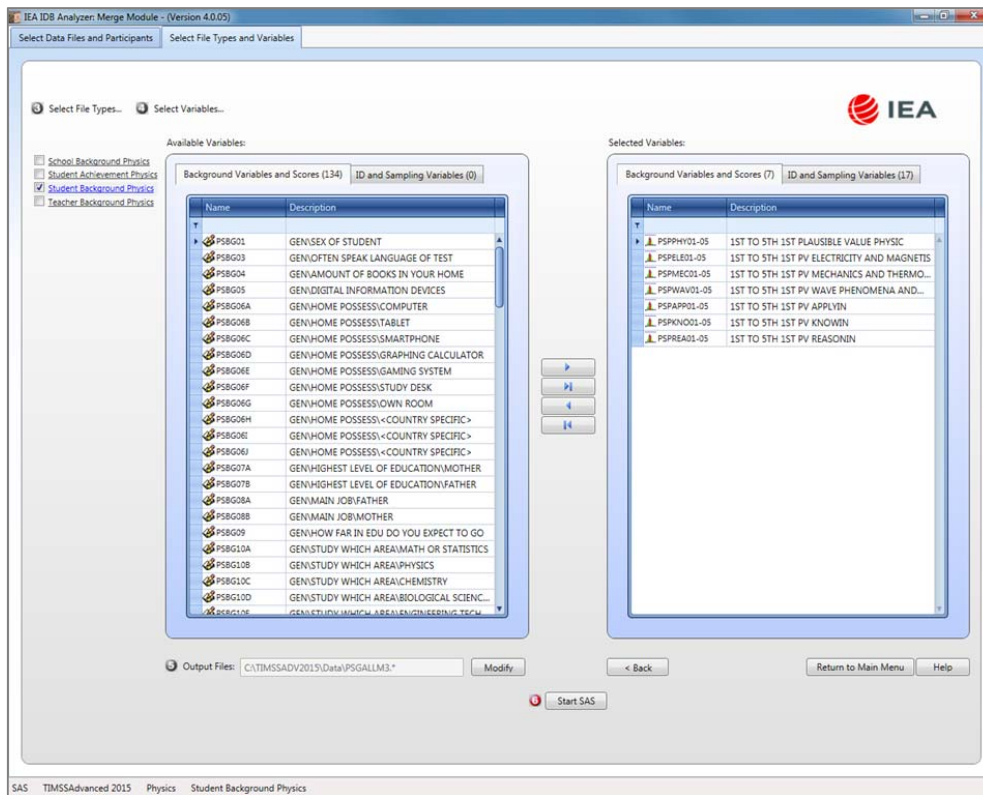
Exhibit 2.2 IEA IDB Analyzer Merge Module: Select Data files and Participants



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 merged data file PSGALLM3.SAS7BDAT and the syntax file PSGALLM3.SAS both will be created and stored in the folder titled “C:\TIMSSADV2015\Data.” In SPSS, it will be the merged data file PSGALLM3.SAV and the syntax file PSGALLM3.SPS. 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.

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.

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



Merging Student and Teacher Data Files

The teachers in the TIMSS Advanced 2015 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 (MST/PST) 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 **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.
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 can then 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 will be analyzed simultaneously.

Merging Student and School Data Files

Because TIMSS Advanced 2015 has 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.

Merged Data Files for the Examples

To conduct the analysis examples presented in this chapter, a number of merged data files were created following the instructions provided above. The following merged data files were created with all available background variables and achievement scores selected:

- PSGALLM3 Merged physics student background data files for all countries
- PTGALLM3 Merged physics teacher background data files for all countries
- PCGALLM3 Merged physics 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 TIMSS Advanced 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 TIMSS Advanced 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.¹ 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 **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.

¹ Starting with TIMSS Advanced 2015, 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 TIMSS Advanced 2015*.

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 may 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.² 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 **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.

² Effect coding and dummy coding of categorical variables are described in the Help manual of the IEA IDB Analyzer.

Achievement Benchmarks

These are the values that will be used as cut points on an achievement scale, selected in the **Plausible Values** section, for computing the percentages of students meeting the specified benchmarks. Multiple cut points can be specified, each separated by a blank space. A drop-down menu is available to select the TIMSS Advanced 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 or school data. MATWGT will be used when analyzing advanced mathematics teacher data and PHYWGT when analyzing physics teacher data. Chapter 4 of this User Guide provides more information on the TIMSS Advanced 2015 sampling weights.

2.5 TIMSS Advanced Analyses with Student Background Data

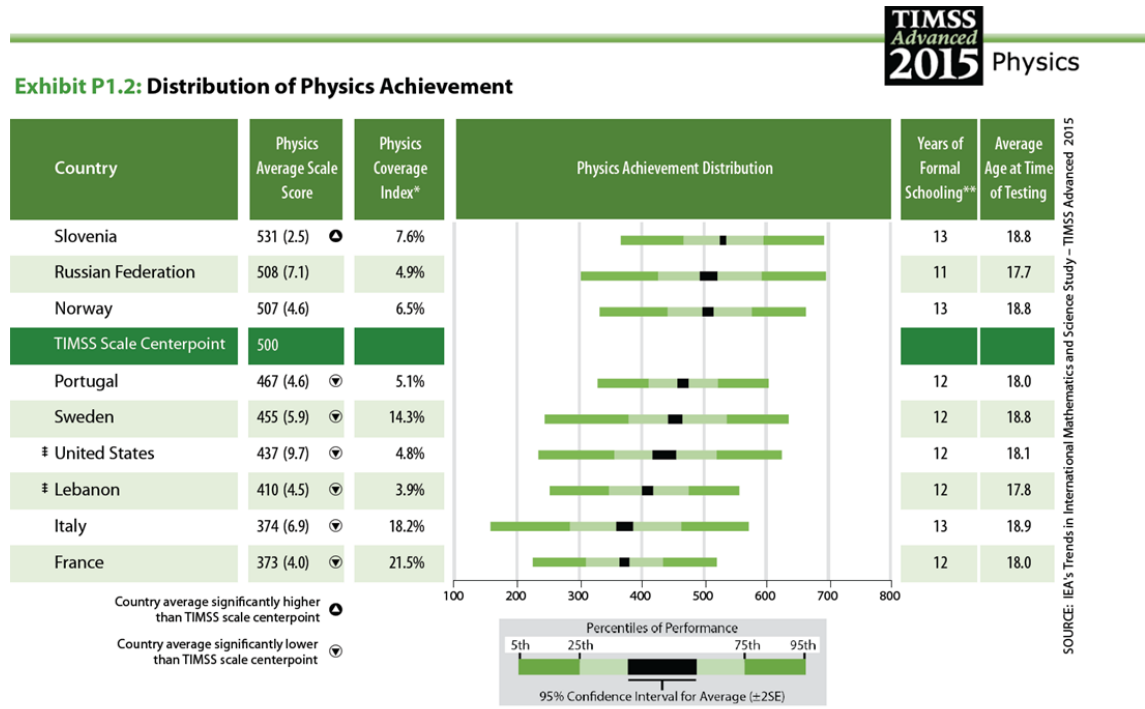
Many analyses of the TIMSS Advanced 2015 International Database can be undertaken using only student background data. This section presents examples of actual analyses used to produce exhibits from *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*. Examples of linear regression analyses and computing percentages of students reaching the TIMSS Advanced International Benchmarks also are included in this section.

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 TIMSS Advanced International Benchmarks. Finally, a fifth example computes the average scale score for a context questionnaire scale, along with the percentage of students, with their average achievement, for the categories of the scale's corresponding index.

Student Background Data Analysis with Achievement

In our first example, we want to replicate the analysis of the overall distribution of physics achievement. These results are presented in Exhibit P1.2 of *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics* 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.

Exhibit 2.4 Exhibit of Example Student Background Data Analysis with Achievement, Taken from TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics (Exhibit P1.2)



* See Appendix PC.2 for a description of the Physics Coverage Index.
 ** Represents years of schooling counting from first year of primary or basic education (first year of ISCED Level 1).
 The TIMSS Advanced achievement scale was established in 1995 based on the combined achievement distribution of all countries that participated in TIMSS Advanced 1995. 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 PC.5 for sampling guidelines and sampling participation notes †, ‡, and £.
 () Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

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 physics achievement, we will create the merged data file PSGALLM3 from the PSG files of all countries that participated in TIMSS Advanced 2015.

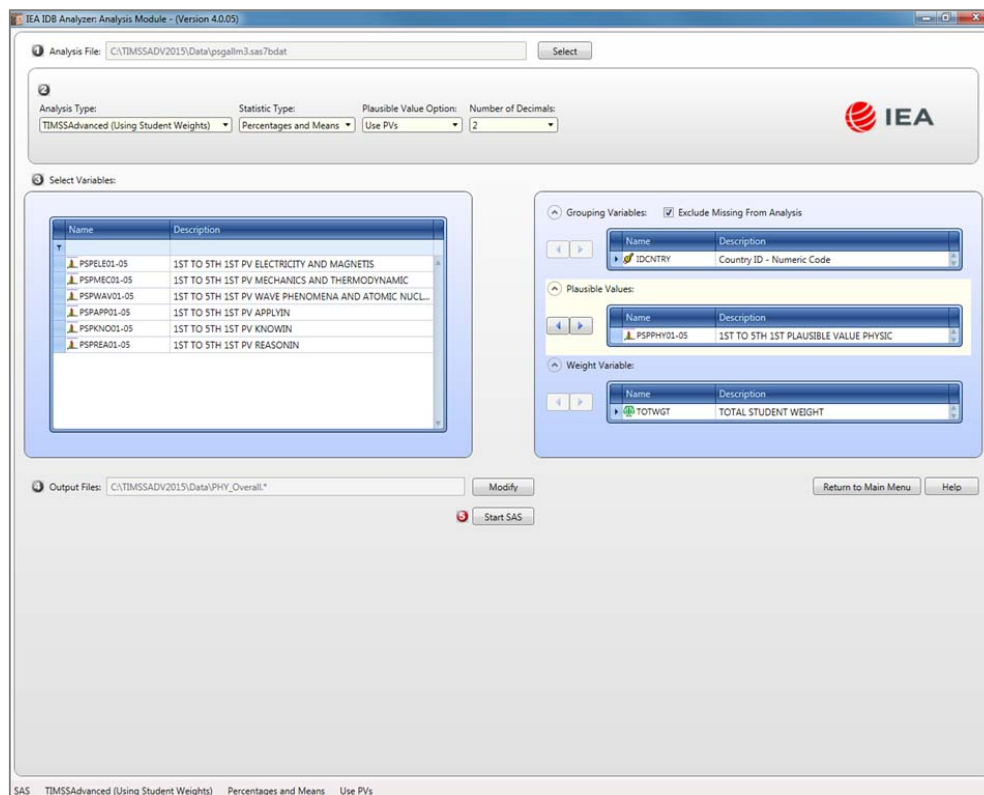
After creating the merged data file PSGALLM3, the analysis module of the IEA IDB Analyzer is used to perform the analysis in the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.
2. Select the merged data file PSGALLM3 as the **Analysis File** by clicking the **Select** button.
3. Select **TIMSS (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 PSPPHY01-05 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 automatically selected 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.
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 2.5 shows the completed analysis module for this example analysis, and Exhibit 2.6 displays the results with our four example countries.

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 category (Country) and its standard error, followed by the estimated average mathematics 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 PSPPHY0 by (IDCNTRY)

Country ID - Numeric Code	n	TOTAL STUDENT WEIGHT	sumw_se	pct	pct_se	mnpv	mnpv_se	sdpv	sdpv_se	pctmiss
France	3958	172177.51	3574.07	29.96	0.89	373.06	3.98	89.68	1.69	0.00
Italy	3424	104649.76	1128.01	18.21	0.55	373.92	6.91	126.36	3.45	0.00
Lebanon	1156	4463.63	130.92	0.78	0.03	410.16	4.50	93.75	2.72	0.00
Norway	2472	4163.14	106.86	0.72	0.02	507.26	4.57	99.58	2.51	0.00

As shown in the first line of Exhibit 2.6, France had valid data for 3,958 physics students and these sampled students represented a population of 172,178 physics students. The average physics achievement in France was 373.06 (standard error of 3.98) and its standard deviation was 89.68 (standard error of 1.69).

Student Background Data Analysis with Achievement by Gender

In our second example, we wish to replicate another set of results presented in *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*. We are interested in investigating the relationship between physics students’ gender and physics achievement. These results, presented in Exhibit P1.6 of *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*, 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.

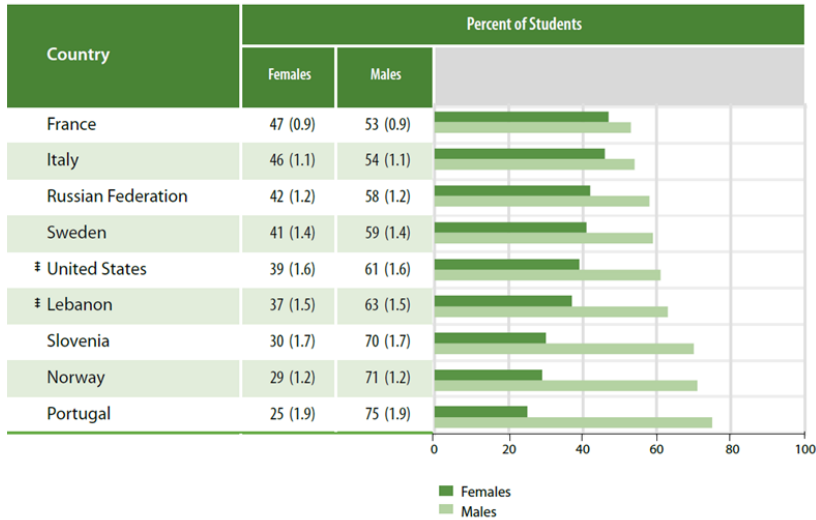
After reviewing the appropriate codebook, we observe that the variable ITSEX contains categorical information on the gender of students, and this variable is found in the student background data files. The **Percentages and Means** statistic type and the **Use PVs** plausible value option will compute 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 the analysis using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.
2. Select the merged data file PSGALLM3 as the **Analysis File** by clicking the **Select** button.
3. Select **TIMSS (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**.

**Exhibit 2.7 Exhibit of Example Student Background Data Analysis with Achievement by Gender, Taken from
TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics (Exhibit P1.6)**

Exhibit P1.6: Physics Participation and Average Achievement by Gender

Participation in Physics by Gender



Average Physics Achievement by Gender



See Appendix PC.5 for sampling guidelines and sampling participation notes †, ‡, and §.

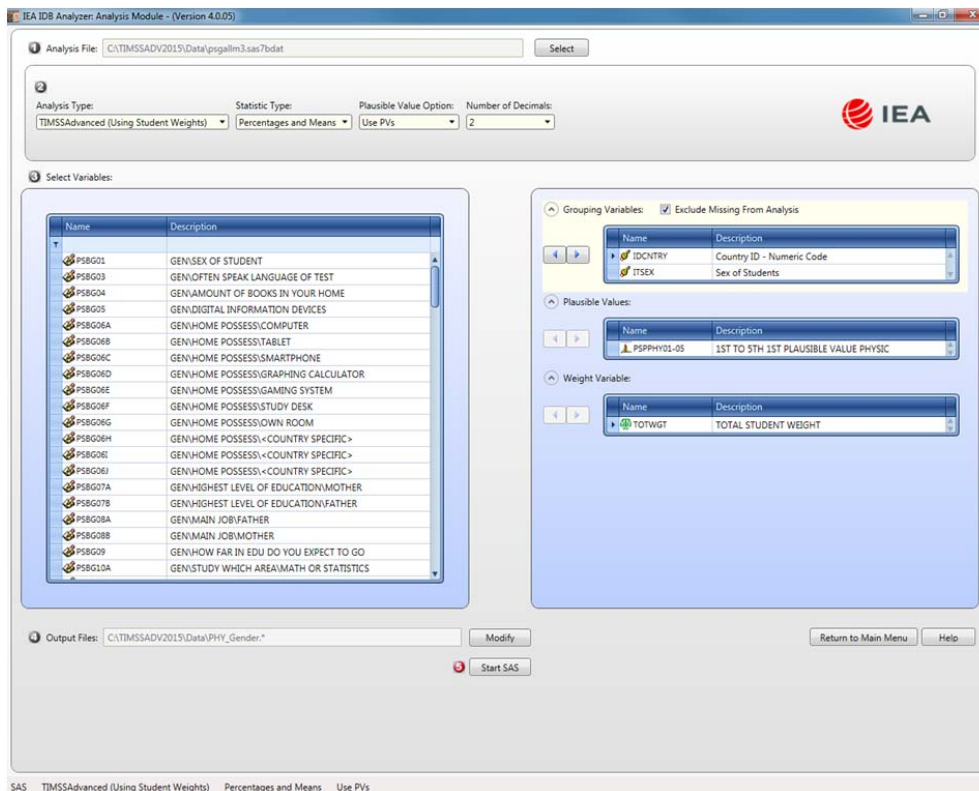
(†) Standard errors appear in parentheses. Because of rounding some results may appear inconsistent.

- 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.
- Specify the achievement scores to be used for the analysis by clicking the **Plausible Values** field to activate it. Select PSPPHY01-05 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 automatically selected 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.
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 mathematics 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 PSPPHY0 by (IDCNTRY ITSEX)

Country ID - Numeric Code	Sex of Students	n	TOTAL STUDENT WEIGHT	sumw_se	pct	pct_se	mnpv	mnpv_se	sdpv	sdpv_se	pctmiss
France	1	1854	81083.61	2401.44	47.09	0.91	354.40	4.16	84.44	2.13	0.00
	2	2104	91093.90	2340.21	52.91	0.91	389.66	4.55	90.93	2.26	0.00
Italy	1	1572	48049.00	1238.70	45.91	1.06	356.41	7.31	120.60	4.76	0.00
	2	1852	56600.76	1247.15	54.09	1.06	388.79	8.40	129.15	4.44	0.00
Lebanon	1	431	1658.69	90.29	37.16	1.48	417.22	5.19	87.47	4.52	0.00
	2	725	2804.94	93.63	62.84	1.48	405.98	6.44	96.95	4.13	0.00
Norway	1	719	1208.77	68.81	29.03	1.23	488.78	6.00	92.18	3.94	0.00
	2	1753	2954.38	74.66	70.97	1.23	514.82	4.82	101.48	2.63	0.00

From the two lines of results for Italy, 45.91% of physics students in Italy were girls and 54.09% were boys. The average physics achievement of girls was 356.41 (standard error of 7.31) and it was 388.79 for boys (standard error of 8.40).

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 physics achievement between girls and boys and determine if it is statistically significant. The results of this example also are presented in Exhibit P1.6 of *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics* and shown in Exhibit 2.7, in the column labeled “Difference.”

It is worth noting that our previous example (Student Background Data Analysis with Achievement by Gender) the IEA IDB Analyzer performed the same gender difference significance test in the background. In the C:\TIMSSADV2015\Data\ folder, there is a CSV file (readable in Excel), PHY_Gender_PSPPHY0_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 physics achievement of girls, and the regression slope will be the estimated increase in average physics 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 PSGALLM3 as the **Analysis File** by clicking the **Select** button.
3. Select **TIMSS (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 IDCNTRY 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 PSPPHY01-05 as the analysis variable. This is done by selecting PSPPHY01-05 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 automatically selected 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.
11. Click the **Start SAS** button to create the SAS syntax file and open it in an 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 of the two lines of results for Italy in Exhibit 2.11 labeled “Intercept” is the estimated average physics achievement of girls in Italy, which was 356.41 with a standard error of 7.31. This estimate concurs with the results obtained in the previous example (Exhibit 2.9). From the second line of results for Italy, labeled “ITSEX_D2”, the physics boys in Italy had a positive average physics achievement difference of 32.38. With an estimated standard error of 7.78, this achievement difference is statistically significant at the 95% confidence level. Adding the two regression coefficients together ($356.41 + 32.38$) yields the estimated average physics achievement of boys in Italy, which was 388.79, the same estimate from the previous example. The IEA IDB Analyzer also computes standardized regression coefficients, whereby the 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

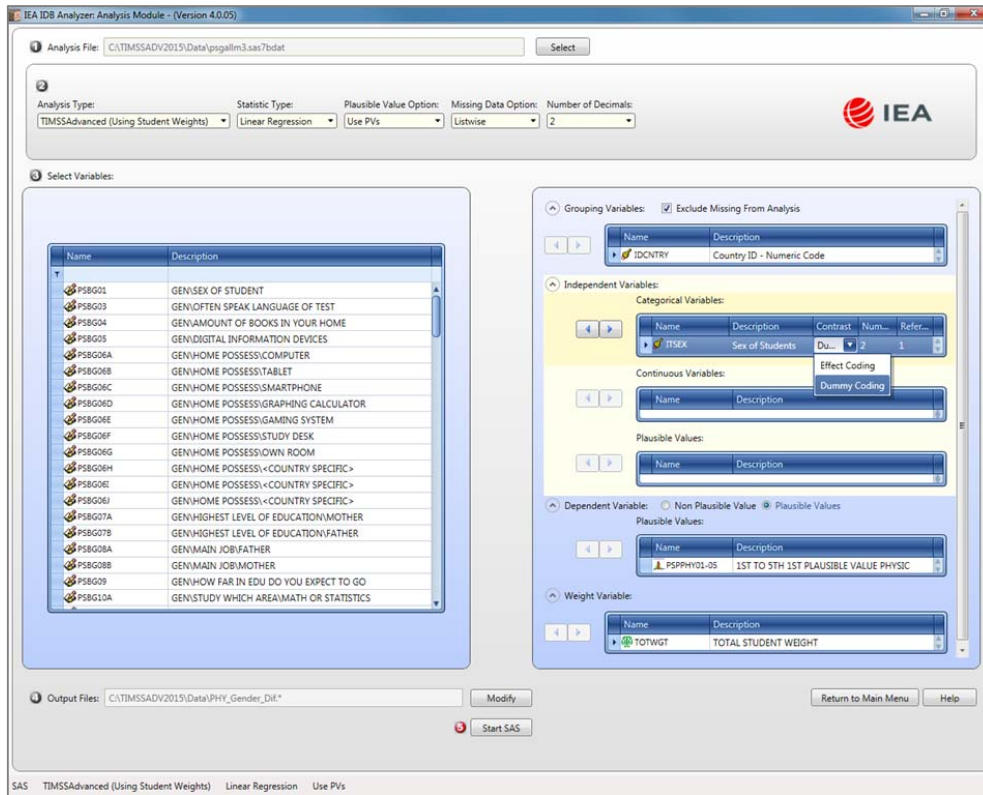


Exhibit 2.11 SAS Output for Example Student Background Data Linear Regression Analysis

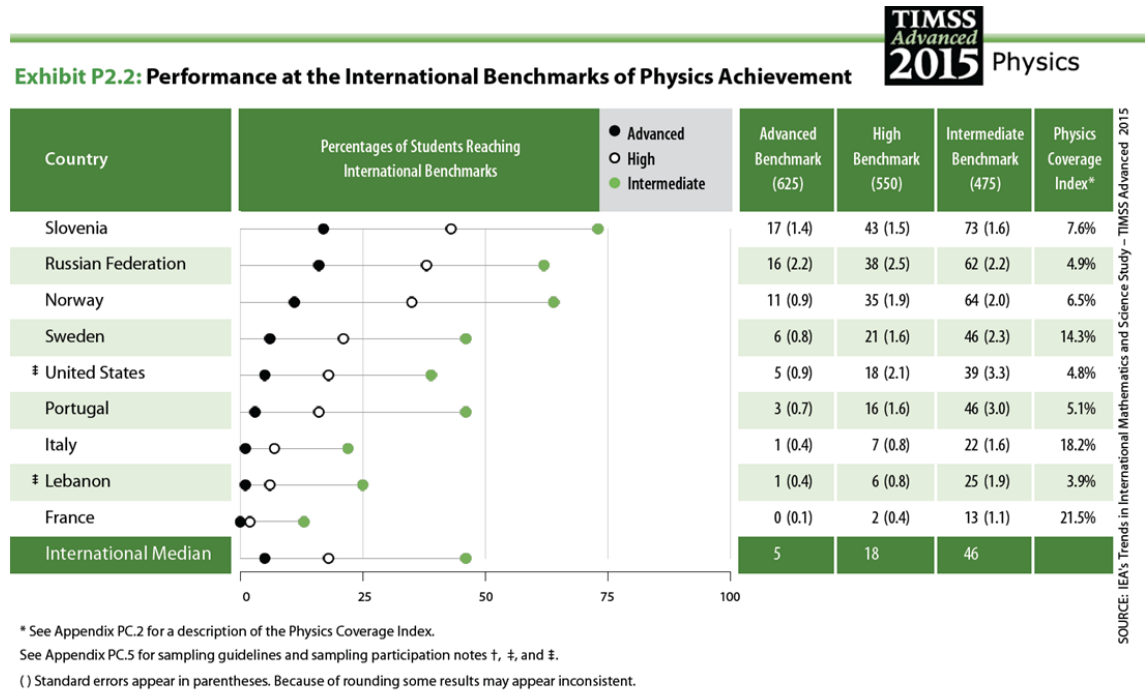
Equation Coefficients of (ITSEX_D2) predicting (PSPPHY0) by (IDCNTRY)

Country ID - Numeric Code	variable	label	Parameter Estimate	Estimate_se	Estimate_t	Standardized Estimate	StandardizedEst_se
France	Intercept	Intercept	354.40	4.16	85.18	0.00	0.00
	ITSEX_D2		35.26	3.77	9.35	0.20	0.02
Italy	Intercept	Intercept	356.41	7.31	48.77	0.00	0.00
	ITSEX_D2		32.38	7.78	4.16	0.13	0.03
Lebanon	Intercept	Intercept	417.22	5.19	80.40	0.00	0.00
	ITSEX_D2		-11.23	8.23	-1.37	-0.06	0.04
Norway	Intercept	Intercept	488.78	6.00	81.44	0.00	0.00
	ITSEX_D2		26.04	5.32	4.90	0.12	0.02

The TIMSS Advanced International Benchmarks

This section describes how to use the IEA IDB Analyzer to perform analyses of student achievement in relation to the TIMSS Advanced International Benchmarks. As an example, we will compute the percentages of students reaching the three TIMSS Advanced 2015 International Benchmarks of physics achievement (advanced, high, and intermediate) using the merged PSGALLM3 data file. These results, presented in Exhibit P2.2 of *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*, are shown here in Exhibit 2.12.

Exhibit 2.12 Example Exhibit of TIMSS Advanced International Benchmarks Analysis, Taken from TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics (Exhibit P2.2)



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 PSGALLM3 as the **Analysis File** by clicking the **Select** button.
3. Select **TIMSS (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 TIMSS international benchmarks.
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. Select PSPPHY01-05 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 TIMSS Advanced 2015 International Benchmarks—475, 550, and 625 (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 automatically selected 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.
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.

Exhibit 2.13 IEA IDB Analyzer Setup for Example TIMSS International Benchmarks Analysis

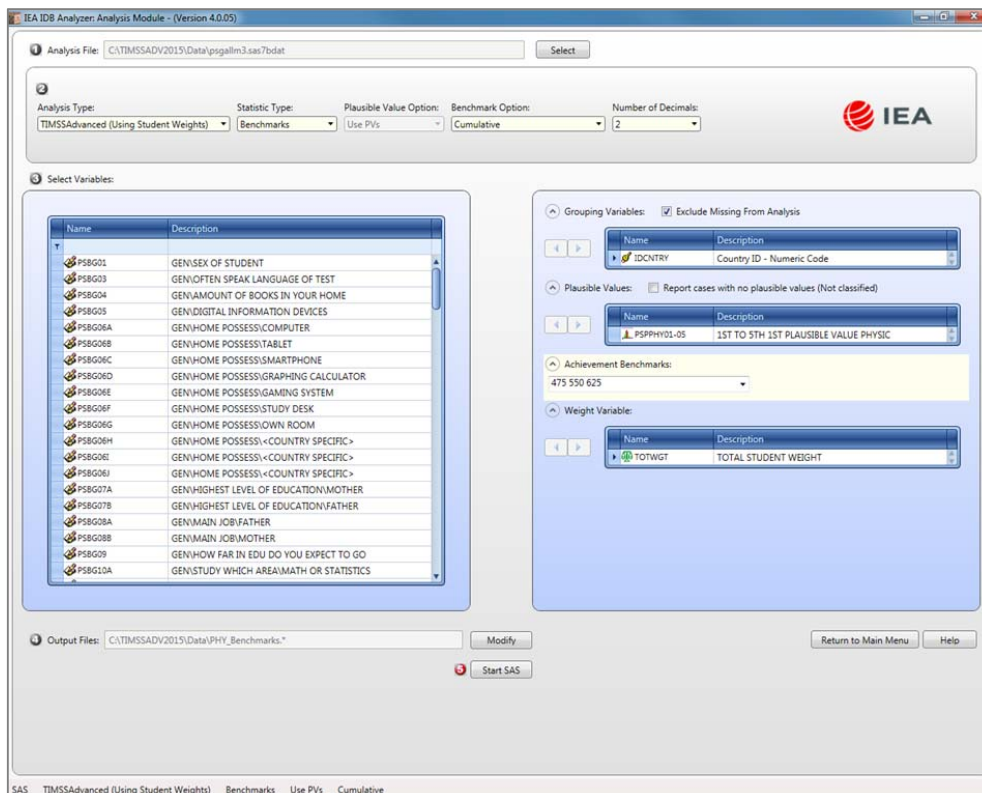


Exhibit 2.14 SAS Output for Example TIMSS Advanced International Benchmarks Analysis

Percent at or Above (475 550 625) of PSPPHY0 by (IDCNTRY)

Country ID - Numeric Code	cutvar	n	TOTAL STUDENT WEIGHT	sumw_se	pct	pct_se
France	At or above 475	541.0	22492.11	1933.19	13.06	1.10
	At or above 550	101.8	4122.84	647.10	2.39	0.37
	At or above 625	5.0	210.48	157.39	0.12	0.09
Italy	At or above 475	791.6	23063.25	1732.40	22.04	1.64
	At or above 550	256.4	7742.59	819.69	7.40	0.78
	At or above 625	47.4	1516.06	387.43	1.45	0.37
Lebanon	At or above 475	279.0	1128.60	95.58	25.28	1.92
	At or above 550	66.6	269.09	38.02	6.03	0.85
	At or above 625	10.8	44.68	16.58	1.00	0.37
Norway	At or above 475	1589.0	2661.69	124.15	63.93	2.04
	At or above 550	871.6	1453.78	98.72	34.92	1.91
	At or above 625	283.8	470.97	43.32	11.31	0.92

As shown in the three lines of results for Lebanon, 25.28% of the physics students in Lebanon performed at or above the intermediate International Benchmark, with a standard error of 1.92; 6.03% of the students reached the High International Benchmark, with a standard error of 0.85; and 1.00% of the students reached the Advanced International Benchmark, with a standard error of 0.37.

Student-level Analysis with a Context Questionnaire Scale

TIMSS Advanced 2015 reports some context questionnaire data by creating context questionnaire scales based on Rasch modeling.³ 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 *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*. Exhibit 2.15 shows one such example, Exhibit P10.2 of *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*, reporting Students Like Learning Physics, based on Students’ responses to a set of 12 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—Very Much Like Learning Physics, Like Learning Physics, and Do Not Like Learning Physics—along with their average physics achievement in each category.

This example will replicate both the average scale score of the Students Like Learning Physics 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 PSGALLM3 data file.

³ The context questionnaire scales are described in Chapter 15 of *Methods and Procedures in TIMSS Advanced 2015*.

**Exhibit 2.15 Example Exhibit of a Context questionnaire Scale, Taken from
TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics (Exhibit P10.2)**

Exhibit P10.2: Students Like Learning Physics

Reported by Physics Students

Students were scored according to their degree of agreement with twelve statements on the *Students Like Learning Physics* scale. Students who **Very Much Like Learning Physics** had a score on the scale of at least 11.4, which corresponds to their “agreeing a lot” with six of the twelve statements and “agreeing a little” with the other six, on average. Students who **Do Not Like Learning Physics** had a score no higher than 8.8, which corresponds to their “disagreeing a little” with six of the twelve statements and “agreeing a little” with the other six, on average. All other students **Like Learning Physics**.

Country	Very Much Like Learning Physics		Like Learning Physics		Do Not Like Learning Physics		Average Scale Score
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Norway	36 (1.2)	560 (3.7)	49 (1.1)	494 (5.4)	15 (1.0)	422 (7.3)	10.7 (0.05)
Lebanon	35 (2.4)	439 (8.7)	52 (1.7)	400 (5.7)	13 (2.1)	392 (10.6)	10.7 (0.12)
Portugal	32 (1.7)	512 (6.0)	51 (1.3)	455 (5.6)	17 (1.1)	416 (5.9)	10.6 (0.08)
Russian Federation	28 (1.6)	568 (9.2)	50 (1.2)	501 (6.2)	22 (1.2)	447 (8.1)	10.3 (0.09)
United States	21 (1.6)	513 (8.7)	48 (1.0)	442 (8.0)	31 (1.9)	380 (11.9)	9.8 (0.10)
Sweden	15 (0.6)	540 (6.8)	46 (1.4)	472 (6.1)	39 (1.4)	403 (6.4)	9.3 (0.05)
Italy	15 (0.9)	467 (10.0)	45 (0.9)	384 (7.4)	40 (1.3)	331 (7.0)	9.4 (0.06)
Slovenia	15 (1.0)	599 (8.5)	63 (1.7)	538 (3.3)	23 (1.6)	472 (6.0)	9.8 (0.05)
France	11 (0.6)	454 (5.4)	54 (0.8)	386 (4.1)	35 (1.0)	329 (4.5)	9.4 (0.04)
International Avg.	23 (0.5)	517 (2.6)	51 (0.4)	452 (2.0)	26 (0.5)	399 (2.6)	

This TIMSS Advanced questionnaire scale was established in 2015 based on the combined response distribution of all countries that participated in TIMSS Advanced 2015. 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.

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

SOURCE: IEA's Trends in International Mathematics and Science Study - TIMSS Advanced 2015

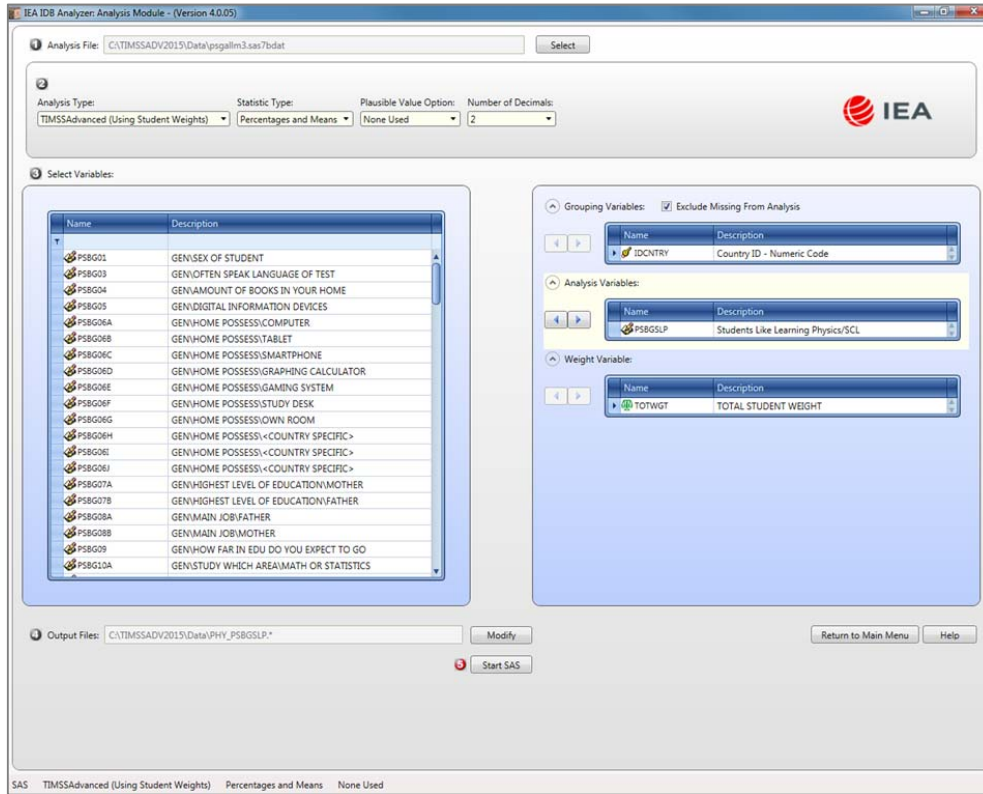
The first step will compute the average scale score using the scale variable PSBGSLP. It is performed by the analysis module of the IEA IDB Analyzer using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.
2. Specify the data file PSGALLM3 as the **Analysis File**.
3. Select **TIMSS (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 PSBGSLP as the analysis variable by clicking the **Analysis Variables** field to activate it. Select PSBGSLP 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 automatically selected by the software; TOTWGT is selected by default because this example analysis uses student background data.
8. Specify the name and the folder of the output files in the **Output Files** field by clicking the **Define/Modify** button.

- 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, students in Norway scored 10.74, with a standard error of 0.05, on the Students Like Learning Physics context questionnaire scale. Note that this is well above 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 PSBGSLP by (IDCNTRY)

Country ID - Numeric Code	n	TOTAL STUDENT WEIGHT	sumw_se	pct	pct_se	mnx	mnx_se	sdx	sdx_se	pctmiss
France	3944	170897.16	3013.41	30.07	0.85	9.41	0.04	1.61	0.03	0.74
Italy	3375	103130.42	1201.90	18.15	0.55	9.37	0.06	1.99	0.04	1.45
Lebanon	1117	4345.68	133.73	0.76	0.03	10.69	0.12	1.93	0.06	2.64
Norway	2450	4126.24	107.26	0.73	0.02	10.74	0.05	1.91	0.04	0.89

In the second step, we will compute the percentages of students—with their average physics achievement—in each category of the corresponding index variable PSDGSLP. 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 PSGALLM3 as the **Analysis File**.
3. Select **TIMSS (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 PSDGSLP.
6. Specify the variable PSDGSLP as a second grouping variable by clicking the **Grouping Variables** field to activate it. Select PSDGSLP 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 PSPPHY01-05 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 automatically selected 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.
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 Norway, 36.40% of physics students in Norway very much like learning physics (standard error of 1.18) and their average physics achievement was 560.31 (standard error of 3.74); 48.90% of students like learning physics (standard error of 1.12) and their average physics achievement was 494.32 (standard error of 5.41); and 14.70% of students do not like learning physics (standard error of 0.99) and their average physics achievement was 422.23 (standard error of 7.34).

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

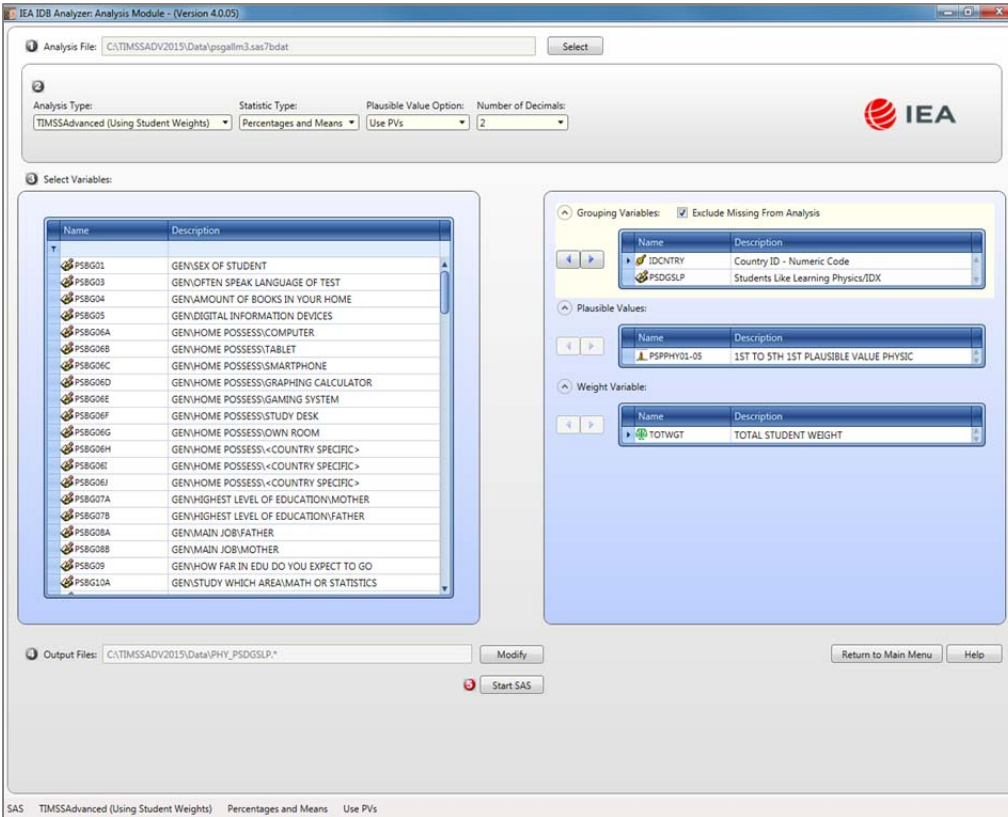


Exhibit 2.19 SAS Output for Example Context Questionnaire Scale Analysis (Step 2)

Average of PSPPHY0 by (IDCNTRY PSDGLP)

Country ID - Numeric Code	Students Like Learning Physics IDX	n	TOTAL STUDENT WEIGHT	sumw_se	pct	pct_se	mnpv	mnpv_se	sdpv	sdpv_se	pctmiss
France	1	441	19129.57	1049.90	11.19	0.58	453.66	5.42	80.85	4.44	0.00
	2	2141	92706.92	2104.92	54.25	0.81	385.61	4.10	82.45	2.07	0.00
	3	1362	59060.66	1952.06	34.56	0.96	328.91	4.47	78.47	2.03	0.00
Italy	1	467	15105.34	996.39	14.65	0.94	466.83	10.00	118.38	6.28	0.00
	2	1546	46744.33	996.70	45.33	0.89	384.44	7.35	116.34	4.16	0.00
	3	1362	41280.74	1473.46	40.03	1.33	330.70	6.97	119.19	4.25	0.00
Lebanon	1	368	1508.88	93.40	34.72	2.42	438.59	8.74	93.47	5.15	0.00
	2	621	2276.27	98.15	52.38	1.71	399.73	5.70	89.41	3.50	0.00
	3	128	560.53	101.05	12.90	2.07	392.10	10.59	87.33	8.92	0.00
Norway	1	897	1502.10	65.37	36.40	1.18	560.31	3.74	84.45	2.62	0.00
	2	1194	2017.55	66.41	48.90	1.12	494.32	5.41	90.76	3.63	0.00
	3	359	606.59	43.41	14.70	0.99	422.23	7.34	87.47	4.35	0.00

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 Learning Physics and Students Value Physics, between achievement scores—such as algebra and calculus, and between a combination of both—such as students’ age and their physics 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 TIMSS Advanced 2015 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, of 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 IDB Analyzer's Help manual.

2.6 TIMSS Advanced 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 physics students according to their physics teachers’ major areas of study. The results of such an analysis are presented in Exhibit P8.3 of *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics* 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 mathematics achievement by reporting categories of teachers’ major areas of study.

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 major areas of study of physics teachers. The variable PTBG05, parts A through K, contains information on teachers’ major areas of study. That information was collapsed into four reporting categories and stored in the

derived variable PTDG05 (see Supplement 3 to this User Guide). The four categories of the PTDG05 variable are described in Exhibit 2.21.

**Exhibit 2.20 Exhibit of Example Teacher Background Data Analysis, Taken from
TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics (Exhibit P8.3)**

**TIMSS
Advanced
2015** Physics

Exhibit P8.3: Physics Teachers Majored in Physics and Education

Reported by Physics Teachers

Country	Major in Physics and Physics Education		Major in Physics but No Major in Physics Education		Major in Physics Education but No Major in Physics		All Other Majors	
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement
France	25 (2.8)	371 (7.2)	71 (2.8)	376 (4.6)	0 (0.0)	~ ~	4 (1.2)	385 (7.5)
Italy	38 (2.9)	369 (12.4)	37 (3.0)	393 (9.2)	1 (0.5)	~ ~	24 (3.3)	379 (14.6)
Lebanon	51 (4.8)	406 (7.2)	48 (4.8)	419 (8.0)	1 (0.4)	~ ~	0 (0.0)	~ ~
Norway	10 (2.9)	513 (11.6)	88 (3.3)	506 (5.1)	1 (0.6)	~ ~	2 (1.0)	~ ~
Portugal	42 (4.5)	470 (7.1)	55 (4.4)	466 (6.4)	0 (0.0)	~ ~	2 (1.2)	~ ~
Russian Federation	64 (3.7)	513 (9.3)	34 (3.9)	498 (9.1)	1 (0.7)	~ ~	1 (0.6)	~ ~
Slovenia	41 (3.1)	531 (4.2)	44 (3.3)	526 (4.8)	14 (3.9)	541 (9.0)	1 (0.0)	~ ~
Sweden	71 (4.3)	463 (7.9)	21 (3.8)	446 (12.9)	6 (1.8)	456 (20.5)	2 (1.0)	~ ~
United States	24 (4.1)	444 (14.7)	31 (4.7)	429 (22.4)	6 (2.1)	457 (18.8)	39 (5.1)	448 (16.3)
International Avg.	41 (1.2)	453 (3.2)	48 (1.3)	451 (3.5)	3 (0.5)	485 (9.7)	8 (0.7)	404 (7.7)

SOURCE: IEA's Trends in International Mathematics and Science Study – TIMSS Advanced 2015

() 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.

Exhibit 2.21 Categories of the PTDG05 Variable

Code	Description
1	Major in Physics and Physics Education
2	Major in Physics but No Major in Physics Education
3	Major in Physics Education but No Major in Physics
4	All Other Majors

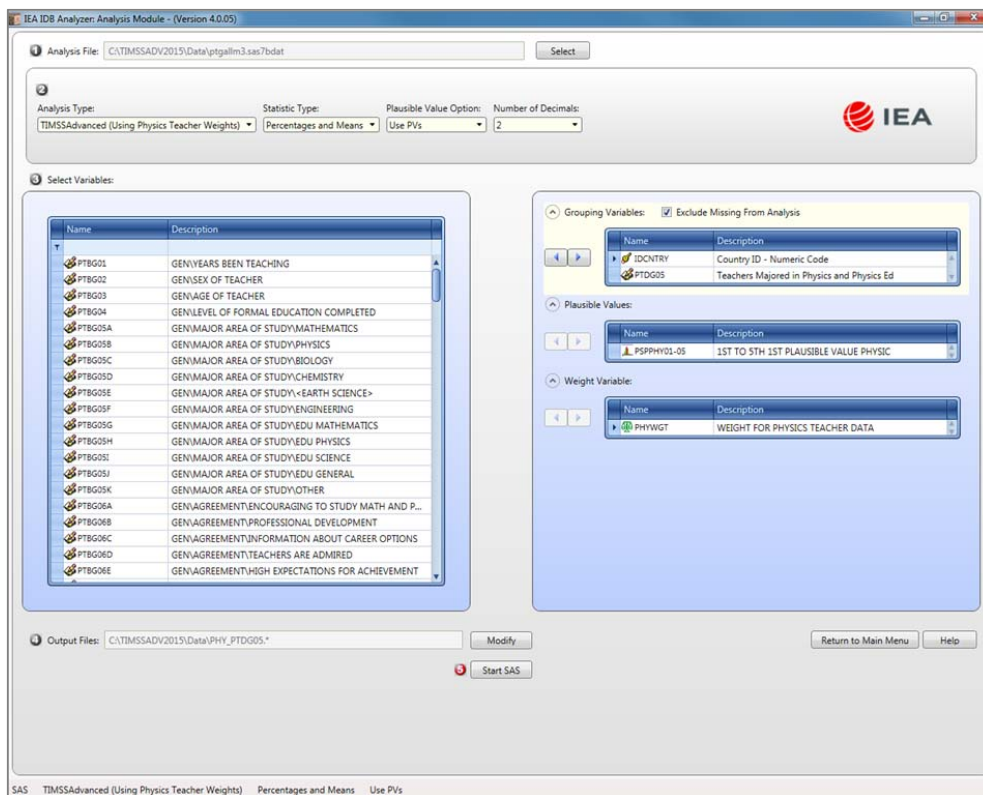
The merged data file PTGALLM3 will be used for this example teacher background data analysis, which will be performed by the analysis module of the IEA IDB Analyzer using the following steps:

1. Open the **Analysis Module** of the IEA IDB Analyzer.
2. Select the merged data file PTGALLM3 as the **Analysis File**.
3. Select **TIMSS (Using Physics Teacher Weights)** as the **Analysis Type** since we want to analyze the responses of the physics teachers.
4. Select **Percentages and Means** as the **Statistic Type**.
5. Select **Use PVs** as the **Plausible Value Option**.

6. Specify the variable PTDG05 as a second grouping variable by clicking the **Grouping Variables** field to activate it. Select PTDG05 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 PSPPHY01-05 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 automatically selected by the software; PHYWGT 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.
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 2.22 shows the completed analysis module for this example analysis, and Exhibit 2.23 displays the results.

Exhibit 2.22 IEA IDB Analyzer Setup for Example Teacher Background Data Analysis



Each country's results are displayed on up to four lines, one for each value of the PTDG05 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 PTDG05.

Exhibit 2.23 SAS Output for Example Teacher Background Data Analysis

Average of PSPPHY0 by (IDCNTRY PTDG05)

Country ID - Numeric Code	Teachers Majored in Physics and Physics Ed	n	WEIGHT FOR PHYSICS TEACHER DATA	sumw_se	pct	pct_se	mnpv	mnpv_se	sdpv	sdpv_se	pctmiss
France	1	1119	39186.42	4762.24	24.93	2.75	370.74	7.17	88.71	3.22	0.00
	2	3279	111680.14	4641.75	71.05	2.81	376.16	4.56	89.81	1.83	0.00
	4	196	6326.44	1971.86	4.02	1.22	385.41	7.53	72.52	4.89	0.00
Italy	1	1187	36258.48	2769.27	38.19	2.88	368.91	12.44	124.01	4.73	0.00
	2	1225	35104.34	3058.64	36.97	3.03	392.82	9.21	118.94	4.68	0.00
	3	12	611.84	432.97	0.64	0.46	441.43	75.01	117.73	34.89	0.00
Lebanon	4	716	22979.73	3206.05	24.20	3.31	378.81	14.60	133.21	7.49	0.00
	1	576	2013.99	175.72	50.84	4.83	406.44	7.15	93.64	3.56	0.00
	2	482	1890.54	236.61	47.73	4.83	418.50	7.98	93.50	4.12	0.00
Norway	3	19	56.60	14.55	1.43	0.37	356.72	23.45	71.54	13.99	0.00
	1	223	384.79	112.06	9.94	2.85	513.31	11.56	97.22	4.28	0.00
	2	2017	3387.81	156.01	87.50	3.32	505.51	5.08	99.34	2.61	0.00
	3	14	24.31	24.31	0.63	0.63	335.37	15.87	89.94	7.90	0.00
	4	45	74.80	39.62	1.93	1.02	487.89	31.41	103.05	14.55	0.00

As shown in the four lines of results for Italy, 38.19% (standard error of 2.88) of physics students were taught by teachers with majors in physics and physics education, 36.97% (standard error of 3.03) were taught by teachers with a major in physics but not in physics education, 0.64% (standard error of 0.46) by teachers with a major in physics education but not in physics, and 24.20% (standard error of 3.31) by teachers with other majors. Also, the estimated average physics achievement was 368.91 (standard error of 12.44) for physics students taught by teachers with majors in physics and physics education, 392.82 (standard error of 9.21) for students taught by teachers with a major in physics but not in physics education, 441.43 (standard error of 75.01) for students taught by teachers with a major in physics education but not in physics, and 378.81 (standard error of 14.60) for students taught by teachers with other majors. The IEA IDB Analyzer also produces the standard deviations of achievement for all categories of PTDG05.

2.7 TIMSS 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 of physics students and their average achievement who attend schools composed of students with different levels of economic background. The results of this analysis are presented in Exhibit P5.1 of *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics* and are replicated here in Exhibit 2.24.

**Exhibit 2.24 Exhibit of Example School Background Data Analysis, Taken from
TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics (Exhibit P5.1)**

Exhibit P5.1: School Composition by Economic Background of the Student Body

Reported by Principals

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
France	50 (4.1)	384 (5.6)	28 (3.9)	376 (6.6)	22 (3.4)	336 (8.0)
Italy	62 (4.4)	378 (8.9)	35 (4.5)	362 (17.7)	2 (1.4)	~ ~
Lebanon	34 (5.1)	442 (11.0)	29 (4.6)	404 (12.1)	38 (3.2)	391 (5.7)
Norway	71 (5.1)	510 (6.1)	27 (5.0)	503 (6.7)	2 (1.2)	~ ~
Portugal	19 (3.4)	502 (7.9)	42 (5.9)	454 (6.9)	39 (5.5)	457 (7.0)
Russian Federation	84 (3.4)	519 (6.9)	13 (2.9)	458 (18.4)	3 (1.2)	439 (46.9)
Slovenia	75 (3.9)	540 (4.0)	18 (3.6)	484 (8.3)	7 (1.5)	562 (15.5)
Sweden	75 (3.1)	468 (7.9)	19 (3.3)	409 (12.0)	6 (2.0)	384 (15.9)
United States	42 (5.9)	473 (15.5)	24 (4.2)	476 (14.8)	34 (4.6)	396 (14.8)
International Avg.	57 (1.5)	468 (2.9)	26 (1.4)	436 (4.1)	17 (1.0)	424 (7.9)

SOURCE: IEA's Trends in International Mathematics and Science Study - TIMSS Advanced 2015

() 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.

We will use the **Percentages and Means** statistic type and select the **Use PVs** option to estimate the percentages of students with their average mathematics 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 PCDG03 (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 PCGALLM3 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 PCGALLM3 as the **Analysis File**.
3. Select **TIMSS (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 PCDG03 as a second grouping variable by clicking the **Grouping Variables** field to activate it. Select PCDG03 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 PSPPHY01-05 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 automatically selected 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.
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 2.25 shows the completed analysis module for this example analysis and the results are presented in Exhibit 2.26.

Exhibit 2.25 IEA IDB Analyzer Set-Up for Example School Background Data Analysis

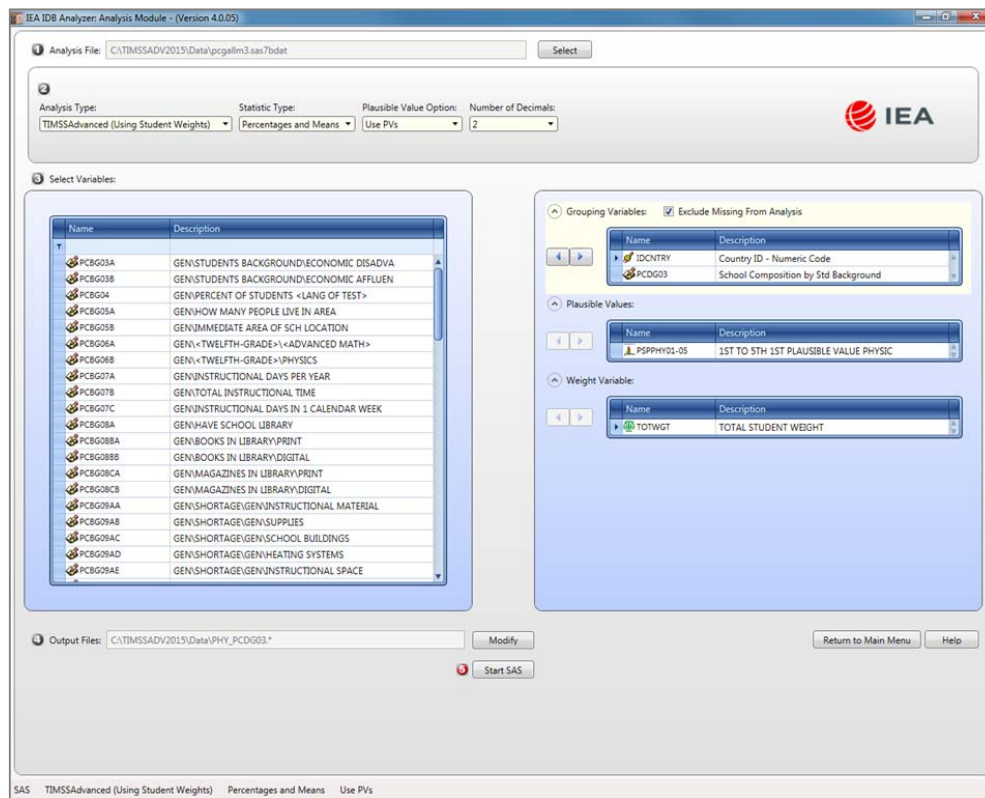


Exhibit 2.26 SAS Output for Example School Background Data Analysis

Average of PSPPHY0 by (IDCNTRY PCDG03)

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
France	1	1876	78414.03	6427.80	49.66	4.11	384.09	5.58	88.17	2.74	0.00
	2	992	44903.77	6349.07	28.44	3.86	375.56	6.61	88.70	2.30	0.00
	3	740	34573.34	5807.95	21.90	3.45	336.28	7.96	84.38	3.69	0.00
Italy	1	2059	58401.91	4332.04	62.44	4.44	377.81	8.86	122.33	3.48	0.00
	2	903	32939.98	4520.25	35.22	4.51	361.82	17.70	135.43	7.70	0.00
	3	49	2184.56	1358.10	2.34	1.44	347.09	33.35	114.65	26.14	0.00
Lebanon	1	298	1389.78	238.45	33.68	5.10	441.79	11.03	90.45	4.60	0.00
	2	295	1180.54	172.27	28.61	4.59	403.99	12.15	94.18	5.85	0.00
	3	465	1556.66	136.63	37.72	3.19	390.75	5.66	87.85	3.33	0.00
Norway	1	1351	2272.77	223.99	70.71	5.09	509.76	6.13	101.26	3.59	0.00
	2	507	876.79	163.41	27.28	4.98	502.90	6.66	95.17	4.21	0.00
	3	38	64.83	37.87	2.02	1.17	449.92	23.54	96.29	10.13	0.00

In this example, each country’s results are presented on three lines, one for each value of the PCDG03 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 PCDG03. As shown in the three lines of results for Lebanon, 33.68% (standard error of 5.10) of physics students attended schools with more affluent than disadvantaged students, 28.61% (standard error of 4.59) attended schools with neither more affluent nor more disadvantaged students, and 37.72% (standard error of 3.19) attended schools with more disadvantaged than affluent students. Also, the estimated average physics achievement was 441.79 (standard error of 11.03) for physics students in schools with more affluent students, 403.99 (standard error of 12.15) for students in schools with neither more affluent nor more disadvantaged students, and 390.75 (standard error of 5.66) for students in schools with more disadvantaged students. The IEA IDB Analyzer also produces the standard deviations of achievement for all categories of PCDG03.

References

IBM Corporation. (2013). *IBM SPSS Statistics* (version 22.0). Somers, NY: IBM Corporation.

International Association for the Evaluation of Educational Achievement. (2017). *IDB Analyzer* (version 4.0). Hamburg, Germany: IEA Hamburg. Available online at <http://www.iea.nl/data.html>.

Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.). (2016). *Methods and Procedures in TIMSS Advanced 2015*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timss.bc.edu/publications/timss/2015-a-methods.html>.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/timss2015/international-results/advanced/>.

Mullis, I.V.S. & Martin, M.O. (Eds.). (2014). *TIMSS Advanced 2015 Assessment Frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

SAS Institute. (2012). *SAS System for Windows* (version 9.4). Cary, NC: SAS Institute.

Chapter 3

Special SPSS and SAS Programs

3.1 Overview

This chapter presents special SPSS (IBM Corporation, 2013) and SAS (SAS Institute, 2012) programs essential to make full use of the TIMSS Advanced 2015 International Database. A SAS program converts the SAS export files into SAS data files, needed as input to the IEA IDB Analyzer (IEA, 2017). Users who intend to run the IEA IDB Analyzer with SAS should execute this program.

There is also a set of SPSS and SAS programs to recode the responses to the individual TIMSS Advanced 2015 items from the achievement data files to their appropriate score levels. Users who intend to run the IEA IDB Analyzer with SAS should execute the SAS version of these programs; users who intend to run the IEA IDB Analyzer with SPSS should execute the SPSS version of these programs.

3.2 Converting the SAS Export Files

The TIMSS Advanced 2015 International Database provides a SAS program—TA15_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 (version 4.0) requires the use of SAS data files; it will not function with the SAS export files. The TA15_CONVERT.SAS program will convert all data file types and will work for both TIMSS Advanced subjects. To convert SAS export files into SAS data files, users should perform the following steps:

1. Open TA15_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 TA15_CONVERT program lists all advanced mathematics and physics data 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, all TIMSS Advanced 2015 countries are listed and the program will automatically select the appropriate list by subject based on the file type specified.
6. Submit the edited code for processing by SAS.

The TA15_CONVERT.SAS program is presented in Exhibit 3.1. This example converts the SAS export files of all advanced mathematics and physics data file types for all countries. For this example, all SAS export files are located in the folder “C:\TIMSSADV2015\Data\SAS_Data,” where the converted SAS data files also will be located.

Exhibit 3.1 The TA15_CONVERT SAS Program

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

. . .

%LET COUNTRY = <List of TIMSS Advanced 2015 countries> ;

. . .

PROC CIMPORT FILE="&INDIR.\&FTYPE&CTRY.M3.EXP"
             DATA=OUTDIR.&FTYPE&CTRY.M3 ;

RUN ;

. . .

%MEND DOIT ;

%DOIT (TYPE = MCG MSA MSG MSR MST MTG
      PCG PSA PSG PSR PST PTG ,
      INDIR = C:\TIMSSADV2015\Data\SAS_Data ,
      OUTDIR = C:\TIMSSADV2015\Data\SAS_Data ) ;
```

Users are advised to run the TA15_CONVERT program for all countries and all file types. The file types for advanced mathematics are MCG, MSA, MSG, MSR, MST, and MTG. For physics, the file types are PCG, PSA, PSG, PSR, PST, and PTG. 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 users do with the TIMSS Advanced 2015 International Database before undertaking any data analyses with SAS and the IEA IDB Analyzer, in particular the data analysis examples in this User Guide.

3.3 Scoring the TIMSS Advanced 2015 Items

Student achievement in TIMSS Advanced 2015 is represented by sets of five plausible values for advanced mathematics and physics and their content and cognitive domains. 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 TIMSS Advanced 2015 International Database be assigned their correctness score levels, rather than the actual response options selected by students for multiple choice items, or the two-digit codes given to students’ responses on constructed response items. The International Database provides SPSS and SAS programs to perform this task.

For multiple choice items, the numbers 1 through 5 are used to represent response options A through E, respectively, in the TIMSS Advanced 2015 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. For constructed response items, worth either a total of one or two points, two-digit codes are used in the achievement data files to represent the students’

written responses. These codes also must be recoded to represent the correct point values of the responses—either zero, one, or two 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 may be recoded in order to carry out specific item-level analyses. By default, the “Not Administered” response code is left as missing and the “Omitted” and “Not Reached” response codes are recoded as incorrect. 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 TIMSS Advanced 2015 items, whereas they were treated as incorrect when deriving achievement scores for students.

The TIMSS Advanced 2015 International Database includes two SPSS programs—MSASCRM3.SPS for the advanced mathematics items and PSASCRM3.SPS for the physics items—and two SAS programs—MSASCRM3.SAS and PSASCRM3.SAS—which will recode the responses to individual items from the achievement data files (MSA/PSA) 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 (MSA/PSA) 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. Exhibit 3.2 shows a condensed version of the SAS program to score the individual TIMSS Advanced 2015 items.

Exhibit 3.2 The MSASCRM3/PSASCRM3 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 ERIGHT = <List of multiple choice items where E is correct> ;
DO OVER ERIGHT ; %SCOREIT (ERIGHT, "MC", 5, .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:\TIMSSADV2015\Data ,
      OUTDIR = C:\TIMSSADV2015\Data ,
      INFILE = PSAALLM3 ) ;
```

To score each individual TIMSS Advanced 2015 item, the program code in the SAS program must be adapted by completing the following steps:

1. Open MSASCRM3.SAS or PSASCRM3.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.

In this example, the merged SAS data file of student achievement data files is called PSAALLM3, located in the folder C:\TIMSSADV2015\Data, and contains the physics items. The resulting merged data file of scored achievement items will be called PSAALLM3_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 (MSA/PSA), 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 conduct analyses with this merged data file.

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 MSASCRM3 program for the advanced mathematics items and the PSASCRM3 program for the physics 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 TIMSS Advanced 2015 items.

Exhibit 3.3 The MSASCRM3/PSASCRM3 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 = MC / ITEM = <List of multiple choice items where E is correct> /
        RIGHT = 5 / 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:\TIMSSADV2015\Data /
      OUTDIR = C:\TIMSSADV2015\Data /
      INFILE = PSAALLM3 .

```

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 */
```

References

IBM Corporation. (2013). *IBM SPSS Statistics* (version 22.0). Somers, NY: IBM Corporation.

International Association for the Evaluation of Educational Achievement. (2017). *IDB Analyzer* (version 4.0). Hamburg, Germany: IEA Hamburg. Available online at <http://www.iea.nl/data.html>.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/timss2015/international-results/advanced/>.

SAS Institute. (2012). *SAS System for Windows* (version 9.4). Cary, NC: SAS Institute.



Chapter 4

The TIMSS Advanced 2015 International Database Files

4.1 Overview

The TIMSS Advanced 2015 International Database contains student achievement data and student, teacher, and school background data collected in the nine countries that took part in TIMSS Advanced 2015. Exhibit 4.1 lists all of the TIMSS Advanced 2015 countries included in the International Database, along with the identification codes used in the International Database. The 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 TIMSS Advanced 2015 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 TIMSS Advanced 2015 International Database

Countries	Identification Codes		Advanced Mathematics	Physics
	Alpha	Numeric		
France	FRA	250	●	●
Italy	ITA	380	●	●
Lebanon	LBN	422	●	●
Norway	NOR	578	●	●
Portugal	PRT	620	●	●
Russian Federation	RUS	643	●	●
Russian Federation 6hr+	RTR	643002	●	
Slovenia	SVN	705	●	●
Sweden	SWE	752	●	●
United States	USA	840	●	●

In TIMSS Advanced 2015, the Russian Federation defined a broader advanced mathematics target population, as described in Exhibit M1.1 of *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics* (Mullis, Martin, Foy, & Hooper, 2016). In order to maintain the ability to report trends back to 1995 and 2008, however, a subset of the Russian Federation target population, where students receive at least 6 hours of advanced mathematics instruction per week,

was identified and set aside in its own set of data files labeled “Russian Federation 6hr+” in Exhibit 4.1, and with the RTR identification code. This sub-sample of students is comparable to the Russian Federation’s target population of past TIMSS Advanced assessments. The full 2015 target population of the Russian Federation is found in the data files with the RUS identification code, which also includes the data from the RTR data files.

4.2 The Supplements to the TIMSS Advanced 2015 User Guide

The TIMSS Advanced 2015 User Guide is an important resource to the TIMSS Advanced 2015 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 TIMSS Advanced 2015 Context Questionnaires— Supplement 1 includes the international version of all context questionnaires administered in TIMSS Advanced 2015, including the Curriculum Questionnaires. 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 TIMSS Advanced 2015 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, Teacher, and School Questionnaire Data— Supplement 3 describes how the derived background variables were computed to produce exhibits in *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*.

4.3 The Restricted Use Version of the TIMSS Advanced 2015 International Database

There are two versions of the TIMSS Advanced 2015 International Database. The public use version is available for immediate access from the TIMSS Advanced 2015 International Database webpage (<http://timssandpirls.bc.edu/timss2015/advanced-international-database/>) 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 TIMSS Advanced 2015 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 through its Study Data Repository (<http://www.iea.nl/data.html>) to obtain permission and access to the restricted use version of the TIMSS Advanced 2015 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 TIMSS Advanced 2015 Variables Removed from Public Use Version of International Database

Variables	Description
ITBIRTHY / ITBIRTHM	Students' year and month of birth from the tracking forms
MSBG02A / MSBG02B PSBG02A / PSBG02B	Students' year and month of birth from student questionnaire
ITDATE	TIMSS Advanced 2015 testing date
MCBG01 / PCBG01	Total school enrollment
MCBG02 / PCBG02	School enrollment in the target grade
INOSS	Data collected through Online Survey System

4.4 TIMSS Advanced 2015 Achievement Items Documentation

A number of documents related to the TIMSS Advanced 2015 achievement items are available for download as part of the TIMSS Advanced 2015 International Database. They include summary information on the TIMSS Advanced 2015 items, the IRT item parameters estimated from the TIMSS Advanced 2015 item calibrations and their associated linear transformation constants, and percent correct statistics for the TIMSS Advanced 2015 items.

Achievement item information files are provided to enable users of the TIMSS Advanced 2015 International Database to readily produce summaries of item characteristics. There are separate achievement item information files for advanced mathematics and physics. These files are available in Excel format and include the following information for each item in the TIMSS Advanced 2015 assessments:

- The item's unique identifier;
- The item's block and its sequential location within the block;
- The item's label;
- The item's content domain and cognitive domain;
- 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 2015 assessment.

The information in the columns labeled “Topic Area” and “Topic,” is described in *TIMSS Advanced 2015 Assessment Frameworks* (Mullis & Martin, 2014).

The International Database includes Excel and PDF files with the IRT item parameters estimated for all TIMSS Advanced 2008 and TIMSS Advanced 2015 items from the concurrent item calibrations. There are separate files for each subject. These same item parameters are presented in Chapter 13 of *Methods and Procedures in TIMSS Advanced 2015* (Martin, Mullis, & Hooper, 2016). The TIMSS Advanced 2015 International Database also includes files with the advanced mathematics and physics linear transformation constants that were used to set the TIMSS Advanced 2015 achievement scores on the TIMSS Advanced trend scales.

The International Database also includes percent correct statistics for all TIMSS Advanced 2015 items in Excel and PDF format. There are separate documents for advanced mathematics and physics.

4.5 The TIMSS Advanced 2015 Data Files

The TIMSS Advanced 2015 International Database includes the actual data from all instruments administered to the students, their teachers, and their school principals. This includes the student responses to the achievement items and the responses to the student, 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 *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*. National Research Coordinators’ responses to the curriculum questionnaires also are part of the International Database and are described later in this chapter.

This section describes the format and contents of the TIMSS Advanced 2015 data files. With the exception of the curriculum data files, the TIMSS Advanced data files are provided in SPSS (IBM Corporation, 2013) format (.SAV) and SAS (SAS Institute, 2012) export format (.EXP). Data files are provided for each country that participated in TIMSS Advanced 2015 and for which internationally comparable data are available. The file names given to the various data file types in TIMSS Advanced 2015 are shown in Exhibit 4.3. For example, PSGPRTM3.SAV is an SPSS file that contains Portugal’s TIMSS Advanced 2015 physics 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 TIMSS Advanced 2015 Data File Names

Assessment	File Name	Description
Advanced Mathematics	MCG●●●M3	Advanced mathematics school background data files
	MSA●●●M3	Advanced mathematics student achievement data files
	MSG●●●M3	Advanced mathematics student background data files
	MSR●●●M3	Advanced mathematics within-country scoring reliability data files
	MST●●●M3	Advanced mathematics student-teacher linkage files
	MTG●●●M3	Advanced mathematics teacher background data files
Physics	PCG●●●M3	Physics school background data files
	PSA●●●M3	Physics student achievement data files
	PSG●●●M3	Physics student background data files
	PSR●●●M3	Physics within-country scoring reliability data files
	PST●●●M3	Physics student-teacher linkage files
	PTG●●●M3	physics 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 TIMSS Advanced 2015 Student Achievement Data Files (MSA/PSA)

The TIMSS Advanced 2015 student achievement data files contain the student responses to the individual achievement items in the TIMSS Advanced 2015 assessments. The student achievement data files are best suited for performing item-level analyses. Achievement scores (plausible values) for all of the TIMSS Advanced 2015 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).

Students who participated in TIMSS Advanced 2015 were administered one of 12 assessment booklets, six of which consisted of advanced mathematics items and the other six of physics items.¹ 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 codes assigned to the constructed response items through the TIMSS Advanced 2015 scoring guides.²

¹ The TIMSS Advanced 2015 assessment design is described in Chapter 4 of *TIMSS Advanced 2015 Assessment Frameworks* (Mullis & Martin, 2014).

² The TIMSS Advanced 2015 scoring guides are described in *TIMSS 2015 Item Writing Guidelines* (Mullis & Martin, 2013).

Item Variable Naming Convention

The achievement item variable names are based on a 9-character alphanumeric code, which observes the following rules:

- The first and second characters are either “MA” for advanced mathematics items, or “PA” for physics items.
- The third character indicates the assessment cycle when the item was first presented in TIMSS Advanced. The items in the TIMSS Advanced 2015 assessment have either “1” for items introduced in 1995, “2” for items introduced in 2008, or “3” for new items in 2015.
- The fourth character is always “3”, reserved to identify TIMSS Advanced items.
- The fifth through seventh characters represent a unique three-digit number used to identify the items.
- The eighth character indicates the item part, and appears only when required. It is generally a letter from “A” to “E,” depending how many parts there are to a particular item.
- A ninth character is present when an item part itself consists of further parts, usually used to define a derived item.

For example, MA33225A is the first part of an advanced mathematics item introduced in 2015 and whose unique sequence number is 225.

Item Response Code Values

A series of conventions also were adopted to code the data included in the TIMSS Advanced 2015 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 5 were used to correspond to the response options A through E, 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 TIMSS Advanced 2015 International Database to derive correctness response codes for these items (see Chapter 3).

Each constructed response item had its own scoring guide that relied on a two-digit coding scheme to provide diagnostic information. The first digit designated the correctness level of the response: 2 for a response worth 2 points, 1 for a response worth 1 point, and 7 for an incorrect response. The second digit, combined with the first, represented a diagnostic code used to identify specific types of approaches, strategies, or common errors and misconceptions in responding to the item. A second digit of 0 through 5 was used for pre-defined international codes at each correctness level, while a second digit of 9 corresponded to “other” types of responses which fell within the appropriate correctness level but which did not fit any of the pre-defined international codes. A special two-digit code, 99, was used for responses left completely blank. Chapter 3 describes SPSS and SAS programs that convert the two-digit response codes into correctness response codes.

Derived Items Code Values

For some items, students were asked to provide more than one answer, each one being scored appropriately. The pattern of responses across these item parts determined the score on the item as a whole. These multi-part items can be worth either 1 point or 2 points, according to their scoring rubric. For two-point derived items, code 20 was used to award 2 points, code 10 was used to award 1 point, and code 70 for no points. One-point derived items were awarded either code 10 for 1 point, or code 70 for no points. The total score for the item is contained in a derived variable, usually identified by word "DERIVED" in the item label.

For example, derived variable MA33225 contains the combined score for its five parts labeled MA33225A through MA33225E. This derived item was awarded 2 points (code 20) if all five parts were answered correctly, 1 point (code 10) if four parts were answered correctly, and no points (code 70) otherwise. All derived items in the TIMSS Advanced 2015 assessments are described in Appendix D of Chapter 11 of *Methods and Procedures in TIMSS Advanced 2015*.

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, or unintentionally when an item was misprinted or otherwise unavailable for a student to respond. The not administered code was used in the following cases:

- Achievement item not assigned to the student—All students participating in TIMSS Advanced 2015 received one of 12 available test booklets. All variables corresponding to items that were not present in a student’s assigned booklet 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, 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 question 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, 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. For most purposes, TIMSS Advanced 2015 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 13 of *Methods and Procedures in TIMSS Advanced 2015*).

TIMSS Advanced Achievement Scores

Achievement scales were produced for advanced mathematics and physics and their content and cognitive domains, as shown in Exhibit 4.4. A total of 7 achievement scales were produced for advanced mathematics and 7 for physics. A detailed description of the TIMSS Advanced 2015 scaling approach and how these achievement scales were created is available in Chapter 13 of *Methods and Procedures in TIMSS Advanced 2015*. For each achievement scale, the TIMSS Advanced 2015 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 TIMSS Advanced 2015 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, 2017), as described in Chapter 2 of this User Guide.

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

- The first character is either "M" for an advanced mathematics score, or "P" for a physics score;
- The second character is always "S" to indicate it is a student score variable;
- The third character is either "M" for an advanced mathematics score, or "P" for a physics score, whether it is an overall score, a content domain score, or cognitive domain score;
- The fourth through sixth characters are a three-character code describing the achievement scale, as shown in Exhibit 4.4; and
- The seventh and eighth characters are a two-digit number indicating the plausible value—01, 02, 03, 04, or 05.

For example, MSMCAL01 is the first plausible value on the advanced mathematics calculus content domain achievement scale.

Exhibit 4.4 TIMSS Advanced 2015 Achievement Scales

Overall	MAT	Advanced Mathematics	PHY	Science
	ALG	Algebra	MEC	Mechanics and Thermodynamics
Content Domains	CAL	Calculus	ELE	Electricity and Magnetism
	GEO	Geometry	WAV	Wave Phenomena and Atomic/Nuclear Physics
Cognitive Domains	KNO	Knowing	KNO	Knowing
	APP	Applying	APP	Applying
	REA	Reasoning	REA	Reasoning

TIMSS Advanced International Benchmarks of Achievement

To help users of the TIMSS Advanced 2015 International Database understand what performance on the overall mathematics and physics achievement scales signifies in terms of the advanced mathematics and physics students know and can do, TIMSS Advanced identified three points on the overall advanced mathematics and physics scales to serve as International Benchmarks. As shown in Exhibit 4.5, the TIMSS Advanced International Benchmark scores are 625, 550, and 475, which correspond to the Advanced International Benchmark, the High International Benchmark, and the Intermediate International Benchmark, respectively. TIMSS Advanced used a technique known as scale anchoring to summarize and describe student achievement at these three points on the scale.³ The *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics* present the results of this scale anchoring, and report the percentage of students in each country reaching each of the TIMSS Advanced International Benchmarks.

The TIMSS Advanced 2015 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 overall advanced mathematics and overall physics scales. The International Benchmark variables follow the achievement score variable naming convention where the fourth through sixth positions have the letters “IBM.” Thus, MSMIBM01-05 are the five benchmark variables for overall advanced mathematics and PSPIBM01-05 for overall physics. The codes defined for all the benchmark variables are described in Exhibit 4.6.

Exhibit 4.5 TIMSS Advanced 2015 International Benchmarks of Advanced Mathematics and Physics Achievement

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

³ The scale anchoring procedure is described in Chapter 14 of *Methods and Procedures in TIMSS Advanced 2015*.

Exhibit 4.6 TIMSS Advanced 2015 International Benchmark Variable Codes

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

4.5.2 TIMSS Advanced 2015 Within-Country Scoring Reliability Data Files (MSR/PSR)

The TIMSS Advanced 2015 within-country scoring reliability data files contain data that can be used to investigate the reliability of the TIMSS Advanced constructed response item scoring. The scoring reliability data files contain one record for each assessment booklet that was double scored during the within-country scoring reliability exercise (see Chapter 11 of *Methods and Procedures in TIMSS Advanced 2015*). For each constructed response item in the assessment, the following three variables are included in the scoring reliability data files:

- Original Score (two-digit score assigned by the first scorer);
- Second Score (two-digit score assigned by the second scorer); and
- Score Agreement (degree of 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 *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*.

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 second character in the variable name differentiates between the three reliability variables:

- The Original Score variable has the letter “A” as the second character, as in the achievement item naming convention (e.g., MA33121);
- The Second Score variable has the letter “R” as the second character (e.g., MR33121); and
- The Score Agreement variable has the letter “I” as the second character (e.g., MI33121).

Scoring Reliability Variable Score Values

The values contained in both the Original Score and Second Score variables are the two-digit diagnostic codes assigned using the TIMSS Advanced 2015 scoring guides. The Score Agreement variable may have one of three values, depending on the degree of agreement between the two scorers, as described in Exhibit 4.7.

Exhibit 4.7 TIMSS Advanced 2015 Score Agreement Variable Codes

Code	Description
0	Identical codes (both digits in the original and second scores)
1	Identical score levels, but different diagnostic codes (first digit of both scores are the same; second digits are different)
2	Different score levels (first digit of both scores are different)

4.5.3 TIMSS Advanced 2015 Context Questionnaire Data Files

There are five usual types of TIMSS Advanced 2015 context questionnaire data files: three (student, teacher, and school) data files correspond to the three types of context questionnaires administered in TIMSS Advanced 2015, the fourth corresponds to the curriculum questionnaire administered to the TIMSS Advanced National Research Coordinators, and the fifth serves to link the student and teacher background data. The first three types of data files contain the responses to the questions asked in their respective context questionnaires.

TIMSS Advanced 2015 Student Background Data Files (MSG/PSG)

Students who participated in TIMSS Advanced 2015 were administered a context questionnaire with questions related to their home background, school experiences, and attitudes towards advanced mathematics and physics. The student background data files contain students’ responses to these questions. They also contain students’ advanced mathematics and physics achievement scores (plausible values) to facilitate analyses of relationships between student background characteristics and achievement.

There were two versions of the student questionnaire: one for advanced mathematics students and one for physics students. Students completed the student questionnaire that was appropriate for the assessment booklet they received—either advanced mathematics or physics. The responses of students to the advanced mathematics student questionnaire are found in the MSG files and the responses of students to the physics student questionnaire are in the PSG files.

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 *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*. These variables are described later in this chapter.

TIMSS Advanced 2015 School Background Data Files (MCG/PCG)

The school background data files contain school principals’ responses to the questions in the TIMSS Advanced 2015 school context questionnaire. 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.

There was a single TIMSS Advanced 2015 school background questionnaire for all schools, whether they were sampled for advanced mathematics, physics, or both. Consequently, the TIMSS Advanced school files for advanced mathematics (MCG) and physics (PCG) have identical structures. The variable names differ, however; the variables with the school principals' responses start with the letter "M" in the MCG files, whereas they begin with the letter "P" in the PCG files.

TIMSS Advanced 2015 Teacher Background Data Files (MTG/PTG)

The advanced mathematics and physics teachers of the students that were sampled for TIMSS Advanced 2015 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 classes taught either by an advanced mathematics or a physics teacher. If a teacher 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 classes were taught by the teacher, and the background information in part A from the completed questionnaire was entered into these teacher records.

There were two types of teacher questionnaires—one for the advanced mathematics teachers of students given an advanced mathematics assessment booklet and one for the physics teachers of the students given a physics assessment booklet. The responses of teachers to the advanced mathematics teacher questionnaire are found in the MTG files and the responses of teachers to the physics teacher questionnaire are in the PTG files.

In the teacher background data files, each teacher has a unique identification number (IDTEACH) and a link number (IDLINK) that is specific to the 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. So, for example, students linked to teachers identified by the same IDTEACH but different IDLINK are taught by the same teacher but in different classes. Consequently, 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.

TIMSS Advanced 2015 Student-Teacher Linkage Data Files (MST/PST)

The TIMSS Advanced 2015 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 sampling and weighting information and achievement scores.

TIMSS Advanced 2015 Curriculum Data Files

The TIMSS Advanced 2015 curriculum questionnaire data files contain the responses provided by the TIMSS Advanced National Research Coordinators of the participating countries to the TIMSS Advanced 2015 curriculum questionnaires. There are two separate curriculum questionnaire data files for the two subjects, advanced mathematics and physics. These files are available as Excel files in the TIMSS Advanced 2015 International Database.

Background Variable Naming Convention

The background variable naming convention for the variables in the student, 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 either “M” for advanced mathematics data, or “P” for physics 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 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 to 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.⁴
- The fourth character is used to indicate the subject or topic to which a background question refers. The following letters are used:

G—General questions (not subject specific);

M—Questions related to advanced mathematics; and

P—Questions related to physics.

- The fifth through eighth characters of all background variables represent the sequential numbering of the questions as presented in their respective questionnaires.

The curriculum data files follow their own variable naming convention whereby the first three characters of a variable name are either “MAA” for the advanced mathematics curriculum variables, or “PHA” for the physics curriculum variables. The remaining characters in the curriculum data

⁴ The context questionnaire scales are described in Chapter 15 of *Methods and Procedures in TIMSS Advanced 2015*.

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 was assigned a unique location code as shown in Exhibit 4.8. 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-05, it refers to general question 5 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 4.8 TIMSS Advanced 2015 Context Questionnaire Variable Location Convention

Questionnaire	Location Variable	
Student Questionnaire	SQG-●●●	General questions
	SQM-●●●	Advanced mathematics questions
	SQP-●●●	Physics questions
Teacher Questionnaire	TQG-●●●	General questions
	TQM-●●●	Advanced mathematics questions
	TQS-●●●	Physics questions
School Questionnaire	SCQ-●●●	Both subjects
Curriculum Questionnaire	CQMA-●●●	Advanced mathematics questions
	CQPA-●●●	Physics questions

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

Data Coding Conventions

A series of conventions also were 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, etcetera. 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, 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, 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 lead 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 TIMSS Advanced 2015 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 TIMSS Advanced 2015 International Database and described in Chapter 15 of *Methods and Procedures in TIMSS Advanced 2015*.

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 were combined into a single variable to report a single educational level. Supplement 3 to the User Guide provides a description of the derived variables included in the TIMSS Advanced 2015 International Database.

Sampling and Weighting Variables

Several sampling and weighting variables are included in the TIMSS Advanced 2015 data files; they are listed and described in Exhibit 4.9. Exhibit 4.10 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.

Exhibit 4.9 TIMSS Advanced 2015 Sampling and Weighting Variables

Variable Names	Descriptions
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
MATWGT	Advanced mathematics teacher weight
PHYWGT	Physics 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, TIMSS Advanced 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.

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

Exhibit 4.10 Location of Sampling and Weighting Variables in the TIMSS Advanced 2015 Data Files

Sampling and Weighting Variables	Data File Types			
	MSG PSG	MSA PSA	MST PST	MCG PCG
JKREP	●	●	●	
JKZONE	●	●	●	
JKCREP				●
JKCZONE				●
TOTWGT	●	●		
SEWGT	●	●		
HOUWGT	●	●		
MATWGT			●	
PHYWGT			●	
SCHWGT				●
WGTFAC1	●			●
WGTADJ1	●			●
WGTFAC2	●			
WGTADJ2	●			
WGTFAC3	●			
WGTADJ3	●			

The weight variables MATWGT and PHYWGT are specifically designed for using teacher background data in student-level analyses and are based on TOTWGT. The weight variables MATWGT and PHYWGT are used for analyses of advanced mathematics and physics teachers, respectively. These teacher weights are located in the student-teacher linkage files (MST and PST), not in the actual teacher background data files (MTG and PTG). Analyses with teacher data will be weighted properly by merging the teacher files with the student-teacher linkage files.

The sampling variables beginning with the letters “JK” are used to compute standard errors based on the jackknife repeated replication methodology.⁵ All weighting variables beginning with the letters “WGT” provide insight into the multi-stage sampling and weighting methodology applied to the TIMSS Advanced data. All weighting variables are described in Chapter 3 of *Methods and Procedures in TIMSS Advanced 2015*.

Structure and Design Variables

Besides the variables used to store responses to the context questionnaires and achievement booklets, the TIMSS Advanced 2015 data files also contain variables meant to store information that identify and describe the respondents and design information required to properly analyze the data.

Identification Variables

In all TIMSS Advanced 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 TIMSS Advanced and are not meant to represent actual school identifiers in the participating countries. They are not unique across countries. Schools across countries can be identified uniquely only with the IDCNTY 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

⁵ Starting with TIMSS Advanced 2015, 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 TIMSS Advanced 2015*.

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 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 TIMSS Advanced 2015, the usual value is “12” for most countries.

IDBOOK

IDBOOK identifies the specific TIMSS Advanced 2015 assessment booklet that was administered to each student. The booklets are given a numerical value from “1” through “6” for the advanced mathematics booklets, and from “7” through “12” for the physics booklets.

Exhibit 4.11 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 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 4.11 Location of Identification Variables in the TIMSS Advanced 2015 Data Files

Identification Variables	Data File Types				
	MSA PSA	MSG PSG	MST PST	MTG PTG	MCG PCG
IDCNTRY	●	●	●	●	●
IDSCHOOL	●	●	●	●	●
IDCLASS	●	●	●	●	
IDSTUD	●	●	●		
IDTEACH			●	●	
IDLINK			●	●	
IDGRADE	●	●	●	●	●
IDBOOK	●	●			

Tracking Variables

Information about students, teachers, and schools provided by the survey tracking forms is stored in the tracking variables.⁶ 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 and student background data files.

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 TIMSS Advanced 2015 International Database.

ITDATE

Testing date for each student. This variable is available only in the restricted use version of the TIMSS Advanced 2015 International Database.

ITLANG

Language of testing for each student. The valid codes for ITLANG are specified in the codebook files and in Supplement 2 to this User Guide.

⁶ Survey tracking forms are lists of students, teachers, and schools used for sampling and administrative purposes.

4.6 TIMSS Advanced 2015 Codebook Files

All information related to the structure of the TIMSS Advanced 2015 data files, as well as the source, format, descriptive labels, and response option codes for all variables, is contained in codebook files. There are two codebook files in Excel format:

TA15_MAT_Codebook TIMSS Advanced 2015 codebook for the advanced mathematics data files

TA15_PHY_Codebook TIMSS Advanced 2015 codebook for the physics data files

In the codebook Excel files, there is a tab for each appropriate data file type in the TIMSS Advanced 2015 International Database. These tabs describe the contents and structure of the TIMSS Advanced 2015 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 TIMSS Advanced 2015 Data Almanac Files

Data almanacs provide weighted summary statistics for all variables in the TIMSS Advanced 2015 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. 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 TIMSS Advanced 2015 assessments. There are separate achievement data almanacs for advanced mathematics and physics.

The achievement data almanacs display for each item its classification in the content and cognitive domains, the item block 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 trend item almanacs provide summary statistics for achievement items used in both the 2008 and 2015 assessments. The achievement data almanac files available in the International Database are listed in Exhibit 4.12.

Exhibit 4.12 TIMSS Advanced 2015 Achievement Data Almanacs

Achievement Data Almanacs	Contents
TA15_MAT_ItemAlmanac	Almanac for advanced mathematics items
TA15_MAT_TrendItemAlmanac	Trend almanac for advanced mathematics items
TA15_PHY_ItemAlmanac	Almanac for physics items
TA15_PHY_TrendItemAlmanac	Trend almanac for physics items

The achievement data almanacs also display the international averages for each item, with each country weighted equally. In the advanced mathematics achievement data almanacs, although there is an explicit row of results for “Russian Federation 6hr+,” this row is not included in the international average since these results are also included in the “Russian Federation” row.

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—Percent of students that responded correctly to a multiple choice item.
- A, B, C, D, and E—The percent of students choosing each one of the response options for a multiple choice item.
- Scoring Guide Codes (e.g., 20, 21, 10, 11, 70, 71)—The percent of student responses assigned each of the codes in the scoring guide for a constructed response item.
- OMITTED—The percent of students that omitted to respond to the item.
- NOT REACHED—The percent of students that did not reach the item.
- V1 and V2—The percent of students that scored 1 point or better on the item (V1) or 2 points (V2) for constructed-response items.
- GIRL PCT RIGHT and BOY PCT RIGHT—The percent 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, teacher, and school context questionnaires, including the context questionnaire scales and their indices and the derived variables based on these background variables. Among the statistics reported are mean advanced mathematics and physics 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 International Database are listed in Exhibit 4.13.

The background data almanacs also display the international averages for each variable, with each country weighted equally. In the advanced mathematics achievement data almanacs, although there is an explicit row of results for “Russian Federation 6hr+,” this row is not included in the international average since these results are also included in the “Russian Federation” row.

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, teachers, or schools included in the sample);
- The number of valid cases (number of students, 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.

Exhibit 4.13 TIMSS Advanced 2015 Background Data Almanacs

Background Data Almanacs	Contents
TA15_MAT_StudentAlmanac	Advanced mathematics student background almanac with achievement
TA15_MAT_TeacherAlmanac	Advanced mathematics teacher background almanac with achievement
TA15_MAT_SchoolAlmanac	Advanced mathematics school background almanac with achievement
TA15_PHY_StudentAlmanac	Physics student background almanac with achievement
TA15_PHY_TeacherAlmanac	Physics teacher background almanac with achievement
TA15_PHY_SchoolAlmanac	Physics school background almanac with achievement

The almanac display for continuous variables includes the following:

- The sample size (number of students, teachers, or schools included in the sample);
- The number of valid cases (number of students, 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 5th, 10th, 25th, 50th, 75th, 90th, and 95th 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 TIMSS Advanced 2015 Test-Curriculum Matching Analysis Data Files

The Test-Curriculum Matching Analysis (TCMA) was conducted to investigate the appropriateness of the TIMSS Advanced 2015 advanced mathematics and physics assessments in the participating countries. To that end, participating countries were asked to indicate which items on the TIMSS Advanced 2015 assessments were included in their national curricula. Thus, based on computations of average percent correct, each country was able to see the performance of all countries on the items appropriate for its curriculum, and also the performance of its students on the items judged appropriate for the curricula in other countries. The analytical method used and the results of the TCMA are presented in Appendices ME and PE of *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*.

The International Database contains two TCMA data files—one for each subject—in Excel format and printable PDF format, showing which items were selected by each participating country.

4.9 TIMSS Advanced 2015 Program Files

The TIMSS Advanced 2015 International Database includes a number of SPSS and SAS programs designed to facilitate the manipulation of the TIMSS Advanced 2015 data files. The International Database contains a SAS program that converts the SAS export files into SAS data files. It includes SPSS and SAS syntax files which will recode the responses to individual items from the achievement data files to their appropriate score levels. All programs are described in Chapter 3 of this User Guide.

References

IBM Corporation. (2012). *IBM SPSS Statistics* (version 22.0). Somers, NY: IBM Corporation.

International Association for the Evaluation of Educational Achievement. (2017). *IDB Analyzer* (version 4.0). Hamburg, Germany: IEA Hamburg. Available online at <http://www.iea.nl/data.html>.

Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.). (2016). *Methods and Procedures in TIMSS Advanced 2015*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timss.bc.edu/publications/timss/2015-a-methods.html>.

Mullis, I.V.S. & Martin, M.O. (2013). *TIMSS 2015 Item Writing Guidelines*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: http://timss.bc.edu/publications/timss/2015-methods/pdf/T15_item_writing_guidelines.pdf.

Mullis, I.V.S. & Martin, M.O. (Eds.). (2014). *TIMSS Advanced 2015 Assessment Frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2016). *TIMSS Advanced 2015 International Results in Advanced Mathematics and Physics*. Retrieved from Boston College, TIMSS & PIRLS International Study Center website: <http://timssandpirls.bc.edu/timss2015/international-results/advanced/>.

SAS Institute. (2013). *SAS System for Windows* (version 9.4). Cary, NC: SAS Institute.

Appendix

Organizations and Individuals Responsible for TIMSS Advanced 2015

Introduction

TIMSS Advanced 2015 was a collaborative effort involving hundreds of individuals around the world. This appendix acknowledges the individuals and organizations for their contributions. Given that work on TIMSS Advanced 2015 has spanned approximately four years and has involved so many people and organizations, this list may not include all who contributed. Any omission is inadvertent. TIMSS Advanced 2015 also acknowledges the students, teachers, and school principals who contributed their time and effort to the study. This report would not be possible without them.

Management and Coordination

TIMSS Advanced was conducted by IEA's TIMSS & PIRLS International Study Center at Boston College, which has responsibility for the 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 the IEA Secretariat in 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 the IEA Data Processing and Research Center in 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; 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 Secretariat and IEA Data Processing and Research Center, 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 TIMSS National Research Coordinator (NRC). The NRCs have the challenging task of implementing TIMSS in their countries in accordance with the TIMSS guidelines and procedures. In addition, the NRCs provide feedback and contributions throughout the development of the TIMSS assessment. The quality of the TIMSS 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 TIMSS, the TIMSS Advanced 2015 NRCs performed their many tasks with dedication, competence, energy, and goodwill, and have been commended by the IEA Secretariat, the TIMSS & PIRLS International Study Center, the IEA Data Processing and Research Center, and Statistics Canada for their commitment to the project and the high quality of their work.

Funding

Funding for TIMSS Advanced 2015 was provided primarily by the participating countries. 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*

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

Marcie Bligh, *Manager, Events and Administration*

Victoria A.S. Centurino, *Assistant Research Director, TIMSS Science*

Kerry Cotter, *Research Specialist, TIMSS Mathematics*

Susan Farrell, *Lead Web and Database Designer*

Bethany Fishbein, *Research Specialist, TIMSS Science*

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*

Jenny Liu, *Graduate Assistant*

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

Yenileis Pardini, *Lead Designer/Developer for eAssessments*

Mario A. Pita, *Lead Graphic Designer*

Jyothsnadevi Pothana, *Statistician/Programmer*

Betty Poulos, *Administrative Coordinator*

Katherine Reynolds, *Graduate Assistant*

Ruthanne Ryan, *Senior Graphic Designer*

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

Amy Semerjian, *Graduate Assistant (through 2015)*

Steven A. Simpson, *Senior Graphic Designer*

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

Liquan Yin, *Research Psychometrician*

IEA Secretariat

Dirk Hastedt, *Executive Director*

Hans Wagemaker, *Executive Director (through 2014)*

Paulína Koršňáková, *Director of the IEA Secretariat*

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

Gabriela Nausica Noveanu, *Senior Research Advisor*
 David Ebbs, *Research Officer*
 Michelle Djekić, *Research Officer*
 Roel Burgers, *Financial Manager*
 Juriaan Hartenberg, *Financial Manager (through 2013)*
 Isabelle Braun-Gémin, *Financial Officer*
 Dana Vizkova, *Financial Officer*
 Gillian Wilson, *Publications Officer*
 Manuel Butty, *Public Outreach Officer*

IEA Data Processing and Research Center

Heiko Sibberns, *IEA DPC Director*
 Oliver Neuschmidt, *Senior Research Analyst, Unit Head, International Studies*
 Milena Taneva, *Senior Research Analyst, Project Co-Manager, TIMSS and TIMSS Advanced Data Processing*
 Juliane Hencke, *Senior Research Analyst, Project Co-Manager, TIMSS and TIMSS Advanced Data Processing*
 Sebastian Meyer, *Research Analyst, Deputy Project Manager, TIMSS and TIMSS Advanced Data Processing*
 Mark Cockle, *Research Analyst, Deputy Project Manager, TIMSS and TIMSS Advanced 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, *Unit Head, Sampling & Research and Analyses Unit*
 Ekaterina Mikheeva, *Research Analyst*
 Dirk Oehler, *Research Analyst*
 Duygu Savaşci, *Research Analyst*
 Sabine Tieck, *Research Analyst*
 Meng Xue, *Unit Head, Software*

Statistics Canada

Jean Dumais, *Chief*
Sylvie LaRoche, *Senior Methodologist*
Craig Seko, *Senior Methodologist*

Educational Testing Service

Matthias Von Davier, *Senior Research Director*
Edward Kulick, *Research Director*
Jonathan Weeks, *Associate Research Scientist*
Zhan Shu, *Psychometrician*
Scott Davis, *Senior Data Analysis and Computational Research Specialist*
Mei-Jang Lin, *Data Analysis and Computational Research Specialist*

Sampling Referee

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

TIMSS Advanced 2015 Science and Mathematics Item Review Committee

Mathematics

Kiril Bankov
Faculty of Mathematics and Informatics
University of Sofia

Bulgaria

Sean Close
Educational Research Centre
St. Patrick's College

Ireland

Khattab M. A. Abulibdeh
National Center for Human Resources
Development

Jordan

Sun Sook Noh
Department of Mathematics Education
Ewha Women's University

Korea, Republic of

Liv Sissel Grønmo
Chief Mathematics Consultant
Department of Teacher Education and
School Research
ILS, University of Oslo

Norway

Torgeir Onstad
Department of Teacher Education and
School Research
ILS, University of Oslo

Norway

Mary Lindquist

United States

Science

Newman Burdett (through 2014)
National Foundation for Educational
Research (NFER)

England

Jouni Viiri
University of Jyväskylä

Finland

Siu Ling Alice Wong
Faculty of Education
University of Hong Kong

Honk Kong SAR

Berenice Michels
Faculty of Science
Utrecht University

The Netherlands

Vitaly Gribov
Physics Faculty
Moscow Lomonosov State University

Russian Federation

Galina Kovaleva
Center for Evaluating the Quality of
Education
Federal Institute of the Strategy of
Education Development of the Russian
Academy of Education
Ministry of Education and Science

Russian Federation

Gabriela Noveanu (through 2013)
Institute of Educational Science
Curriculum Department

Romania

Gorazd Planinsic
Faculty of Mathematics & Physics
University of Ljubljana

Slovenia

Wolfgang Dietrich
National Agency for Education

Sweden

Lee Jones
Chief Science Consultant

United States

Christopher Lazzaro
The College Board

United States

Gerry Wheeler

United States

TIMSS Advanced 2015 Item Development Task Forces

Advanced Mathematics

Ina V.S. Mullis, *TIMSS & PIRLS International Study Center Executive Director*
Kerry Cotter, *Research Specialist, TIMSS Mathematics*
Liv Sissel Grønmo, *Chief Mathematics Consultant (ILS, University of Oslo)*
Mary Lindquist, *Mathematics Consultant (United States)*
Torgeir Onstad, *Mathematics Consultant (ILS, University of Oslo)*
Ray Philpot, *Mathematics Consultant (ACER)*

Physics

Victoria A.S. Centurino, *Assistant Research Director, TIMSS Science*

Lee R. Jones, *Chief Science Consultant (United States)*

Ron Martin, *Science Consultant (ACER)*

Gerry Wheeler, *Science Consultant (United States)*

Questionnaire Item Review Committee

Sue Thomson

Australian Council for Educational
Research

Australia

Josef Basl

Czech School Inspectorate

Czech Republic

Wilfried Bos

Institute for School Development
Research (IFS)

TU Dortmund University

Germany

Martina Meelissen

Department of Research Methodology,
Measurement, and Data Analysis

Faculty of Behavioral Sciences

University of Twente

Netherlands

Chew Leng Poon

Planning Division, Research and
Evaluation Section

Ministry of Education

Singapore

Peter Nyström

National Center for Mathematics
Education

University of Gothenburg

Sweden

Jack Buckley

The College Board

United States

TIMSS Advanced 2015 National Research Coordinators

France

Franck Salles

Ministère de l'éducation nationale

Direction de l'évaluation, de la prospective
et de la performance (DEPP)

Italy

Laura Palmerio

Elisa Caponera (through 2014)

Istituto Nazionale per la Valutazione del
Sistema Educativo di Istruzione e di
Formazione (INVALSI)

Lebanon

Nada Oweijane

Leila Maliha Fayad (through 2015)

Educational Center for Research and
Development

Ministry of Education

Norway

Liv Sissel Grønmo

Department of Teacher Education and
School Research

ILS, University of Oslo

Portugal

João Maroco
Instituto de Avaliação Educativa, I. P.

Russian Federation

Galina Kovaleva
Center for Evaluating the Quality of
Education
Federal Institute of the Strategy of
Education Development of the Russian
Academy of Education
Ministry of Education and Science

Slovenia

Barbara Japelj Pavesic
Educational Research Institute

Sweden

Carola Borg
Swedish National Agency for Education

United States

Stephen Provasnik
National Center for Education Statistics
U.S. Department of Education





© IEA, 2017
International Association
for the Evaluation of
Educational Achievement



BOSTON
COLLEGE
timss.bc.edu