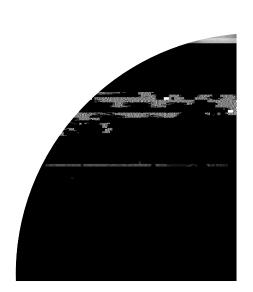


Edited by Er Ingl. I





Contributors

Jay Campbell Pierre Foy Eugenio J. Gonzalez Dirk Hastedt Ursula Itzlinger Marc Joncas Dana L. Kelly Barbara Malak Michael O. Martin Ina V.S. Mullis Marian Sainsbury Knut Schwippert

Edited by Eugenio J. Gonzalez Ann M. Kennedy



2003109101

₹ 1-889938-31-9

+1-617-552-1600 +1-617-552-1203

, // :// · · · · · · · · ·

111/1 (1 (1 - 1 (11 / 11 - - 1 + 1 - 1))

Table of Contents

1 Overview of User Guide

1.1	Overview of User Guide	1-2
1.2	Overview of PIRLS	1-3
1.3	Participating Countries	1-4
1.4	Student Population Assessed	1-4
1.5	Assessment Dates	1-5
1.6	Study Management and Organization	1-5
1.7	Overview of Assessment Framework	1-6
	Exhibit 1.1 Percentages of Reading Assessment Devoted to Reading Purposes and Processes	1-7
1.8	PIRLS Reading Assessment	1-7
	Assessment Design	1-7
	Passages	1-8
	Items and Scoring Guides	1-8
	Releasing Assessment Material to the Public	1-8
1.9	PIRLS Background Questionnaires	1-8
	Student Questionnaire	1-9
	Learning to Read (Home) Survey	1-9
	Teacher Questionnaire	1-9
	School Questionnaire	1-9
1.10	Translation and Verification of Instruments	1-10
1.11	Data Collection	1-10
1.12	Scoring the Constructed-Response Items	1-11
1.13	Data Processing	1-11
1.14	IRT Scaling	1-11
1.15	Data Analysis and Reporting	1-12
1.16	The Trends in IEA's Reading Literacy Study	1-13

2 PIRLS Test Instruments and Booklet Design

2.1	Overview	,	2-2	
2.2	The PIRL	S Assessment Framework	2-2	
	Text Types		2-3	
	Processes and Strategies			
	Test Booklet Design		2-5	
	Exhibit 2.1	Distribution of Literacy and Informational Blocks Across Booklets	2-6	
2.3	Released	Status for PIRLS Passages	2-6	

3 PIRLS Background Questionnaires

3.1	Overview		3-2	
3.2	Framework for the Questionnaires		3-2	
	Exhibit 3.1	Contexts for the Development of Reading Literacy	3-3	
	Exhibit 3.2	Factors within the Homs	2-5 .5cho	dingl,LitNks Acri

5.6	Other Sampling Variables Included in the Student and Student-Teacher Link Files	5-14
5.7	Using Sampling Variables in the Analysis	5-15

6 Scaling Methodology and Achievement Scores

6.1	Overview	6-2
6.2	PIRLS 2001 Scaling Methodology	6-3
	Two- and Three-Parameter IRT Models for Dichotomous Items	6-3
	The IRT Model for Polytomous Items	6-4

7Estimating Sampling and Imputation Variance7.1Overview7-27.2Estimating Sampling Variance7-2

PIRLS 2001 Database User Guide

9

Exhibit 9.13	SPSS Control Statement for Performing Teacher Level Analysis
	Using Plausible Values

Exhibit 9.14 Extract of SPSS Output for Performing Teacher Level Analysis

9-37

Overview of User Guide

1.1	Overview of User Guide	1-2
1.2	Overview of PIRLS	1-3
1.3	Participating Countries	1-4
1.4	Student Population Assessed	1-4
1.5	Assessment Dates	1-5
1.6	Study Management and Organization	1-5
1.7	Overview of Assessment Framework	1-6
	Exhibit 1.1 Percentages of Reading Assessment Devoted to Reading Purposes and Processes	1-7
1.8	PIRLS Reading Assessment	1-7
	Assessment Design	1-7
	Passages	1-8
	Items and Scoring Guides	1-8
	Releasing Assessment Material to the Public	1-8
1.9	PIRLS Background Questionnaires	1-8
	Student Questionnaire	1-9
	Learning to Read (Home) Survey	1-9
	Teacher Questionnaire	1-9
	School Questionnaire	1-9
1.10	Translation and Verification of Instruments	1-10
1.11	Data Collection	1-10

1

1.1 Overview of User Guide

This User Guide accompanies the PIRLS 2001 International Database. The database comprises achievement data in addition to student, parent, teacher, and school background data for the 35 countries that participated in PIRLS 2001. The PIRLS 2001 International Database includes responses from over 150,000 students, 130,000 parents, about 7,000 teachers, and 6,000 school principals. All participating countries gave the IEA permission to release their national data.

The database also contains achievement data and student background data for the nine countries that participated in the Trends in IEA's Reading Literacy Study, in which countries that participated in the IEA's 1991 Reading Literacy Study administered the reading test and student questionnaire to measure trends in reading literacy.

The PIRLS 2001 International Database includes the following for each country for which internationally comparable data are available:

- Students' responses to cognitive reading items
- Students' responses to the background student questionnaire
- Parents' responses to the background home questionnaire
- Teachers' responses to the background teacher questionnaire
- Principals' responses to the background school questionnaire
- Data almanacs.

This user guide includes background information about the design, implementation, and reporting of the PIRLS 2001 assessment in addition to information specific to the data files contained in the international database. Chapters 1–3 describe the design of the PIRLS reading test and background questionnaires. Chapters 4–7 detail the sample design, data collection procedures, scaling, and analysis and reporting issues. Further details of the contents of Chapters 1–7 are available from the *PIRLS 2001 Technical Report* (Martin, Mullis, & Kennedy, 2003). Chapter 8 documents the content and format of the data files in the database, discusses the appropriate use of the files and variables, and draws attention to special considerations that arise from the complex design of the assessment. Example an

1.3 Participating Countries

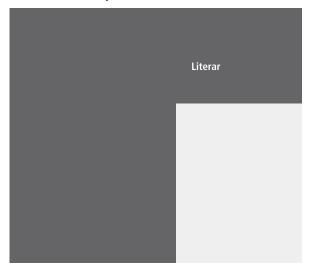
Thirty-five countries joined together to conduct the first PIRLS assessment in 2001:

Argentina	Latvia
Belize	Lithuania
Bulgaria	Macedonia
Canada (O, Q)*	Moldova
Colombia	Morocco
Cyprus	Netherlands
Czech Republic	New Zealand
England	Norway
France	Romania
Germany	Russian Federation
Greece	Scotland
Hong Kong	Singapore
Hungary	Slovak Republic
Iceland	Slovenia
Iran	Sweden
Israel	Turkey
Italy	United States
Kuwait	

* Canada is represented by the provin

type) as stratification variables. At the second stage, one or two fourth-grade classes

Exhibit 1.1 Percentages of Reading Assessment Devoted to Reading Purposes and Processes¹



Passages

The reading passages form the basis for the reading literacy test. In accordance with the framework, four of the assessment blocks contain literary texts and four contain informational texts, and the passages are authentic texts drawn from children's storybooks and informational sources. Submitted and reviewed by the PIRLS countries, the passages represent a range of types of literary and informational texts. The literary passages include realistic stories and traditional tales; while the informational texts include chronological and non-chronological articles, a biographical article, and an informational leaflet.

Items and Scoring Guides

Two item formats were used to assess children's reading literacy – multiple-choice and constructed-response. Each type of item was used to assess both reading purposes and all four reading processes. Multiple-choice items provided students with four possible answers, one of which was correct. Each multiple-choice item was worth one point. Constructed-response items required students to construct their answers rather than select from among possible answers. Short-answer constructedresponse items were worth one or two points, and extended-response items were worth three points.

Scoring guides for constructed-response items were developed together with the items. Each scoring guide is unique to that item. It describes the essential features of appropriate and complete responses – including the kind of evidence of understanding required and example student responses to help scorers determine the score for a particular response. Actual student responses were used to develop the

is a set of questionnaires targeting factors related to reading literacy. PIRLS administered four questionnaires: to the tested students, to their parents, to their reading teachers, and to their school principals.

1.10 Translation and Verification of Instruments

1.12 Scoring the Constructed-Response Items

Because almost two-thirds of the score points came from constructed-response items, PIRLS needed to develop procedures for reliably evaluating student responses within and across countries. The International Study Center prepared detailed guides containing the PIRLS scoring rubrics, and explanations of how to implement them, together with example student responses for the various rubric categories. These guides, along with training packets containing extensive examples of student responses for practice in applying the rubrics, were used as a basis for intensive training of national scorers.

To gather and document empirical information about the within-country agreement among scorers, PIRLS arranged to have a sample of 200 students' responses to each item in each country scored independently by two readers. Scoring reliability within countries was high – the percentage of exact agreement, on average, across countries, was more than 90 percent. PIRLS also conducted a study of scoring reliability across countries, asking countries with scorers proficient in English to score a reference set of student responses chosen from students in English-speaking countries. This study revealed a high level of agreement between scorers also (85% on average).

1.13 Data Processing

To ensure the availability of comparable, high-quality data for analysis, PIRLS took rigorous quality control steps to create the international database. Countries used manuals and software provided by PIRLS to create and check their data files, so that the information would be in a standardized international format before being forwarded to the IEA Data Processing Center. Upon arrival at the DPC, the data underwent an exhaustive cleaning process involving several steps and procedures designed to identify, document, and correct deviations from the international instruments, file structures, and coding schemes. The process also emphasized consistency of information within national data sets, and appropriate linking among the student, parent, teacher, and school data files.

1.14 IRT Scaling

The general approach to reporting the PIRLS achievement data was based primarily on item response theory (IRT) scaling methods. Student reading achievement was summarized using a family of IRT models (2-parameter, 3-parameter, and generalized partial credit models). The IRT methodology was preferred for developing comparable estimates of performance for all students, since students responded to different passages and items depending upon which of the test booklets they received (Booklet 1 through 9 or the PIRLS Reader). This

PIRLS 2001 collected a wide array of information about the home and school context in which students learned to read (from parents, students, teachers, and school principals). The *PIRLS 2001 International Report* (Mullis, et al., 2003) summarizes much of this information, combining data into composite indices, and showing an association with achievement where appropriate. In particular, student reading achievement is described in relation to literacy-related activities in the home, the school curriculum and organization for teaching reading, teachers and reading instruction, school contexts, and students' reading attitudes, self-concepts, and out-of-school activities.

Additional information about the countries participating in PIRLS 2001 may be found in the *PIRLS 2001 Encyclopedia* (Mullis, Martin, Kennedy, & Flaherty, 2002), a

Countries sampled every other PIRLS school for the trend study, resulting in a sample size of at least 75 schools. In each school, one target-grade classroom was sampled and administered the 1991 test and student questionnaire. For some countries, the4 TD- studei arok0.1(st(sampy1(e)-0st(3-a)5.(2(thsuntri9Tw{sok0.1it0vGRr(e)-wnister})5 n-a)5.(2(thsuntri9Tw{sok0.1it0vGRr(e)-wnister})5 n-a)5.(2(thsuntri9Tw{sok0.1it0vGRr(e)-wnis

PIRLS Test Instruments and Booklet Design

2.1	Overview	2-2
2.2	The PIRLS Assessment Framework	2-2
	Text Types	2-3
	Processes and Strategies	2-4
	Test Booklet Design	2-5
	Exhibit 2.1: Distribution of Literacy and Informational Blocks Across Booklets	2-6
2.3	Released Status for PIRLS Passages	2-6

2

2.1 Overview

convey information. The PIRLS assessment included both chronologically and logically structured informational texts, some of which incorporated various types of adjunct aids.

Processes and Strategies

Across text types and purposes for reading, the reader engages in a variety of comprehension processes and strategies to gain and construct meaning from text. The PIRLS assessment framework described four specific processes of comprehension, which vary in terms of the degree of inference or interpretation required and in the focus on text content or structural features of the text. This description of comprehension processes in the framework served as a guide for developing the comprehension questions used to assess students' understandings of texts. Each question was written to engage students in one of four processes: 1) focus on and retrieve explicitly stated information, 2) make straightforward inferences, 3) interpret and integrate ideas and information, and 4) examine and evaluate content, language, and textual elements. A brief description of each process is provided below.

In focusing on and retrieving explicitly stated information, the reader locates specific information or an idea in the text that is relevant to understanding the text's meaning. Little or no inference is required to understand the meaningd b information – it is explicit, and may be viewed as existingdat the surface level of the text. Most often, the retrieved information resides locally in the text, within a specific sentence or phrase. A competent reader's understanding of the retrieved information is typically immediate or automatic.

In making straightforward inferences, the reader goes beyond what is stated explicitly in the text and infers some implied meaning or connection between textually-based ideas. Although not stated explicitly, the inference is very much constrained by the text. The text provides fairly obvious cues to guide the reader in making this type of inference. As skilled readers will often make such an inference automatically as they become engaged in constructing meaning wit TTain a specific part of the text, or as they develop a more global understandttu in a particular a more guient, lan(tu i) arid

3 PIRLS Background Questionnaires

3.1	Overview	3-2
3.2	Framework for the Questionnaires	3-2

Exhibit 3.1 Contexts for the Development of Reading Literacy

N			
			,

3.3 PIRLS Main Survey Questionnaires

The contents of the PIRLS main survey questionnaires used to collect information about home, school, and community contexts for learning to read are described below.

Student Questionnaire

Each student taking the PIRLS reading assessment completed the student questionnaire. The questionnaire asks about aspects of students' home and school experiences – including instructional experiences and reading for homework, self-perceptions and attitudes towards reading, out-of-school reading habits, computer use, home literacy resources, and basic demographic information. The questionnaire was designed to take 15-30 minutes to complete. Exhibit 3.3 presents details regarding the items in the questionnaire.

	Student Questionnaire					
Item	Number	Item Content	Description			
	1	Gender	Whether student is a boy or girl			
	2	Date of birth	Month and year of student's birth			
	3	Out-of-school				

Learning-to-Read Survey (Home)

The learning-to-read survey was completed by the parents or primary caregivers of each student taking the PIRLS reading assessment. It deals with child-parent literacy interactions, home literacy resources, parents' reading habits and attitudes, home-school connections, and basic demographic and socioeconomic indicators. This questionnaire was designed to take 10-15 minutes to complete. Exhibit 3.4 presents details regarding the items in the questionnaire.

Exhibit 3.4 Content of the PIRLS Learning-to-Read Survey (Home Questionnaire)

Learning-to-Read Survey

Exhibit 3.5 Content of the PIRLS Teacher Questionnaire (Continued)

Teacher Questionnaire				
Item Number	Item Content	Description		

Exhibit 3.6 Content of the PIRLS School Questionnaire (Continued)

School Questionnaire

3.4 Reporting Questionnaire Data

As described in this chapter, PIRLS 2001 used four questionnaires to gather

were more than 100 books, more than 25 children's books, and at least three of the educational aids in the home, and at least one parent who completed a university education. Students at the low level had 25 or fewer books, 25 or fewer children's books, no more than two educational aids, and the highest level of education for either parent was some secondary or less. Students with all other response combinations were assigned to the middle category.

The 10 indices constructed for the *PIRLS 2001 International Report* are listed in Exhibit 3.7.

Exhibit 3.7 Summary Indices from Background Data in the *PIRLS 2001* International Report

Name of Index	Label	Exhibit ^a	Analysis Method
Index of Early Home Literacy Activities	EHLA	4.10	ndex based on parents' responses to the frequency of the following activities they engaged in with their child prior to entry into primary school: read books, tell stories, sing songs, play with alphabet toys (e.g., plocks with letters of the alphabet), play word games, pr read aloud signs and labels. Average is computed across the 6 items based on a 3-point scale: Never or almost never = 1, Sometimes = 2, and Often = 3. High evel indicates an average of greater than 2.33 through 3. Medium level indicates 0.949 808(r)8.3(ags)4.3(of 8(t)5.0)-10.ow less

		-	
Name of Index	Label	Exhibit ^a	Analysis Method
Index of Reading for Homework	RFH	6.34	Index based on teachers' responses to two questions: How often do you assign reterT.2(6 67946.3(3ngu a partx ofhHomework)

Exhibit 3.7 Summary Indices from Background Data in the *PIRLS 2001 International Report* (Continued)

PIRLS 2001 Database User Guide

Exhibit 3.7	Summary Indices from Background Data in the PIRLS 2001
	International Report (Continued)

Name of Index	Label	Exhibit ^a	Analysis Method
Index of Principals' Perceptions of School Safety	PPSS	7.17	Index based on principals' responses about the degree each was a school problem: classroom disturbances, cheating, profanity, vandalism, theft, intimidation or verbal abuse of other students, and physical conflicts among students. Average is computed on a 4-point scale: Not a problem = 1, Minor problem = 2, Moderate problem = 3, and Serious problem = 4. High level indicates an average of 1 to less than 2. Medium level indicates an average of 2 through 3. Low level indicates an average of greater than 3 through 4.
Index of Availability of School Resources	ASR	7.18	Index based on principals' responses to how much the school's capacity to provide instruction is affected by a shortage or inadequacy of the following: instructional staff, teachers qualified to teach reading, instructional materials, supplies (e.g., paper, pencils), school buildings and grounds, heating/cooling and lighting systems, instructional space (e.g., classrooms), special equipment for physically disabled students, computers for instructional purposes, computer software for instructional purposes, computer support staff, library books, and audiovisual resources. Average is computed on a 4-point scale: Not at all = 1, A little = 2, Some = 3, and A lot = 4. High level indicates an average of 1 to less than 2. Medium level indicates an average of greater than 3 through 4.

y5uw[Stud)-H(ud)-9.5(en)-7.3(ts)-6d0 369.84 Tm0 0 0 b8.969.88d0008 Tc71IW*n20A I 369.84o20A I 369.84o;

4 PIRLS Survey Operations Procedures

4.1	Overview		4-2

4.2 Responsibilities of the National Research Coort2p2Itor

4.1 Overview

The PIRLS 2001 data collection was a very complex undertaking in each country, requiring close cooperation between the National Research Coordinator (NRC) and school personnel – including students and their parents, school principals, and teachers. Survey operations and procedures for administering the PIRLS 2001 assessment were developed by the international project team, and documented in a series of manuals provided to the national centers. Each country was responsible for implementing the procedures according to the international standards.

The PIRLS 2001 survey operations were designed collaboratively by the International Study Center (ISC) at Boston College, the IEA Data Processing Center, and Statistics Canada. They were based on procedures used successfully in TIMSS and other IEA studies, and refined on the basis of the PIRLS 2001 field-test experience. As well as providing data to inform the instrument development process, the field test, which was conducted in 30 countries in September 2000, allowed participating countries to gain practical experience with the procedures described in the manuals, and provided an opportunity to identify areas in need of improvement.

This chapter describes the survey operations used to collect the PIRLS data, including: the procedure for sampling classrooms within schools and tracking students and teachers, the steps involved in administering the achievement tests and background questionnaires, and the requirements for monitoring the quality of the data collection. It also describes the activities involved in preparing the data files at the national center, particularly those for scoring the constructed-response items,

required to track and account for all excluded students, and were cautioned that excluding an excessive proportion would lead to their results being annotated in international reports. The conditions under which students could be excluded were carefully delineated, because the definition of "disabled" students varied considerably from country to country.

Assigning Instruments to Students and Teachers

The PIRLS reading assessment was packaged into nine student booklets and one magazine-like booklet known as the PIRLS Reader, which came with its own answer booklet (see Chapter 2). Each student was asked to complete just one of the 10 assessment booklets. Students recorded their answers directly in the booklet, except in the case of the Reader, where the accompanying answer booklet was used. All students completed the same student questionnaire. The assessment booklets were numbered sequentially 1 through 9, with R for the Reader answer booklet. Booklets were assigned to students following a systematic procedure that ensured an even distribution of booklets throughout each class. Each student's booklet assignment was recorded in advance on the Student Tracking Form, and the Test Administrator was expected to ensure that the correct booklet was given to each student. To facilitate proper booklet distribution, each booklet was individually labeled. Within each class, a number from one to twelve was randomly selected and used as a starting point to cycle through the list. For example, if 10 was chosen as the starting number in a particular class, the first student in the tracking form was assigned Booklet 8, the second Booklet 9, the third Booklet R, the fourth Booklet 1, and so on until all students in the class were assigned a booklet. It was critical that the test booklets be assigned to the students before the testing day to ensure their correct distribution.

4.4 Within-School Assessment Activities

The School Coordinator in each school was responsible for organizing the administration of the PIRLS 2001 test. The coordinator could be the principal, the principal's designee, or an outsider appointed by the NRC with the approval of the principal. The NRC was responsible for ensuring that the School Coordinator was familiar with his or her responsibilities.

The tasks of the School Coordinator were detailed in the *School Coordinator Manual*. Before the test administration, the School Coordinator, working with the NRC, had to arrange a testing date, select a Test Administrator to conduct the testing sessions, and ensure that the Test Administrator was fully acquainted with the assessment procedures. In some countries, the School Coordinator assumed the roles and responsibilities of the Test Administrator as well as those of the School Coordinator.

Arranging the Testing Sessions

In preparation for the testing day, the School Coordinator worked with the school principal, Test Administrator, and the teacher, to plan the testing sessions by arranging rooms, classes, and materials. In countries where obtaining parental permission for testing was required, the School Coordinator ensured that permission forms were signed and returned in time. Once the testing materials arrived from the national center, the School Coordinator checked that they were for the appropriate students and teachers, that there were enough copies, and that the materials would be kept in a secure place until the testing day.

Distributing Materials

The School Coordinator distributed Teacher Questionnaires to the teachers listed in the Teacher Tracking Form, and a School Questionnaire to the school principal, and ensured that they were completed and returned. Teacher participation was recorded on the Teacher Tracking Form when the questionnaires were returned. In some countries, it was also the responsibility of the School Coordinator to collect the completed Learning to Read Surveys from the schools, record parent participation on the Student Tracking Forms, and return the questionnaires to the NRC.

Test Administration

The Test Administrator was responsible for administering the PIRLS test and student questionnaire. Specific responsibilities of the Test Administrator were described in the *Test Administrator Manual*. The Test Administrator distributed the test booklets and questionnaires according to the assignment documented on the Student Tracking Form,

ensuring that each student received the correct testing materials. The Test Administrator conducted the testing sessions in accordance with a script provided in the *Test Administrator Manual*, and recorded the timing of the testing sessions on the Test Administration Form. After the testing session, the Test Administrator recorded student participation on the Student Tracking Form, and returned the testing materials to the School Coordinator.

Timing of the Testing Sessions

Testing was conducted in two consecutive sessions one for administering the PIRLS achievement test booklets, and one for administering the Student Questionnaires. The first session was conducted in two parts, one for each part of the test booklet. During the field test, 30 minutes were provided for students to answer each part of the booklet. At the end of the 30 minutes, up to 10 extra minutes were allowed if less than 90 percent of the students had not completed answering the questions. The Test Administrator was required to document the timing on the Test Administration Form. As part of the analysis of the field test results, the time requirements were analyzed to determine what amount of time was sufficient across all countries, and, as a consequence, the timing for the main survey was adjusted. For the main survey, the allotted time was increased to 40 minutes – and, of course, no additional time was permitted beyond what was specified. The timing of the session was as follows:

- Approximately 10 minutes for preparation (i.e., reading instructions, distributing test booklets)
- 40 minutes for answering Part 1 of the test booklet
- Approximately 15 minutes for a break
- 5 minutes for preparing students for Part 2
- 40 minutes for answering Part 2 of the test booklet
- At least 20 minutes for the completion of the Student Questionnaire
- 5 minutes for distributing the Learning to Read Surveys.

The Student Questionnaire was to be administered on the same day as the achievement test, following the testing session, or, if this was not possible, on the following day.

4.6 Trends in IEA's Reading Literacy Study

The Trends in IEA's Reading Literacy Study was an option for countries that participated in the 1991 IEA Reading Literacy Study, allowing them to administer the 1991 test booklets and questionnaires again in 2001 – to compare reading achievement over time. A list of countries participating in the Trends in IEA's Reading Literacy Study is provided in Chapter 5.

Documentation for implementing the trend study was incorporated into the manuals for PIRLS 2001. The *Survey Operations Manual* provided instructions for preparing tracking forms, packaging materials for the schools, preparing, administering, and returning the test booklets and questionnaires, and for data entry. *The School Coordinator Manual* and *Test Administrator Manual* included sections specific to the trends study wherever necessary.

The School Sampling Guide for the 10-Year Trend Study (PIRLS, 2000) documented procedures for sampling schools for the Trends in IEA's Reading Literacy Study. Essentially, the procedure was to sample an additional class from the target grade in half of the schools in the PIRLS sample. The survey tracking forms used in PIRLS were also used for tracking trend study schools, classes, and students.

Countries were expected to use test booklets identical to those used in 1991. Schools in which classes were selected for the trend study received two packages from the NRC. One contained the test booklets with the Student Tracking Forms, and the other the Student Questionnaires. Teachers of the classes selected for the trend study were not given a Teacher Questionnaire, nor was the School Questionnaire administered.

The data collection was conducted under the same conditions as in 1991. There were three data-collection sessions, with the same time limitations as in 1991. There was a short break between sessions:

- 1 minute and 30 seconds for answering questions in the Word Test (word recognition), followed by 35 minutes for answering questions in the first part of the Reading Test
- 40 minutes for answering the questions in the second part of the Reading Test
- At least 25 minutes for the completion of the Student Questionnaire.

The Test Administrator and School Coordinator followed the PIRLS procedures for collecting the test instruments, checking for proper documentation on the survey tracking forms, and for calculating student response rates. Once any necessary

makeup sessions had been held, the School Coordinator returned the materials to the NRC for data preparation. Since the trend study test booklets did not include constructed-response items for scoring, they were sorted separately for data entry – along with the trend study Student Questionnaires.

5

Sampling and Sampling Weights

5.1	Overview	5-2
5.2	PIRLS Target Population	5-3
	Sampling from the Target Population	5-3
5.3	Sample Design	5-5
	Units of Analysis and Sampling Units	5-5
	Sampling Precision and Sample Size	5-6
	Stratification	5-7
5.4	Trends in IEA's Reading Literacy Study	5-8
	Trends in IEA's Reading Literacy Study Target Population	5-8
	Sample Design	5-9
5.5	Sampling Weights	5-9
	Weight Variables Included in the Student Data Files	5-10
	Weight Variables Included in the Student-Teacher Linkage Files	5-13
	Weight Variables Included in the School Data Files	5-13
5.6	Other Sampling Variables Included in the Student and	
	Student-Teacher Link Files	5-14
5.7	Using Sampling Variables in the Analysis	5-15

5.1 Overview

This chapter describes the PIRLS 2001 procedures for sampling from the student population in each participating country. To be acceptable for PIRLS, national

Where the local situation required it, NRCs were permitted to adapt the sample design for their educational systems, using more sampling information, and more sophisticated designs and procedures than the base design required. However, these solutions had to be approved by the International Study Center (ISC) at Boston College, and by Statistics Canada.

5.2 PIRLS Target Population

In IEA studies, the target population for all countries is known as the *international desired target population*. This is the grade or age level that each country should address in its sampling activities. The international desired target population for PIRLS was the following:

All students enrolled in the upper of the two adjacent grades that contain the largest proportion of 9-year-olds at the time of testing.

The PIRLS target grade was usually the fourth grade of primary school. Because fourth grade generally signals the completion of formal reading instruction, countries for which the target grade would have been the third grade (based on the international desired target population) were permitted to retain the fourth grade as their target grade. The PIRLS target population was derived from that used by TIMSS in 1995, and identical to that used by TIMSS 2003 at the primary school level.

Sampling from the Target Population

PIRLS expected all participating countries to define their *national desired population* to correspond as closely as possible to its definition of the international desired population. For example, if fourth grade was the upper of the two adjacent grades containing the greatest proportion of 9-year-olds in a particular country, then fourth grade should be the national desired population for that country. Although countries were expected to include all students in the target grade in their definition of the population, sometimes they had to reduce their coverage. Lithuania, for example, planned to collect data only about students in Lithuanian-speaking schools, so their national desired population fell short of the international desired population. The international report documents such deviations from the international definition of the PIRLS target population.

Using its national desired population as a basis, each participating country had to define its population in operational terms for sampling purposes. This definition, known in IEA terminology as the *national defined population*, is essentially the sampling frame from which the first stage of sampling takes place. Ideally, the national defined population should coincide with the national desired population, although in reality there may be some school types or regions that cannot be included;

consequently, the national defined population is usually a very large subset of the national desired population. All schools and students in the desired population not included in the defined population are referred to as the *excluded population*.

PIRLS participants were expected to ensure that the national defined population included at least 95 percent of the national desired population. Exclusions (which should be kept to a minimum) could occur at the school level, within the sampled schools, or both. Because the national desired population was restricted to schools that contained the required grade, schools not containing the target grade were considered to be outside the scope of the sample and therefore not part of the target population.

Although countries were expected to do everything possible to maximize coverage

•

Sampling Precision and Sample Size

Sampling sizes for the two stages of the PIRLS sampling had to be specified so as to meet the sampling precision requirements of the study. Since students were the

applied to each school-sampling frame, to select the sample of schools. In PIRLS, the main reason for considering explicit stratification was to ensure disproportionate allocation of the school sample across strata. For example, a country stratifying by school size might require a specific number of schools from each stratum, regardless of the relative size of the stratum.

• **Implicit stratification** makes use of a single school-sampling frame, but sorts the schools in this frame by a set of stratification variables. This type of stratification is a simple way of ensuring proportional sample allocaconsideJbshple.Gou.0001

Weight Variables Included in the Student Data Files¹

Each student's sampling weight is a composite of five factors: the school weighting factor, the school weighting adjustment, the class weighting factor, the student

WGTADJ3 Student Weighting Adjustment

This is an adjustment applied to the variable WGTFAC3 to account for nonparticipating students in the selected school and/or classroom. Multiplying the variables WGTFAC2, WGTFAC3, and WGTADJ3 and adding them up within each school gives an estimate of the number of students within the sampled school.

The five variables listed above are combined to give a student's overall sampling weight. The probability of selecting an individual student is the product of three independent events: selecting the school, the classroom, and the student. To obtain the probability of selection for an individual student, multiply three selection probabilities – school, classroom, and student – and their respective adjustment factors. Inverting this probability gives the sampling weight for the student.

Three versions of the students' sampling weight are provided in the user database. All three give the same figures for statistics such as means and proportions, but vary for statistics such as totals and population sizes. Each one has particular advantages in certain circumstances. These three versions are as follows:

TOTWGT Total Student Weight

This is obtained by simply multiplying the variables WGTFAC1, WGTADJ1, WGTFAC2, WGTFAC3, and WGTADJ3 for the student. The sum of these weights within a sample provides an estimate of the size of the population. Although this is a commonly used sampling weight, it sometimes adds to a very large number, and to a different number within each country. This is not always desirable. For example, if you want to compute a weighted estimate of the mean achievement in the population across all countries, using the variable TOTWGT as your weight variable will lead each country to contribute proportionally to its population size, with the large countries counting more than small countries. Although this is desirable in some circumstances (e.g., when computing the 75th percentile for mathematics achievement for students around the world), in general TOTWGT is not the student weight of choice for cross-country analyses, since it does not treat countries equally, and gives inflated results in significance tests when the proper adjustments are not used.

A key property of the sampling weights is that the same population estimates for means and proportions will be obtained as long as you use a weight variable proportional to the original weights (TOTWGT). For example, you could take the sampling weights for a large country and divide them by a constant to make them smaller. You could also take the weights of a smaller country and multiply them by a constant to make them bigger. Regardless of which constant is used within a country, the weighted estimates obtained from each of these proportional transformations of the weights will be exactly the same. Two other weight variables are computed and included in the student data files. Each of these is computed for a specific purpose and will yield exactly the same results within each country, but will have some desirable properties when estimates across countries are computed or significance tests are performed.

SENWGT Senate Weight

This variable is computed as

$$SENWGT_{g_{i}i} = TOTWGT_{g_{i}i} \left(\frac{500}{\sum_{i=1}^{II} TOTWGT_{g_{i}i}} \right)$$

for each student, by grade within each country, where *i* is the individual student and *g* is the grade of the student. The transformation of the weights will be different within each country, but in the end the sum of the variable SENWGT within each country will be 500 per grade or 1,000 for both grades. The variable SENWGT, within each country, is proportional to TOTWGT multiplied by the ratio of 500 divided by the sum of the weights over all students in the grade. These sampling weights can be used when international estimates are sought and you want to have each country contribute the same amount to the international estimate. When this variable is used as the sampling weight for international estimates, the contribution of each country is the same, regardless of the size of the population.

HOUWGT House Weight

This variable is computed as

HOUWGTmts

WGTADJ1 School Weighting Adjustment

This is an adjustment that is applied to WGTFAC1 to account for nonparticipating schools in the sample. If you were to multiply WGTFAC1 by WGTADJ1 you would obtain the sampling weight for the school, adjusted for non-participation.

SCHWGT School-level Weight

The school sampling weight is the inverse of the probability of selection for the school, multiplied by its corresponding adjustment factor. It is computed as the product of WGTADJ1 and WGTFAC1. Although this weight variable can be used to estimate the number of schools with certain characteristics, it is important to keep in mind that the sample selected for PIRLS is a good sample of students, but not necessarily an optimal sample of schools. Schools are selected with probability proportional to their size, so it is expected that there is a greater number of large schools in the sample. For countries that sampled by track within school, the SCHWGT is based on the track size rather than the total school size. This may lead to invalid school-weighted analyses.

5.6 Other Sampling Variables Included in the Student and Student-Teacher Link Files

With complex sampling designs that involve more than simple random sampling, as in the case of PIRLS where a multi-stage cluster design was used, there are several methods for estimating the sampling error of a statistic that avoid the assumption of simple random sampling. One such method is the jackknife repeated replication (JRR) technique (Wolter, 1985). The particular application of the JRR technique used in PIRLS is termed a paired selection model because it assumes that the sampled population can be partitioned into strata, with the sampling in each stratum consisting of two primary sampling units (PSU), selected independently.

The following variables capture the information necessary to estimate correct standard errors using the JRR technique:

JKZONE

The variable JKZONE indicates the sampling zone or stratum to which the student's school is assigned. The sampling zones can have values from 1 to 75 in

PIRIS 2001 Database User Guide

6 Scaling Methodology and Achievement Scores

6.1	Overview	6-2
6.2	PIRLS 2001 Scaling Methodology	6-3
	Two- and Three-Parameter IRT Mo	

6.2 PIRLS 2001 Scaling Methodology³

The scaling approach used by PIRLS was developed originally by Educational Testing Service for use in the U.S. National Assessment of Educational Progress. It is based on psychometric models that were first used in the field of educational measurement in the 1950s, and have become popular since the 1970s for use in large-scale surveys, test construction, and computer adaptive testing.⁴

Three distinct scaling models, depending on item type and scoring procedure, were used in the analysis of the PIRLS 2001 assessment data. Each is a "latent variable" model that describes the probability that a student will respond in a specific way to an item in terms of the respondent's proficiency, which is an unobserved or "latent"

where

- x_i is the response to item *i*, 1 if correct and 0 if incorrect;
- θ_k is the proficiency of a person on a scale *k* (note that a person with higher proficiency has a greater probability of responding correctly);
- *a_i* is the slope parameter of item *i*, characterizing its discriminating power;
- b_i is the location parameter for the item, characterizing its difficulty;
- *ci* is the lower asymptote parameter for the item, reflecting the chances of respondents of very low proficiency selecting the correct answer.

The probability of an incorrect response to the item is defined as:

(2)

$$P_{i0} \equiv P\left(x_i = 0 \middle| \theta_k, a_i, b_i, c_i\right) = 1 - P_{i1}\left(\theta_k\right)$$

The two-parameter (2PL) model was used for the short constructed-response items that were scored as correct or incorrect. The form of the 2PL model is the same as Equation 1, with the *c*_i parameter fixed at zero.

The IRT Model for Polytomous Items

In PIRLS 2001, constructed-response items requiring an extended response were scored for partial credit (with 0, 1, 2, and 3 as the possible score levels). These polytomous items were scaled using a generalized partial credit model (Muraki, 1992). The fundamental equation of this model gives the probability that a person with proficiency θ_k on scale k will have, for the *i*-th item, a response x_i that is scored in the *i*-th of m_i ordered score categories:

(3)

$$P(x_{i} = l | \theta_{k}, a_{i}, b_{i}, d_{i,l}, \mathsf{K} | d_{i,m_{i}-l}) = \frac{\exp\left[\sum_{v=0}^{l} 1.7a_{i}(\theta_{k} - b_{i} + d_{i,v})\right]}{\sum_{c=0}^{m_{i}-1_{j}} \exp\left[\sum_{v=0}^{c} 1.7a_{i}(\theta_{k} - b_{i} + d_{i,v})\right]} = P_{il}(\theta_{k})$$

where:

 m_i is the number of response categories for item *i*;

 x_i is the response to item *i*, possibilities ranging between 0 and mi - 1;

 θ_k is the proficiency of person on a scale *k*;

$$t^{*}(x, y) = E[t(\theta, y)x, y]$$
$$= \int t(\theta, y)p(\theta|x, y)d\theta$$

It is possible to approximate t^* using random draws from the conditional distribution of the scale proficiencies given the student's item responses x_{j_t} the student's background variables y_{j_t} and model parameters for the student. These values are referred to as imputations in the sampling literature, and as plausible values in large-scale surveys such as TIMSS, NAEP, NALS, and IALLS. The value of θ for any respondent that would enter into the computation of *t* is thus replaced by a randomly selected value from his or her conditional distribution. Rubin (1987) proposed repeating this process several times so that the uncertainty associated with imputation can be quantified. For example, the average of multiple estimates of *t*, each computed from a different set of plausible values, is a numerical approximation of *t** of Equation 4; the variance among them reflects uncertainty due to not observing θ

Conditioning

A multivariate normal distribution was assumed for $P(\theta_i | x_i, y_i, \Gamma, \Sigma)$, with a common variance, Σ , and with a mean given by a linear model with regression parameters, Γ . Since, in large-scale studies like PIRLS, there are many hundreds of background variables, it is customary to conduct a principal components analysis to reduce the number to be used in Γ . Typically, components representing 90 percent of the variance in the data are selected. These principal components are referred to as the conditioning variables, and denoted as *y*. The following model is then fit to the data:

(6)

$$\theta = \Gamma' y^c + \varepsilon$$

In Equation 6, ε is normally distributed with mean zero and variance Σ . As in a regression analysis, Γ is a matrix each of whose columns is the effects for each scale, and Σ is the matrix of residual variance between scales.

Note that, in order to be strictly correct for all functions Γ of θ , it is necessary that $P(\theta | y)$ be correctly specified for all background variables in the survey. Estimates of

For respondents with an insufficient number of responses, the Γ and Σ s described in the previous paragraph are fixed. Hence, all respondents – regardless of the number of items attempted – are assigned a set of plausible values.

The plausible values can then be employed to evaluate an arbitrary statistic T as follows:

- 1. Using the first vector of plausible values for each respondent, evaluate T as if the plausible values were the true values of θ . Denote the result T_1 .
- 2. As in step 1 above, evaluate the sampling variance of T, or $Var(T_1)$, with respect to respondents' first vectors of plausible values.
- 3. Carry out steps 1 and 2 for the second through fifth vectors of plausible values, thus obtaining T_u and Var_u for $u=2, \ldots, M$, where M is the number of imputed values.
- 4. The best estimate of *T* obtainable from the plausible values is the average of the five values obtained from the different sets of plausible values:

$$T. = \frac{\sum_{i=1}^{n} T_{u}}{5}$$

5.

If θ values were observed for all sampled respondents, the statistic $(t-T)/U^{1/2}$ would follow a *t*-distribution with *d* degrees of freedom. Then the incomplete-data statistic $(t^*-T)/(Var(t^*))^{1/2}$ is approximately t-distributed, with degrees of freedom (Johnson & Rust, 1993) given by:

$$v = \frac{1}{\frac{f_{M}^{2}}{M-1} + \frac{(1-f_{M})^{2}}{d}}$$

where *d* is the degrees of freedom for the complete-data statistic, and *f* is the proportion of total variance due to not observing θ values:

$$f_M = \frac{\left(1 + M^{-1}\right)B_M}{V_M}$$

where B_M is the variance among M imputed values and V_M is the final 78333 -4.75.6 505.92 Tm0Tc6.553

6.3 Student Achievement Scores

The PIRLS international database contains several student-level achievement scores, including the plausible values described in this chapter. These scores were computed at different stages of the study to serve specific purposes. This chapter presents a description of these achievement scores, how they were derived, how they were used by PIRLS, and how users of the database can use them. For identification purposes, the first letter for the variable name identifies the population for which the score was computed. The scores computed for Population 1 (4th Grade) have the letter A as the first character. This convention has been followed with other background and derived variables and with the files included in the database.

Achievement Scores in the Student Files

Four types of achievement scores are included in the student data files: plausible values, raw scores, standardized raw scores, and national Rasch scores. Each type is described below.

Reading Scores: Plausible Values

As described earlier in this chapter, PIRLS made use of multiple imputation or plausible values methodology to provide estimates of student proficiency in reading. Because of the uncertainty involved in the imputation process, PIRLS produced not one but five imputed values for each student. The plausible values available for each of the PIRLS 2001 reading scales are listed in Exhibit 6.1; those for the trends in IEA's Reading Literacy Study reading scales are listed in Exhibit 6.2.

The need for plausible values arises from the fact that any student was administered only a fraction of the items in the assessment, as described in Chapter 2. Time constraints did not allow for all the items to be administered to each student. A plausible value is an estimate of how the individual student would have performed on a test that included all possible items in the assessment. Since no student responded to all items, this estimate is based on the responses to the items that were included in the test booklet that the student actually took, and the performance of students with similar characteristics based on their responses to the background questionnaire.

Overall reading plausible values were standardized so that the mean is equal to 500 and the standard deviation equals 100 when all countries are weighted equally. One set of plausible values can be considered as good as another. Each of these sets is equally well-designed to estimate population parameters, although the estimates will differ somewhat. This difference is attributable to imputation variance, which contributes to the uncertainty of the estimate. Five sets of plausible values are provided so that analyses can be replicated as many as five times. Results which vary

Exhibit 6.5 Standardized Scores Included in the PIRLS Data Files

Variable	Description
ASASTDR	Standardized raw score points on the Antarctica Block
ASCSTDR	Standardized raw score points on the Clay Block
ASFSTDR	Standardized raw score points on the Flowers Block
ASHSTDR	Standardized raw score points on the Hare Block
ASLSTDR	Standardized raw score points on the Leonardo Block
ASMSTDR	Standardized raw score points on the Mice Block
ASNSTDR	Standardized raw score points on the Pufflings Block
ASRSTDR	Standardized raw score points on the River Trail Block

Exhibit 6.6 Standardized Scores Included in the Reading Literacy Study Data Files

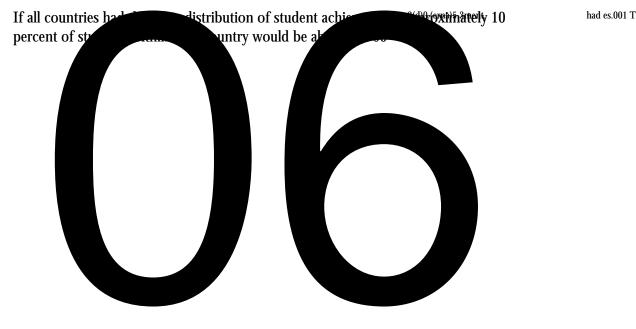
Variable	Description
ASASTDR	Standardized raw score points on all the items
ASDSTDR	Standardized raw score points on the document items
ASESTDR	Standardized raw score points on the expository items
ASNSTDR	Standardized raw score points on the narrative items

National Rasch Scores

The national Rasch scores were also computed for preliminary analyses. These provided a basic Rasch score for preliminary analyses within countries, but cannot be used for international comparisons, since each block within each country has been assigned the same mean score. The national Rasch scores were computed by standardizing block logit scores to have a weighted mean of 150 and a standard deviation of 10 within each country. The logit scores were computed using the Quest Rasch analysis software; Quest provides maximum likelihood (ML) estimates of a scaled score, based on the one-parameter Rasch model, for the performance of students on a set of items. These logit scores were obtained using item difficulties that were computed for each country using all available item responses for the country and centering the item difficulty around zero. When computing the item difficulties, responses marked as "not reached" were treated as items that were not administered. This avoided giving inflated item difficulties to the items located at the end of the test in cases where students systematically do not reach the end of the test. These item difficulties were then used to compute logit scores for each student.

6.4 International and National Benchmarks of Achievement

In order to provide more information about student achievement, PIRLS identified four points on the reading scales for use as international benchmarks. The Top 10 percent Benchmark was defined as the 90th percentile on the PIRLS scale, computed across all students in all participating countries, with countries weighted in proportion to the size of their fourth-grade population. This point on each scale is the point above which the top 10 percent of students in the PIRLS assessment scored. The Upper Quarter Benchmark is the 75th percentile on the scale, above which the top 25 percent of students scored. The Median Benchmark is the 50th percentile, above which the top half of students scored. Finally, the Lower Quarter Benchmark is the 25th percentile, the point reached by the top 75 percent of students.



	25 th	50 th	75 th	90 th
Proficiency Score	Percentile	Percentile	Percentile	Percentile
Plausible Value 1	434.335	509.255	571.206	622.816
Plausible Value 2	435.964	510.496	571.238	620.814
Plausible Value 3	435.295	509.981	571.043	622.122
Plausible Value 4	434.653	509.567	571.388	621.975
Plausible Value 5	434.424	509.711	571.775	620.954

Exhibit 6.10 PIRLS International Benchmarks of Reading for Literary Purposes Achievement

Exhibit 6.11 PIRLS International Benchmarks of Reading to Acquire and Use Information Achievement

Proficiency Score	25 th	
Tronciency Score	Percentile	

7.1 Overview

When analyzing data from complex designs such as PIRLS, it is important to compute correct error variance estimates for the statistics of interest. In PIRLS this error variance comes from two sources: the sampling process and the imputation process. This chapter describes the methods used to estimate each of these error variance components.

7.2 Estimating Sampling Variance

With complex sampling designs that involve more than simple random sampling, as in the case of PIRLS where a multi-stage cluster design was used, there are several methods for estimating the sampling error of a statistic that avoid the assumption of simple random sampling. One such method is the jackknife repeated replication (JRR) technique (Wolter, 1985). The particular application of the JRR technique used in PIRLS is termed a paired selection model because it assumes that the sampled population can be partitioned into strata, with the sampling in each stratum consisting of two primary sampling units (PSU), selected independently. Following this first-stage sampling, there may be any number of subsequent stages of selection that may involve equal or unequal probability selection of the corresponding elements. The PIRLS design called for a total of 150 schools for the target population. These schools constituted the PSUs in most countries, and were paired sequentially after sorting by a set of implicit stratification variables. This resulted in the implicit creation of 75 strata, with two schools selected per stratum.

The jackknife repeated replication (JRR) method is suitable for estimating sampling errors in the PIRLS design because it provides approximately unbiased estimates of the sampling error arising from the complex sample selection procedure for estimates such as percentages, totals, and means. In addition, this method can also be readily adapted to the estimation of sampling errors for parameters estimated using other statistical modeling procedures, such as percent-correct technology. The general use of the JRR entails systematically assigning pairs of schools to sampling zones, and the random selection of one of these schools to have its contribution doubled, and the other zeroed, so as to construct a number of "pseudo-replicates" of the original sample. The statistic of interest is computed once for all of the original sample, and once more for each of the pseudo-replicate samples. The variation between the original sample estimate and the estimates from each of the replicate samples is the jackknife estimate of the sampling error of the statistic.

Computing Sampling Variance Using the JRR Method

When implementing the JRR method in PIRLS, for each country, it was assumed that there were up to 75 strata or zones (H) within each country, each one containing two sampled schools selected independently. When computing a statistic "t" from the sample for a country, the formula for the JRR variance estimate of the statistic t is then given by the following equation:

(1)

$$Var_{jrr}(t) = \sum_{h=1}^{H} \left[t \left(J_h - t \left(S \right) \right) \right]^2$$

where *H* is the number of pairs in the entire sample for the country. The term t(S) corresponds to the statistic computed for the whole sample (computed with any specific weights that may have been used to compensate for the unequal probability of selection of the different elements in the sample or any other post-stratification weight). The element $t(J_h)$ denotes the same statistic using the *h*th jackknife replicate, computed for all cases except those in the *h*th stratum of the sample, removing all cases associated with one of the randomly selected units of the pair within the *h*th stratum, and including, twice, the elements associated with the other unit in the *h*th stratum. In practice, this is effectively accomplished by recoding to zero the weights for the cases of the element of the pair to be excluded from the replication, and multiplying by two the weights of the remaining element within the *h*th pair. This results in a set of *H* replicate weights that may be used in computing the JRR variance estimate.

As can be seen from the above formula, the computation of the JRR variance estimate for any statistic from the PIRLS database requires the computation of any statistic up to 76 times for any given country: once to obtain the statistic for the full sample, and up to 75 times to obtain the statistics for each of the jackknife replicates (J_{j}) . The number of times a statistic needs to be computed for a given country will depend on the number of implicit strata or sampling zones defined for the sample.

Note that when using the JRR technique for the estimation of sampling variability, the approach will appropriately reflect the combined effect of the between- and within-sampling zone contributions to the variance.

Doubling and zeroing the weights of the selected units within the strata or "zones" is accomplished effectively with the creation of replicate weights that ished 0ation onDoubling and zer3..1302 TD appr

7.5 Comparing Achievement with the International Mean

Many of the data exhibits in the *PIRLS 2001 International Report* (Mullis et al., 2003) show countries' mean achievement compared with the international mean, and you might be interested in replicating some of these analyses or doing some analysis of your own.

When comparing a country's mean with the international average, PIRLS took into account the fact that the country contributed to the international standard error. To correct for this contribution the sampling component of the standard error of the difference for country *j* was computed as follows:

(6)

N-1 se_{s_dif_j} =

differences between boys and girls once for each of the 75 replicate samples, and five more times, once for each plausible value. These differences are then summarized into an error of the difference using the formula presented earlier in this chapter.

Chapters 9 and 10 include macros that can be easily used to compute these differences by using contrast variables and regression models.

Content and Format of Database Files

8.1	Overview	8-2
8.2	The Data Files	8-3
	Data File and Variable Naming Conventions	
	Exhibit 8.1 PIRLS and Reading Literacy Data Files	8-5
	Exhibit 8.2 Country Identification and Availability of Data Files	8-6
	Exhibit 8.3 Background Questionnaire Variable Location Conventions	8-8
	Coding Convention	8-9
	Types of Variables Included in the Data Files	8-13
	PIRLS Background Data Files	8-17
	PIRLS Student Achievement Data Files	8-18
	PIRLS Scoring Reliability Data Files	8-18
	PIRLS Student-Teacher Linkage Files	8-20
	Reading Literacy Student Background and Achievement Data Files	8-20
8.3	Data Almanacs	8-20
	Exhibit 8.4 Data Almanac Files for PIRLS 2001 and Reading Literacy Study	8-22
	Background Item Data Almanacs	8-22

8

8.1 Overview

The database contains achievement data and student, home, teacher, and school background data collected in the 35 countries that participated in PIRLS. It also contains data for the nine countries that took part in the repeat of the 1991 Reading Literacy Study. In total, the PIRLS database contains responses of over 150,000 students, 130,000 parents, 7,000 teachers, and over 5,700 school principals collected between 1991 and 2001. This chapter describes the content and format of the PIRLS International Database. This chapter is organized in seven major sections corresponding to the types of files included in the database. Within each section, the contents of the files are described. These file types are:

Data Files

The PIRLS and Reading Literacy Study data files reflect the result of an extensive series of data management and quality control steps taken to insure the international comparability, quality, accuracy, and general utility of the database in order to provide a strong foundation for secondary analyses. They contain responses to background questionnaires administered to students, their parents, their teachers, and the principals of their schools. As part of the international data files, variables derived for reporting in the international reports are also included.

Data Almanacs

Data Almanac files contain weighted summary statistics for each participating country, on each achievement item and each variable in the student, home, teacher, and school background questionnaires.

Codebook Files

These specifically document the structure of each of the data files as well as information about the format of all variables in each of the data files.

Program Files

Several program files are provided for use in secondary analyses, including Data Access Control files for converting the raw data files provided in ASCII format into SAS or SPSS files, macro programs for computing statistics using the jackknife repeated replication method and using the plausible values discussed in Chapter 7, and macro programs for converting achievement item response codes to score values used in the computation of international scores.

- The first character of the files is always "A." This indicates that the file refers to 4th grade data. In other IEA studies the letter "B" has been used for 8th grade data and the letter "C" for 12th grade data.
- The second character indicates the source or level of the information in the file:
 - C indicates a school-level file
 - T indicates a teacher-level file
 - S indicates a student-level file.
- The third character indicates the type of the data in the file. The following abbreviations are used:
 - A Student Achievement Booklets for PIRLS
 - C Student Achievement Booklets and Background Questionnaires for the Trends in Reading Literacy Study
 - G General Background Questionnaires (School and Student Questionnaires) for PIRLS
 - T Teacher Background Questionnaires for PIRLS
 - R Constructed-Response Scoring Reliability (scoring reliability sample of student test booklets) for PIRLS
 - H Home Background Questionnaires for PIRLS
 - T Student-Teacher Link Files for PIRLS
- The next three digits identify the country with the three-character alphanumeric country abbreviation following the ISO coding scheme. Exhibit 8.2 lists the ISO code for each participating country.
- The seventh and eighth characters indicate the study cycle:
 - R1 stands for the 2001 PIRLS files
 - T1 stands for the 1991 Reading Literacy Study files
 - T2 stands for the 2001 Reading Literacy Study files.
- The three-character file extension used for the data files is .DAT.

For each file type, a separate data file is provided for each participating country. Countries for which data is not available have an empty file. The file names in Exhibit 8.1 illustrate the naming conventions for each of the data files.

All PIRLS and Reading Literacy Study data files are provided in ASCII format, and are also available in SAS and SPSS export format. All details of the file structure are provided in codebook files related to each of the data files. The use of these codebooks is described later in this chapter.

Background Variable Naming Convention

International background variables obtained from the student, home, teacher, and school questionnaires are provided in the corresponding background data files. In general, the background variables are provided for all countries where the data are considered internationally comparable. In a few cases, some slightly modified specific country options were retained in the international variables. The Codebook files contain the international background variable names, descriptive labels, response code definitions, formats, and field locations corresponding to each questionnaire item.

In addition to the background variables contained in the questionnaires, a number of derived variables were computed for use in the international report. These derived variables, many of which use data from more than one source, are also included in the International Database for use in secondary analyses. The derived variables based on background variables can be used to reproduce the values shown in the international report tables by applying the appropriate school, teacher or student filters and weights.

The naming convention for the background variables permits the identification of the survey population and data source based on 7- or 8-character codes. These are the variables that capture the responses given to the survey instruments (background questionnaires). The following convention is followed in naming these variables:

- The first character indicates the population. "A" is used to indicate fourth grade.
- The second character indicates the type of respondent. The following abbreviations are used:
 - C School Principal
 - T Teacher
 - S Student or Parent.
- The third character indicates the type of question. The following abbreviations are used:
 - B Background Questionnaire

- The second and third character indicates the assessment year when the item was first used in PIRLS. The code "01" is used for PIRLS 2001.
- The fourth character indicates the population for which the item was developed originally. The code "1" is always used in PIRLS.
- The fifth character indicates the passage where the item was used. The following codes are used:
 - A Antarctica
 - C Clay
 - F Flowers
 - H Hare
 - L Leonardo
 - M Mice
 - N Pufflings
 - R River Trail.
- The sixth and seventh character is a sequential number that identifies the sequence of the item within the block.
- The eighth character indicates the item type. The following codes were used:
 - C Constructed-Response Question (open ended)
 - M Multiple-Choice Question.

For the Reading Literacy Study test items, the following convention was used:

- The first character indicates the population for which the item was developed. The code "A" is used to indicate the fourth grade.
- The second character indicates the reading type for the text that accompanies the item. The following codes were used:

Background Item Response Code Values

The values assigned to each of the background item variables depend on the item format and the number of options available. For the multiple-choice items, one-digit numerical values are used to correspond to the response option. This number corresponds to the sequence of the letter in the alphabet. For example, response option A is represented with a 1, response option B with a 2, etc. Open-ended items such as "number of students in a class" are coded with the actual number given as a response to the question.

Achievement Item Response Code Values

The values assigned to each of the achievement item variables also depend on the item format. For the multiple-choice items, one-digit numerical values of 1-4 are used to correspond to the response options A through D. For these items, the correct response is included as part of the item-variable label in the Codebook files, and program code is included as part of the database to score these items.

For the constructed-response achievement test items, one-digit numerical codes are used that correspond to fully-correct, partially-correct, and incorrect responses for each item.

In addition to the correctness score information, specific missing codes are also defined that are described in the section discussing missing codes. Since all achievement item variables are included for all students in the achievement files regardless of which test

booklet they completed, a "Not Administered" code is given to all items that were not included in the test booklet assigned to each student.² However, only the variables for the items administered in a particular year are included in the corresponding files.

Missing Code Values

All values assigned to variables in the PIRLS international data files are numeric, and a subset of the numeric values for each of the variable types is reserved for specific

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

Omitted response codes are used for questions/ items that a student, parent, teacher, or school principal should have answered but did not answer. These are coded as "omitted" in the codebooks. For questionnaire data, no differentiation has been 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 way were also recoded to "omitted." For achievement items, an omitted response code was given only in cases in which the item was left blank; a special code was used for invalid answers as described below. The length of the omitted response code given to a variable in the ASCII file depends on the number of characters needed to represent the variable. In all cases the space necessary to represent the variable was filled with 9's.

Uninterpretable Response Codes (ASCII: 7; SAS: .N; SPSS: 7)

For the achievement test items, separate codes were established to distinguish between totally blank responses (omitted) and uninterpretable or invalid responses. For achievement items, cases where more than one response option was checked, or where an uninterpretable response was given, were coded with a 7.

Not Administered Codes (ASCII: 8, 98, 998, ...; SAS: .A; SPSS: 8, 98, 998, ...)

Special codes were given for items that were "not administered" to distinguish these cases from data that are missing due to non-response. The specific not administered code value given depends on the length of the field for the variable. In general, the Not Administered Code was used when an item was not administered as part of the questionnaire or test instruments either by design,

- Student absent from session If a student or individual was not present for a particular testing session, then all variables referring to that session have been coded to "Not Administered." However, if a student participated in a session and did not answer any of the items, these questions have been coded to "Omit."
- Item left out or misprinted If a particular question or item (or a whole page) was misprinted or not available to the student, teacher, or school, the corresponding variables have been coded to "Not Administered."
- Achievement items omitted or mistranslated in student test booklets Any items identified during the translation verification or item analysis processes that were mistranslated, such that the nature of the question was altered, were removed for a country.
- Background questionnaire items were omitted questions in the student, home, teacher, or school background questionnaires that were considered not applicable in some countries were not included in their questionnaires.
- Background questionnaire items were mistranslated or not internationally comparable – In some cases, questions in the international version of the questionnaires were mistranslated or modified to fit the national situation.
 Whenever possible, modified background questionnaire items were recoded to match as closely as possible the international version of the items. This

IDCLASS

Identification number that uniquely identifies the sampled class within the school.

IDBOOK

Identifies the specific test booklet (0 through 9) that was administered to the student. The code "0" is used for the reader.

Additional identification variables in the teacher files include the following:

IDTEACH

Identification number that uniquely identifies the selected teacher within the school. It is a hierarchical identification number formed by the combination IDSCHOOL and a two-digit sequential number within each school. This variable is unique to each teacher within each country but is not unique in the teacher file.

IDLINK

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

In the Student Background file, the IDSTUD variable provides a unique identification number to identify each student within each country. Snce teachers may teach more than one class, the IDTEACH and IDLINK combinations in the Teacher Background files provide a unique identification for each teacher teaching a specific class. Teacher background variables are linked to appropriate students using the Student-Teacher Linkage file. The variable IDSCHOOL, contained in all three background files, is a unique identification number for each school within a country that may be used to link school background data to corresponding students or teachers.

Linking and Tracking Variables

Information about students, teachers, and schools provided on the survey tracking forms⁴ is included in linking or tracking variables. These variables have prefixes of IL or IT. Some of the important linking and tracking variables are listed below.

. . .

⁴ Survey tracking forms are listings of students, teachers, or schools used for sampling and administration purposes.

Linking and tracking variables in the Student Background Files include the following:

ITSEX Gender of each student as stated in the student tracking form.

ITBIRTHM and ITBIRTHY

Month and year of birth of each student as stated in the student tracking forms.

ITDATEM and ITDATEY

Month and year of testing for each student.

ITLANG⁵

Sampling and Weighting Variables

Several sampling and weighting variables are included in the student data files. These variables are described in Chapter 5. The variables are the following:

MOTELOA	
WGTFAC1	School Weighting Factor. This variable is included in the Student and in the School Background data files.
WGTADЛ	School Weighting Adjustment. This variable is included in the student and in the School Background data files.
WGTFAC2	Class Weighting Factor. This variable is included in the Student data files.
WGTFAC3	Student Weighting Factor. This variable is included in the Student data files.
WGTAD	Student Weighting Adjustment. This variable is included in the Student data files.
TOTWGT	Total Student Weight. This variable is included in the Student data files.
SENWGT	Student Senate Weight. This variable is included in the Student data files.
HOUWGT	Student House Weight. This variable is included in the Student data files.
TCHWGT	Overall Teacher Weight. This variable is included in the Student-Teacher Linkage data files.
SCHWGT	School-level Weight. This variable is included in the School Background data file.
J KZONE	The sampling zone or stratum to which the student's school is assigned. This variable is included in the Student and Student-Teacher Linkage data files.
J KREP	The PSU to which the student is assigned. This variable is included in the Student and Student-Teacher Linkage data files.
J KCZONE	The sampling zone or stratum to which the school is assigned. This variable is included in the School Background data files.
JKCREP	The PSU to which the school is assigned. This variable is included in the School Background data files.

Achievement Variables

Several achievement variables are also included in the Student data files. These variables are described in Chapter 6.

PIRLS Background Data Files

There are four different types of PIRLS background files – Student, Home, Teacher, and School. These are described below.

Student Background File

Students who participated in PIRLS were administered a background questionnaire with questions related to home background and school experiences. The Student Background file contains students' responses to these questions. The Student Background files also contain a series of identification variables, link variables, sampling variables, achievement variables, and the derived variables that were used for the creation of the international reports.

Home Background File

The parents or primary caregivers of each student taking the PIRLS reading assessment completed a learning to read survey or home questionnaire. This questionnaire addresses child-parent literacy interactions, home literacy resources, parents' reading habits and attitudes, home-school connections, and b8Tj/TTilet lind 6ographTild schocioenneomT

country, including the reliability booklet set to which each student was assigned (ITBSET) and the identification of the first and second coders (IDSCORA and IDSCORR).

PIRLS Student-Teacher Linkage Files

The Student-Teacher Linkage files for PIRLS contain information required to link the Student and Teacher files and to compute appropriately weighted teacher-level data using the student as the unit of analysis. The Student-Teacher Linkage files contain one entry per student-teacher linkage combination in the data. In some cases, students are linked to more than one teacher, and in these cases there will be one record for each student-teacher link. In addition, in some countries some students may also have more than one teacher. For instance, if three teachers are linked to a student, there are three entries in the file corresponding to that student.

Reading Literacy Student Background and Achievement Data Files

The students that participated in the Reading Literacy study were administered a reading test and a student background questionnaire. The information from both these instruments is included already combined into one file with the responses to both instruments.

8.3 Data Almanacs

Data Almanac files contain weighted summary statistics for each participating country on each achievement item and on each variable in the student, home, teacher, and school background questionnaires. Separate data almanacs are included for the background variables and for the achievement item variables. The Data Almanac files corresponding to each variable type are listed in Exhibit 8.4 and are described in the following sections.

The naming convention of the almanac files is as follows:

- The first character of the files is always "A." This indicates that the file refers to fourth grade data.
- The second character indicates the source or level of the information in the file:
 - S indicates a background, student level almanac
 - I indicates an achievement item almanac.
- The third, fourth and fifth character identify the file as an almanac and it is always "ALM."

- The sixth character identifies the information contained in the almanacs. The following codes are used for the sixth character:
 - 1 Student Background Questionnaire with Reading Achievement
 - 3 School Background Questionnaire with Reading Achievement
 - 5 Teacher Background Questionnaire with Reading Achievement
 - 9 Home Background Questionnaire with Reading Achievement
 - I Reading to Acquire and Use Information achievement items
 - L Reading for Literary Purposes achievement items

There are two types of displays in the almanacs, depending on whether the item is a multiple-choice item or a free-response item. The statistics displayed in these

the following sections. One Codebook file is provided for each of the data files listed earlier.

• The following three-character file extensions are used:

SAS for a SAS control file

SPS for an SPSS control file.

Chapters 9 and 10 further describe the SAS and SPSS program files and how they are applied through the use of specific example analyses using the PIRLS Student, Home, Teacher, and School Data files.

8.6 Item Information Files

Item Information files are provided to enable users of the database to readily produce summaries of item characteristics and cross-reference the different item identification numbers. The Item Information files include the following information:

- Permanent ID number for the item
- Item Name and Cluster Location in the corresponding year
- Test Subject (Mathematics or Science)
- Item Type (Multiple choice or Free Response)
- Response Key (correct response option for multiple-choice items only)
- Maximum Points (maximum score points possible on the item)
- Whether the item was used for scaling
- Release status for the year
- Number of options
- Content Area Subscale or Reporting Category
- Item Label
- Performance Expectation Category.

There are two Item Information files containing this information for the test items. These files are located in the subdirectory called ITEMS for the corresponding study. The two Item Information files are:

AITINFR1.TXTItem Information for the PIRLS 2001 Test ItemsAITINFT1.TXTItem Information for the Reading Literacy Study Test Items

These files are in text format, with their fields separated by tabs, and there is one record for each achievement item part. They may, therefore, be opened and manipulated directly with a spreadsheet program.

Exhibi	Stu	SS Control Statements for Performing Analyses with Ident-Level Variables Involving Plausible Values (AMPLE2.SPS)	9-33
Exhibi		ract of SPSS Computer Output for Performing Analyses with dent-Level Variables Involving Plausible Values (EXAMPLE 2)	9-34
Performin		ses with Teacher-Level Variables	~ ~ ~
i enternim	ig Analys	es with reacher-Lever variables	9-35
Exhibi	t9.12 Sar	nple Exhibit for Teacher-Level Analysis Involving Plausible ues Taken From the PIRLS 2001 International Report	9-35 9-36

9.7

9.1 Overview

This chapter presents some basic examples of analyses that can be performed with the PIRLS International Database using the sampling weights and scores discussed in previous chapters. It also provides details on selected SPSS programs to conduct such analyses, and the results of these analyses. The analyses presented here are simple in nature, and are designed primarily to familiarize you with the different files and their structure, as well as the relevant variables that need to be used in most analyses. The programs compute the percentage of students in specified subgroups, the mean reading achievement for those groups, and the corresponding standard errors for the percentage and mean statistics. Additionally, some examples of regression statistics are presented. These analyses that are included in the *PIRLS 2001 International Report* (Mullis et al., 2003). You are encouraged to practice analyzing the PIRLS data by replicating the analyses used to produce the exhibits that are presented in the international report.

In our examples we use macros written for SPSS that can be used to perform any of the analyses that are described in this chapter. These are general procedures that can be used for many purposes, provided you have some basic knowledge of the SPSS

The following programs also can be found in this sub-directory:

ASASCRR1.SPS, ASASCRT1.SPS and ASASCRT2.SPS

These files contain SPSS programs that can be used to convert the response codes to the cognitive items to their corresponding correctness score levels. The use of these programs is described in this chapter.

JACKPV.SPS

This macro program in SPSS can be used to compute weighted percentages of students within defined groups, and their mean achievement scores on an achievement scale using plausible values. This macro makes use of the plausible values in computing the mean achievement scores. This macro also generates replicate weights and computes the jackknife repeated replication (JRR) sampling variances for the percentages of students within the group, and the JRR and imputation variances for the mean achievement scores. This macro should only be used when multiple plausible values are used in the analysis.

JACKGEN.SPS

This macro program in SPSS can be used to compute weighted percentages of students within defined groups, and their means on a specified continuous variable. This macro also generates replicate weights and computes JRR sampling variances for the percentages and mean estimates. The variable can be any continuous variable in the file. How to use each of these macro programs is described later in this chapter. If computing with plausible values, you will need to use the macro JACKPV.SPS.

JACKREG.SPS

This macro program in SPSS can be used to compute regression coefficients and their corresponding standard errors within defined groups. This macro can be used with any variable in the analysis but it does not make use of plausible values.

JACKREGP.SPS

This macro program in SPSS can be used to compute regression coefficients and their corresponding standard errors when using plausible values as the dependent variables within defined groups.

Each of the macros above has a corresponding sample program that calls the macro and prints out the results. These programs are discussed later in the chapter.

EXAMPLE1.SPS, EXAMPLE2.SPS, EXAMPLE3.SPS, EXAMPLE4.SPS, EXAMPLE5.SPS These are the programs used in the samples presented later in this chapter. 5.

Code 6	Not reached
Code 7	Invalid response (e.g., a student chose more than one of the options available)
Code 8	Not administered
Code 9	No response although the item was administered and was reached (i.e., item was omitted).

The remaining items were constructed-response type, where the students were asked to construct a response to a question, rather than choosing an answer from a list of options. Responses were either short-answer or extended-response, depending upon the complexity of the task. Constructed-response items were worth a total of one, two, or three points. Scorers trained to use the scoring rubrics described in Chapter 2 of this guide scored the answers to these questions. The codes used to represent the responses to these items are the following:

Code 3	Three-point answer
Code 2	Two-point answer
Code 1	One-point answer
Code 0	Zero-point answer or uninterpretable response
Code 6	Not reached
Code 8	Not administered.

responses) for the purpose of calibrating the items. But these same responses were then recoded as incorrect when scoring the item for the individual countries, and for the purpose of calculating the scale scores for individuals. By default, the scoring program provided with the database recodes the items coded as not reached and those left blank as incorrect responses.

To use the SCOREIT macro you need to include it as part of the SPSS programs

```
Item = <list items where option B is the correct one> /
RIGHT = 2 / nr = 6 / na = 8 / om = 9 / other = 7.
SCOREIT Type = MC /
Item = <list items where option C is the correct one> /
RIGHT = 3 / nr = 6 / na = 8 / om = 9 / other = 7.
SCOREIT Type = MC /
Item = <list items where option D is the correct one> /
RIGHT = 4 / nr = 6 / na = 8 / om = 9 / other = 7.
SCOREIT Type = OE /
Item = <list open-ended items> /
RIGHT = 0 / nr = 96 / na = 98 / om = 99 / other = 90.
```

9.5 Basic Analyses with the PIRLS Data: Means, Percentages, Regression Coefficients, and Their JRR Standard Errors

In this section, four macros that can be used to compute the correct standard errors of sampling and imputation are described – including examples in which these macros are used to replicate exhibits in the *PIRLS 2001 International Report* and almanacs.

Computing Sampling and Imputation Variance for Plausible Values Using SPSS (JACKPV.SPS)

This section presents example SPSS code that can be used to compute the JRR standard errors for mean plausible values and percentages. This code is provided in the form of an SPSS macro that computes the percentages of students within subgroups defined by a set of classification variables, the JRR standard errors of these percentages, the means for the groups on one of the achievement scales using plausible values, and the standard errors of these means including the sampling and imputation variance components.

When using this macro, you need to specify a set of classification variables, the name of the plausible values and how many there are, the number of replicate weights to be generated, the variables that contain the sampling information such as JKZONE and JKREP, and the sampling weight that is to be used for the analysis. You will also need to specify the data file that contains the data to be processed.

You need to know some basic SPSS macro language in order to use JACKPV.SPS. The macro should be included in the program file where it is going to be used. If you are operating in batch mode, then the macro should be called in every batch. If you are using SPSS interactively, then the macro should be called once at the beginning of the session and it will remain active throughout the session. If the session is terminated and restarted at a later time the macro should be called once again. Once the macro is included in a specific session the word "JACKPV" should not be used within that session because doing so will call the macro. The macro is included in the program

JKR	The variable that captures whether the case is to be dropped or have its weight doubled for the corresponding replicate weight. The name of this variable in all PIRLS files is JKREP.
NJKZ	This indicates the number of replicate weights to be generated when computing the JRR error estimates. When conducting analyses using the data from all countries, the value of NJKZ should be set to 75, the maximum possible value. When you are working with the data for only one country, you should set the NJKZ argument to as many replicates as are needed in the country. If the data from two or more countries are being used for an analysis, then the largest number of jackknife zones should be used. When in doubt about what number to set the NJKZ parameter to, set it to 75. The error variance will always be estimated correctly if more replicate weights than necessary are computed, but will be underestimated if you specify fewer.
WGT	The sampling weight to be used in the analysis, generally TOTWGT when using the student files, or TCHWGT when using the teacher files.

The simplest and most straightforward way to call the macro is by using the conventional SPSS notation for calling macros. This involves listing the macro name followed by the corresponding list of arguments for the analysis, each separated by a slash. For example, if the macro is called using the following code

```
include "c:\pirls\programs\jackpv.sps".
jackpv infile= "c:\pirls\data\asgallrl.sav" /
    cvar = idcntry itsex /
    pvs = ASRREA01 to ASRREA05 /
    npv = 5 /
    jkz = JKZONE /
    jkr = JKREP /
    njkz = 75 /
```

The file that contains these results is then called FINAL and is saved to the default directory being used by SPSS. The variables that are contained in this file are:

Classification Variables

Each of the classification variables is kept in the resulting file. In the above example there are two classification variables in the FINAL data set, IDCNTRY and ITSEX. There is one unique occurrence for each combination of the categories for these variables.

Weight Variable

Contains the estimate of the population size of the groups defined by each specific combination of the classification variable categories. In our example this

MNPV

Contains the means of the plausible values used in the analysis.

MNPV_SE

Contains the standard errors for the means of the plausible values. These standard errors contain the sampling and the imputation components of the errors of the estimates.

The file resulting from using this macro can be printed using an SPSS procedure of choice. An example call to this macro and a printout of the resulting file is presented in Exhibit 9.3. This code is included in the file SampleJackPV.SPS on the CD that accompanies the User Guide.

Exhibit 9.3 SPSS Control Code and Extract of Output File for Using the Macro JACKPV.SPS

```
get file = "c:\pirls\data\asgallr1.sav"
 / keep=idcntry idstud idgrader jkrep jkzone totwgt
        itsex asrrea01 to asrrea05 inprl01.
select if (itsex=1 or itsex=2) and inprl01=1.
save outfile = student.
include "c:\pirls\programs\jackpv.sps".
jackpv infile= student
                          1
       cvar = idcntry itsex /
       pvs = ASRREA01 TO ASRREA05 /
             = 5
       npv
                          /
              = JKZONE
       jkz
                          /
             = JKREP
       jkr
                          /
       njkz = 75
       WGT
             = TOTWGT
print formats identry itsex n (F6.0) totwgt (f10.0)
      mnpv mnpv_se mnx mnx_se pct pct_se (f6.2).
report format=list
 / var = identry itsex n totwgt mnpv mnpv_se mnx mnx_se pct pct_se.
TDCNTRY
                        TOTWGT
                                        MNPV_SE
                                                                            PCT_SE
         TTSEX
                                  MNPV
                                                MNX
427.63
407
                                                           MNX SE
                                                                     PCT
                   N
                 1647
                                                                   50.68
  32
                         358324 428.18
                                          6.24
           1
                                                            6.19
                                                                              1.06
   32
           2
                 1645
                         348747 410 04
                                                                   49 32
                                           6 50
                                                            6 17
                                                                              1 06
  84
                 1415
                          3573 341.73
                                                 342 03
           1
                                           5.44
                                                            5.27
                                                                   49 71
                                                                               .93
  84
           2
                 1414
                          3615
                                314.56
                                           5.37
                                                 312.68
                                                            5.20
                                                                   50.29
                                                                               .93
  100
           1
                 1783
                         49229
                                562 29
                                           3.69
                                                 562.42
                                                            3.58
                                                                   51.47
                                                                               .91
                          46418 538.06
 100
           2
                 1675
                                           4.67
                                                 537.63
                                                            4.63
                                                                   48.53
                                                                               .91
                        109940 552 89
                                                                               .67
  124
           1
                 4121
                                           2.62
                                                 553 12
                                                            2.56
                                                                    49 53
                                                                               .67
           2
                                535.58
 124
                 4131
                        112039
                                           2.55
                                                 535.67
                                                            2.52
                                                                   50.47
  705
                         10551 512 34
                                                                   50 09
                                                                               .82
           1
                 1468
                                           2.52
                                                 511.39
                                                            2.44
  705
           2
                 1484
                         10515
                                490.65
                                           2.38
                                                 490.97
                                                            2.28
                                                                    49.91
                                                                               .82
  752
           1
                 2965
                          57583
                                572 27
                                           2.62
                                                 572 06
                                                            2 53
                                                                    48 74
                                                                               .73
                                                                               .73
           2
                 3079
                         60551
                                                            2.45
  752
                                550.31
                                           2.53
                                                 550.83
                                                                   51.26
  792
           1
                 2426
                         470383
                                459.37
                                           3.96
                                                 458.57
                                                            3.90
                                                                    48.13
                                                                               .92
  792
           2
                 2699
                         506933 440.06
                                           3.73
                                                 440.15
                                                            3.69
                                                                   51.87
                                                                               .92
  807
           1
                 1814
                         13392
                                452.24
                                           5.14
                                                 452.15
                                                            5.06
                                                                    48.94
                                                                               .88
  807
           2
                 1897
                          13973
                                431.37
                                           4.76
                                                 431.15
                                                            4.73
                                                                   51.06
                                                                                .88
```

840	1	1911	1927005	551.06	3.80	551.97	3.73	50.68	.82
840	2	1852	1875552	532.99	4.89	531.22	4.66	49.32	.82

In this example, the mean of all five plausible values and the mean of the first plausible value and their corresponding standard errors for reading achievement are calculated separately for boys and girls by country. In the listing of the results we can see that there are entries for each country corresponding to the results for females (ITSEX = 1) and males (ITSEX = 2). The first column has the country code, the second column indicates the gender of the students, the third column has the number of students in each sample, and the fourth column has the total weight of this sample in the population. This is followed by the mean of all five plausible values in reading achievement and the corresponding standard error. Following this is the mean of the first plausible value in reading achievement and its corresponding standard error. The last two columns indicate the percentage of girls and boys in each country's sample and their corresponding standard errors.

For example, Argentina (IDCNTRY=32) sampled 1,647 girls representing 358,324 students in the whole population. The mean of the five plausible values for these girls is 428.18 with a standard error of 6.24. Girls made up 50.68 percent of Argentina's sampled students. The mean of the first plausible value for the girls is 427.63 with a standard error of 6.19. Additionally, Argentina sampled 1,645 boys representing 348,747 students in the whole population. The mean of the five

plausible values for these boys is 410.04 with a standard error of 6.50. Boys made up 49.32 percent of Argentina's sampled students. The mean of the first plausible value for the boys is 407.13 with a standard error of 6.17.

Computing Sampling Variance for Variables Other Than Plausible Values Using SPSS (JACKGEN.SPS)

You need to know some basic SPSS macro language in order to use the macro. First, it needs to be included in the program file where it is going to be used. If you are operating in batch mode, then the macro needs to be called in every batch. If you are using SPSS interactively, then the macro needs to be called only once at the beginning of the session and it will remain active throughout the session. If the session is terminated and restarted at a later time the macro needs to be called once again. Once the macro is included in a specific session, the word "JACKGEN" should not be used within that program because doing so will call the macro.

The macro is included in the program file where it will be used by issuing the following command under SPSS:

include `<path>jackgen.sps'.

where <path> points to the specific drive and directory where the macro JACKGEN.SPS can be found. The macro requires that several parameters be specified when it is called. These parameters are:

INFILE	The name of the data file that contains the variables necessary for the analysis. If the path location is included as part of the file name the name of the file has to be enclosed in quotes. It is important to emphasize that this system file must include only those cases that are of interest in the analysis. If you want to have specific cases excluded from the analysis (e.g., students with missing variables or selected students from a specific grade), this should be done prior to calling the macro.
CVAR	This lists the variables that are to be used to classify the students in the data file. This can be a single variable, or a list of variables. The maximum number of variables will depend mostly on the computer resources available to you at the time. It is recommended to always include the variable that identifies the country. At least one variable has to be specified, usually IDCNTRY.
DVAR	This is the variable for which means are to be computed. Only one variable can be listed here. If you want to examine, for example, results in two different variables, then the macro needs to be called separately to generate each table.
NJKZ	This indicates the number of replicate weights to be generated when computing the JRR error estimates. When conducting analyses using the data from all countries, the value of NJKZ should be set to 75, the maximum possible value. When you are working with the data for only one country, you should set the

	NJKZ argument to as many replicates as are needed in the
	country. If the data from two or more countries are being used
	for an analysis, then the largest number of jackknife zones
	should be used. When in doubt about what number to set the
	NJKZ parameter to, set it to 75. The error variance will always
	be estimated correctly if more replicate weights than necessary are computed, but will be underestimated if you specify fewer.
JKZ	The variable that captures the assignment of the student to a particular sampling zone. The name of this variable in all PIRLS and Reading Literacy Study files is JKZONE.
JKR	The variable that captures whether the case is to be dropped or have its weight doubled for the corresponding replicate weight. The name of this variable in all PIRLS and Reading Literacy Study files is JKREP.
WGT	The sampling weight to be used in the analysis, generally TOTWGT when using the student files, or TCHWGT when using the teacher files.

The simplest and most straightforward way to call the macro is by using the conventional SPSS notation for invoking macros. This involves listing the macro name followed by the corresponding list of arguments for the analysis, each separated by a slash. For example, if the macro is called using the following code:

```
include "c:\pirls\programs\jackgen.sps".
jackgen
      infile = asgallr1
                                     /
      cvar = idcntry
dvar = asdgage
                                     /
                                     /
      njkz = 75
                                     /
      jkz
jkr
              = jkzone
= jkrep
                                     /
                                     /
      wgt
              = totwgt
                                     .
```

Weight Variable

Contains the estimate of the population size of the groups defined by each specific combination of the classification variable categories. In our example this variable is called TOTWGT.

MNX

Contains the means of the variable DVAR for the groups defined by the corresponding combinations of classification variable categories.

MNX_SE

Contains the standard errors of the MNX values computed using the JRR method for computing the standard error.

РСТ

Contains the percentages of people in the groups for the classification variable listed last, within the specific combination of the categories defined by the groups initially. In our example, we would obtain the percentage of students by grade for each country.

PCT_SE

Contains the standard errors of PCT computed using the JRR method for computing the standard error.

The file resulting from using this macro can then be printed using the SPSS procedure of choice. An example call to this macro, and a subset of the resulting file, is presented in Exhibit 9.4. In this example, the macro will compute the percentages

Exhibit 9.4 SPSS Control Code and Extract of Output File for Using the Macro JACKGEN.SPS

```
get file = "c:\pirls\data\asgallr1.sav"
    / keep=idcntry idstud inprl01 jkrep jkzone totwgt
    asdgage.
select if not(missing(asdgage)) and inprl01=1.
save outfile = student.
include "c:\PIRLS\programs\jackgen.sps".
jackgen
    infile = student /
    cvar = idcntry /
    dvar = asdgage /
    njkz = 75 /
    jkz = jkzone /
    jkr = jkrep /
    wgt = totwgt .
print formats idcntry n (f6.0) totwgt (f10.0) mnx mnx_se pct pct_se (F6.2).
```

Computing Regression Coefficients and Their JRR Standard Errors for Variables Other Than Plausible Values (JACKREG.SPS)

In this section, example SPSS code that may be used to compute regression coefficients and their JRR standard errors is described. The CD containing the PIRLS International Database contains the SPSS macro program called JACKREG.SPS. The macro computes the multiple correlation between the specified dependent and independent variables within a subgroup defined by a set of classification variables, as well as the regression coefficients and the JRR standard error of the regression coefficients.

If you wish to conduct regression analyses using plausible values as the dependent variable, please refer to the macro JACKREGP.SPS described in the next section.

When using this macro, you need to specify a set of classification variables, the dependent and independent variable, the number of replicate weights to be generated, the variables that contain the sampling information such as JKZONE and JKREP, and the sampling weight that is to be used for the analysis. You will also need to specify the data file that contains the data that is to be processed.

You need to know some basic SPSS macro language in order to use the macro. First, it needs to be included in the program file where it is going to be used. If you are operating in batch mode, then the macro needs to be called in every batch. If you are using SPSS interactively, then the macro needs to be called once at the beginning of the session and it will remain active throughout the session. If the session is terminated and restarted at a later time the macro needs to be called once again. Once the macro is included in a specific session the word "JACKREG" should not be used within that program because doing so will call the macro.

The macro is included in the program file where it will be used by issuing the following command under SPSS:

include `<path>jackreg.sps'.

where <path> points to the specific drive and directory where the macro JACKREG.SPS can be found. The macro requires that several parameters be specified when it is called. These parameters are:

INFILE The name of the data f5.1(ii6ac)5.14tf527.4cTsto te is cals mmand ero

CVAR This lists the variables that are to be used to classify the students in the data file. This can be a single variable, or a list of variables. The maximum number of variables will depend mostly on the computer resources available to you at the time. It is recommended to always include the variable that identifies the The simplest and most straightforward way to call the macro is by using the conventional SPSS notation for calling macros. This involves listing the macro name followed by the corresponding list of arguments for the analysis, each separated by a slash. For example, if the macro is called using

```
include "c:\pirls\programs\jackreg.sps".
jackreg infile = asgallr1 /
    cvar = idcntry idgrader /
    xvar = regsex /
```

of the estimate. The first regression coefficient (B01) is the difference in the average value of the variable ASBGTVDY between the boys and the girls, and the standard error of this coefficient (B01_SE) is the jackknifed standard error of this difference.

For example, from the fourth line in the output we can say that in Canada (IDCNTRY = 124), data is available for 8,185 cases. The coefficient of determination between gender and hours of watching TV or video is .022 with girls, on average, reporting watching less TV or videos every day (difference in average score of 0.35). This difference is statistically significant (0.35 / .04 = 8.75). Overall, we can say that on average, and across most countries (except Argentina), girls report watching less TV or videos than boys. This difference is significant in most cases, which can be determined by dividing the value of B01 by its standard error and comparing it to the appropriate critical value.

Computing Regression Coefficients and Their JRR Standard Errors with Plausible Values (JACKREGP.SPS)

In this section, example SPSS code that can be used to compute the JRR standard errors for regression coefficients using plausible values as the dependent variable is described. This code is provided in the form of an SPSS macro called JACKREGP.S monhe JRmacro computes the average multiple correlation coefficient

between the specified plausible values and independent variables within subgroups defined by a set of classification variables, as well as the regression coefficients and their JRR standard errors.

When using th JRmacro, you need to specify a set of classification variables, the independent and dependent variables, the number of replicate weights to be generated, the variables that contain the sampling information such as JKZONE and JKREP, and the sampling weight that is to be used for the analysis. You will also need to specify the data file that contains the data that is to be processed.

You need to know some basic S mSRmacro language in order to use the macro. First, it needs to be included in the program file where it JRgoing to be used. If you are operating in batch mode, then the macro needs to be called in every batch. If you are using S mS nteractively, then the macro needs to be called once at the beg nning of the session, and will remain active thrn3(c3t)5.3(mpa.5(uiy2eSR)5.2(in t)5.1s som)5.5(e)-.3(ess1(e is)**T**J0 -

	for an analysis, then the largest number of jackknife zones should be used. When in doubt about what number to set the NJKZ parameter to, set it to 75. The error variance will always be estimated correctly if more replicate weights than necessary are computed, but will be underestimated if you specify fewer.
JKZ	The variable that captures the assignment of the student to a particular sampling zone. The name of this variable in all PIRLS files is JKZONE.
JKR	The variable that captures whether the case is to be dropped or have its weight doubled for the corresponding replicate weight. The name of this variable in all PIRLS files is JKREP.
WGT	The sampling weight to be used in the analysis, generally TOTWGT when using the student files, or TCHWGT when using the teacher files.

The simplest and most straightforward way to call the macro is by using the conventional SPSS notation for invoking macros. This involves listing the macro name followed by the corresponding list of arguments for the analysis, each separated by a slash. For example, if the macro is called using

```
include "c:\pirls\programs\jackregP.sps".
jackregp infile = asgallr1 /
    cvar = idcntry idgrader /
    xvar = regsex /
    rootpv = asrrea0 /
    npv = 5 /
    njkz = 75 /
    jkz = jkzone /
    jkr = jkrep /
    wqt = totwqt .
```

it will compute the regression equation for the variable REGSEX as a predictor of the plausible values in reading. The data will be read from the data set ASGALLR1 and the standard error of the statistics will be computed based on 75 replicate weights.

The file that contains these results is then called REG and is saved to the default directory being used by SPSS. The variables that are contained in this file are:

Classification Variables

Each of the classification variables is kept in the resulting file. In our example above there would be two variables in the resulting system file, IDCNTRY and IDGRADER. There is one unique occurrence for each combination of the categories for these variables.

- Use the macro JACKGEN with the corresponding arguments and parameters.
- Print out the result file.

Exhibit 9.7 SPSS Control Statements for Performing Analyses with Student-Level Variables NOT Using Plausible Values (EXAMPLE1.SPS)

```
title "Example 1: PIRLS 2001 User Guide".
add files
/ file = "c:\pirls\data\asgargr1.sav"
/ file = "c:\pirls\data\asgbgrrl.sav"
/ file = "c:\pirls\data\asgswerl.sav"
/ file = "c:\pirls\data\asgturr1.sav"
 / file = "c:\pirls\data\asgusar1.sav".
select if not(missing(asbgtvdy)).
value labels
identry 32 'Argentina' 100 'Bulgaria' 752 'Sweden' 792 'Turkey' 840 'United
States'.
recode asbgtvdy (1=0.0) (2=0.5) (3=2.0) (4=4.0) (5=7.0) (else=copy) into
tvwatch.
save outfile = student.
* Now use the macro to get the results.
include "c:\pirls\programs\jackgen.sps".
jackgen
     infile = student
                                /
      cvar = idcntry
                                /
     dvar = tvwatch
njkz = 75
                                /
                                /
      jkz = jkzone
                                /
      jkr = jkrep
wgt = totwgt
                         .
sort cases by identry .
print format n (f6.0) totwgt (f10.0) mnx mnx_se pct pct_se (f6.2).
report format=list automatic
/ var
       = n totwgt mnx mnx_se pct pct_se
 / break = idcntry .
```

```
sort cases by identry.
print formats n (f6.0) totwgt (f10.0) mnpv mnpv_se pct pct_se (f6.2).
report format=list automatic
   / var = asbgbook (label) n totwgt mnpv mnpv_se pct pct_se
   / break = identry .
```

Exhibit 9.11 Extract of SPSS Computer Output for Performing Analyses

while the 144 students (5.93 percent of the sampled students) who reported having "More than 200 books" have a mean reading achievement of 473.98 with a standard error of 12.58.

9.7 Performing Analyses with Teacher-Level Variables

The PIRLS teacher files do not contain representative samples of teachers within a country. Rather, these are the teachers for a representative sample of students within a country. Therefore, it is appropriate that statements about the teachers be made only in terms of how many students are taught by teachers of one kind or another, and not in terms of how many teachers in the country do one thing or another.

To analyze the information provided by the students' teachers, it is first necessary to link it with the student information. Each student record in the Student Background data file can be linked to one or more different teachers in the Teacher Background data file. To facilitate the linking between students and their teachers in the teacher file, the Student-Teacher Linkage file was created and is part of the International Database. These files are called AST<COUNTRY>R1. The Student-Teacher Linkage file contains one record for each student-by-teacher combination, with the corresponding identification variables.

Each record also contains the number of teachers for the student and a set of weights that can be used when conducting analyses with these data. Student achievement plausible values, sampling weights, and JRR replication information have been added to the Student-Teacher Linkage file in order to simplify the merging process for analyses that link teacher variables to student achievement. For such analyses it is necessary to merge only the Teacher Background file with the Student-Teacher Linkage file. For analyses linking teacher variables to other student variables, it is necessary also to merge the Student Background files with the Teacher Background file after it has been combined with the Student-Teacher Linkage file.

Conducting analyses with the teacher data requires some extra steps that are not required when analyzing the student or school background data.

For our example, we want to find out about the frequency of reading instruction reported by the teachers who teach the fourth-grade students in the PIRLS countries. In particular, we want to find out what percentage of fourth-grade students are taught by teachers who report having reading instruction "Every day," "3-4 days per week," or "Fewer than 3 days per week." We are also interested in the mean reading achievement of the students taught by these teachers. These results are reported in Exhibit 5.14 of the international report (replicated results are presented in Exhibit 9.12).

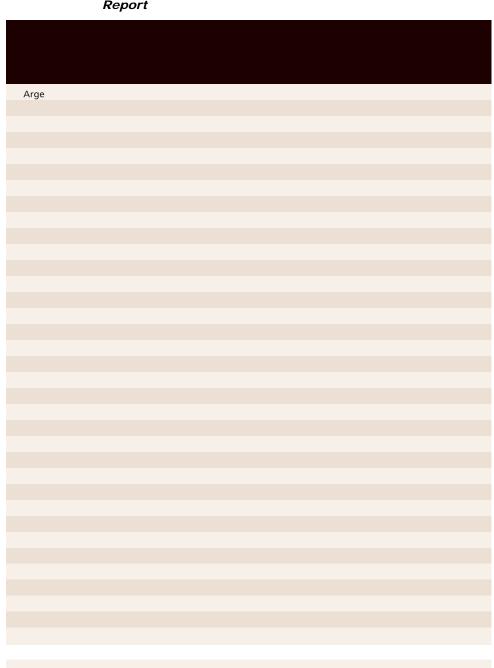


Exhibit 9.12 Sample Exhibit for Teacher-Level Analysis Involving Plausible Values Taken From *the PIRLS 2001 International Report* As before, we first proceed to identify the variables relevant to the analysis in the corresponding files, and review the documentation on the specific national adaptations to the questions of interest (on the CD). Since we are using teacher-level variables, we need to look into the teacher file and the Student-Teacher Linkage files to find the variables.

In our example, our teacher variable of interest (ATBGRACT) is a categorical variable with three categories.

We then proceed to read the necessary information from the Student-Teacher Linkage file and the Teacher Background file.

The two files are then merged or matched into one file that will then be used with the JACKPV macro. These two files will be merged using the variables IDCNTRY, IDTEACH, and IDLINK. The combination of values for these three variables is unique within the teacher data, but is repeated in the Student-Teacher Linkage file as many times as the specific teacher teaches students in a class. After the files are merged, the macro JACKPV is used and the results can be printed. The code in SPSS for this example is presented in Exhibit 9.13. Selections of the results obtained from this program are displayed in Exhibit 9.14.

Exhibit 9.13 SPSS Control Statement for Performing Teacher Level Analysis Using Plausible Values

```
title "Example 3: PIRLS 2001 User Guide".
add files
/ file = "c:\pirls\data\astargr1.sav"
/ file = "c:\pirls\data\astbgrr1.sav"
 / file = "c:\pirls\data\astswer1.sav"
 / file = "c:\pirls\data\astturr1.sav"
 / file = "c:\pirls\data\astusar1.sav"
sort cases by identry idteach idlink.
save outfile = studteac.
add files
/ file = "c:\pirls\data\atgargr1.sav"
 / file = "c:\pirls\data\atgbgrrl.sav"
/ file = "c:\pirls\data\atgswerl.sav"
/ file = "c:\pirls\data\atgturr1.sav"
 / file = "c:\pirls\data\atgusar1.sav".
sort cases by identry idteach idlink atbgract.
save outfile = teacher.
* Now merge the two files.
match files
 / file=studteac
  / table=teacher
  / by identry idteach idlink.
select if not(missing(atbgract)) and inprl01=1.
* Define the format for the variables used.
value labels
```

```
identry 32 'Argentina' 100 'Bulgaria' 752 'Sweden' 792 'Turkey' 840 'United
States' /
atbgract 1 'Every Day' 2 '3-4 Days per week' 3 'Fewer than 3 days per week'.
save outfile = merged.
* Now use the macro to get the results.
include "c:\pirls\programs\jackpv.sps".
jackpv infile= merged
                                  /
      cvar = identry atbgract /
       pvs = asrreall to asrrea05 /
npv = 5 /
jkz = JKZONE /
       jkr = JKREP /
njkz = 75 /
WGT = TCHWGT .
sort cases by identry.
print formats n (f6.0) tchwgt (f10.0) mnpv mnpv_se pct pct_se (f6.2).
report format=list automatic
  / var = atbgract (label) n tchwgt mnpv mnpv_se pct pct_se
  / break = idcntry .
```

Exhibit 9.14 Extract of SPSS Output for Performing Teacher Level Analysis Using Plausible Values (EXAMPLE 3)

```
Example 3: PIRLS 2001 User Guide

PAGE 1

GEN/RDG INSTR WITH STD/HOW

* 0.00 rg3 0./ rgT.98 0 0 7.98 144 544.86 TmHOW
```

For example, we can say that in Argentina the 931 students (27.90 percent of the sampled students) with a teachers who report having reading instruction "Every day" have a mean reading achievement of 414.82 with a standard error of 11.36, while the 471 students (19.10 percent of the sampled students) with a teacher holding reading instruction "Fewer than 3 days per week" have a mean reading achievement of 431.83 with a standard error of 13.63. Perhaps these results might point to the fact that more instruction is taking place for students who are not reading as well.

In summary, to perform analyses such as those using the Student and Teacher Background data files you need to do the following:

- Identify the variable or variables of interest in the corresponding teacher file and find out about any specific national adaptations to the variable.
- Retrieve the relevant variable or variables from the corresponding teacher data files.
- Retrieve the relevant variables from the Student-Teacher Linkage file. This includes the identification information for the country and teacher (IDCNTRY, IDTEACH, and IDLINK), the achievement score, JRR replication information, and the sampling weight.
- Merge the variables from the teacher data files into the Student-Teacher Linkage files using the variables IDCNTRY, IDTEACH, and IDLINK.
- Use the macro JACKPV or JACKGEN with the corresponding arguments and parameters.
- Print out the result file.

```
jackpv infile= merged /
    cvar = idcntry acbgcomm /
    pvs = asrrea01 to asrrea05 /
    npv = 5 /
    jkz = JKZONE /
    jkr = JKREP /
    njkz = 75 /
    WGT = TOTWGT .
sort cases by idcntry.
print formats n (f6.0) totwgt (f10.0) mnpv mnpv_se pct pct_se (f6.2).
report format=list automatic
    / var = acbgcomm (label) n totwgt mnpv mnpv_se pct pct_se
    / break = idcntry .
```

Exhibit 9.17 Extract of SPSS Computer Output for Performing Student-Weighted Analyses with School-Level Variables(EXAMPLE 4)

Example 4: P PAGE 1	IRLS 2001 User G	Guide					
	GEN/CHARACTER OF SCH						
COUNTRY ID	LOCATION AREA	N	TOTWGT	MNPV	MNPV_SE	PCT	PCT_SE
Argentina	Urban	2498	508261	427.13	6.46	75.42	4.50
5	Suburban	472	129481	395.02	15.18	19.21	4.04
	Rural	156	36134	412.00	15.75	5.36	2.03
Bulgaria	Urban	2382	59461	566.33	4.15	64.12	2.77
	Suburban	384	10742	542.38	17.14	11.58	2.59
	Rural	594	22529	515.61	9.14	24.29	2.29
Sweden	Urban	1276	17377	567.15	5.39	15.17	3.02
	Suburban	3962	76118	560.33	2.84	66.47	3.76
	Rural	695	21021	560.23	5.41	18.36	3.51
Turkey	Urban	1887	341684	464.38	5.96	36.18	4.24

Suburban 3962

As in the previous example, the first step in our analysis is to locate the variables of interest in the specific codebook and file. We find the variable ASDHPART in the Home Background file, and the student weights and plausible values in the Student Background file. We then proceed to read the Home Background file that has the variables that are relevant to our analysis. We then read the variables of interest from the student data file. Next, we proceed to merge the home information with the student information using the variables IDCNTRY and IDSTUD as merge variables, and then use the macro JACKPV to obtain the corresponding percentages of students within each group, and their mean reading achievement scores. The computer code used to run this analysis in SPSS can be found in Exhibit 9.19 and an extract of the results is shown in Exhibit 9.20.

```
= 5
      npv
                       /
       jkz = JKZONE
                       /
      jkr = JKREP
                       /
      njkz = 75
                       /
            = TOTWGT
      WGT
sort cases by identry.
print formats n (f6.0) totwgt (f10.0) mnpv mnpv_se pct pct_se (f6.2).
report format=list automatic
  / var
        = asdhpatr (label) n totwgt mnpv mnpv_se pct pct_se
  / break = idcntry .
```

Exhibit 9.20 Extract of SPSS Computer Output for Performing Home Background Variables (EXAMPLE 5)

Example 5: PI PAGE 1	RLS 2001 Use	r Guide					
	INDEX/PARENT	ſS					
	ATTITUDE TO						
	READING						
COUNTRY ID	(PATR)	N	TOTWGT	MNPV	MNPV_SE	PCT	PCT_SE
Argentina	High	689	138668	462.12	7.11	32.87	1.59
	Medium	1234	263009	416.76	6.81	62.35	1.66
	Low	99	20151	409.70	11.05	4.78	.50
Bulgaria	High	1776	47355	574.10	3.73	54.34	1.77
	Medium	1136	32371	538.07	4.99	37.15	1.37
	Low	230	7417	496.82	11.58	8.51	1.24
Sweden	High	4022	76513	571.35	2.13	71.24	.83
	Medium	1204	25579	546.77	3.40	23.82	.85
	Low	237	5303	536.14	4.53	4.94	.40
Turkey	High	1351	256311	482.51	4.99	29.34	1.42
	Medium	2854	552066	439.75	3.48	63.20	1.26
	Low	323	65102	436.43	6.84	7.45	.76

The results are then presented by country for each value of the variable IDCNTRY after selecting only those cases with IDGRADER=2. The country and the three levels of attitude towards reading are presented in the first two columns. The third column has the number of students sampled in each category. The fourth column corresponds to the weight given these sampled students in the whole population, followed by the mean plausible value for reading achievement and its corresponding

10.2 Contents of the CD

A CD accompanies this User Guide, containing the data from both the PIRLS and Reading Literacy Studies. The CD has the following internal file structure:

- A main directory identifying the study and year (PIRLS2001)
- Within the main directory, there are six sub-directories:

DATA:	Contains data files in ASCII format
PROGRAMS:	Contains SAS and SPSS programs
CODEBOOK:	Contains Codebook files with a description of each

The following programs also can be found in this sub-directory:

ASASCRR1.SAS, ASASCRT1.SAS and ASASCRT2.SAS

These files contain SAS programs that can be used to convert the response codes to the cognitive items and to their corresponding correctness score levels. The use of these programs is described in this chapter.

JACKPV.SAS

This macro program in SAS can be used to compute weighted percentages of students within defined groups, and their mean achievement scores on an achievement scale using plausible values. This macro makes use of the plausible values in computing the mean achievement scores. This macro also generates replicate weights and computes the jacknife repeated replication (JRR) sampling variances for the percentages of students within the group, and the JRR and imputation variances for the mean achievement scores. This macro should only be used when multiple plausible values are used in the analysis.

JACKGEN.SAS

This macro program in SAS can be used to compute weighted percentages of students within defined groups, and their means on a specified continuous variable. This macro also generates replicate weights and computes JRR

10.4 Scoring the Items

There were two types of items administered as part of the PIRLS and Reading Literacy Study achievement tests. About half the items were multiple-choice type, in which the student was asked to select one of four options as the correct response. The responses to these items are coded with one digit. The codes used to represent the responses to these items are as follows:

Code 1	Option A
Code 2	Option B
Code 3	Option C
Code 4	Option D
Code 6	Not reached
Code 7	Invalid response (e.g., a student chose more than one

```
THEN SCORE = 0;
      IF &ITEM = &OM
     IF &ITEM = \&OTHER THEN SCORE = 0 ;
   &ITEM = SCORE;
 %END;
%MEND SCOREIT;
%LET ARIGHT = <list items where option A is the correct one>;
%LET BRIGHT = <list items where option B is the correct one>;
%LET CRIGHT = <list items where option C is the correct one>;
%LET DRIGHT = <list items where option D is the correct one>;
%LET OPENEND = <list open-ended items>;
ARRAY ARIGHT &ARIGHT;
ARRAY BRIGHT &BRIGHT;
ARRAY CRIGHT &CRIGHT;
ARRAY DRIGHT &DRIGHT;
ARRAY OPENEND & OPENEND;
DO OVER ARIGHT ; %SCOREIT(ARIGHT, "MC", 1, 6, 8, 9, 7); END;
DO OVER BRIGHT ; %SCOREIT(BRIGHT, "MC", 2, 6, 8, 9, 7); END;
DO OVER CRIGHT ; %SCOREIT(CRIGHT, "MC", 3, 6, 8, 9, 7); END;
DO OVER DRIGHT ; %SCOREIT(DRIGHT, "MC", 4, 6, 8, 9, 7); END;
DO OVER OPENEND; %SCOREIT(OPENEND, "OE", ,96,98,99,90); END;
```

10.5 Basic Analyses with the PIRLS Data: Means, Percentages, Regression Coefficients, and Their JRR Standard Errors

In this section four macros that can be used to compute the correct standard errors of sampling and imputation are described, including examples in which these macros are used to replicate exhibits in the *PIRLS 2001 International Report* and almanacs.

Computing Sampling and Imputation Variance for Plausible Values Using SAS (JACKPV.SAS)

This section presents example SAS code that can be used to compute the JRR standard errors for mean plausible values and percentages. This code is provided in psJR pluitag -1dc 1 T

you are using SAS interactively, then the macro should be called once at the

JKZ	The variable that captures the assignment of the student to a particular sampling zone. The name of this variable in all PIRLS files is JKZONE.
JKR	The variable that captures whether the case is to be dropped or have its weight doubled for the corresponding replicate weight. The name of this variable in all PIRLS files is JKREP.
NJKZ	This indicates the number of replicate weights to be generated when computing the JRR error estimates. When conducting analyses using the data from all countries, the value of NJKZ should be set to 75, the maximum possible value. When you are working with the data for only one country, you should set the NJKZ argument to as many replicates as are needed in the country. If the data from two or more countries are being used for an analysis, then the largest number of jackknife zones should be used. When in doubt about what number to set the NJKZ paod-ebtr to ,set t T JT 0 .0001 Tc-0.000

MNPV

Contains the means of the plausible values used in the analysis.

MNPV_SE

Contains the standard errors for the means of the plausible values. These standard errors contain the sampling and the imputation components of the errors of the estimates.

The file resulting from using this macro can be printed using an SAS procedure of choice. An example call to this macro and a printout of the resulting file is presented in Exhibit 10.3 below. This code is included in the file SampleJackPV.SAS on the CD that accompanies the User Guide.

Exhibit 10.3 SAS Control Code and Extract of Output File for Using the Macro JACKPV.SAS

libname ar	ibname arl "c:\pirls\data";								
data stude set arl.a where it	sgall		;						
	<pre>%include "c:\pirls\programs\jackpv.sas"; %jackpv (totwgt , jkzone, jkrep, 75, idcntry itsex, asrrea0, 5, student);</pre>								_);
var i	<pre>proc print data=final noobs; var identry itsex N totwgt mnpv mnpv_se mnx mnx_se pct pct_se; format identry itsex N 6.0 totwgt 10.0 mnpv mnpv_se mnx mnx_se pct pct_se 6.2;</pre>								
run;									
IDCNTRY pct_se	ITSE	EX N	TOT	NGT mn	pv mnp	ov_se	mnx	mnx_se	pct
32	1	1647	358324	428.18	6.24	427.63	6.19	50.68	1.06
32	2	1645	348747	410.04	6.50	407.13	6.17	49.32	1.06
84	1	1415	3573	341.73	5.44	342.03	5.27	49.71	0.93
84	2	1414	3615	314.56	5.37	312.68	5.20	50.29	0.93
100	1	1783	49229	562.29	3.69	562.42	3.58	51.47	0.91
100	2	1675	46418	538.06	4.67	537.63	4.63	48.53	0.91
124	1	4121	109940	552.89	2.62	553.12	2.56	49.53	0.67

see that there are entries for each country corresponding to the results for females (ITSEX = 1) and males (ITSEX = 2). The first column has the country code, the second column indicates the gender of the students, the third column has the number of students in each sample, and the fourth column has the total weight of this sample in the population. This is followed by the mean of all five plausible values in reading achievement and the corresponding standard error. Following this is the mean of the first plausible value in reading achievement and its corresponding standard error. The last two columns indicate the percentage of girls and boys in each country's sample and their corresponding standard errors.

For example, Argentina (IDCNTRY=32) sampled 1,647 girls representing 358,324 students in the whole population. The mean of the five plausible values for these Use5.2(rd e0.1)5.4(A 3(n2SsE7b5.1rife0 Tc0 Tw0on()5.1(e)s)40.1a)5..5(e)sfd49004sto t,lsan of 18 (n)-0rahe 1(e who1)

terminated and restarted at a later time the macro needs to be called once again. Once the macro is included in a specific session, the string "%JACKGEN" should not be used within that program because doing so will call the macro.

The macro is included in the program file where it will be used by issuing the following command under SAS:

%include `<path>jackgen.sas'.

where <path> points to the specific drive and directory where the macro JACKGEN.SAS can be found. The macro requires that several parameters be specified when it is called. These parameters are:

INFILE The name of the data file that contains the variables necessary for the analysis. If the path location is included as part of the file name the name of the file has to be enclosed in quotes. It is important to emphasize that this dataset must include only those cases that are of interest in the analysis. If you want to have specific cases excluded from the analysis (e.g., students with missing variables or selected students from a specific grade), this should be done prior to calling the macro.

Weight Variable

Contains the estimate of the population size of the groups defined by each specific combination of the classification variable categories. In our example this variable is called TOTWGT.

Exhibit 10.4 SAS Control Code and Extract of Output File for Using the Macro JACKGEN.SAS

libname ar1 "c:\pirls\d	ata";					
data student;						
<pre>set ar1.asgallr1;</pre>						
where inprl01=1 and a	sdgage > 0	;				
_						
<pre>%include "c:\pirls\prog</pre>	rams\jackg	en.sas";				
\$ jackgen (totwgt , jkzo	ne, jkrep,	75, identry	ldgrader,	asdgage, s	student);	
proc print data=final n						
var identry N totw	-			C D .		
format identry n 6	• totwgt	10.0 mnx mnx_s	se pct pct	_se 6.2;		
run;						
IDCNTRY	N	TOTWGT	mnx	mnx_se	pct	pct_se
					1	1
32	3284	705891	10.18	0.03	100.00	0.00
84	2853	7255	9.79	0.06	100.00	0.00
100	3450	95447	10.93	0.02	100.00	0.00
124	8253	222012	10.02	0.01	100.00	0.00
•						
705	2952	21066	9.80	0.01	100.00	0.00
752	6043	118083	10.80	0.01	100.00	0.00
792	5125	977316	10.23	0.02	100.00	0.00
807	3711	27365	10.65	0.01	100.00	0.00
840	3763	3802557	10.22	0.02	100.00	0.00
926	3155	592507	10.21	0.01	100.00	0.00
927	2717	64375	9.78	0.01	100.00	0.00

In this example, ASDGAGE is used to calculate the mean age of the sampled students in each country. In the listing of the results we can see that there is one entrF-0.48on(ts)5ine,ch co feauntof the samplvaluestof the samplvariable Im001 Tw5 In [Compu1(n6.1(ea7(g Reg)5.s(n6.on Coefficie)4.9.1(ea7(.2(and)3TJT(at 3(eir JRR n6.Stn6.andardlE(y

JACKREG.SAS. The macro computes the multiple correlation between the

of variables will depend mostly on the computer resources available to you at the time. It is recommended to always include the variable that identifies the country. At least one variable has to be specified, usually IDCNTRY.

The file resulting from using this macro can then be printed using the SAS procedure of choice. An example call to this macro, and a subset of the resulting file is presented in Exhibit 10.5. This code is included in the file SampleJackREG.SAS.

Exhibit 10.5 SAS Control Code and Extract of Output File for Using the Macro JACKREG.SAS

libname arl "c:\pi	irls\data";							
<pre>data student; set arl.asgallr1; where itsex in(1,2) and inprl01=1; regsex = itsex - 1;</pre>								
<pre>%include "c:\pirls</pre>	s\programs\jackreg.sa	s";						
<pre>%jackreg (totwgt ,</pre>	, jkzone, jkrep, 75,	idcntry, regsex, a	sbgtvdy, student);					
var identry M format identr	<pre>proc print data=reg noobs; var identry N Mult_RSQ SS_Res SS_Reg SS_Total b00 b00_se b01 b01_se; format identry n 6.0 Mult_RSQ 5.3 SS_Total SS_Reg SS_Res 10.0 B00 B00_SE B01 B01_SE 6.2; run;</pre>							
IDCNTRY N Mult_	RSQ SS_Res SS_Reg	SS_Total b00	b00_se b01 b01_se					
32 3054 0.0	00 1144576 249	1144825 3.21	0.05 -0.04 0.05					
84 2695 0.0	04 12940 58	12998 2.51	0.07 0.18 0.07					
100 3413 0.0	07 109164 781	109945 3.04	0.03 0.18 0.04					
124 8185 0.0	22 293554 6635	300190 2.67	0.03 0.35 0.04					
705 2939 0.0	37 20786 792	21578 2.68	0.04 0.39 0.04					
752 5891 0.0	08 84716 693	85409 2.96	0.02 0.15 0.02					
792 5046 0.0	04 947853 3744	951597 2.44	0.03 0.12 0.03					
807 3559 0.0	08 26911 230	27141 2.61	0.04 0.19 0.04					
840 3740 0.0	14 5733356 83710	5817066 2.83	0.05 0.30 0.04					

In this example, the variable REGSEX is created by subtracting one from the variable ITSEX. As a result, the girls receive a code of 0 and the boys receive a code of 1 on this variable. In this particular model the variable REGSEX is used to predict the values of the variable ASBGTVDY (the number of hours spent watcg.b8m

For example, from the fourth line in the output we can say that in Canada (IDCNTRY = 124), data is available for 8,185 cases. The coefficient of

	NJKZ argument to as many replicates as are needed in the country. If the data from two or more countries are being used for an analysis, then the largest number of jackknife zones should be used. When in doubt about what number to set the NJKZ parameter to, set it to 75. The error variance will always be estimated correctly if more replicate weights than necessary are computed, but will be underestimated if you specify fewer.
JKZ	The variable that captures the assignment of the student to a particular sampling zone. The name of this variable in all PIRLS files is JKZONE.
JKR	The variable that captures whether the case is to be dropped or have its weight doubled for the corresponding replicate weight. The name of this variable in all PIRLS files is JKREP.
WGT	The sampling weight to be used in the analysis, generally TOTWGT when using the student files, or TCHWGT when using the teacher files.

The simplest and most straightforward way to call the macro is by using the conventional SAS notation for invoking macros. This involves listing the macro name followed by the corresponding list of arguments for the analysis, each separated by a slash. For example, if the macro is called using:

%jackregP(totwgt,jkzone,jkrep,75,idcntry idgrader,regsex,asrrea0,5,asgallr1);

it will compute the regression equation for the variable REGSEX as a predictor of the plausible values in reading. The data will be read from the data set ASGALLR1 and the standard error of the statistics will be computed based on 75 replicate weights.

The file that contains these results is then called REG and is saved to the default directory being used by SAS. The variables that are contained in this file are:

Classification Variables

Each of the classification variables is kept in the resulting file. In our example above there would be two variables in the resulting dataset, IDCNTRY and IDGRADER. There is one unique occurrence for each combination of the categories for these variables.

Mult_RSQ

The squared multiple correlation coefficient for the model.

SS_Res, SS_Reg, SS_Total

The residual, regression, and total sum of squares for the model within each group as defined by the classification variables.

Regression Coefficients and Standard Errors (B## and B##.SE)

These are the regression coefficients for each of the predictor variables in the

Exhibit 10.9 Sample Exhibit for Student-Level Analysis Involving Plausible Values Taken From *the PIRLS 2001* International Report



Exhibit 10.11 Extract of SAS Computer Output for Performing Analyses with Student-Level Variables Involving Plausible Values (EXAMPLE 2)

Example 2: PIRLS 2001 User Guide									
COUNTRY ID=Argent	*COUNTRY ID*=Argentina								
ASBGBOOK	N	TOTWGT	mnpv	mnpv_se	pct	pct_se			
	964	213557	416.09	5.52	42.23	2.48			
11-25 BOOKS	610	121633	445.83	6.79	24.05	1.16			
26-100 BOOKS 101-200 BOOKS	499	107115	465.93	9.13	21.18	1.61			
	181	121633 107115 33400	483.59	11.11	6.60	0.67			
MORE THAN 200 BOOKS	144	29995	473.98	12.58	5.93	0.73			
COUNTRY ID=Bulgaria									
ASBGBOOK	N	TOTWGT	mnpv	mnpv_se	pct	pct_se			
0 -10 BOOKS	583	18440 11986 21738 14772 25005	498.15	8.11	20.06	1.78			
11-25 BOOKS	412	11986	537.18	5.67	13.04	1.01			
26-100 BOOKS	821	21738	565.01	3.76	23.64	1.09			
101-200 BOOKS	561	14772	579.21	4.94	16.07	0.86			
MORE THAN 200 BOOKS	962	25005	582.60	3.79	27.20	1.44			
COUNTRY ID=Sweden									
ASBGBOOK	N			mnpv_se					
0 -10 BOOKS	163 505 1704 1527	3794	503.47	6.92 4.05 2.27 2.54	3.29	0.51			
11-25 BOOKS	505	10812	530.15	4.05	9.39	0.66			
26-100 BOOKS	1704	35470	552.30	2.27	30.80	1.10			
101-200 BOOKS	1527	29900	572.06	2.54	25.96	0.68			
MORE THAN 200 BOOKS	1970	35198	576.76	2.68	30.56	1.46			
COUNTRY ID=Turkey									
ASBGBOOK	N	TOTWGT	mnov	mnov se	pct	pct se			
0 -10 BOOKS 11-25 BOOKS 26-100 BOOKS 101-200 BOOKS	1976	376947	421.68	3.84	39.73	1.95			
11-25 BOOKS	1482	284996	455.46	3.66	30.04	1.08			
26-100 BOOKS	996	190694	478.28	4.25	20.10	1.07			
101-200 BOOKS	264	50261	498.85	8.29	5.30	0.54			
MORE THAN 200 BOOKS	223	45775	494.30	10.81	4.83	0.64			
COUNTRY ID=United S	tates								
ASBGBOOK	N	TOTWGT	mnpv	mnpv_se	pct	pct_se			
0 -10 BOOKS	314	334096	476.66	6.98	8.93	0.95			
11-25 BOOKS	654	334096 653582	521.51	6.98 4.23	17.47	0.96			
26-100 BOOKS	1167	1171921	551.13	3.84	31.33				
101-200 BOOKS	799	804711	568.38	4.29	21.51 20.76	1.10			
MORE THAN 200 BOOKS	767	1171921 804711 776508	555.70	5.82	20.76	1.24			

In this example, each country's mean plausible value for reading achievement is reported for each response category in the variable ASBGBOOK. The results are presented by country for each value of the variable IDCNTRY after selecting only those cases with IDGRADER=2. The country and the five response options are presented in the first two columns. The third column has the number of students sampled in each category. The fourth column corresponds to the sum of the weights given these sampled students in the whole population, followed by their mean plausible values for reading achievement and the corresponding standard errors. The

last two columns represent the percentages of students sampled responding within each category and the corresponding standard errors.

For example, from the first line of the report, we can say that in Argentina the 964 students (42.23 percent of the sampled students) who reported having more than "0-10 Books" have a mean reading achievement of 416.09 with a standard error of 5.52, while the 144 students (5.93 percent of the sampled students) who reported having "More than 200 books" have a mean reading achievement of 473.98 with a standard error of 12.58.

10.7 Performing Analyses with Teacher-Level Variables

The PIRLS teacher files do not contain representative samples of teachers within a country. Rather, these are the teachers for a representative sample of students within a country. Therefore, it is appropriate that statements about the teachers be made only in terms of how many students are taught by teachers of one kind or another, and not in terms of how many teachers in the country do one thing or another.

To analyze the information provided by the students' teachers, it is first necessary to link it with the student information. Each student record in the Student Background data file can be linked to one or more different teachers in the Teacher Background data file. To facilitate the linking between students and their teachers in the teacher file, the Student-Teacher Linkage filed errr"as created and is part of the International Database. These files are called AST<COUNTRY>R1. The Student-Teacher Linkage file contains one record for each student-by-teacher combination, with the corresponding identification variables.

Each record also contains the number of teachers for the student and a set of weights that can be used when conducting analyses with these data. Student achievement plausible values, sampling weights, and JRR replication information have been added to the Student-Teacher Linkage file in order to simplify the merging process for analyses that link teacher variabled errs to student achievement. For such analyses it is necessary to merge only the Teacher Background file with the Student-Teacher Linkage file. For analyses linking teacher variables to other student variables, it is necessary also to merge the Student Background files with the Teacher Background file after it has been combined with the Student-Teacher Linkage file.

Conducting analyses with the teacher data requires some extra stepd errs that are not required when analyzing the student or school background data.

For our example, we want to find out about the frequency of reading instruction reported by the teachers who teach the fourth-grade students in the PIRLS

n 3 days per week." We are also interested in the udents taught by these teachers. These results are *LS International Report* (replicated results are

or Teacher-Level Analysis Involving Taken From *the PIRLS 2001* port



In our example, our teacher variable of interest (ATBGRACT) is a categorical variable with three categories.

We then proceed to read the necessary information from the Student-Teacher Linkage file and the Teacher Background file.

The two files are then merged or matched into one file that will then be used with the JACKPV macro. These two files will be merged using the variables IDCNTRY, IDTEACH, and IDLINK. The combination of values for these three variables is unique within the teacher data, but is repeated in the Student-Teacher Linkage file as many times as the specific teacher teaches students in a class. After the files are merged, the macro JACKPV is used and the results can be printed. The code in SAS for this example is presented in Exhibit 10.13. Selections of the results obtained from this program are displayed in Exhibit 10.14.

Exhibit 10.13 SAS Control Statement for Performing Teacher-Level Analysis Using Plausible Values

```
title "Example 3: PIRLS 2001 User Guide";
libname ar1 "c:\pirls\data";
%include "c:\pirls\Programs\jackpv.sas";
data studteac;
 set arl.astargrl
    arl.astbgrrl
     arl.astswerl
     ar1.astturr1
     arl.astusarl;
proc sort data=studteac;
     by identry idteach idlink;
data teacher;
 set arl.atgargr1
     arl.atgbgrrl
     ar1.atgswer1
     arl.atgturr1
     arl.atgusarl;
proc sort data=teacher;
     by identry idteach idlink;
* Now merge the two files;
data merged;
     merge studteac teacher;
     by identry idteach idlink;
      if nmiss(atbgract) = 0 and inprl01 = 1;
* Define the format for the variables used;
proc format
```

```
by identry ;
var atbgract N tchwgt mnpv mnpv_se pct pct_se;
format identry country. atbgract ract.
        N 6.0 tchwgt 10.0 mnpv mnpv_se pct pct_se 6.2;
```

run;

Exhibit 10.14 Extract of SAS Output for Performing Teacher-Level Analysis Using Plausible Values (EXAMPLE 3)

	- 13					
Example 3: PIRLS 2001 Use	er Guide	2				
COUNTRY ID=Argentina						
ATBGRACT	N	TCHWGT	mnpv	mnpv_se	pct	pct_se
Every Day	931	180294	414.82	11.36	27.90	4.10
3-4 Days per week	1535	342494	415.79	9.18	53.00	4.80
Fewer than 3 days per week		123414	431.83	13.63	19.10	3.64
rewei than 5 days per week	4/1	123414	431.03	13.05	19.10	3.04
COUNTRY ID=Bulgaria						
ATBGRACT	N	TCHWGT	mnpv	mnpv_se	pct	pct_se
			-		-	
Every Day	1071	28227	550.79	8.22	30.75	4.37
3-4 Days per week	1942	55286	553.90	5.03	60.22	4.45
Fewer than 3 days per week	309	8288	542.86	10.44	9.03	2.27
COUNTRY ID=Sweden						
ATBGRACT	N	TCHWGT	mnpv	mnpv_se	pct	pct_se
Every Day	3486	60112	562.50	3.28	56.97	3.56
3-4 Days per week	2154	31905	555.74	3.55	30.24	3.36
Fewer than 3 days per week	841	13504	568.21	3.41	12.80	2.22
tommou tot multi						
COUNTRY ID=Turkey ATBGRACT	37	mautuam				
AIBGRACI	N	TCHWGT	mnpv	mnpv_se	per	pct_se
Every Day	2600	487931	453.83	5.20	53.62	4.97
3-4 Days per week	1748	335216	439.26	6.67	36.84	4.47
Fewer than 3 days per week		86759	459.33	9.93	9.53	2.70
COUNTRY ID=United States						
ATBGRACT	N	TCHWGT	mnpv	mnpv_se	pct	pct_se
			-		-	—
Every Day	3323	3340658	541.37	4.23	93.20	2.05
3-4 Days per week	264	225928	555.06	7.54	6.30	2.01
Fewer than 3 days per week	3	18002	539.93	13.98	0.50	0.50

The results are presented by country for each value of the variable IDCNTRY. The country and the three categories are presented in the first two columns. The third column has the number of students sampled in each category. The fourth column corresponds to the weight given these sampled students in the whole population, followed by the mean plausible value for reading achievement and its corresponding standard error. The last two columns represent the percentage of students sampled within each category and its corresponding standard error.

For example, we can say that in Argentina the 931 students (27.90 percent of the sampled students) with a teachers who report having reading instruction "Every day" have a mean reading achievement of 414.82 with a standard error of 11.36, while the 471 students (19.10 percent of the sampled students) with a teacher holding reading instruction "Fewer than 3 days per week" have a mean reading

achievement of 431.83 with a standard error of 13.63. Perhaps these results might point to the fact that more instruction is taking place for students who are not reading as well.

In summary, to perform analyses such as those using the Student and Teacher Background data files you need to do the following:

- Identify the variable or variables of interest in the corresponding teacher file and find out about any specific national adaptations to the variable.
- Retrieve the relevant variable or variables from the corresponding teacher data files.
- Retrieve the relevant variables from the Student-Teacher Linkage file. This includes the identification information for the country and teacher (IDCNTRY, IDTEACH, and IDLINK), the achievement score, JRR replication information, and the sampling weight.
- Merge the variables from the teacher data files into the Student-Teacher Linkage files using the variables IDCNTRY, IDTEACH, and IDLINK.
- Use the macro JACKPV or JACKGEN with the corresponding arguments and parameters.
- Print out the result file.

10.8 Performing Analyses with School-Level Variables

Although the students in the PIRLS samples were selected from within a sample of schools, the school sample was designed to optimize the resulting sample of students, rather than to give an optimal sample of schools. For this reason, it is

information contained in the student-level file to make the desired statements. The examples presented in this section describe how this can be accomplished using SAS.

Let us say that we want to find out the percentage of fourth graders who attend schools located in a certain area of the country as captured by the question asked of the school principals: "How would you characterize the area in which your school is



Exhibit 10.16 SAS Control Statements for Performing Student-Weighted Analyses with School-Level Variables (EXAMPLE4.SAS)

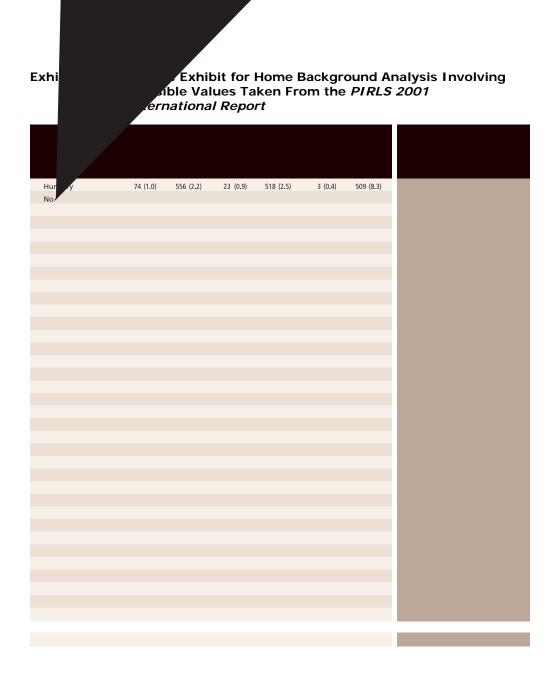
title "Example 4: PIRLS 2001 User Guide";

- Merge the variables from the school database onto the student database using the variables IDCNTRY and IDSCHOOL.
- Use the macro JACKGEN or JACKPV with the corresponding arguments and parameters.
- Print out the result file.

10.9 Performing Analyses with Home Background Variables

Students in the PIRLS samples were also given questionnaires to be completed by their parents or guardian. The responses to these questionnaires are included in the home background questionnaire file. Like any other variable in the PIRLS database, home background variables need to be analyzed as attributes of students, rather than as elements in their own right.

Let us say that we want to find out the percentage of fourth graders who's parents have a high, medium, or low attitude towards reading. These results, presented in Exhibit 10.18, also can be found in Exhibit 4.17 of the *PIRLS International Report*. This exhibit reports the results for the Index of Parents Attitude Towards Reading (ASDHPATR). The options for this question were "High," "Medium," and "Low."



the student data file. Next, we proceed to merge the home information with the student information using the variables IDCNTRY and IDSTUD as merge variables, and then use the macro JACKPV to obtain the corresponding percentages of students within each group, and their mean reading achievement scores. The computer code used to run this analysis in SAS can be found in Exhibit 10.19 and an extract of the results is shown in Exhibit 10.20.

Exhibit 10.19 SAS Control Statements for Performing Home Background Analyses (EXAMPLE5.SAS)

title "Example 5: PIRLS 2001 User Guide";

A

Acknowledgements

The design and development of PIRLS 2001 was achieved through the collaborative efforts of individuals around the world. Staff from the national research centers in each participating country, members of advisory committees, the International Association for the Evaluation for Educational Achievement (IEA), funding agencies, and the International Study Center (ISC) at Boston College worked together to develop and implement the PIRLS 2001 assessment. This appendix acknowledges the individuals and organizations for their contributions. Given that the development and implementation of PIRLS 2001 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. PIRLS 2001 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.

A.1 Funding Agencies

Funding for the development of PIRLS 2001 was provided by the National Center for Education Statistics of the U.S. Department of Education (NCES), The World Bank, and the participating countries. Valena Plisko, Eugene Owen, Dawn Nelson and Lawrence Ogle of NCES and Vincent Greaney of the World Bank were instrumental in making PIRLS 2001 possible and for ensuring the quality of the study. Each participating country was responsible for funding national project costs and implementing PIRLS in accordance with the international procedures.

A.2 Management and Operations

PIRLS 2001 was conducted under the auspices of the IEA. The study is directed by Ina V.S. Mullis and Michael O. Martin, and managed centrally by the staff of the International Study Center at Boston College, Lynch School of Education. The PIRLS International Study Center worked closely with organizations that were responsible for particular aspects of the study, the PIRLS advisory committees, and representatives of the participating countries. In the IEA Secretariat, Hans Wagemaker was responsible for overseeing fundraising and country participation. Barbara Malak, also of the IEA Secretariat, was responsible for managing the ambitious translation verification effort and for recruiting international quality control monitors. Statistics Canada worked with countries to ensure that the international sampling procedures were followed, adapted the international design to national conditions, documented the national samples, and computed sampling weights. The National Foundation for Educational Research in England and Wales had major responsibility for developing the reading test, including collecting reading passages, developing items and scoring guides, and conducting scoring training. The IEA Data Processing Center was responsible for processing and verifying the data from the 35 countries, and for constructing the international database.

IEA Secretariat

Alejandro Tiana, IEA Chair Hans Wagemaker, Executive Director Barbara Malak, Manager Membership Relations Juriaan Hartenberg, Financial Manager

PIRLS and TIMSS International Study Ce Tm1(cre. 1 Tf3(na))H) fight in the study Ce Tm1(cre. 1 Tf3(na)) H) fight is the study of the study Ce Tm1(cre. 1 Tf3(na)) H) fight is the study of the study Ce Tm1(cre. 1 Tf3(na)) H) fight is the study of the stud

PIRta pllase. 5r1(alor.) PIRta pllase. 5r1(alor.) PIRta pllase. 5r1(al

Christine Hoage, Manager, Finance Rita Holmes, Administrative Coordinator Laura Misas, Administrative Coordinator

IEA Data Processing Center

Pierre Foy, Senior Researcher Ursula Itzlinger, Senior Researcher Juliane Barth, Research Assistant

Statistics Canada

Marc Joncas, Senior Methodologist

National Foundation for Educational Research in England and Wales

Chris Whetton, Assistant Director Marian Sainsbury, Principal Research Officer Jenny Bradshaw, Senior Research Officer Anne Kispal, Senior Research Officer Jane Sowerby, Senior Research Officer Jenny Phillips, Research Officer

Educational Testing Services

Jay Campbell, PIRLS Reading Coordinator Ed Kulick, Psychometric Scaling Consultant Matthias von Davier, Psychometric Scaling Consultant

Westat

Keith Rust, Sampling Referee

American Institutes for Research

Eugene Johnson, Psychometric Design Consultant

PIRLS Questionnaire Development Group

Annette Lafontaine Université de Liège Belgium

Michael Marshall University of British Columbia Canada

Ivana Prochazkova Institute for Information on Education Czech Republic

Monica Rosén Göteborg University Sweden

Graham Ruddock National Foundation for Educational Research in England and Wales England

Maurice Walker Ministry of Education New Zealand

National Research Coordinators

The PIRLS 2001 National Research Coordinators (NRCs) were responsible for the crucial task of implementing the study in their countries. They participated in every aspect of the work to ensure that the study was of high quality. All the PIRLS 2001 NRCs and their staff members are to be commended for their professionalism and their dedication in conducting all aspects of the project.

Argentina

Lilia Toranzos Ministerio de Educación

Belize

Rosalind Bradley Denise Robateau Belize Teachers' Training College

Bulgaria

Georgi Bishkov Felyanka Kaftandjieva University of Sofia

Canada

Francine Jaques Robert Deschênes Education Quality & Accountability Office (EQAO)

Louis-Philippe Gaudreault Jean-Louis Lebel Ministère de l'Éducation

Michael Marshall University of British Columbia

Colombia

Martha Rocha Claudia Saenz Servicio Nacional de Pruebas

Cyprus

Mary Koutselini Constantinos Papanastasiou University of Cyprus

Czech Republic

Iveta Kramplova Ivana Prochazkova Institute for Information on Education

England

Liz Twist National Foundation for Educational Research (NFER)

France

Marc Colmant Ministère de l'Éducation Nationale

Germany

Wilfried Bos Knut Schwippert Eva-Maria Lankes

Kuwait

Mansour Hussein Ministry of Education

Latvia Ieva Johansone University of Latvia

Lithuania Aiste Mackeviciute Ministry of Education and Science

Republic of Macedonia Bojana Naceva

Pedagogical Institute of Macedonia

Moldova

Ilie Nasu Ministry of Education & Science

Morocco Abdellah Belachkar Ministère de L'Education Nationale

The Netherlands Mieke Van Diepen University of Nijmegen

New Zealand

Megan Chamberlain Maurice Walker Ministry of Education

Norway

Finn Egil Tønnessen Ragnar Gees Solheim Stavanger College

Romania

Gabriela Noveanu Institute for Educational Sciences

Russian Federation

Galina Kovalyova Institute of General Secondary Education

Scotland

Liz Levy Brian Semple Scottish Office Education and Industry Department

Singapore

Siow-Chin Ng Ministry of Education

Slovak Republic

Zuzana Lukacková SPU–National Institute for Education

Slovenia

Marjeta Doupona-Horvat Center for IEA Studies Educational Research Institute

Sweden

Monica Rosén Göteborg University

Turkey

Yurdanur Atlioglu Educational Research and Development Directorate (ERDD)

United States of America

Laurence Ogle Dawn Nelson National Center for Education Statistics (NCES)