



PIRLS

PIRLS 2001 User Guide for the International Database

Edited by
Eugenio J. Gonzalez
Ann M. Kennedy





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

PIRLS

PIRLS 2001 User Guide

for the International Database

Edited by

Eugenio J. Gonzalez
Ann M. Kennedy



© 2003 International Association for the
Evaluation of Educational Achievement (IEA)

PIRLS 2001 User Guide for the International
Database / Edited by Eugenio J. Gonzalez
and Ann M. Kennedy

Publisher: International Study Center
Lynch School of Education, Boston College

Library of Congress
Catalog Card Number: 2003109101

ISBN 1-889938-31-9

For more information about PIRLS contact:

PIRLS International Study Center
Lynch School of Education
Boston College
Chestnut Hill, MA 02467
United States

tel: +1-617-552-1600
fax: +1-617-552-1203

e-mail: pirls@bc.edu

<http://www.pirls.bc.edu>

Boston College is an equal opportunity,
affirmative action employer.

Printed and bound in the United States.

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
	<i>Exhibit 1.1 Percentages of Reading Assessment Devoted to Reading Purposes and Processes</i>	<i>1-7</i>
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 PIRLS Assessment Framework	2-2
	Text Types	2-3
	Processes and Strategies	2-4
	Test Booklet Design	2-5
	<i>Exhibit 2.1 Distribution of Literacy and Informational Blocks Across Booklets</i>	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
	<i>Exhibit 3.1 Contexts for the Development of Reading Literacy</i>	3-3
	<i>Exhibit 3.2 Factors within the Home, School, and National and Community Contexts Addressed by PIRLS</i>	3-3
3.3	PIRLS Main Survey Questionnaires	3-4
	Student Questionnaire	3-4
	<i>Exhibit 3.3 Content of the PIRLS Student Questionnaire</i>	3-5
	Learning-to-Read Survey (Home)	3-6
	<i>Exhibit 3.4 Content of the PIRLS Learning-to-Read Survey (Home Questionnaire)</i>	3-6
	Teacher Questionnaire	3-7
	<i>Exhibit 3.5 Content of the PIRLS Teacher Questionnaire</i>	3-7
	School Questionnaire	3-9
	<i>Exhibit 3.6 Content of the PIRLS School Questionnaire</i>	3-9
3.4	Reporting Questionnaire Data	3-11
	Summary Indices from Background Data	3-11
	<i>Exhibit 3.7 Summary Indices from Background Data in the PIRLS 2001 International Report</i>	3-13
	Reporting Student Questionnaire Data	3-16
	Reporting Teacher Questionnaire Data	3-16
	Reporting Parents' Questionnaire Data	3-17
	Reporting School Questionnaire Data	3-17

4 PIRLS Survey Operations Procedures

4.1	Overview	4-2
4.2	Responsibilities of the National Research Coordinator	4-2
4.3	Within-School Sampling Procedures	4-3
	Survey Tracking Forms	4-4
	Excluding Students from Testing	4-4
	Assigning Instruments to Students and Teachers	4-5
4.4	Within-School Assessment Activities	4-6
	Arranging the Testing Sessions	4-6
	Distributing Materials	4-6
	Test Administration	4-6
	Timing of the Testing Sessions	4-7
	Activities Following the Test Administration	4-8
4.5	Scoring the Constructed-Response Items	4-8
	Preparing Materials to Train the Scorers	4-8
	Documenting the Reliability of the Constructed-Response Scoring	4-9
4.6	Trends in IEA's Reading Literacy Study	4-10

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

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
	Proficiency Estimation Using Plausible Values	6-6
	Generating Proficiency Scores	6-8
	Working with Plausible Values	6-9
6.3	Student Achievement Scores	6-11
	Achievement Scores in the Student Files	6-11
	<i>Exhibit 6.1 Plausible Values Based on PIRLS 2001 Content Area</i>	6-12
	<i>Exhibit 6.2 Plausible Values Based on Reading Literacy Study Content Area</i>	6-12
	<i>Exhibit 6.3 Raw Scores Included in the PIRLS Data Files</i>	6-13
	<i>Exhibit 6.4 Raw Scores Included in the Reading Literacy Study Data Files</i>	6-13
	<i>Exhibit 6.5 Standardized Scores Included in the PIRLS Data Files</i>	6-14
	<i>Exhibit 6.6 Standardized Scores Included in the Reading Literacy Study Data Files</i>	6-14
	<i>Exhibit 6.7 National Rasch Scores Included in the PIRLS Data Files</i>	6-15
	<i>Exhibit 6.8 National Rasch Scores Included in the Reading Literacy Study Data Files</i>	6-15
6.4	International and National Benchmarks of Achievement	6-16
	<i>Exhibit 6.9 PIRLS International Benchmarks of Reading Achievement</i>	6-16
	<i>Exhibit 6.10 PIRLS International Benchmarks of Reading for Literary Purposes Achievement</i>	6-17
	<i>Exhibit 6.11 PIRLS International Benchmarks of Reading to Acquire and Use Information Achievement</i>	6-17
	<i>Exhibit 6.12 Variables Indicating Achievement of International Benchmarks of Reading to Acquire Information Achievement</i>	6-17
	<i>Exhibit 6.13 Variables Indicating Achievement of National Benchmarks of Reading to Acquire Information Achievement</i>	6-19

7 Estimating Sampling and Imputation Variance

7.1	Overview	7-2
7.2	Estimating Sampling Variance	7-2
	Computing Sampling Variance Using the JRR Method	7-3
	Construction of Sampling Zones for Sampling Variance Estimation	7-4
	Computing the JRR Replicate Weights	7-4
7.3	Estimating Imputation Variance	7-6
7.4	Combining Sampling and Imputation Variance	7-6
7.5	Comparing Achievement with the International Mean	7-7
7.6	Gender Differences	7-7

8 Content and Format of Database Files

8.1	Overview	8-2
8.2	The Data Files	8-3
	Data File and Variable Naming Conventions	8-3
	<i>Exhibit 8.1 PIRLS and Reading Literacy Data Files</i>	8-5
	<i>Exhibit 8.2 Country Identification and Availability of Data Files</i>	8-6
	<i>Exhibit 8.3 Background Questionnaire Variable Location Conventions</i>	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
	<i>Exhibit 8.4 Data Almanac Files for PIRLS 2001 and Reading Literacy Study</i>	8-22
	Background Item Data Almanacs	8-22
	Achievement Item Data Almanacs	8-23
8.4	Codebook Files	8-24
8.5	Program Files	8-26
8.6	Item Information Files	8-27

9 Performing Analyses with the PIRLS Data Using SPSS

9.1	Overview	9-3
9.2	Contents of the CD	9-4
9.3	Creating SPSS System Files	9-6
	<i>Exhibit 9.1</i> Extract from SPSS Control Code for Creating a Student Background SPSS System File	9-7
9.4	Scoring the Items	9-7
	<i>Exhibit 9.2</i> Extracted Sections of SPSS Macro SCOREIT Used to Convert Cognitive Item Response Codes to Correctness-Score Levels	9-9
9.5	Basic Analyses with the PIRLS Data: Means, Percentages, Regression Coefficients, and Their JRR Standard Errors	9-10
	Computing Sampling and Imputation Variance for Plausible Values Using SPSS (JACKPV.SPS)	9-10
	<i>Exhibit 9.3</i> SPSS Control Code and Extract of Output File for Using the Macro JACKPV.SPS	9-14
	Computing Sampling Variance for Variables Other Than Plausible Values Using SPSS (JACKGEN.SPS)	9-15
	<i>Exhibit 9.4</i> SPSS Control Code and Extract of Output File for Using the Macro JACKGEN.SPS	9-19
	Computing Regression Coefficients and Their JRR Standard Errors for Variables Other Than Plausible Values (JACKREG.SPS)	9-20
	<i>Exhibit 9.5</i> SPSS Control Code and Extract of Output File for Using the Macro JACKREG.SPS	9-23
	Computing Regression Coefficients and Their JRR Standard Errors with Plausible Values (JACKREGP.SPS)	9-24
	<i>Exhibit 9.6</i> SPSS Control Code and Extract of Output File for Using the Macro JACKREGP.SPS	9-27
9.6	Conducting Analyses with the PIRLS 2001 International Database: Student Level	9-28
	Example Analysis with Student-Level Variables Not Using Plausible Values	9-28
	<i>Exhibit 9.7</i> SPSS Control Statements for Performing Analyses with Student-Level Variables NOT Using Plausible Values (EXAMPLE1.SPS)	9-30
	<i>Exhibit 9.8</i> Extract of SPSS Computer Output for Performing Analyses with Student-Level Variables Not Using Plausible Values (EXAMPLE 1)	9-31
	Example Analysis with Student-Level Variables Using Plausible Values	9-31
	<i>Exhibit 9.9</i> Sample Exhibit for Student-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report	9-32
	<i>Exhibit 9.10</i> SPSS Control Statements for Performing Analyses with Student-Level Variables Involving Plausible Values (EXAMPLE2.SPS)	9-33
	<i>Exhibit 9.11</i> Extract of SPSS Computer Output for Performing Analyses with Student-Level Variables Involving Plausible Values (EXAMPLE 2)	9-34
9.7	Performing Analyses with Teacher-Level Variables	9-35
	<i>Exhibit 9.12</i> Sample Exhibit for Teacher-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report	9-36

Exhibit 9.13	SPSS Control Statement for Performing Teacher Level Analysis Using Plausible Values	9-37
Exhibit 9.14	Extract of SPSS Output for Performing Teacher Level Analysis Using Plausible Values (EXAMPLE 3)	9-38
9.8	Performing Analyses with School-Level Variables	9-40
Exhibit 9.15	Sample Exhibit for School-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report	9-41
Exhibit 9.16	SPSS Control Statements for Performing Student-Weighted Analyses with School-Level Variables (EXAMPLE4.SPS)	9-42
Exhibit 9.17	Extract of SPSS Computer Output for Performing Student-Weighted Analyses with School-Level Variables(EXAMPLE 4)	9-43
9.9	Performing Analyses with Home Background Variables	9-44
Exhibit 9.18	Sample Exhibit for Home Background Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report	9-45
Exhibit 9.19	SPSS Control Statements for Performing Home Background Analyses (EXAMPLE5.SPS)	9-46
Exhibit 9.20	Extract of SPSS Computer Output for Performing Home Background Variables (EXAMPLE 5)	9-47

10 Performing Analyses with the PIRLS Data Using SAS

10.1	Overview	10-3
10.2	Contents of the CD	10-4
10.3	Creating SAS Datasets	10-6
Exhibit 10.1	Extract from SAS Control Code for Creating a Student Background SAS Dataset	10-7
10.4	Scoring the Items	10-8
Exhibit 10.2	Extracted Sections of SAS Macro SCOREIT Used to Convert Cognitive Item Response Codes to Correctness-Score Levels	10-9
10.5	Basic Analyses with the PIRLS Data: Means, Percentages, Regression Coefficients and Their JRR Standard Errors	10-10
	Computing Sampling and Imputation Variance for Plausible Values Using SAS (JACKPV.SAS)	10-10
Exhibit 10.3	SAS Control Code and Extract of Output File for Using the Macro JACKPV.SAS	10-14
	Computing Sampling Variance for Variables Other Than Plausible Values Using SAS (JACKGEN.SAS)	10-15
Exhibit 10.4	SAS Control Code and Extract of Output File for Using the Macro JACKGEN.SAS	10-19
	Computing Regression Coefficients and Their JRR Standard Errors for Variables Other Than Plausible Values (JACKREG.SAS)	10-19
Exhibit 10.5	SAS Control Code and Extract of Output File for Using the Macro JACKREG.SAS	10-23
	Computing Regression Coefficients and Their JRR Standard Errors with Plausible Values (JACKREGP.SAS)	10-24

<i>Exhibit 10.6</i>	<i>SAS Control Code and Extract of Output File for Using the Macro JACKREGP.SAS</i>	10-27
10.6	Conducting Analyses with the PIRLS 2001 International Database: Student Level	10-28
	Example Analysis with Student-Level Variables Not Using Plausible Values	10-28
<i>Exhibit 10.7</i>	<i>SAS Control Statements for Performing Analyses with Student-Level Variables NOT Using Plausible Values (EXAMPLE1.SAS)</i>	10-30
<i>Exhibit 10.8</i>	<i>Extract of SAS Computer Output for Performing Analyses with Student-Level Variables Not Using Plausible Values (EXAMPLE 1)</i>	10-30
	Example Analysis with Student-Level Variables Using Plausible Values	10-31
<i>Exhibit 10.9</i>	<i>Sample Exhibit for Student-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report</i>	10-32
<i>Exhibit 10.10</i>	<i>SAS Control Statements for Performing Analyses with Student-Level Variables Involving Plausible Values (EXAMPLE2.SAS)</i>	10-33
<i>Exhibit 10.11</i>	<i>Extract of SAS Computer Output for Performing Analyses with Student-Level Variables Involving Plausible Values (EXAMPLE 2)</i>	10-34
10.7	Performing Analyses with Teacher-Level Variables	10-35
<i>Exhibit 10.12</i>	<i>Sample Exhibit for Teacher-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report</i>	10-36
<i>Exhibit 10.13</i>	<i>SAS Control Statement for Performing Teacher-Level Analysis Using Plausible Values</i>	10-37
<i>Exhibit 10.14</i>	<i>Extract of SAS Output for Performing Teacher-Level Analysis Using Plausible Values (EXAMPLE 3)</i>	10-38
10.8	Performing Analyses with School-Level Variables	10-39
<i>Exhibit 10.15</i>	<i>Sample Exhibit for School-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report</i>	10-41
<i>Exhibit 10.16</i>	<i>SAS Control Statements for Performing Student-Weighted Analyses with School-Level Variables (EXAMPLE4.SAS)</i>	10-42
<i>Exhibit 10.17</i>	<i>Extract of SAS Computer Output for Performing Student-Weighted Analyses with School-Level Variables (EXAMPLE 4)</i>	10-42
10.9	Performing Analyses with Home Background Variables	10-44
<i>Exhibit 10.18</i>	<i>Sample Exhibit for Home Background Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report</i>	10-45
<i>Exhibit 10.19</i>	<i>SAS Control Statements for Performing Home Background Analyses (EXAMPLE5.SAS)</i>	10-46
<i>Exhibit 10.20</i>	<i>Extract of SAS Computer Output for Performing Home Background Variables (EXAMPLE 5)</i>	10-47

Appendix A: Acknowledgements

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
	<i>Exhibit 1.1 Percentages of Reading Assessment Devoted to Reading Purposes and Processes</i>	<i>1-7</i>
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

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 analyses are illustrated for the use of both SPSS and SAS statistical software in Chapters 9 and 10. Several electronic documents supplement the data files – the international version of each of the four PIRLS 2001 background questionnaires, and the documentation of the national adaptations to each questionnaire.

1.2 Overview of PIRLS

IEA has been conducting cross-national studies of educational achievement for more than 40 years – including periodic assessments of children’s reading literacy. In 1973, reading was one of the subjects in IEA’s six-subject study that was conducted in 15 countries (Thorndike, 1973; Walker, 1976). In 1991, IEA’s Reading Literacy Study was conducted in 32 educational systems (Elley, 1992; 1994). Most recently, PIRLS (Progress in International Reading Literacy Study) was established as the IEA’s latest study to monitor progress in children’s reading literacy into the future (Mullis, Martin, Gonzalez, & Kennedy, 2003).

In 1998, the IEA General Assembly formally agreed that PIRLS would be part of the IEA’s regular cycle of assessments, which also includes mathematics and science. At that point, some basic principles were established:

- PIRLS would begin in 2001 with an assessment of children in fourth grade.
- PIRLS would focus on reading literacy achievement, as well as home and school contexts for learning to read.
- Reading literacy would be measured through a comprehensive assessment based on authentic reading materials requiring students to engage in a range of reading processes.
- The reading test would be designed so that future assessments could measure trends in achievement.
- Questionnaires would be administered to the tested students, their current reading teachers, and their school principals – to collect contextual data with which to interpret achievement.

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 provinces of Ontario and Quebec only.

1.4 Student Population Assessed

In 2001, PIRLS assessed the reading literacy of children in “the upper of the two grades with the most 9-year-olds at the time of testing” (PIRLS, 1999). This corresponds to the fourth grade in most countries. This population was chosen because it represents an important transition point in children’s development as readers. In most countries, by the end of fourth grade, children are expected to have learned how to read, and are now reading to learn. This grade is also assessed in the IEA’s Trends in International Mathematics and Science Study (TIMSS), to provide countries participating in both studies with achievement and background data for three subjects at the same grade level.

In each country, representative samples of students were selected using a two-stage sampling design. In the first stage, at least 150 schools were selected using probability-proportional-to-size sampling. Countries could incorporate in their sampling design important reporting variables (for example, urbanicity or school

type) as stratification variables. At the second stage, one or two fourth-grade classes were randomly sampled in each school. This resulted in a sample size of at least 3,750 students in each country. Some countries opted to include more schools and classes, enabling additional analyses, which resulted in larger sample sizes.

1.5 Assessment Dates

PIRLS was administered near the end of the school year in each country. In countries in the Northern Hemisphere (where the school year typically ends in May or June) the assessment was conducted in April, May, or June 2001. In the Southern Hemisphere, the school year typically ends in November or December; so in these countries, the assessment was conducted in October or November 2001.

1.6 Study Management and Organization

PIRLS is directed in the United States by Ina V. S. Mullis and Michael O. Martin, at the International Study Center (ISC) at Boston College. The PIRLS International Study Center was responsible for the design, development, and implementation of the study – including developing the instruments and survey procedures, ensuring quality in data collection, and analyzing and reporting the study results. The International Study Center worked closely with the organizations responsible for particular aspects of the study, the PIRLS advisory committees, and representatives of the participating countries.

Each country appointed a National Research Coordinator (NRC) who, together with staff members at the PIRLS National Center, was responsible for all aspects of the study within that country. The PIRLS ISC organized meetings of the NRCs several times throughout the study to review study materials and procedures, to receive training in scoring constructed-response items, and to enter the data using the prescribed software.

The IEA Secretariat provided guidance in all aspects of the study, and was responsible for managing the ambitious translation verification effort conducted for the field test and main assessment. Statistics Canada was responsible for all aspects of sampling – including working with countries to ensure that the international procedures are followed, adapting the international design to national conditions, documenting the national samples, and computing 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 (DPC) in Hamburg was responsible for processing and verifying

the data from the 35 countries, and for constructing the international database. Educational Testing Service provided software and support for scaling the achievement data.

The study directors and representatives from the International Study Center, IEA Secretariat, IEA DPC, Statistics Canada, the National Foundation for Educational Research in England and Wales, and Educational Testing Service met periodically to review the study's progress, procedures, and schedule.

The PIRLS Reading Development Group (see Appendix A) contributed their invaluable expertise to the framework and reading test. Committee members reviewed various drafts of the framework and assessment blocks, and reviewed and endorsed the final reading test. The PIRLS Questionnaire Development Group (see Appendix A) helped develop the PIRLS questionnaires (including writing items and reviewing drafts of all questionnaires).

1.7 Overview of Assessment Framework

At the heart of the PIRLS assessment is the definition of reading literacy established by the Reading Development Group, and refined by the National Research Coordinators. The PIRLS definition of reading literacy builds on the definition used in the IEA 1991 study, but elaborates on that definition by making specific reference to reading by children. PIRLS defines reading literacy as:

...the ability to understand and use those written language forms required by society and/or valued by the individual. Young readers can construct meaning from a variety of texts. They read to learn, to participate in communities of readers, and for enjoyment (Campbell, Kelly, Mullis, Martin, & Sainsbury, 2001).

Growing out of this definition are the three aspects of reading literacy assessed by PIRLS:

- Processes of comprehension
- Purposes for reading, and
- Reading behaviors and attitudes.

Processes of comprehension and purposes for reading are the foundation of the written assessment of reading comprehension. The purposes for reading and processes of comprehension, as well as the percentages of the assessment devoted to each, are shown in Exhibit 1.1. Each process is assessed with each purpose for reading. Reading behaviors and attitudes are assessed through a questionnaire completed by the students.

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

		Purposes for Reading		
		Literary Experience	Acquire and Use Information	
	Focus on and Retrieve Explicitly Stated Information	9%	13%	13%
	Make Straightforward Inferences	14%	9%	24%
	Interpret and Integrate Ideas and Information	20%	20%	36%
	Examine and Evaluate Content, Language, and Textual Elements	6%	8%	26%
		50%	50%	

¹ Because numbers are rounded to the nearest whole number, some totals may appear inconsistent.

1.8 PIRLS Reading Assessment

PIRLS had ambitious goals for covering the domain of reading literacy. The Reading Development Group felt that at least eight passages and items (four for each reading purpose) were needed to provide a valid and reliable measure of reading achievement. Since it would not be possible to administer the entire test to any one child, PIRLS used a matrix sampling technique to distribute the assessment material among students, yet retain linkages necessary for scaling the achievement data.

Assessment Design

The material was divided into 40-minute “blocks,” each comprising a passage (a story or article) and items representing at least 15 score points. There are eight such blocks, four for each reading purpose. The eight assessment blocks are distributed across ten test booklets, and each student completed one booklet in an 80-minute testing session. Each booklet contains two blocks, and most blocks appeared in three booklets. One of the ten booklets is the PIRLS Reader, a color booklet containing two reading passages; the test items are located in a separate booklet. The two blocks comprising the Reader appear only in that booklet. The distribution of blocks across booklets “links” the booklets to enable the achievement data to be scaled using item response theory methods.

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 constructed-response 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 guides, and illustrate the kinds of responses garnering different points.

Releasing Assessment Material to the Public

The PIRLS test design provides for the release of half of the assessment material into the public domain after data collection, including the entire PIRLS Reader. The remaining half were kept secure and included in future PIRLS assessments so trends in achievement can be measured. After data collection in 2001, four blocks, one literary and one informational block (as well as the blocks in the Reader), were released to the public.

1.9 PIRLS Background Questionnaires

By gathering information about children's experiences together with reading achievement on the PIRLS test, it is possible to identify the factors or combinations of factors that relate to high reading literacy. An important part of the PIRLS design

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.

Student Questionnaire

Each student taking the PIRLS reading assessment completes 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.

Learning to Read (Home) Survey

The learning to read survey is completed by the parents or primary caregivers of each student taking the PIRLS reading assessment. It addresses child-parent literacy interactions, home literacy resources, parents' reading habits and attitudes, home-school connections, and basic demographic and socioeconomic indicators.

Teacher Questionnaire

The reading teacher of each fourth-grade class sampled for PIRLS completes a questionnaire designed to gather information about classroom contexts for developing reading literacy. This questionnaire asks teachers about characteristics of the class tested (such as size, reading levels of the students, and the language abilities of the students). It also asks about instructional time, materials and activities for teaching reading and promoting the development of their students' reading literacy, and the grouping of students for reading instruction. Questions about classroom resources, assessment practices, and home-school connections also are included. The questionnaire also asks teachers for their views on opportunities for professional development and collaboration with other teachers, and for information about their education and training.

School Questionnaire

The principal of each school sampled for PIRLS responds to the school questionnaire. It asks school principals about enrollment and school characteristics (such as where the school is located, resources available in the surrounding area, and indicators of the socioeconomic background of the student body), characteristics of reading education in the school, instructional time, school resources (such as the availability of instructional materials and staff), home-school connections, and the school climate.

1.10 Translation and Verification of Instruments

The PIRLS reading test and questionnaires were prepared in English, then translated into 31 other languages for use in the 35 participating countries. Countries were responsible for translating the instruments into their local language (or languages) following internationally prescribed procedures. To ensure standardization of instruments across countries, PIRLS undertook an extensive verification process, whereby each country's data collection instruments were independently reviewed and verified by an external translation company engaged by the IEA. The verifiers' reviews of the translated documents were used to improve the translations.

Instruments were verified twice, once before the field test and once before the main data collection. In addition to the external review, the International Study Center also reviewed the countries' instruments against the verifiers' comments to ensure that all necessary corrections were made. Finally, statistical analyses of item data were conducted to check for evidence of differences in student performance across countries that could indicate a translation problem.

1.11 Data Collection

Each country was responsible for carrying out all aspects of the data collection, using standardized procedures developed for the study. Manuals provided explicit instructions to the NRCs and their staff members on all aspects of the data collection – from contacting sampled schools to packing and shipping materials to the IEA Data Processing Center for processing and verification. Manuals were also prepared for test administrators and for individuals in the sampled schools who work with the national centers to arrange for the data collection within the schools. These manuals addressed all aspects of the assessment administration within schools (including test security, distribution of booklets, timing and conduct of the testing session, and returning materials to the national center).

The PIRLS International Study Center placed great emphasis on monitoring the quality of the PIRLS data collection. In particular, the International Study Center implemented an international program of site visits, whereby international quality control monitors visited a sample of 15 schools in each country and observed the test administration. In addition to the international program, NRCs were also expected to organize an independent national quality control program based upon the international model. The latter program required national quality control observers to document data collection activities in their country. The national quality control observers visited a random sample of 10 percent of the schools (in addition to those visited by the international quality control monitors), and recorded their observations from the testing sessions for later analysis.

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 methodology produces a score by averaging the responses of each student to the items that he or she took, in a way that takes into account the difficulty and discriminating power of each item. The approach followed in PIRLS uses information from the background questionnaires to provide improved estimates of student performance (a process known as conditioning) and multiple imputation to generate student scores (or “plausible values”) for analysis and reporting.

The IRT analysis provides a common scale on which performance can be compared across countries. In addition to providing a basis for estimating mean achievement, scale scores permit estimates of how students within countries vary and provide information on percentiles of performance. The PIRLS scale score average across countries was set to 500 and the standard deviation to 100. Since the countries varied in size, each country was weighted to contribute equally to the mean and standard deviation of the scale. The average and standard deviation of the scale scores are arbitrary and do not affect scale interpretation.

In the PIRLS analysis, achievement scales were produced for each of the two reading purposes, reading for literary experience and reading for information, as well as for reading overall.

1.15 Data Analysis and Reporting

The *PIRLS 2001 International Report* (Mullis, Martin, Gonzalez & Kennedy, 2003) summarizes fourth-grade students’ student reading achievement in each country. This report presents average student achievement in reading overall, as well as in reading for literary experience and reading to acquire and use information, together with standard errors and tests of significance as appropriate. Average achievement is reported separately for girls and boys.

To provide additional information about reading achievement among high- and low-achieving students, PIRLS reported the percentage of students in each country performing at each of four international benchmarks of student achievement – corresponding to the 25th, 50th, 75th, and 90th percentiles of the international distribution of reading achievement. To enhance this reporting approach, PIRLS conducted a scale anchoring analysis to describe student performance at the international benchmarks in terms of the kind of reading students performing at each benchmark can do, and the skills and strategies they use to demonstrate their comprehension of the text. Complementing this approach further, the PIRLS international report presents examples of questions from both literary and informational passages that anchor at each of the benchmarks (providing another perspective on the reading demands of the benchmarks), and also displays student performance in each country on the example questions.

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 volume providing general information on the cultural, societal, and economic situation in each country, as well as a more focused perspective on the structure and organization of their respective educational systems as it pertains specifically to the promotion of reading literacy. Consisting of a chapter from each country, the *PIRLS 2001 Encyclopedia* describes primary/elementary schooling as it pertains to classroom organization and instruction, and assessment practices. As such, it is an extremely valuable companion publication to the international report, providing insights and detailed information about the policies, practices, and resources within each country.

1.16 The Trends in IEA's Reading Literacy Study

While PIRLS 2001 is the first in a cycle of assessments designed to measure trends in reading achievement, some countries also measured achievement trends from 1991 to 2001. Countries that participated in the IEA's 1991 Reading Literacy Study were eligible to administer the 1991 reading test and student questionnaire to a sample of students in 2001 so that they could obtain information about how their children's reading literacy today compares with that of ten years ago. The following countries participated in the trend study:

Greece	Singapore
Hungary	Slovenia
Iceland	Sweden
Italy	United States
New Zealand	

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, the 1991 target grade and the PIRLS target grade were not the same. Statistics Canada worked with these countries to tailor the design so as to achieve a representative sample of students. The IRT scaling methodology used with PIRLS 2001 also was applied in the trends in reading literacy study. The results of the trend study are presented in the *Trends in Children's Reading Literacy Achievement 1991-2001: IEA's Repeat in Nine Countries of the 1991 Reading Literacy Study* (Martin, Mullis, Gonzalez, & Kennedy, 2003).

2

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
	<i>Exhibit 2.1: Distribution of Literacy and Informational Blocks Across Booklets</i>	2-6
2.3	Released Status for PIRLS Passages	2-6

2.1 Overview

The development of the PIRLS reading assessment took place over a two-year period, from 1999 to 2001. The work was undertaken by a team from the National Foundation for Educational Research in England and Wales, with support and advice at all stages from the PIRLS Reading Coordinator, the Reading Development Group, the National Research Coordinators, the PIRLS Project Management Team, and staff of the PIRLS International Study Center at Boston College. Test development was based firmly on the *Framework and Specifications for the PIRLS Assessment 2001* (Campbell, et al., 2001). The framework presents a view of reading literacy as a complex interactive process. It identifies two main purposes for reading relevant to the age group selected for the assessment: reading for literary experience, and reading to acquire and use information. The framework specifies four principal comprehension processes that readers use to construct meaning that are the same for both reading purposes. The assessment requires passages that offer students an authentic engagement with text, and items that draw upon the central qualities of that engagement.

2.2 The PIRLS Assessment Framework

The PIRLS assessment development effort was guided by the description of reading literacy in the PIRLS assessment framework. Central to the framework is its definition of reading literacy:

The ability to understand and use those written language forms required by society and/or valued by the individual. Young readers can construct meaning from a variety of texts. They read to learn, to participate in communities of readers, and for enjoyment.

The view of literacy embodied in this definition – and described in more detail throughout the framework – is derived from and informed by numerous theories of reading. The framework was not intended to reflect any single theory of reading or approach to reading instruction. Rather, it was based on a multinational consensus about the nature of reading literacy, the goals of reading instruction, and the expectations for developing readers in a literate society.

Development of a thorough and theoretically cohesive framework was a necessary first step in the instrument development process. The framework provided explicit descriptions of the types of reading material that were to be represented in the assessment, and the types of comprehension questions that were to be developed to measure students' understandings of the reading material. In describing the types of reading materials to be used in the assessment, the focus was on purposes for reading. Because readers often approach different types of texts for different reading

purposes, and because it is expected that students by age 9 should have developed the ability to read for a variety of purposes, the characterization of text types by purposes for reading provided assurance of broad construct coverage in the assessment. While reading for different purposes, readers engage in a variety of processes to comprehend text. As such, a description of comprehension processes was included in the framework to guide the development of test questions.

The following sections provide a description of the text types (purposes for reading) and the item types (processes and strategies) that were included in the framework – and that guided the instrument development process.

Text Types

Readers interact with text in different ways to construct meaning. Their approach to constructing meaning varies by the purpose for reading and the type of text being read. Certain purposes for reading are associated with certain types of text. For nine- and ten-year-old students, the two most common purposes for reading are reading for enjoyment and reading to learn. As such, the PIRLS framework specifies the inclusion of two broad types of text in the assessment: literary texts read for *literary experience or enjoyment*, and informative texts read to *acquire and use information*.

In reading for literary experience, readers engage with the text in order to become immersed in the world portrayed by the author. Readers may vicariously experience a world unfamiliar to them, or make connections and find similarities between the text and their own experiences. Young readers by age 9 have already developed an awareness of narrative text structures and use of language, upon which they draw to construct meaning and to react to the text. The PIRLS framework called for the inclusion of literary texts that represent the types of narrative structures and language usages most common to 9-year-old readers. The main form of literary text used in the assessment was narrative fiction.

In reading to acquire and use information, the reader is mostly focused on understanding the aspects of the real world described in the text. In addition, depending on the nature of the text and the reader's orientation, the text may evoke an action or response – as in following a set of directions or reacting to a persuasive argument or appeal. The type of texts that fall into this category may be structured chronologically or logically. Examples of texts that may be structured in a chronological manner include biographical accounts of the lives of contemporary or historical figures and procedural documents that detail step-by-step directions to be followed in sequence. Examples of texts that are structured logically many include those that are written to provide information about a given topic and those that are intended to persuade or convince the reader to think and act in a certain manner. Often, these texts include adjunct aids (such as charts, pictures, and graphs) to

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 meaning of such information – it is explicit, and may be viewed as existing at 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 such, skilled readers will often make such an inference automatically as they become engaged in constructing meaning within a specific part of the text, or as they develop a more global understanding of the text's overall meaning.

In order to construct a more complete and richer understanding of the text, readers must be able to interpret and integrate ideas and information. With this type of process, the reader moves beyond the phrase or sentence level of text to make connections between textual ideas, synthesize information, or consider the broader implications of textual meaning. In doing so, readers often draw on their background knowledge and experiences to develop interpretations, which may vary slightly – depending upon the reader's perspective.

Readers shift from constructing meaning to a critical consideration of the text as they examine and evaluate content, language, and textual elements. The reader recognizes that the text has been written to convey ideas, feelings, and information. The textual content may be evaluated for its overall value, believability, or relevance to the reader. Its structural and linguistic features may also be judged for its effectiveness, completeness, or impact. In examining and evaluating the text, readers may draw upon their understanding of the world, and on their past reading experiences.

Test Booklet Design

The test booklet design used in the operational PIRLS assessment was based on several considerations. First, in order to ensure broad coverage of reading comprehension (as described in the PIRLS framework) a total of eight reading blocks – each block consisting of a single passage or set of passages accompanied by comprehension questions – were developed. Each block was developed to assess a single purpose for reading; the assessment comprised a total of four literary blocks and four informational blocks. Secondly, it was acknowledged that the burden required for each student to take the entire assessment (a total of more than five hours testing time) would be too great. Consequently, a matrix sampling technique was employed so that each student would take only a portion of the assessment (two reading blocks), and that an appropriately representative sample of students would be assigned to each portion. Finally, it was important to ensure adequate linking of results across blocks, since each student would not be administered the entire assessment.

With these considerations in mind, the four literary blocks and four informational blocks were distributed across 10 assessment booklets. Each student participating in the assessment was administered one of the 10 booklets. Because students were given 40 minutes to complete each block, the total assessment time was 80 minutes. (An additional 15 to 30 minutes was devoted to having students complete a background and instructional experience questionnaire.)

Exhibit 2.1 illustrates the distribution of literary and informational blocks across the 10 test booklets. The block designations L1, L2, L3, and L4 refer to the four literary blocks. The block designations I1, I2, I3, and I4 refer to the four informational blocks.

Although this booklet design does not provide for all possible combinations of literary and informational blocks (which would have resulted in twice the number of test booklets), it was determined that the block combinations represented here were more than adequate to provide for suitable linking between blocks. Each block appears in three booklets, and each block is combined with at least one block

assessing the same purpose for reading, and at least one block assessing the other purpose for reading. Note that the nature of booklet 10 (the PIRLS Reader), which links one specific literary and informational block, and made it impossible to link these blocks to others in the design without substantially increasing student assessment time. Consequently, booklet 10 was distributed across sampled students at three times the rate of the other booklets.

Exhibit 2.1: Distribution of Literacy and Informational Blocks Across Booklets

	Booklet 1	Booklet 2	Booklet 3	Booklet 4	Booklet 5	Booklet 6	Booklet 7	Booklet 8	Booklet 9	Booklet R (Reader)
Assessment Block	L1	L2	L3	I1	I2	I3	L1	I2	I3	L4
	L2	L3	I1	I2	I3	L1	I1	L2	L3	I4

2.3 Released Status for PIRLS Passages

As part of its design, PIRLS releases half of the assessment material into the public domain after data collection, including the entire PIRLS Reader. After data collection in 2001, four blocks were released to the public. The item information sheets contained in the international database indicate whether or not the item has been released for public use. The released reading passages and items are available on the internet at the PIRLS Website (<http://pirls@bc.edu>), and also are available in the *PIRLS 2001 International Report* (Mullis et al., 2003).

3

PIRLS Background Questionnaires

3.1	Overview	3-2
3.2	Framework for the Questionnaires	3-2
	<i>Exhibit 3.1 Contexts for the Development of Reading Literacy</i>	3-3
	<i>Exhibit 3.2 Factors within the Home, School, and National and Community Contexts Addressed by PIRLS</i>	3-3
3.3	PIRLS Main Survey Questionnaires	3-4
	Student Questionnaire	3-4
	<i>Exhibit 3.3 Content of the PIRLS Student Questionnaire</i>	3-5
	Learning-to-Read Survey (Home)	3-6
	<i>Exhibit 3.4 Content of the PIRLS Learning-to-Read Survey (Home Questionnaire)</i>	3-6
	Teacher Questionnaire	3-7
	<i>Exhibit 3.5 Content of the PIRLS Teacher Questionnaire</i>	3-7
	School Questionnaire	3-9
	<i>Exhibit 3.6 Content of the PIRLS School Questionnaire</i>	3-9
3.4	Reporting Questionnaire Data	3-11
	Summary Indices from Background Data	3-11
	<i>Exhibit 3.7 Summary Indices from Background Data in the PIRLS 2001 International Report</i>	3-13
	Reporting Student Questionnaire Data	3-16
	Reporting Teacher Questionnaire Data	3-16
	Reporting Parents' Questionnaire Data	3-17
	Reporting School Questionnaire Data	3-17

3.1 Overview

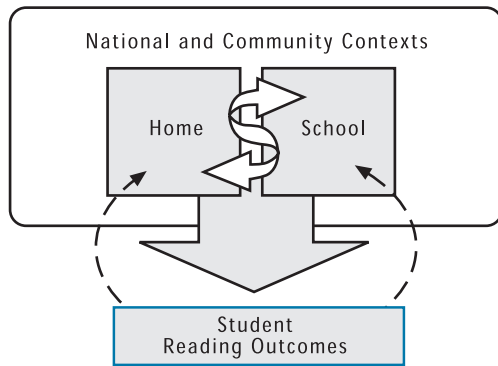
Children are exposed to language and print at home and at school, receive formal reading instruction, and see others reading for recreation and to perform tasks. These and other experiences and activities at home and school influence how well and how often children read, and how they feel about reading by the end of fourth grade. To be sure, not all children have the same experiences. Children have varying levels of home support for reading, and different levels of exposure to language and print throughout their lives. They also attend schools with different approaches to learning and resources with which to teach. Beyond influences within the home and at school are those in the local contexts in which children live and learn. Community size and resources, organization of the educational system, and educational decision-making affect homes and schools, and thus children's literacy development.

By gathering information about children's experiences in learning to read together with reading achievement on the PIRLS test, it is possible to identify the factors or combinations of factors associated with a high degree of reading literacy. The PIRLS design includes a set of questionnaires targeting important factors related to reading literacy. PIRLS administered four questionnaires to the tested students and their parents, reading teachers, and school principals.

3.2 Framework for the Questionnaires

The PIRLS questionnaires are grounded in a conceptual model relating reading outcomes – students' reading literacy achievement and attitudes – to home, school, and community and national contexts. Exhibit 3.1 illustrates how PIRLS conceptualizes the influences on children's reading by depicting the relationship between home and school, and how both are situated within the community and the country.

Exhibit 3.1 Contexts for the Development of Reading Literacy



The PIRLS questionnaires address factors within each of the aspects that are deemed important for the development of reading literacy. In addition to reading achievement, reading outcomes include students’ reading attitudes and behaviors. The factors within the home, school, and national and community contexts addressed by PIRLS are shown in Exhibit 3.2.

Exhibit 3.2 Factors within the Home, School, and National and Community Contexts Addressed by PIRLS

Factors
National and Community Contexts
Demographics and resources
Governance and organization of educational system
Curriculum characteristics and policies
Home Contexts
Activities fostering reading literacy
Language in the home
Home resources
Home/school connection
Students out-of-school literacy activities
School Contexts
School environment and resources
Teacher training and preparation
Classroom environment and structure
Instructional strategies and activities
Instructional materials and technology

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.

Exhibit 3.3 Content of the PIRLS Student 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 activities	Frequency student does various reading-related activities and watches television
4	Reading outside of school	Frequency student reads different types of texts outside school
5	Use of library	Frequency student borrows books from library for fun
6	Television watching	Frequency student watches television on a normal school day
7-8	Instructional activities	Frequency student does certain reading instructional activities in school
9-10	Homework	Frequency reading for homework is assigned and amount of time student spends on reading homework
11	Computer use	Frequency computer is used for different literacy activities, and where computer is used
12	Attitude toward reading	Student's attitude towards reading
13	Reading self-concept	Student's self-concept regarding his/her reading ability
14	Feelings about school	Student's feelings about school – safety, perception of students and teachers
15	School environment	Student's reports of problematic behavior by other students at school
16-18	Language in the home	Student's use of the language of the test at home (used as indicator of home environment and home support for reading in language of the test)
19	Books in the home	Number of books in student's home (used as indicator of home environment and socioeconomic status)
20	Home possessions	The presence of various socioeconomic indicators (used as indicator of home environment and socioeconomic status)
21-22	Persons living in home	Number of people and children living in the home (used as indicator of home environment and socioeconomic status)
23-25	Student and parents born in country	Provides information on immigrant status (used as indicator of home environment and home support for reading in language of the test)

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		
Item Number	Item Content	Description
1	Respondent	Who completed the survey
2	Parent/child literacy interactions	Frequency parents engaged in different literacy activities with child during early childhood
3	Attend kindergarten	Whether, and for how long, child attended kindergarten (or equivalent)
4	Age began school	Age when child began formal schooling
5	Literacy skills when began school	Child's literacy skills when he/she began formal schooling
6	Home literacy activities	Frequency parent engages in different literacy activities with child (now)
7	Home/school connection	Parents' perception of school's connection with home
8	View of school	Parents' opinion of school
9-10	Parents' literacy activities	Time parent spends reading for enjoyment, and frequency reading for different purposes
11	Parents' attitude towards reading	Parents' attitude towards reading
12-13	Books in home	Number of books (total and children's) in the home (used as an indicator of home environment and support for literacy, and for socioeconomic indicator)
14	Parents' education	Highest level of education completed by both parents (used as an indicator of home environment and socioeconomic status)
15-16	Occupational status	Employment status and type of profession of each parent
17-18	Wealth	Perception of wealth relative to others and annual income
19	Time	Amount of time required to complete questionnaire

Teacher Questionnaire

The reading teacher of each fourth-grade class sampled for PIRLS completed a teacher questionnaire, which was designed to gather information about classroom contexts for developing reading literacy. This questionnaire asks teachers to describe the general characteristics of the class tested, such as class size, and the reading levels and language abilities of the students. Several questions in the questionnaire focus on factors related to reading instruction, such as instructional time, materials, grouping of students for instruction, and activities to teach reading and promote the development of the students' reading literacy. The questionnaire also asks teachers about classroom resources, assessment practices, and efforts to maintain home-school connections. It also asks teachers for their views about opportunities provided for cooperation and collaboration with other teachers, for professional development, and for information about themselves and their education and training. This questionnaire requires about 30 minutes of the teacher's time. Exhibit 3.5 presents details regarding the items in the questionnaire.

Exhibit 3.5 Content of the PIRLS Teacher Questionnaire

Teacher Questionnaire		
Item Number	Item Content	Description
1	Class size	Number of students total, and in the grade tested in the class
2-5	Students in class	Describes the students in the class with respect to reading level, language ability, and reading/language services received
6-7	Language instruction	Whether language instruction is conducted as part of instruction in different curriculum areas or as a separate subject, how much time is spent on language instruction, and how frequently language homework is assigned
8	Reading instruction	Whether reading instruction is conducted as part of instruction in different curriculum areas or as a separate subject
9	Reading instructional time	Amount of time spent on reading instruction, and if that time is for formal reading instruction
10	Reading instruction frequency	Number of days per week reading instruction is provided
11	Instructional grouping	Whether, and how, students are grouped for reading instruction
12	Instructional material	Frequency teacher uses different materials in reading instruction (worksheets, textbooks, etc.)
13	Reading material	Frequency teacher uses different types of texts in reading instruction

Continued

Exhibit 3.5 Content of the PIRLS Teacher Questionnaire (Continued)

Teacher Questionnaire		
Item Number	Item Content	Description
14	Reading instructional materials and different abilities	How teacher uses reading instructional materials for students at different reading levels
15-17	Instructional activities	Frequency teacher has students do different reading instructional activities
18	Instructional media	Frequency, and how teacher uses media in reading instruction
19	Computer use	Availability and use of computers and internet for literacy activities
20	Classroom library	Availability, size, and use of classroom library or reading corner
21	School library use	Frequency teacher takes or sends students to the school library
22-23	Homework	Frequency teacher assigns reading for homework and how much time is expected to be spent on reading homework
24-25	Reading difficulties	Resources available to the teacher to deal with students' reading difficulties, and teacher's approach to dealing with reading difficulties
26-28	Assessment	Teacher's use of different assessment methods to monitor students' progress and performance in reading
29	Professional development	Teacher's perception of opportunities for professional development in school
30	Cooperation and collaboration	Frequency teacher meets with other teachers to discuss and plan reading curriculum or teaching approaches
31	Home/school connection	Frequency teacher meets with parents or sends students' work home
32	Expectations for success	Teacher's expectations for students' future success as readers
33-34	Teaching experience	Number of years teacher has been teaching altogether and teaching the grade tested in particular
35-36	Age and gender	Teacher's age and gender
37-39	Education/training	Teacher's highest level of education, teaching certification, and academic preparation for teaching reading
40	Professional development	Time teacher has spent in professional development in the last two years

Continued

Exhibit 3.5 Content of the PIRLS Teacher Questionnaire (Continued)

Teacher Questionnaire		
Item Number	Item Content	Description
41	Reading habits	Frequency teacher reads different material and reads for different purposes
42	Teach class	Whether the class is taught by the teacher only, or by a team of teachers teaching different subjects
43	Time	Amount of time required to complete questionnaire

School Questionnaire

The principal of each school sampled for PIRLS responded to the school questionnaire. It asks school principals about enrollment and school demographic characteristics, such as school location, resources available in the surrounding area, and indicators of the socioeconomic background of the student body. The school questionnaire also asks principals about reading curriculum policies and total instructional time for the school year. It also includes questions about resources, the availability of materials and staff, and perceptions of the school climate, as well as the interaction between the schools and the students' parents and families. The school questionnaire was designed to be completed in about 30 minutes. Exhibit 3.6 presents details on the items in the questionnaire.

Exhibit 3.6 Content of the PIRLS School Questionnaire

School Questionnaire		
Item Number	Item Content	Description
1	Grades	Grades below the grade test that are present in the school
2-3	Enrollment	Number of boys and girls in school, and in grade tested
4-5	Community	Size and type of community in which the school is located
6	Community resources	Resources available in the community in which the school is located
7-9	Student body	Describes students in the school with respect to stability of student body, family socioeconomic status, and academic abilities
10	Tracking	Whether classes are formed on the basis of ability/performance
11	Instructional time	Number of instructional days per year, amount of instructional time per week, and number of days per week school is open

Continued

Exhibit 3.6 Content of the PIRLS School Questionnaire (Continued)

School Questionnaire		
Item Number	Item Content	Description
12	Years with same teacher	Number of years students typically stay with the same teacher
13	Influence on curriculum	Influence on the school's curriculum at grade tested by national or regional curriculum and examinations/assessments, standardized tests, and parents' and students' wishes
14	Literacy skills of students beginning formal schooling	Literacy skills of the students in the school when they begin formal schooling
15-16	Reading literacy emphasis	Relative emphasis placed on reading, writing, and oral language skills by school, presence of own reading curriculum and programs in support of reading education
17	Coordination of reading instruction	Whether school has a policy to coordinate reading instruction across primary school grades
18	Instructional materials	Emphasis on different types of reading material in reading instruction for primary school students
19	Instructional emphasis	Emphasis on different literacy skills and activities at different grades in primary school
20	Reading instruction and different abilities	How reading instructional program is implemented for students at different reading levels
21	School library	Availability, size, and staffing of school library
22	Classroom libraries	Availability of classroom libraries in school
23	Computer availability and internet access	Availability of computers and access to the internet for instructional purposes
24	Instructional resources	Material factors affecting school's capacity to provide instruction
25-27	Home/school connection	Availability of programs offered by the school to families, frequency of activities involving parents, and percentage of parents involved in school activities
28-29	School climate	Principal's perception of teachers', parents', and students' attitudes and the severity of students' problem behavior
30-31	Teacher collaboration	Existence of a school policy to promote cooperation and collaboration among teachers and the frequency with which teachers meet to share or develop instructional materials and approaches
32	Principal's time	Percentage of time the principal spends on various roles and functions
33	Time	Amount of time required to complete questionnaire

3.4 Reporting Questionnaire Data

As described in this chapter, PIRLS 2001 used four questionnaires to gather information about students' home and school environments and their experiences in learning to read:

- Students answered questions pertaining to their home and school experiences in learning to read, including instructional experiences, self-perception and attitudes towards reading, out-of-school reading habits, computer use, home literacy resources, and basic demographic information.
- Parents or caregivers of the sampled students responded to questions about the students' early reading experiences, child-parent literacy interactions, parents' reading habits and attitudes, home-school connections, and demographic and socioeconomic indicators.
- The teachers of the sampled students were asked about characteristics of the class tested, instructional activities for teaching reading, classroom resources, assessment practices, and about their education, training, and opportunities for professional development.
- The principals of schools reported on enrollment and school characteristics, school organization for reading instruction, school staffing and resources, home-school connections, and the school environment.

The *PIRLS 2001 International Report* devotes five chapters to the questionnaire data, dealing with 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-concept, and out-of-school activities.

Summary Indices from Background Data

To summarize the information obtained from the background questionnaires concisely, and focus attention on educationally relevant support and practice, PIRLS sometimes combined information from a number of questions to form an index that was more global and reliable than the component questions. According to the responses of students, their parents, teachers, or school principals, students were placed in a "high," "medium," or "low" category for each index, with the high level being set so that it corresponds to conditions or activities generally associated with higher academic achievement. For example, a three-level index of home educational resources was constructed from students' responses to two questions about home educational resources: number of books in the home and educational aids in the home (computer, study desk/table for own use, books of their own, access to a daily newspaper); and parents' responses to two questions: number of children's books in the home, and parents' education. Students were assigned to the high level if there

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	Index 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., blocks with letters of the alphabet), play word games, or 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 level indicates an average of greater than 2.33 through 3. Medium level indicates an average of 1.67 through 2.33. Low level indicates an average of 1 to less than 1.67.
Index of Home Educational Resources	HER	4.60	Index based on students' responses to two questions about home educational resources: number of books in the home, and educational aids in the home (computer, study desk/table for own use, books of their own, access to a daily newspaper); and parents' responses to two questions: number of children's books in the home, and parents' education. High level indicates more than 100 books in the home, more than 25 children's books, 3 or 4 educational aids, and highest level of education for either parent is finished university. Low level indicates 25 or fewer books in the home, 25 or fewer children's books, 2 or fewer educational aids, and highest level of education for either parent is some secondary or less. Medium level includes all other combinations of responses.
Index of Parents' Attitudes Toward Reading	PATR	4.17	Index based on parents' agreement with the following: I read only if I have to, I like talking about books with other people, I like to spend my spare time reading, I read only if I need information, and Reading is an important activity in my home. Average is computed across the 5 items based on a 4-point scale: Disagree a lot = 1, Disagree a little = 2, Agree a little = 3, and Agree a lot = 4. Responses for negative statements were reverse-coded. High level indicates an average of greater than 3 through 4, Medium level indicates an average of 2 through 3, and Low level indicates an average of 1 to less than 2.

Continued

a Exhibit number in the international report where data based on the index were presented.

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 Reading for Homework	RFH	6.34	Index based on teachers' responses to two questions: How often do you assign reading as part of homework (for any subject)? In general, how much time do you expect students to spend on homework involving reading (for any subject) each time you assign it? High level indicates students are expected to spend more than 3 minutes at least 1-2 times a week. Low level indicates students are never assigned homework or are expected to spend no more than 30 minutes less than once a week. Medium level indicates all other combinations of frequencies.
Index of Home-School Involvement	HSI	7.90	Index based on principals' responses to how often and what percentage of students' parents participate in the following provided by the school: teacher-parent conferences; letters, calendars, newsletters, etc., sent home to provide information about school; written reports (report cards) of child's performance sent home; and events at school to which parents are invited. High level indicates that 4 or more times a year schools hold teacher-parent conferences and events at school attended by more than half of the parents; send home letters, calendars, newsletters, etc., with information about the school 7 or more times a year; and send written reports (report cards) of child's performance 4 or more times a year. Low level indicates schools never hold teacher-parent conferences, or if they do, only 0-25% of parents attend; schools never hold events, or do so only yearly, attended by 0-25% of parents; send home letters, calendars, newsletters, etc., no more than 3 times a year; and send home written reports of children's performance never or only once a year. Medium level indicates all other combinations.
Index of Principals' Perceptions of School Climate	PPSC	7.14	Index based on principals' characterization in their school: teachers' job satisfaction, teachers' expectations for student achievement, parental support for student achievement, students' regard for school property, and students' desire to do well in school. Average is computed on a 5-point scale: Very high = 1, High = 2, Medium = 3, Low = 4, and Very low = 5. High level indicates an average of 1 to less than 2.33. Medium level indicates an average of 2.33 through 3.67. Low level indicates an average of greater than 3.67 through 5.

Continued

a Exhibit number in the international report where data based on the index were presented.

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 2 through 3. Low level indicates an average of greater than 3 through 4.
Index of Students' Attitudes Toward Reading	SATR	8.1 and 8.2	Index based on students' agreement with the following: I read only if I have to, I like talking about books with other people, I would be happy if someone gave me a book as a present, I think reading is boring, and I enjoy reading. Average is computed on a 4-point scale: Disagree a lot = 1, Disagree a little = 2, Agree a little = 3, and Agree a lot = 4. Responses for negative statement were reverse-coded. High level indicates an average greater than 3 through 4. Medium level indicates an average of 2 through 3. Low level indicates an average of 1 to less than 2.
Index of Students' Reading Self Concept	SRSC	8.3 and 8.4	Index based on students' agreement with the following: reading is very easy for me, I do not read as well as other students in my class, and reading aloud is very hard for me. Average is computed on a 4-point scale: Disagree a lot = 1, Disagree a little = 2, Agree a little = 3, and Agree a lot = 4. Responses for negative statement were reverse-coded. High indicates an average of greater than 3 through 4. Medium indicates an average of 2 through 3. Low indicates an average of 1 to less than 2.

Continued

a Exhibit number in the international report where data based on the index were presented.

The exhibit that displays each index shows the percentages of students at each level of the index, together with their reading achievement. In addition, the percentage at the high level was displayed graphically, with the countries ranked in order.

Reporting Student Questionnaire Data

Reporting the data from the student questionnaire was fairly straightforward. Most of the exhibits in the international report that include data from the student questionnaire present weighted percentages of students in each country for each response category, together with the mean reading achievement of those students. International averages are also displayed for each category. In general, jackknife standard errors accompany the statistics reported. In addition to the exhibits showing percentages of students overall, the international report includes some information separately by gender. For gender-based exhibits, the percentages of boys and girls in each category were displayed, and the statistical significance of the difference between genders was indicated.

Reporting Teacher Questionnaire Data

The teacher of each PIRLS fourth-grade class was asked to complete a questionnaire to provide information about the students in the class, reading instruction for those students, their use of computers and library facilities, homework and assessment, and about the teacher's own education and professional training and development. Because the sampling for the teacher questionnaires was based on participating students, the teachers that responded do not necessarily represent all of the teachers of the target grade in each of the PIRLS countries. Rather, they represent teachers of the representative samples of students assessed. It is important to note that in the international report, the student was always the unit of analysis, even when information from the teacher questionnaires was being reported. That is, the data presented are the percentages of *students* whose teachers reported various characteristics or instructional strategies. Using the student as the unit of analysis makes it possible to describe the instruction received by representative samples of students. Although this approach may provide a different perspective from that obtained by simply collecting information from teachers, it is consistent with the PIRLS goals of illuminating students' educational contexts and performance.

Although the vast majority of the PIRLS classes were taught by a single teacher, in Sweden each class had two teachers, each of which completed a teacher questionnaire. For reporting in these cases, the student's sampling weight was divided between the teachers, so that the student's contribution to student population estimates thus remained constant regardless of the number of teachers. This was consistent with the policy of reporting attributes of teachers and their classrooms in terms of the percentages of students taught by teachers with these attributes.

Reporting Parents' Questionnaire Data

The PIRLS Learning to Read Survey was completed by the parents or primary caregivers of the students participating in the study. Like the teacher questionnaire, the data from the parents' questionnaire were linked to the student, who was always the unit of analysis, even when information from the parents' questionnaires was being reported. That is, the data presented are the percentages of students whose parents reported various characteristics or instructional strategies.

Reporting School Questionnaire Data

The principals of the selected schools in PIRLS completed questionnaires on the school contexts in which the learning and teaching of reading occur. Although schools constituted the first stage of sampling, the PIRLS school sample was designed to optimize the student sample, not to provide an optimal sample of schools. Therefore, like the teacher data, the school-level data were reported using the student as the unit of analysis to describe the school contexts for the representative samples of students. In general, the exhibits based on the school data present percentages of students in schools with different characteristics for each country and for the international average.



4

PIRLS Survey Operations Procedures

4.1	Overview	4-2
4.2	Responsibilities of the National Research Coordinator	4-2
4.3	Within-School Sampling Procedures	4-3
	Survey Tracking Forms	4-4
	Excluding Students from Testing	4-4
	Assigning Instruments to Students and Teachers	4-5
4.4	Within-School Assessment Activities	4-6
	Arranging the Testing Sessions	4-6
	Distributing Materials	4-6
	Test Administration	4-6
	Timing of the Testing Sessions	4-7
	Activities Following the Test Administration	4-8
4.5	Scoring the Constructed-Response Items	4-8
	Preparing Materials to Train the Scorers	4-8
	Documenting the Reliability of the Constructed-Response Scoring	4-9
4.6	Trends in IEA's Reading Literacy Study	4-10

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, creating and checking data files for the achievement test and questionnaire responses, and dispatching the completed files to the IEA Data Processing Center in Hamburg.

In addition to administering the PIRLS 2001 instruments, the Trends in IEA's Reading Literacy Study was an option for PIRLS countries that also participated in the 1991 IEA Reading Literacy Study. A separate section in this chapter describes the survey operations for the trend study.

4.2 Responsibilities of the National Research Coordinator

The NRC for each country had primary responsibility for carrying out the survey operations. The NRC was responsible for collecting and preparing the data for the PIRLS assessment according to the procedures specified internationally. Earlier

chapters of this report describe the tasks of the NRC with regard to choosing a sample of schools, and translating the achievement tests and questionnaires. This chapter focuses on activities associated with the data collection itself.

An important responsibility of the NRC was to identify a School Coordinator for each of the sampled schools to act as a liaison between the national center and the school. The primary role of the School Coordinator was to assist the NRC in the assessment activities within the school, such as the sampling and identification of classes, and ensuring that the administrative and testing materials were correctly distributed, completed, and collected. The School Coordinator also was responsible for identifying and training a Test Administrator to conduct the testing sessions. In some countries, the School Coordinator assumed the roles and responsibilities of the Test Administrator. Both the School Coordinator and Test Administrator received training materials that described their responsibilities in detail.

4.3 Within-School Sampling Procedures

The goal of the PIRLS sampling procedures was to select a nationally representative sample of students in each country. Sampling intact classrooms within randomly selected schools offered the simplest solution from an operational perspective, while optimizing the information gathered about the students and their teachers. Although this was the standard procedure, it could only be implemented where classes in a school constituted an exhaustive and mutually exclusive partition of the students in the grade. In order for a random sample of classes to result in a representative sample of students, every student in the target grade in each country had to belong to one (and only one) of the classes in the school.

A key step in the PIRLS within-school sample selection was the correct identification of classes and teachers. Before classes could be sampled, all eligible classes within the sampled school had to be identified. To that end, the NRC asked the School Coordinator to list (on a Class Listing Form) all classes in the target grade, along with the names of the teacher or teachers responsible for teaching reading to the students in those classes. From the list, the NRC then prepared the Class Sampling Form and applied a prescribed sampling algorithm to select at least one class at random. Within each school, a class identification number was assigned to each class in the target grades listed on the Class Sampling Form. The six-digit

class ID consisted of the four-digit school ID plus the two-digit identification number for the class within the school. All students in this class were then selected to participate in the testing.

Survey Tracking Forms

PIRLS 2001 relied on a set of tracking forms to implement and record the sampling of classes, teachers, and students. It was essential that these were used and completed accurately, since they indicated the particular test booklet that each student was to receive. They were also used to record participation in each school. In addition to facilitating the data collection, the tracking forms provided essential information for computing sampling weights, and in evaluating the quality of the sampling procedures. Although the tracking forms could be produced manually, NRCs were urged to use the W3S sampling software, which automated many of the sampling and tracking procedures. Once completed, all tracking forms were retained for review by staff at the ISC.

For each sampled class, the NRC asked the School Coordinator to provide a list of students, giving their names, dates of birth, and genders. The School Coordinator also confirmed the name(s) of the teacher(s) of the sampled class. Once the NRC received the list and verified that the requested information had been supplied, he or she transcribed it onto a Student Tracking Form. Each student listed was assigned an eight-digit student identification number consisting of the six-digit class ID plus a two-digit number corresponding to the student's sequential position on the Student Tracking Form. Three extra student records were created at the end of the list to provide for additional students, or for assigning replacement materials in the event that a student received a damaged test booklet or questionnaire. These three extra records were also assigned a student ID.

The NRC produced a Teacher Tracking Form listing each teacher of the students in the sampled class. Each teacher of the selected class was assigned a teacher ID that consisted of the four-digit school ID, followed by a two-digit number for the teacher within the school, and a two-digit running number that sequentially numbered the entries on the Teacher Tracking Form. The two-digit running number was referred to as the Teacher Link Number, which identified each unique occurrence of a teacher in the Teacher Tracking Form.

During the test administration, the Test Administrator and School Coordinator used the tracking forms to record student, parent, and teacher participation, then returned them to the NRC after the test – together with the completed test booklets and questionnaires.

Excluding Students from Testing

Although all students enrolled in the target grade were part of the target population and were eligible to be selected for testing, PIRLS recognized that some students in some schools would be unable to take part in the 2001 assessment because of a physical or mental disability. Accordingly, the sampling procedures provided for the exclusion of students with specified disabilities (see Chapter 5). Countries were

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.

A student questionnaire was prepared for each entry in the Student Tracking Form, including the three extra entries. The student IDs and names were identified on the student questionnaires so that students’ responses could be linked with their assessment booklets.

The teachers who taught the selected students were each given a teacher questionnaire. This questionnaire focused on the teacher’s instructional practices as they applied to the target classes in the grade tested for PIRLS. The Teacher Tracking Form indicated the target grade and class with respect to which the teacher should have responded.

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.

Activities Following the Test Administration

After the test administration, the School Coordinator was responsible for calculating the student response rate in each class and, if the participation rate in any class in the school was below 90 percent, for arranging for a makeup session. The School Coordinator then returned to the NRC all testing materials, including the completed Student Tracking Form, the Test Administration Form, and any unused materials. Any relevant information about the test administration in the school was communicated to the NRC by means of a Test Administration Form that was completed by the Test Administrator and School Coordinator.

4.5 Scoring the Constructed-Response Items

Reliable application of the scoring guides to the constructed-response questions and empirical documentation of the reliability of that process were essential to the success of PIRLS 2001. The PIRLS *Survey Operations Manual* contained information about arranging for staff and facilities for the constructed-response scoring effort required for the main survey. The manual outlined how to select and train the scorers, and specified the procedures for scoring the constructed-response items and double-scoring a random sample of at least 200 responses per item, to document scoring reliability.

In selecting those who were to do the scoring, NRCs took care to arrange for persons who were conscientious and attentive to detail, knowledgeable in reading, and willing to apply the scoring guides as stated – even if they disagreed with a particular definition or category. Good candidates for scoring included teachers, retired teachers, college or graduate students, and staff members from educational agencies, ministries, or research centers.

Preparing Materials to Train the Scorers

The success of assessments containing constructed-response questions depends upon the reliability of scoring student responses. In PIRLS 2001, reliability was assured through the provision of scoring guides (manuals), extensive training in their use, and monitoring of the quality of the work. In addition, PIRLS 2001 provided training packets for training in selected questions, along with practice papers to help scorers achieve a consistent level of scoring.

Each scorer received a copy of the PIRLS Scoring Guides for Constructed-Response Items. This document explained the PIRLS scoring system, which was designed to produce a rich and varied profile of the range of students' competencies in reading literacy. A description of the development and content of the scoring guides is provided in Chapter 2.

At international scoring training meetings for both the field test and the main survey, NRCs received training packets containing “anchor papers” (example student papers) and practice papers to help them achieve accuracy and consistency in scoring. About 10 to 15 responses were sufficient for most items, but the complexity of some scoring guides made additional practice papers necessary.

Documenting the Reliability of the Constructed-Response Scoring

In order to demonstrate the quality of the PIRLS 2001 scoring, it was important to document the agreement between scorers. To establish the scoring reliability, NRCs were required to have a random sample of at least 200 responses for each constructed-response item independently scored by two scorers. This number is equal to approximately 25 percent of the responses based on the typical sample size. With the exception of the test booklet that accompanied the Reader, each item appeared in three booklets, meaning that the scorers needed to score a sample of 67 of each booklet. The items in the booklet accompanying the Reader do not appear in any other booklet; the scorers needed to score a sample of 200 of these booklets.

The *Survey Operations Manual* provides a procedure for interleaving the double scoring of the sample of reliability booklets with the regular booklet scoring, so that the reliability sample is scored in the same way and at the same time as the other booklets.

The activity described above provides evidence of the extent of agreement among scorers within each country, but does not address the question of scoring consistency across countries. Since PIRLS is administered in each country’s own language, and since scorers generally do not know other countries’ languages, cross-country scoring reliability is much more difficult to establish. As a partial solution to this problem, PIRLS took samples of student responses to a selection of constructed-response questions from each of the PIRLS countries that tested in English, and had this common set of English responses scored by two scorers in each country where scorers could operate through English. The common set consisted of 200 student responses to 25 questions from four of the PIRLS reading passages, two of which were literary and two informational. The cross-country reliability scoring was conducted in each country after the scoring of the national PIRLS data had been completed.

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 sample designs had to result in probability samples that gave accurate weighted estimates of population parameters such as means and percentages, and for which estimates of sampling variance could be computed. The PIRLS sample design is derived from the design of IEA's TIMSS (see Foy & Joncas, 2000), with minor refinements. Since sampling for PIRLS was to be implemented by the National Research Coordinator (NRC) in each participating country – often with limited resources – it was essential that the design be simple and easy to implement while yielding accurate and efficient samples of both schools and students. The design that was chosen for PIRLS strikes a good balance, providing accurate sample statistics while keeping the survey simple enough for all participants to implement.

The international project team provided manuals and expert advice to help NRCs adapt the PIRLS sample design to their national system, and to guide them through the phases of sampling. The *School Sampling Manual* (PIRLS, 1999) describes how to implement the international sample design to select the school sample; and offers advice on initial planning, adapting the design to national situations, establishing appropriate sample selection procedures, and conducting fieldwork. The *Survey Operations Manual* and *School Coordinator Manual* (PIRLS, 2001b, 2001a) provide information on sampling within schools, assigning assessment booklets and questionnaires to sampled students, and tracking respondents and non-respondents. To automate the rather complex within-school sampling procedures, NRCs were provided with sampling software jointly developed by the IEA Data Processing Center and Statistics Canada (IEA, 2001).

As well as administering the PIRLS 2001 instruments, countries that had participated in IEA's 1991 Reading Literacy Study had the option of using their national 1991 Reading Literacy Study instruments to measure trends in reading achievement between 1991 and 2001. This component of PIRLS 2001 was known as the Trends in IEA's Reading Literacy Study. The *School Sampling Guide for the 10-Year Trend Study* (PIRLS, 2000) describes how to implement the international sample design for the trend study.

In addition to sampling manuals and software, expert support was made available to help NRCs with their sampling activities. Statistics Canada (in consultation with the PIRLS sampling referee) reviewed and approved the national sampling plans, sampling data, sampling frames, and sample implementation. Statistics Canada also provided advice and support to NRCs at all stages of the sampling process, drawing national school samples for more than half of the PIRLS participants.

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 of the population by the sampling plan, schools could be excluded, where necessary, from the sampling frame for the following reasons:

- They were in geographically remote regions.
- They were extremely small in size.
- They offered a curriculum or a school structure that was different from the mainstream educational system(s).
- They provided instruction only to students in the categories defined as “within-school exclusions.”

Within-school exclusions were limited to students who, because of some disability, were unable to take the PIRLS tests. NRCs were asked to define anticipated within-school exclusions. Because these definitions can vary internationally, they were also asked to follow certain rules adapted to their jurisdictions. In addition, they were to estimate the size of the included population so that their compliance with the 95 percent rule could be projected.

The general PIRLS rules for defining within-school exclusions included the following three groups:

- **Educable mentally-disabled students.** These are students who were considered, in the professional opinion of the school principal or other qualified staff members, to be educable mentally disabled – or who had been so diagnosed in psychological tests. This category included students who were emotionally or mentally unable to follow even the general instructions of the PIRLS test. It did not include students who merely exhibited poor academic performance or discipline problems.

- **Functionally-disabled students.** These are students who were permanently physically disabled in such a way that they could not perform in the PIRLS tests. Functionally-disabled students who could perform were included in the testing.
- **Non-native-language speakers.** These are students who could not read or speak the language of the test, and so could not overcome the language barrier of testing. Typically, a student who had received less than one year of instruction in the language of the test was excluded, but this definition was adapted in different countries.

A major objective of PIRLS was that the effective target population, the population actually sampled by PIRLS, be as close as possible to the international desired population. Each country had to account for any exclusion of eligible students from the international desired population. This applied to school-level exclusions as well as within-school exclusions. See Appendix B of the *PIRLS 2001 Technical Report* for a detailed account of sample implementation in each country.

5.3 Sample Design

The international sample design for PIRLS is generally referred to as a two-stage stratified cluster sample design. The first stage consists of a sample of schools, which may be stratified; the second stage consists of a sample of one or more classrooms from the target grade in sampled schools.

Units of Analysis and Sampling Units

The PIRLS analytical focus was on the cumulative learning of students, as well as on instructional characteristics affecting learning. The sample design, therefore, had to address the measurement both of characteristics thought to influence cumulative learning, and of those specific to the instructional settings. As a consequence, schools, classrooms, and students were all potential units of analysis; all had to be considered as sampling units in the sample design in order to meet specific requirements for data quality and sampling precision at all levels.

Although the second stage sampling units were intact classrooms, the ultimate sampling elements were students – making it important that each student from the target grade be a member of one (and only one) of the classes in a school from which the sampled classes would be selected.

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 principal units of analysis, the reliability of estimates of student characteristics was paramount. However, PIRLS planned to report extensively on school, teacher, and classroom characteristics, so it was necessary also to have sufficiently large samples of schools and classes. The PIRLS standard for sampling precision requires that all student samples have an effective sample size of at least 400 students for the main criterion variables. In other words, all student samples should yield sampling errors that are no greater than would be obtained from a simple random sample of 400 students.

An effective sample size of 400 students results in the following approximate 95 percent confidence limits for sample estimates of population means, percentages, and correlation coefficients:

- Means: $m \pm 0.1s$ (where m is the mean estimate, and s is the estimated standard deviation for students)
- Percentages: $p \pm 5\%$ (where p is a percentage estimate)
- Correlations: $r \pm 0.1$ (where r is a correlation estimate).

Furthermore, since PIRLS planned to conduct analyses at the school and classroom levels, at least 150 schools were to be selected from the target population. A sample of 150 schools yields 95 percent confidence limits for school-level and classroom-level mean estimates that are precise to within 16 percent of their standard deviations. To ensure sufficient sample precision for school-level analyses, some participants had to sample more schools than would have been selected otherwise.

The precision of multistage cluster sample designs is generally affected by the so-called clustering effect. Students are clustered in schools, and are also clustered in classrooms within the schools. A classroom – as a sampling unit – constitutes a cluster of students who tend to be more like each other than like other members of the population. The *intra-class correlation* is a measure of this within-class similarity. Sampling 30 students from a single classroom when the intra-class correlation is high will yield less information than a random sample of 30 students spread across all classrooms in a school. Such sample designs are less efficient, in terms of sampling precision, than a simple random sample of the same size. This clustering effect was considered in determining the overall sample size for PIRLS.

The size of the cluster (classroom) and the size of the intra-class correlation determine the magnitude of the clustering effect. For planning the sample size, therefore, each country had to identify a value for the intra-class correlation and a

value for the expected cluster size (this was known as the minimum cluster size). For PIRLS, the intra-class correlation for each country was estimated from past studies (such as TIMSS) or from national assessments. In the absence of these sources, an intra-class correlation of 0.3 was assumed. Since participants were sampling intact classrooms, the minimum cluster size was in fact the average classroom size.

Sample-design tables were produced and included in the *PIRLS School Sampling Manual*. These tables illustrate the number of schools necessary to meet the PIRLS sampling precision requirements for a range of values of intra-class correlations and minimum cluster sizes. PIRLS participants could refer to the tables to determine how many schools they should sample. Whenever the estimated number of schools to sample fell below 150, participants were asked to sample at least 150 schools.

The sample-design tables could be used also to determine sample sizes for more complex designs. For example, a number of strata could be constructed for which different minimum cluster sizes could be specified, thereby refining the national sample design in a way that might avoid special treatment of small schools (see Chapter 5 of the *PIRLS 2001 Technical Report* for more details).

Stratification

Stratification is the grouping of sampling units (e.g., schools) in the sampling frame according to some attribute or variable prior to drawing the sample. It is generally used for the following reasons:

- To improve the efficiency of the sample design, thereby making survey estimates more reliable
- To apply different sample designs or disproportionate sample-size allocations to specific groups of schools (such as those within certain states or provinces)
- To ensure adequate representation in the sample of specific groups from the target population.

Examples of stratification variables for school samples are: geography (such as states or provinces), school type (such as public and private), and level of urbanization (such as rural and urban). Stratification variables in the PIRLS sample design could be used explicitly, implicitly, or both:

- **Explicit stratification** consists of building separate school lists, or sampling frames, according to the stratification variables under consideration. Where, for example, geographic regions are an explicit stratification variable, separate school-sampling frames would be constructed for each region. Different sample designs, or different sampling fractions, would then be

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 allocation without the complexity of explicit stratification. It can also improve the reliability of survey estimates – provided the variables are related to school mean student achievement in reading literacy.

5.4 Trends in IEA’s Reading Literacy Study

PIRLS countries that had earlier participated in the 1991 IEA Reading Literacy Study had the option of undertaking the Trends in IEA’s Reading Literacy Study, which measured trends in reading achievement using IEA’s 1991 reading test and student questionnaire. Since the target population for the Trends in IEA’s Reading Literacy Study was similar (but not identical to) the PIRLS target population, it was possible to use the PIRLS school sample as the basis for the trend study sample. Accordingly, the sampling plan for the Trends in IEA’s Reading Literacy Study was simple: select every second school sampled for PIRLS, and from each of these, sample one additional classroom from the target grade. Since the sample of schools for the Trends in IEA’s Reading Literacy Study is essentially a sub-sample of the PIRLS sample of schools, most of the required sampling tasks were carried out during the PIRLS school sampling.

Trends in IEA’s Reading Literacy Study Target Population

The target population in 1991 was the grade with the greatest number of nine-year-olds at the time of testing, and to maintain comparability, the same population was targeted by the trend data collection in 2001. However, the PIRLS 2001 target population differs somewhat from the 1991 population in that PIRLS targeted the upper of the *two* grades with most nine-year-olds, and so the target grade in each country was not always the same for the two studies. These definitions yield the same target grade in Greece, Iceland, Italy, New Zealand, Slovenia, and the United States, but different ones in Hungary, Singapore, and Sweden.

Sample Design

In general, the sample for the trend study consisted of half of the PIRLS school sample, with one classroom chosen at random from the target grade in each of the sampled schools. The procedure was as follows:

- Select every second school sampled for PIRLS, starting randomly with the first or second school.
- Sample an extra classroom (in addition to the PIRLS classroom already sampled) within these selected schools.
- If a school sampled for both studies has only one classroom, assign that school and classroom to PIRLS and use the first replacement school for that school as the sampled school for the trend study.

5.5 Sampling Weights

Appropriate estimation of population characteristics based on the PIRLS samples requires that the PIRLS sample design be taken into account in all analyses. This is accomplished in part by assigning a weight to each respondent, where the sampling weight properly accounts for the sample design, takes into account any stratification or disproportional sampling of subgroups, and includes adjustments for non-response (Foy, 2003).

The students within each country were selected using probability sampling. This means that the probability of each student being selected as part of the sample is known. The inverse of this selection probability is the sampling weight. In a properly selected and weighted sample, the sum of the weights for the sample approximates the size of the population. As is the case in PIRLS, the sum of the sampling weights for a sample is an estimate of the size of the population of students within the country in the sampled grades. The sampling weights must be used whenever population estimates are required. The use of the appropriate sampling weights ensures that the different subgroups that constitute the sample are properly and proportionally represented in the computation of population estimates.

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 weighting factor, and the student weighting adjustment. In addition, three versions of each student's weight are provided – the “total student” weight, the “senate” weight, and the “house” weight – each with its own particular uses.

The variables described in this section are included in the Student Background and Student Achievement files. The meaning and interpretation of the weights in each of the files is the same. The weighting factors included in the student-level data files and their adjustment factors are as follows:

WGTFAC1 School Weighting Factor

This variable corresponds to the inverse of the probability of selection for the school where the student is enrolled.

WGTADJ1 School Weighting Adjustment

This is an adjustment that is applied to WGTFAC1 to account for non-participating schools in the sample. Multiplying WGTFAC1 by WGTADJ1 gives the sampling weight for the school, adjusted for non-participation.

WGTFAC2 Class Weighting Factor

This is the inverse of the probability of selection of the classroom within the school. Since, in general, only one classroom was selected per grade within each school, there was no need to compute an adjustment factor for the classroom weight.

WGTFAC3 Student Weighting Factor

This is the inverse of the probability of selection for the individual student within a classroom. In cases where an intact classroom was selected, the value is set to 1 for all members of the classroom.

■ ■ ■

¹ For the user not familiar with the data files included in the PIRLS International Database reading Chapter 8 is recommended before proceeding with this section.

WGTADJ3 Student Weighting Adjustment

This is an adjustment applied to the variable WGTAC3 to account for non-participating students in the selected school and/or classroom. Multiplying the variables WGTAC2, WGTAC3, 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 WGTAC1, WGTADJ1, WGTAC2, WGTAC3, 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

$$SENWGT_{g,i} = TOTWGT_{g,i} \left(\frac{500}{\sum_{i=1}^I TOTWGT_{g,i}} \right)$$

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

$$HOUWGT_{g,i} = TOTWGT_{g,i} \left(\frac{n}{\sum_{i=1}^I TOTWGT_{g,i}} \right)$$

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 variables HOUWGT over all schools within each country will add up to the sample size for that country. The variable HOUWGT is proportional to TOTWGT multiplied by the ratio of the sample size (n) divided by sum of the weights over all students in the grade. These sampling weights can be used when you want the actual sample size to be used in performing significance tests. Although some statistical computer software packages allow you to use the sample size as the

divisor in the computation of standard errors, others will use the sum of the weights, and this results in severely deflated standard errors for the statistics if TOTWGT is used as the weighting variable. When performing analyses using such software, it is recommended to use the variable HOUWGT as the weight variable. Because of the clustering effect in most PIRLS samples, it may also be desirable to apply a correction factor such as a design effect to the HOUWGT variable.

Weight Variables Included in the Student-Teacher Linkage Files

The individual student sampling weights generally should be used when you want to obtain estimates at the student level. The exception is when student and teacher data are to be analyzed together. In this case, a separate set of weights have been computed to account for the fact that a student could have more than one teacher. This set of weights is included in the Student-Teacher Linkage file and is listed below.

TCHWGT

This weight is computed by dividing the sampling weight for the student by the number of teachers that the student has. This weight should be used whenever you want to obtain estimates regarding students and their teachers.

The Student-Teacher Linkage file also includes variables that indicate the number of teachers the student has.

Weight Variables Included in the School Data Files

The PIRLS samples are samples of students within countries. Although they are made up of a sample of schools within the countries, the samples of schools are selected so that the sampling of students, rather than the sampling of schools, is optimized. In particular, the probability-proportional-to-size sampling methodology causes large schools to be oversampled. Several weight variables are included in the school files, as follows:

WGTFAC1 School Weighting Factor

This variable corresponds to the inverse of the probability of selection for the school where the student is enrolled.

WGTADJ1 School Weighting Adjustment

This is an adjustment that is applied to WGTFAC1 to account for non-participating 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 the Student Background and Student Achievement data files. This variable is included in the Student Background and the Student Achievement data files.

JKREP

The variable JKREP indicates the PSU and its value are used to determine how the student is to be used in the computation of the replicate weights. This variable can have values of either 1 or 0. Those student records with a value of 0 should be excluded from the corresponding replicate weight, and those with a value of 1 should have their weights doubled. This variable is included in the Student Background and the Student Achievement data files. For each individual student, this variable is identical in these two files.

Additionally, the variables JKCZONE and JKCREP are included in the school file.

JKCZONE

The variable JKCZONE indicates the sampling zone to which the school is assigned. The sampling zones can have values from 1 to 75 in the School data files. Schools selected with certainty receive the value of 0 for this variable.

JKCREP

The variable JKCREP can have values of either 1 or 0. It indicates whether this school is to be dropped or have its weight doubled when estimating standard errors. Those school records with a value of 0 should be excluded from the corresponding replicate weight, and those with a value of 1 should have their weights doubled.

5.7 Using Sampling Variables in the Analysis

It is necessary to use the information captured by the sampling variables whenever you are conducting analyses using the PIRLS data. Chapter 7 describes how to use these variables and information, and Chapters 9 and 10 provide further examples.



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
	Proficiency Estimation Using Plausible Values	6-5
	Generating Proficiency Scores	6-8
	Working with Plausible Values	6-9
6.3	Student Achievement Scores	6-11
	Achievement Scores in the Student Files	6-11
	<i>Exhibit 6.1 Plausible Values Based on PIRLS 2001 Content Area</i>	6-12
	<i>Exhibit 6.2 Plausible Values Based on Reading Literacy Study Content Area</i>	6-12
	<i>Exhibit 6.3 Raw Scores Included in the PIRLS Data Files</i>	6-13
	<i>Exhibit 6.4 Raw Scores Included in the Reading Literacy Study Data Files</i>	6-13
	<i>Exhibit 6.5 Standardized Scores Included in the PIRLS Data Files</i>	6-14
	<i>Exhibit 6.6 Standardized Scores Included in the Reading Literacy Study Data Files</i>	6-14
	<i>Exhibit 6.7 National Rasch Scores Included in the PIRLS Data Files</i>	6-15
	<i>Exhibit 6.8 National Rasch Scores Included in the Reading Literacy Study Data Files</i>	6-15
6.4	International and National Benchmarks of Achievement	6-16
	<i>Exhibit 6.9 PIRLS International Benchmarks of Reading Achievement</i>	6-16
	<i>Exhibit 6.10 PIRLS International Benchmarks of Reading for Literary Purposes Achievement</i>	6-17
	<i>Exhibit 6.11 PIRLS International Benchmarks of Reading to Acquire and Use Information Achievement</i>	6-17
	<i>Exhibit 6.12 Variables Indicating Achievement of International Benchmarks of Reading to Acquire Information Achievement</i>	6-17
	<i>Exhibit 6.13 Variables Indicating Achievement of National Benchmarks of Reading to Acquire Information Achievement</i>	6-19

6.1 Overview

To achieve its goal of broad coverage of the reading purposes and processes specified in the *PIRLS 2001 Framework and Specifications* (Campbell et al., 2001), the PIRLS 2001 assessment included a range of reading passages and items arranged into eight 40-minute assessment blocks. Each student participating in the assessment completed one student booklet made up of just two of these blocks, keeping individual student response burden to a minimum. PIRLS used a matrix-sampling design¹ to assign assessment blocks to student booklets so that a comprehensive picture of the reading achievement of fourth-grade students in each country could be assembled from the components completed by individual students. PIRLS relied on Item Response Theory (IRT) scaling to combine the student responses to provide accurate estimates of reading achievement in the student population in each country. The PIRLS IRT scaling also uses multiple imputation or “plausible values” methodology to obtain proficiency scores in reading for all students, even though each student responded to only a part of the assessment item pool.

This chapter first reviews the psychometric models and the multiple imputation or plausible values methodology used in scaling the PIRLS 2001 data, and then describes how this approach was applied to the PIRLS 2001 data and to the data from IEA’s Trends in Reading Literacy Study. The PIRLS scaling was conducted at the PIRLS International Study Center (ISC) at Boston College, with software and psychometric support from Educational Testing Service.²



¹ The PIRLS 2001 achievement test design is presented in Chapter 2.

² PIRLS is indebted to Matthias von Davier, Ed Kulick, and John Barone of Educational Testing Service for their advice and support.

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” trait, and various characteristics (or “parameters”) of the item. A three-parameter model was used with multiple-choice items, which were scored as correct or incorrect, and a two-parameter model for those constructed-response items with just two score points – which also were scored as correct or incorrect. Since each of these item types has just two response categories, they are known as dichotomous items. A partial credit model was used with polytomous constructed-response items (i.e., those with more than two score points).

Two- and Three-Parameter IRT Models for Dichotomous Items

The fundamental equation of the three-parameter (3PL) model gives the probability that a person whose proficiency on a scale k is characterized by the unobservable variable θ will respond correctly to item i :

(1)

$$P(x_i = 1 | \theta_k, a_i, b_i, c_i) = c_i + \frac{(1 - c_i)}{1.0 + \exp(-1.7a_i(\theta_k - b_i))}$$

■ ■ ■

³ This section describing the PIRLS scaling methodology has been adapted with permission from the *TIMSS 1999 Technical Report* (Yamamoto & Kulick, 2000).

⁴ For a description of IRT scaling see (Birnbaum, 1968); (Lord & Novick, 1968); (Lord, 1980); and (Van Der Linden & Hambleton, 1996). The theoretical underpinning of the imputed value methodology was developed by Rubin (1987), applied to large-scale assessment by Miserly (1991), and studied further by Miserly, Johnson, & Muraki (1992), and Beaton & Johnson (1992). Procedures used in PIRLS have been used in several other large-scale surveys, including Trends in International Mathematics and Science Study (TIMSS), the U.S. National Assessment of Educational Progress (NAEP), the U.S. National Adult Literacy Survey (NALS), the International Adult Literacy Survey (IALS), and the International Adult Literacy and Life Skills Survey (IALLS).

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;
- c_i 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)

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

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 l -th of m_i ordered score categories:

(3)

$$P(x_i = l | \theta_k, a_i, b_i, d_{i,l}, \dots, d_{i,m_i-l}) = \frac{\exp\left[\sum_{v=0}^l 1.7a_i(\theta_k - b_i + d_{i,v})\right]}{\sum_{c=0}^{m_i-1} \exp\left[\sum_{v=0}^c 1.7a_i(\theta_k - b_i + d_{i,v})\right]} = P_{il}(\theta_k)$$

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 $m_i - 1$;
- θ_k is the proficiency of person on a scale k ;

a_i is the slope parameter of item i , characterizing its discrimination power;

b_i is the location parameter of item i , characterizing its difficulty;

$d_{i,l}$ is category l threshold parameter.

Indeterminacy of model parameters of the polytomous model are resolved by setting $d_{i,0} = 0$, and setting the sum of the threshold parameters equal to 0.

For all of the IRT models there is a linear indeterminacy of the values of item parameters and proficiency parameters (i.e., mathematically equivalent but different values of item parameters can be estimated on an arbitrarily linearly transformed proficiency scale). This linear indeterminacy can be resolved by setting the origin and unit size of the proficiency scale to arbitrary constants, (such as a mean of 500 and a standard deviation of 100). The indeterminacy is most apparent when the scale is set for the first time.

IRT modeling relies on a number of assumptions, the most important being conditional independence. Under this assumption, item response probabilities depend only on θ_k (a measure of person proficiency) and the specified parameters of the item, and are assumed unaffected by the demographic characteristics or unique experiences of the respondents, the data collection conditions, or the other items presented in the test. Under this assumption, the joint probability of a particular response pattern x across a set of n items is given by:

$$P(x|\theta_k, \text{item parameters}) = \prod_{i=1}^n \prod_{l=0}^{m_i-1} P_{il}(\theta_k)^{u_{il}}$$

where $P_{il}(\theta_k)$ is of the form appropriate to the type of item (dichotomous or polytomous), m_i is equal to 2 for the dichotomously scored items, and u_{il} is an indicator variable defined by:

$$u_{il} = \begin{cases} 1 & \text{if response } x_i \text{ is in category } l \\ 0 & \text{otherwise} \end{cases}$$

Replacing the hypothetical response pattern with the real scored data, the above function can be viewed as a likelihood function to be maximized by a given set of item parameters. Once items were calibrated in this manner, a likelihood function for the proficiency θ_k was induced from student responses to the calibrated items. This likelihood function for the proficiency θ_k is called the posterior distribution of the θ s for each respondent.

Proficiency Estimation Using Plausible Values

Most cognitive skills testing is concerned with accurately assessing the performance of individual respondents for the purposes of diagnosis, selection, or placement. Regardless of the measurement model used, whether classical test theory or item

response theory, the accuracy of these measurements can be improved – that is, the amount of measurement error can be reduced – by increasing the number of items given to the individual. Thus, it is common to see achievement tests designed to provide information on individual students that contain more than 70 items. Since the uncertainty associated with each θ in such tests is negligible, the distribution of θ or the joint distribution of θ with other variables can be approximated using individual θ 's.

For the distribution of proficiencies in large populations, however, more efficient estimates can be obtained from a matrix-sampling design like that used in PIRLS 2001. This design solicits relatively few responses from each sampled respondent while maintaining a wide range of content representation when responses are aggregated across all respondents. With this approach, however, the advantage of estimating population characteristics is more efficiently offset by the inability to make precise statements about individuals. The uncertainty associated with individual θ estimates becomes too large to be ignored. In this situation, aggregations of individual student scores can lead to seriously biased estimates of population characteristics (Wingersky, Kaplan, & Beaton, 1987).

Plausible values methodology was developed as a way to address this issue by using all available data to estimate directly the characteristics of student populations and subpopulations, and then generating multiple imputed scores (called plausible values) from these distributions, which can be used in analyses with standard statistical software. A detailed review of plausible values methodology is given by Mislevy (1991).⁵

What follows is a brief overview of the plausible values approach. Let y represent the responses of all sampled students to background questions or background data of sampled students collected from other sources, and let θ represent the proficiency of interest. If θ were known for all sampled students, it would be possible to compute a statistic $t(\theta, y)$ – such as a sample mean or sample percentile point – to estimate a corresponding population quantity T .

Because of the latent nature of the proficiency, however, θ values are not known even for sampled respondents. One solution to this problem is to follow Rubin (1987) by considering θ as “missing data” and approximate $t(\theta, y)$ by its expectation given (x, y) , the data that actually were observed, as follows:

(4)

■ ■ ■

⁵ Along with theoretical justifications, Mislevy presents comparisons with standard procedures, discusses biases that arise in secondary analyses, and offers numerical examples.

$$t^*(x, y) = E[t(\theta, y) | x, y]$$

$$= \int t(\theta, y) p(\theta | x, y) d\theta$$

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 , the student's background variables y_j , 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 IALIS. 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 . It should be noted that this variance does not include the variability of sampling from the population.

Plausible values are not test scores for individuals in the usual sense, but rather are imputed values that may be used to estimate population characteristics correctly. When the underlying model is correctly specified, plausible values will provide consistent estimates of population characteristics – even though they are not generally unbiased estimates of the proficiencies of the individuals with whom they are associated.⁶

Plausible values for each respondent j are drawn from the conditional distribution $P(\theta_j | x_j, y_j, \Gamma, \Sigma)$, where Γ is a matrix of regression coefficients for the background variables, and Σ is a common variance matrix for residuals. Using standard rules of probability, the conditional probability of proficiency can be represented as Equation 5, where θ_j is a vector of scale values, $P(x_j | \theta_j)$ is the product over the scales of the independent likelihoods induced by responses to items within each scale, and $P(\theta_j | y_j, \Gamma, \Sigma)$ is the multivariate joint density of proficiencies of the scales, conditional on the observed value y_j of background responses and parameters Γ and Σ . Item parameter estimates are fixed, and regarded as population values in the computations described in this equation.

(5)

$$P(\theta_j | x_j, y_j, \Gamma, \Sigma) \propto P(x_j | \theta_j, y_j, \Gamma, \Sigma) P(\theta_j | y_j, \Gamma, \Sigma) = P(x_j | \theta_j) p(\theta_j | y_j, \Gamma, \Sigma)$$

■ ■ ■

⁶ For further discussion, see (Mislevy, Beaton, Kaplan, & Sheehan, 1992).

Conditioning

A multivariate normal distribution was assumed for $P(\theta | x_i, y_i, \Gamma, \Sigma)$, 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^c . 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 functions Γ involving background variables not conditioned on in this manner are subject to estimation error due to misspecification. The nature of these errors was discussed in detail in Mislevy (1991). In PIRLS 2001, however, principal component scores based on nearly all background variables were used. Those selected variables were chosen to reflect high relevance to policy, and to education practices. The computation of marginal means and percentile points of θ for these variables is nearly optimal.

The basic method for estimating Γ and Σ with the expectation and maximization (EM) procedure is described in Mislevy (1985) for a single scale case. The EM algorithm requires the computation of the mean θ , and variance Σ , of the posterior distribution in Equation 6.

Generating Proficiency Scores

After completing the EM algorithm, plausible values for all sampled students are drawn from the joint distribution of the values of Γ in a three-step process. First, a value of Γ is drawn from a normal approximation to $P(\Gamma, \Sigma | x_i, y_i)$ that fixes Σ at the value $\hat{\Sigma}$ (Thomas, 1993). Second, conditional on the generated value of Γ (and the fixed value of $\Sigma = \hat{\Sigma}$), the mean θ , and variance Σ_p of the posterior distribution in Equation 6 are computed using the methods applied in the EM algorithm. In the third step, the proficiency values are drawn independently from a multivariate normal distribution with mean θ and variance Σ_p . These three steps are repeated five times, producing five imputations of θ for each sampled respondent.

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 $u=2, \dots, M$, 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:

$$T. = \frac{\sum T_u}{5}$$

5. An estimate of the variance of T . is the sum of two components: an estimate of $Var(T_u)$ obtained as in step 4 and the variance among the T_u :

$$Var(T.) = \frac{Var_u}{M} + \left(1 + M^{-1}\right) \frac{(T_u - T.)^2}{M - 1}$$

The first component in V_M reflects uncertainty due to sampling respondents from the population; the second reflects uncertainty due to the fact that sampled respondents' θ s are not known precisely, but only indirectly through x and y .

Working with Plausible Values

Plausible values methodology was used in PIRLS 2001 to ensure the accuracy of estimates of the proficiency distributions for the PIRLS population as a whole, and particularly for comparisons between subpopulations. A further advantage of this method is that the variation between the five plausible values generated for each respondent reflects the uncertainty associated with proficiency estimates for individual respondents. However, retaining this component of uncertainty requires that additional analytical procedures be used to estimate respondents' proficiencies, as follows:

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:

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

where B_M is the variance among M imputed values and V_M is the final estimate of the variance of T . When B is small relative to U^* , the reference distribution for the incomplete-data statistic differs little from the reference distribution for the corresponding complete-data statistics. If, in addition, d is large, the normal approximation can be used instead of the t -distribution.

For k -dimensional t , such as the k coefficients in a multiple regression analysis, each U and U^* is a covariance matrix, and B is an average of squares and cross-products rather than simply an average of squares. In this case, the quantity $(T-t^*)V^{-1}(T-t^*)'$ is approximately F distributed with degrees of freedom equal to k and v , with v defined as above but with a matrix generalization of f_M :

$$f = \frac{(1 - M^{-1})Trace(BV^{-1})}{k}$$

A chi-square distribution with k degrees of freedom can be used in place of the above quantity $(T-t^*)V^{-1}(T-t^*)'$ for the same reason that the normal distribution can approximate the t distribution.

Statistics t^* , the estimates of ability conditional on responses to cognitive items and background variables, are consistent estimates of the corresponding population values T , as long as background variables are included in the conditioning variables. The consequences of violating this restriction are described by Beaton & Johnson (1990), Mislevy (1991), and Mislevy & Sheehan (1987). To avoid such biases, the PIRLS 2001 analyses included effectively all background variables in the conditioning.

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

considerably from replication to replication may be influenced by unreliability in the achievement measures, and considered to be suspect. When reporting international achievement all five plausible values are used. Chapter 7 details the procedure by which the information from the five plausible values is summarized. The plausible values are included in the Student Background, Student Achievement, and Student-Teacher Link files.

Exhibit 6.1 Plausible Values Based on PIRLS 2001 Content Area

Variable	Description
ASRREA01 – ASRREA05	Reading Overall Plausible Values
ASRLIT01 – ASRLIT05	Reading for Literary Purpose Plausible Values
ASRINF01 – ASRINF05	Reading to Acquire Information Plausible Values

Exhibit 6.2 Plausible Values Based on Reading Literacy Study Content Area

Variable	Description
ASRREA01 – ASRREA05	Reading Overall Plausible Values
ASRNAR01 – ASRNAR05	Reading Narrative Text Plausible Values
ASRDOC01 – ASRDOC05	Reading Documents Plausible Values
ASREXP01 – ASREXP05	Reading Expository Text Plausible Values

Raw Scores in the Student Files

After the items were recoded as right or wrong, or scored to their level of correctness in the case of the open-ended items, raw scores were computed by adding the number of points obtained across the items for a block. Multiple-choice items received a score of either 1 or 0. Constructed-response items received score points from 0 to 3 depending on their scoring guide. A description of the algorithm used to score the items can be found in Chapters 9 and 10 in the section “Scoring the Items.”

Although these scores can be used to compare students’ performances on the same block, they should not be used to compare students’ performances across different blocks. Different blocks contain different numbers of items, and the specific items contained in one block have varying difficulties. It is recommended that these scores be used only to verify whether the items have been recoded correctly when a user decides to recode the items to their level of correctness. Raw scores for the individual students can be found in the Student Achievement data files. Variable names for the PIRLS 2001 and Trends in IEA’s Reading Literacy Study data files are listed in Exhibit 6.3 and 6.4, respectively.

Exhibit 6.3 Raw Scores Included in the PIRLS Data Files

Variable	Description
ASASCPT	Number of raw score points obtained on the Antarctica Block
ASCSCPT	Number of raw score points obtained on the Clay Block
ASFSCPT	Number of raw score points obtained on the Flowers Block
ASHSCPT	Number of raw score points obtained on the Hare Block
ASLSCPT	Number of raw score points obtained on the Leonardo Block
ASMSCPT	Number of raw score points obtained on the Mice Block
ASNSCPT	Number of raw score points obtained on the Pufflings Block
ASRSCPT	Number of raw score points obtained on the River Trail Block

Exhibit 6.4 Raw Scores Included in the Reading Literacy Study Data Files

Variable	Description
ASASCPT	Number of raw score points obtained on all the items
ASDSCPT	Number of raw score points obtained on the document items
ADESCPT	Number of raw score points obtained on the expository items
ASNSCPT	Number of raw score points obtained on the narrative items

Standardized Raw Scores

Because of the difficulty in making any comparisons across the test blocks using only the number of raw score points obtained on a set of items, raw scores were standardized by booklet to provide a simple score that could be used in comparisons across blocks for preliminary analyses. The standardized score was computed so that the weighted mean score within each block is equal to 50, and the weighted standard deviation is equal to 10.

The standardized raw scores were used in the initial item analysis for computing the discrimination coefficients for each of the items in the test. This initial item analysis was conducted prior to scaling the test items. Listed in Exhibits 6.5 and 6.6, the standardized raw scores for the individual students can be found in the Student Achievement data files.

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.

When computing the student logit scores the responses marked as “not reached” were treated as incorrect responses. This avoided unfairly favoring students who started answering the test and stopped as soon as they did not know the answer to a question. Logit scores for the students generally ranged between -4 and +4. When using maximum likelihood estimation procedures it is not possible to obtain finite logit scores for students who correctly answered all or none of the items. Scores for these students were set to +5 and -5 logits, respectively. These logit scores were then standardized to have a weighted mean of 150 and a standard deviation of 10 for each country.

The national Rasch scores should not be used for international comparisons for two reasons: they were computed with a different set of item difficulties for each country, and the weighted mean score within each country is always equal to 150. National Rasch scores, listed in Exhibits 6.7 and 6.8, can be found in the Student Achievement data files.

Exhibit 6.7 National Rasch Scores Included in the PIRLS Data Files

Variable	Description
ASANRSC	National Rasch Score on the Antarctica Block
ASCNRSC	National Rasch Score on the Clay Block
ASFNRSC	National Rasch Score on the Flowers Block
ASHNRSC	National Rasch Score on the Hare Block
ASLNRSC	National Rasch Score on the Leonardo Block
ASMNRSC	National Rasch Score on the Mice Block
ASNNRSC	National Rasch Score on the Pufflings Block
ASRNRSC	National Rasch Score on the River Trail Block

Exhibit 6.8 National Rasch Scores Included in the Reading Literacy Study Data Files

Variable	Description
ASANRSC	National Rasch Score on the all the Items
ASDNRSC	National Rasch Score on the Document items
ASENRSC	National Rasch Score on the expository items
ASNNRSC	National Rasch Score on the narrative items

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.

If all countries had the same distribution of student achievement, approximately 10 percent of students within each country would be above the 90th percentile in the international distribution, regardless of the country's population size.

Because of the imputation technology used to derive the proficiency scores, the international benchmarks had to be computed once for each of the five plausible values and the results averaged to arrive at the final figure. The international benchmarks are presented in Exhibits 6.9, 6.10, and 6.11 for overall reading achievement and for reading achievement in each of the two reading purposes.

Exhibit 6.9 PIRLS International Benchmarks of Reading Achievement

Proficiency Score	25 th Percentile	50 th Percentile	75 th Percentile	90 th Percentile
Plausible Value 1	435.178	509.777	569.542	615.925
Plausible Value 2	435.061	510.318	570.008	615.568
Plausible Value 3	434.094	510.755	569.749	615.315
Plausible Value 4	434.538	509.698	570.179	614.598
Plausible Value 5	435.013	510.236	569.040	615.140

Exhibit 6.10 PIRLS International Benchmarks of Reading for Literary Purposes Achievement

Proficiency Score	25 th Percentile	50 th Percentile	75 th Percentile	90 th 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 6.11 PIRLS International Benchmarks of Reading to Acquire and Use Information Achievement

Proficiency Score	25 th Percentile	50 th Percentile	75 th Percentile	90 th Percentile
Plausible Value 1	434.366	508.756	566.086	610.699
Plausible Value 2	435.184	509.512	567.211	611.919
Plausible Value 3	435.459	509.504	566.590	610.547
Plausible Value 4	434.628	509.585	566.982	611.051
Plausible Value 5	433.962	510.192	566.463	611.418

The database for the PIRLS assessment contains a set of variables that indicate which international benchmark the student met. Because of the plausible value technology used in PIRLS, five variables are included for each scale score indicating, for each plausible value, which benchmark was reached. These variables are provided in Exhibit 6.12.

Exhibit 6.12 Variables Indicating Achievement of International Benchmarks of Reading to Acquire Information Achievement

Variable	Description
ASRIBM01-5	International Overall Reading Achievement Benchmark reached with each of the five plausible values
ASIIBM01-5	International Reading to Acquire and Use Information Achievement Benchmark reached with each of the five plausible values.
ASLIBM01-5	International Reading for Literary Purposes Achievement Benchmark reached with each of the five plausible values.

The following codes are used for these variables:

Code 1	Student performed below the 25 th international percentile.
Code 2	Student performed at or above the 25 th international percentile, but below the 50 th international percentile.
Code 3	Student performed at or above the 50 th international percentile but below the 75 th international percentile.
Code 4	Student performed at or above the 75 th international percentile but below the 90 th international percentile.
Code 5	Student performed at or above the 90 th international percentile.

Additionally, four points on the reading achievement scales were identified within each country for use as national benchmarks. The Top 10 percent national benchmark for each country was defined as the 90th percentile on the PIRLS scale, computed across all students within each country. This point on each scale (overall, and for each purpose) is the point above which the top 10 percent of students within the country scored. The upper quarter national benchmark is the 75th percentile on the scale, above which the top 25 percent of students within the country scored. The median national benchmark is the 50th percentile, above which the top half of students within the country scored. Finally, the lower quarter national benchmark is the 25th percentile, the point reached by the top 75 percent of students within the country.

Because of the nature of these national benchmarks, approximately 10 percent of students within each country would be above the 90th percentile in the national distribution, 25 percent would be above the 75th percentile, and so on. Because of the imputation technology used to derive the proficiency scores, the national benchmarks also had to be computed once for each of the five plausible values.

The database for the PIRLS assessment also contains a set of variables that indicate the national benchmarks met by each student. Because of the plausible value technology used in PIRLS, five variables are included for each subject indicating, for each plausible value, which national benchmark was reached. These variables are listed in Exhibit 6.13.

Exhibit 6.13 Variables Indicating Achievement of National Benchmarks of Reading to Acquire Information Achievement

Variable	Description
ASRNBM01-5	National Overall Reading Achievement Benchmark reached with each of the five plausible values.
ASINBM01-5	National Reading to Acquire and Use Information Achievement Benchmark reached with each of the five plausible values.
ASLNBM01-5	National Reading for Literary Purposes Achievement Benchmark reached with each of the five plausible values.

The following codes are used for these variables:

- Code 1 Student performed below the 25th national percentile.
- Code 2 Student performed at or above the 25th national percentile, but below the 50th national percentile.
- Code 3 Student performed at or above the 50th national percentile but below the 75th national percentile.
- Code 4 Student performed at or above the 75th national percentile but below the 90th national percentile.
- Code 5 Student performed at or above the 90th national percentile.



7

Estimating Sampling and Imputation Variance

7.1	Overview	7-2
7.2	Estimating Sampling Variance	7-2
	Computing Sampling Variance Using the JRR Method	7-3
	Construction of Sampling Zones for Sampling Variance Estimation	7-4
	Computing the JRR Replicate Weights	7-4
7.3	Estimating Imputation Variance	7-6
7.4	Combining Sampling and Imputation Variance	7-6
7.5	Comparing Achievement with the International Mean	7-7
7.6	Gender Differences	7-7

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)

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

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_h). 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 are then used in the calculations. Chapters 9 and 10 show how this approach allows standard statistical software such as SAS or SPSS to be used to compute JRR estimates of sampling variability. The replicate weight approach requires you to temporarily create a new set of weights for each pseudo-replicate sample. Each replicate weight is equal to k times the

overall sampling weight, where k can take values of zero, one, or two, depending on whether or not the case is to be removed from the computation, left as it is, or have its weight doubled. The value of k for an individual student record for a given replicate depends on the assignment of the record to the specific PSU and zone.

Construction of Sampling Zones for Sampling Variance Estimation

An important step in applying the JRR technique to the estimation of sampling variability is the assignment of the schools to implicit strata known as sampling zones. Most PIRLS sample designs in the participating countries called for a total of 150 sampled schools per target population. Each of these 150 schools was assigned to one of 75 sampling zones. These zones were constructed by pairing the sequentially sampled schools and assigning these pairs to a sampling zone. Since schools were generally sorted by a set of implicit stratification variables, the resulting assignment of sequentially sampled schools to sampling zones takes advantage of any benefit due to this implicit stratification. In cases when more than 75 pairs of schools were sampled within a country, schools were then assigned to sampling zones in a way such that no more than 75 sampling zones were defined. In some cases this meant assigning more than two schools to the same sampling zone.

Sampling zones were constructed within explicit strata. In cases when there was an odd number of schools in an explicit stratum, either by design or because of school-level non-response, the students in the remaining school were randomly divided into two “quasi” schools for the purposes of calculating the jackknife standard error. Each zone then consisted of a pair of schools or “quasi” schools. When computing replicate weights for the estimation of JRR sampling error, one member of each pair of schools is randomly selected to have its weights doubled, while the weights of the other member are set to zero.

The variable JKZONE indicates the sampling zone to which the student’s school is assigned. The sampling zones can have values from 1 to 75 in the Student Background and Student Achievement data files.

The variable JKREP indicates how the student is to be used in the computation of the replicate weights. This variable can have values of either 1 or 0. Those student records with a value of 0 should be excluded from the corresponding replicate weight, and those with a value of 1 should have their weights doubled. This variable is included in the Student Background and the Student Achievement data files.

Computing the JRR Replicate Weights

Having assigned the schools to zones, if it is desired to use standard statistical software such as SAS or SPSS for sampling variance estimation, a convenient way to proceed is to construct a set of replicate weights for each pseudo-replicate sample. In PIRLS, the

schools in the sample were assigned in pairs to one of the 75 zones indicated by the variable JKZONE, and within each zone the pair members were randomly assigned an indicator (u_i) represented by the variable JKREP. Next, the pair members were coded randomly to 1 or 0 so that one of the members of each pair had values of 1 on the variable u_i , and the remaining one a value of 0. This indicator determined whether the weights for the elements in the school in this zone were to be doubled or zeroed.¹ The replicate weight ($W_h^{g,i,j}$) for the elements in a school assigned to zone b is computed as the product of k_b times their overall sampling weight, where k_b could take values of zero, one, or two depending on if the case was not to contribute, was to be left unchanged, or to be counted twice for the computation of the statistic of interest. In the PIRLS database, the replicate weights are not permanent variables. They need to be created by the sampling variance estimation program. An example program that makes use of replicate weights in computing JRR estimates is provided in Chapters 9 and 10.

When creating the replicate weights the following procedure is followed:

1. Each sampled student is assigned a vector of 75 weights or $W_h^{g,i,j}$, where b takes values from 1 to 75.
2. The value of $W_0^{g,i,j}$ is the overall sampling weight which is simply the product of the final school weight, the appropriate final classroom weight, and the appropriate final student weight.
3. The replicate weights for a single case are then computed as:

(2)

$$W_b^{g,i,j} = W_0^{g,i,j} * k_{bi}$$

where the variable k_b for an individual i takes the value $k_{bi}=2*u_i$ if the record belongs to zone b , and $k_{bi}=1$ otherwise.

In PIRLS, a total of 75 replicate weights were computed regardless of the number of actual zones within a country. If a country had fewer than 75 zones, then the replicate weights W_b , where b was greater than the number of zones within the country, were each the same as the overall sampling weight. Although this involves some redundant computation, having 75 replicate weights for each country has no effect on the size of the error variance computed using the jackknife formula, but facilitates the computation of standard errors for a number of countries at one time.

■ ■ ■

¹ Please note that, depending on the program, this information can be used differently. For example, in the examples described in this User Guide, students with a value of JKREP = 0 have their weights zeroed and those with value of 1 have their weights doubled. Programs such as Wesvar do the inverse. Both yield unbiased estimates.

7.3 Estimating Imputation Variance

The general procedure for estimating the imputation variance using plausible values is as follows: first estimate the statistic (t), each time using a different set of the plausible values (M). The statistics t_m can be anything estimable from the data, such as a mean, the difference between means, percentiles, etc. If all of the ($M=5$) plausible values in the PIRLS database are used, the parameter will be estimated five times, once using each set of plausible values. Each of these estimates will be called t_m , where $m=1,2,\dots,5$.

Once the statistics are computed, the imputation variance is computed as:

(3)

$$Var_{imp} = \left(1 + \frac{1}{M}\right) \times Var(t_m)$$

where M is the number of plausible values used in the calculation, and $Var(t_m)$ is the variance of the estimates computed using each plausible value.

7.4 Combining Sampling and Imputation Variance

When reporting standard errors for proficiency estimates using plausible values you will need to combine the sampling and imputation components of the error variance. Under ideal circumstances, and with unlimited computing resources, you would compute the imputation variance for the plausible values, and the JRR sampling variance for each of the plausible values. This would be equivalent to computing the same statistic up to 380 times (once for each of the five plausible values, and 75 times for each set of replicate weights for each plausible value). However, an acceptable shortcut is to compute the JRR variance component using one plausible value, and then the imputation variance using the five plausible values. Using this approach, the same statistic needs to be computed only 80 times. Under this procedure the error variance component for a statistic computed using plausible values is computed using the following formula:

(4)

$$Var(t_{pv}) = Var_{jrr}(t_1) + Var_{imp}$$

where $Var_{jrr}(t_1)$ is the sampling variance for the first plausible value.

This User Guide contains programs in SAS and SPSS that compute each of these variance components for the PIRLS data.

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)

$$se_{s_dif_j} = \sqrt{\frac{\left((N-1)^2 - 1 \right) se_j^2 + \sum_{k=1}^N se_k^2}{N}}$$

where $se_{s_dif_j}$ is the standard error of the difference due to sampling when country j is compared to the international mean; N is the number of countries used to compute the international mean; se_k^2 is the sampling standard error for country k ; and se_j^2 is the sampling standard error for country j .

The imputation component of the standard error was computed by taking the square root of the imputation variance calculated as follows:

(7)

$$se_{s_dif_j} = \sqrt{1 + \frac{1}{NPV} Var(d_{1..5})}$$

where NPV is the number of plausible values used in the analysis, and d_l is the difference between the international mean and the country mean for plausible value l .

Finally, the standard error of the difference is calculated as:

(8)

$$se_{dif_j} = \sqrt{se_{i_dif_j}^2 + se_{s_dif_j}^2}$$

7.6 Gender Differences

Because in most countries boys and girls attend the same schools, the samples of boys and girls cannot be treated as independent for the purpose of statistical tests.

Accordingly, a jackknife procedure applicable to correlated samples for estimating the standard error of the difference should be applied. This involves computing the

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.

8

Content and Format of Database Files

8.1	Overview	8-2
8.2	The Data Files	8-3
	Data File and Variable Naming Conventions	8-3
	<i>Exhibit 8.1 PIRLS and Reading Literacy Data Files</i>	8-5
	<i>Exhibit 8.2 Country Identification and Availability of Data Files</i>	8-6
	<i>Exhibit 8.3 Background Questionnaire Variable Location Conventions</i>	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
	<i>Exhibit 8.4 Data Almanac Files for PIRLS 2001 and Reading Literacy Study</i>	8-22
	Background Item Data Almanacs	8-22
	Achievement Item Data Almanacs	8-23
8.4	Codebook Files	8-24
8.5	Program Files	8-26
8.6	Item Information Files	8-27

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.

Item Information Files

The Item Information files contain information about the cognitive test items including item identifications, item type, scoring key, number of points, label, reading purpose, and reading process assessed.

The following sections describe each of the file types and how they can be used to access and analyze the PIRLS international data.

8.2 The Data Files

Data files are provided for each country that participated in PIRLS and for which internationally comparable data are available.

There are five basic types of data files in the PIRLS International Database:

- Background files with information from students, their parents, their teachers, and the principals of their schools¹
- Achievement files with responses to the PIRLS test
- Scoring Reliability files
- Student Teacher-Link files
- Reading Literacy files with background information from students and the test items administered as part of the IEA's Trends in Reading Literacy Study.

Data File and Variable Naming Conventions

Before proceeding to describe the different data files, this section will describe the file and variable-naming conventions used in the PIRLS and Trends in Reading Literacy Study databases. It also describes how the variables are organized in the database.

Data File Naming Convention

The data files included with this database are named using the following conventions:

■ ■ ■

¹ In PIRLS, questionnaires were administered to students, their teachers, school principals, and parents or guardians. In the Reading Literacy Study, questionnaires were only administered to the students.

- 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.

Exhibit 8.1 PIRLS and Reading Literacy Data Files

File Type	1995 Files
Student Achievement Test File	ASA<Country>R1
Student Background File	ASG<Country>R1
Teacher Background File	ATG<Country>R1
Learning to Read (Home) Survey File	ASH<Country>R1
School Background File	ACG<Country>R1
Student – Teacher Linkage File	AST<Country>R1
Student Achievement Scoring Reliability File	ASR<Country>R1
Reading Literacy Study Student File (1991)	ASC<Country>T1
Reading Literacy Study Student File (2001)	ASC<Country>T2

The three-character abbreviations used for each PIRLS country are listed in Exhibit 8.2, along with a numeric code for each country that is used to identify the country in the data files (see the following section discussing identification variables). Each file that is available within a country is annotated with a “•” in Exhibit 8.2.

Exhibit 8.2 Country Identification and Availability of Data Files

Country	ISO Code	Numeric Code	ASG 2001	ASA 2001	ASH 2001	ATG 2001	AST 2001	ACG 2001	ASC 1991	ASC 2001
Argentina	ARG	032	•	•	•	•	•	•		
Belize	BLZ	084	•	•	•	•	•	•		
Bulgaria	BGR	100	•	•	•	•	•	•		
Canada	CAN	124	•	•	•	•	•	•		
Colombia	COL	170	•	•	•	•	•	•		
Cyprus	CYP	196	•	•	•	•	•	•		
Czech Republic	CZE	203	•	•	•	•	•	•		
England	ENG	926	•	•	•	•	•	•		
France	FRA	250	•	•	•	•	•	•		
Germany	DEU	276	•	•	•	•	•	•		
Greece	GRC	300	•	•	•	•	•	•	•	•
Hong Kong, SAR	HKG	344	•	•	•	•	•	•		
Hungary	HUN	348	•	•	•	•	•	•	•	•
Iceland	ISL	352	•	•	•	•	•	•	•	•
Iran, Islamic Rep. Of	IRN	364	•	•	•	•	•	•		
Israel	ISR	376	•	•	•	•	•	•		
Italy	ITA	380	•	•	•	•	•	•	•	•
Kuwait	KWT	414	•	•	•	•	•	•		
Latvia	LVA	428	•	•	•	•	•	•		
Lithuania	LTU	440	•	•	•	•	•	•		
Macedonia, Rep. Of	MKD	807	•	•	•	•	•	•		
Moldova, Rep. of	MDA	498	•	•	•	•	•	•		
Morocco	MAR	504	•	•		•	•	•		
Netherlands	NLD	528	•	•	•	•	•	•		
New Zealand	NZL	554	•	•	•	•	•	•	•	•
Norway	NOR	578	•	•	•	•	•	•		
Romania	ROM	642	•	•	•	•	•	•		
Russian Federation	RUS	643	•	•	•	•	•	•		
Scotland	SCO	927	•	•	•	•	•	•		
Singapore	SGP	702	•	•	•	•	•	•	•	•
Slovak Republic	SVK	703	•	•	•	•	•	•		
Slovenia	SVN	705	•	•	•	•	•	•	•	•
Sweden	SWE	752	•	•	•	•	•	•	•	•
Turkey	TUR	792	•	•	•	•	•	•		
United States	USA	840	•	•		•	•	•	•	•

“•” Indicates that file is available.

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
 - D Derived Variable.

- The fourth character indicates the subject or topic to which a questionnaire item refers. The following abbreviations are used:

G General Questionnaire item

H Home Questionnaire item.

- For Background Questionnaire variables in PIRLS data files, the fifth through eighth characters indicate a question name and are unique to each variable.
- For Background Questionnaire variables in Reading Literacy Study data files, the third through eighth characters indicate a question name and are unique to each variable.

Background Variable Location Convention

To help identify the location of the background variables in the corresponding background questionnaire, a convention has been developed that identifies the questionnaire and location within the questionnaire for each variable. This convention is used in the Data Almanacs and in the description of the variables included in the Supplement CD of this User Guide. Following this questionnaire convention, each questionnaire has been assigned a unique identification code as shown in Exhibit 8.3. This unique code is followed by the sequence number of the question within the questionnaire. For example, if the location of a variable is indicated as SQ1-06, this refers to the Question 6 in the student background questionnaire.

Exhibit 8.3 Background Questionnaire Variable Location Conventions

Questionnaire	Location Code
Student Questionnaire	SG1-***
Home Questionnaire	SH1-***
Teacher Questionnaire	TG1-***
School Questionnaire	CG1-***

*** = the sequential numbering of the question location in the questionnaire.

Achievement Item Naming Convention

The achievement item variable names are based on 7- or 8-character alphanumeric codes according to the general definitions given below.

For the PIRLS test items, the following convention was used:

- The first character indicates the subject of the item. The code “R” is used for Reading.

- 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:

D	Reading Documents
E	Reading Expository text
N	Reading Narrative text.
- The last characters identify the text on which the items are based, and the sequence of the item after the text.

Coding Convention

A series of conventions were also adopted to code the data included in the data files. This section describes such conventions.

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 codes related to different categories of missing data.³ The missing categories defined below are assigned different values depending on the field width of the variable and the variable type. We recommend that you read the following section with particular care since the way in which you make use of these missing codes will have serious consequences for your analysis.

■ ■ ■

² See Chapter 2 for a presentation of the PIRLS test design.

³ The SAS and SPSS control statement files will recode these missing categories to special numeric missing codes in SAS and explicit missing codes in SPSS.

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, such as in the case of some of the test items, or unintentionally, such as when the item was left out of the instrument or misprinted. The Not Administered Codes are used in the following cases:

- Test item not assigned to the student – only one of the ten rotated booklets used in the PIRLS study was assigned to each student. All variables corresponding to items that were not given to a student have been coded to “Not Administered.”
- Booklet not received / booklet lost – If a respondent did not receive the instruments assigned to him/her, or the instruments were lost after administration, all items have been coded to “Not Administered.”

- 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 could not be done in all cases, however, and some national data were recoded to “not administered” in order to include only the internationally comparable data.

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

The Not Applicable Response Codes are used only for the background questionnaire items in which responses are dependent on a filter question. The specific Not Applicable Code given depends on the number of valid codes available for the item, as described above for the Omitted Response Codes.

Not Reached Item Codes (ASCII: 6 ; SAS: .R ; SPSS: 6)

The Not Reached Item Codes are used only for achievement items. Test items at the end of each test booklet in each testing session which were left blank were considered “not reached” due to the fact that the student did not complete the test. These responses are distinguished from the “missing” responses, as they are handled differently during the item calibration process. They are treated as incorrect responses, however, in computing achievement scores.

Types of Variables Included in the Data Files

Identification Variables

In all background files, several identification variables are included that provide information used to identify students, teachers, or schools. These variables are used to link cases between the different data files. The identification variables have the prefix ID and are listed below.

The identification variables included in Student, Teacher, and School files are the following:

IDCNTRY

Five-digit country identification code. This variable should always be used as one of the link variables whenever files are linked within and across countries. For each country the 1st two digits are always zeroes (00).

IDPOP

Identifies the population. The value is always set to 1 in the PIRLS 2001 and Reading Literacy Study databases.

IDSCHOOL

Identification number that uniquely identifies the school within each country. These codes for the school are not unique across countries. Schools across countries can be uniquely identified only by the IDCNTRY and IDSCHOOL combination.

Additional identification variables in the student files include the following:

IDSTUD

Identification number that uniquely identifies each student in the country sampled. The variable IDSTUD is a hierarchical identification number. It is formed by the combination of the variables IDSCHOOL and IDCLASS, followed by a two-digit sequential number within each classroom. Students can be uniquely identified in the database by the combination of IDSTUD and IDCNTRY.

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. Since 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.
ITLANG ⁵	Language of testing for each student. Set to 1 for all countries that tested in a single language. For countries that administered the test in more than one language, codes of 1, 2 and 3 were used to correspond to the order of the languages shown in Exhibit 1.4.
ILRELIAB	Linking variable indicating the inclusion status of each student in the reliability file containing double-coded constructed-response items. The following codes are used: Code 1 Student questionnaire used for reliability analysis Code 0 Student questionnaire NOT used for reliability analysis

Selection Variables

These variables are used to select cases for particular analyses. In the student files and in the Student-Teacher link files the following variables are included:

INPRL01	Students with INPRL01 set to “1” are those who participated in the PIRLS 2001 assessment. Otherwise this variable is set to zero.
INRLS91	Students with INRLS91 set to “1” are those who participated in the 1991 Reading Literacy assessment. Otherwise this variable is set to zero.
INRLS01	Students with INRLS01 set to “1” are those who participated in the 2001 Reading Literacy assessment. Otherwise this variable is set to zero.

■ ■ ■

⁵ See Exhibit 1.4 for a list of the languages included in each country.

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:

WGTFAC1	School Weighting Factor. This variable is included in the Student and in the School Background data files.
WGTADJ1	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.
WGTADJ3	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.
JKZONE	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.
JKREP	The PSU to which the student is assigned. This variable is included in the Student and Student-Teacher Linkage data files.
JKCZONE	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 basic demographic and socioeconomic indicators.

Teacher Background File

The teachers of the students who were sampled for PIRLS were administered one questionnaire with questions pertaining to their background and their teaching practices in the classes of the sampled students. Each teacher was asked to respond to a questionnaire for each class taught that contained sampled students. The Teacher Background files contain one record for each teacher who teaches the sampled students. If a teacher taught more than one class, they were expected to complete two questionnaires, one for each class they taught. In some cases, although the teacher was to respond to more than one questionnaire, responses to only one were obtained. In these cases, there were as many records entered in the teacher file as classes were taught by the teacher. In the Teacher Background data files each teacher was assigned a unique

identification number (IDTEACH) and a Teacher Link Number (IDLINK) that is specific to the class taught by the teacher and to which the information in the data record corresponds.

The IDTEACH and IDLINK combination uniquely identifies a teacher teaching one specific class. So, for example, students linked to teachers identified by the same IDTEACH but different IDLINK are taught by the same teacher but in different classes. The Teacher Background files cannot be merged directly with the Student Data files and they do not contain sampling information or achievement scores. It is important to note that the Teacher Background data files do not constitute a

representative sample of teachers in a country, but rather consist of the teachers who teach a representative sample of students. The teacher data should therefore be analyzed only in conjunction with the Student-Teacher Linkage file. The Teacher Background data file contains a series of other identification variables, and link variables, as well as the responses of the teachers to the background questions. No teacher questionnaire was administered as part of the Reading Literacy Study.

School Background File

The principals or administrators of the schools in the PIRLS sample were administered a school background questionnaire with questions about school policy and school environment. The School Background data file contains the responses given to the questions in this questionnaire. That file also contains a series of

identification variables, link variables, and sampling variables. The school data files can be merged with the student data files by using the country and school identification variables.

PIRLS Student Achievement Data Files

Student Achievement files contain the student response data for the individual achievement items in the PIRLS achievement test. Students who participated in PIRLS were administered one of ten test booklets with questions. Some of these questions were multiple-choice questions and some were open-ended. The responses to the open-ended questions were coded using a one-digit coding system. The achievement test data files contain the answers to the multiple-choice questions and the codes assigned by the coders to the student responses for open-ended questions. Since under the PIRLS test design a student received only a fraction of the total test item pool, the variables for the items that were not included in the test booklet that was administered to the student are coded as not administered. The specific test booklet that was administered to the student is coded in the variable IDBOOK. The Student Achievement data files also contain a series of identification variables, sampling variables, and achievement variables. The data contained in this file can be linked to the Student Background data files by using the variables IDCNTRY and IDSTUD.

PIRLS Scoring Reliability Data Files

The Scoring Reliability files contain data that can be used to investigate the reliability of the PIRLS constructed-response item scoring. The Scoring Reliability file contains one record for each booklet that was double scored during the scoring reliability exercise of the free-response items.

For each constructed-response item in the achievement test, the following three variables are included:

- Original Code Variable (achievement item response codes obtained from the first coder)
- Second Code Variable (achievement item response codes obtained from second coder)
- Response Code Agreement Variable (degree of agreement between the two codes).

Scoring Reliability Variable Naming Convention

The variable names for the Original Code, Second Code, and Agreement Code variables are based on the same general naming system as that for the achievement item variables shown earlier. The second character in the variable name differentiates the three reliability variables:

- O Original Code Variable (e.g., R011M04C)
- R Second Code Variable (e.g., RR11M04C)
- I Agreement Code Variable (e.g., RI11M04C).

Reliability Variable Code Values

The values contained in both the Original Code and Second Code variables are the scores assigned using the PIRLS scoring guides. The Agreement Code variable has the value of the difference between the original score and the score assigned by the reliability scorer.

In general, the response score data contained in the Original Code variables are identical to those contained in the Student Achievement test files. In some cases, however, the response scores for specific items were recoded after a review of the international item statistics revealed inconsistencies in the original scoring guides or showed that the original scores were not functioning as desired. The recoded diagnostic score values were used in computing the achievement scores reflected in the international reports.

The response codes in the Student Achievement test files reflect the recoded values. In contrast, the Original Code variables in the scoring reliability files contain the original unrecoded response codes. This was done so that the coding reliability measure indicated in the Agreement Code variables was based on the original scoring guides used during the constructed-response scoring sessions conducted in each country.

In addition to the scoring reliability variables, the reliability files also include identification variables to aid in case identification. Some tracking variables are also included that were used in conducting the coding reliability study within each

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
 - R Reading Literacy Study items.
- The last two characters in the file name indicates the study cycle. The following codes are used:
 - R1 stands for the PIRLS 2001 almanac files
 - T2 stands for the Reading Literacy Study almanac files.⁶
- The three character file extensions used for the Codebook files indicate the file type. The following codes are used:
 - .DOC Almanacs in MS Word format
 - .LST Almanacs in ASCII format
 - .PDF Almanacs in PDF format.



⁶ The Reading Literacy Data Almanacs contain both the 1991 and 2001 data.

Exhibit 8.4 Data Almanac Files for PIRLS 2001 and Reading Literacy Study

Almanac File	Contents
ASALM1_R1	Student Background with Reading Achievement
ASALM3_R1	School Background with Reading Achievement
ASALM5_R1	Teacher Background with Reading Achievement
ASALM9_R1	Home Background with Reading Achievement
AIALMI_R1	Percent of Responses by Item Category for the Reading to Acquire and Use Information Items
AIALML_R1	Percent of Responses by Item Category for the Reading for Literary Purposes Items
ASALM1_T2	Student Background with Reading Achievement (Reading Literacy Study)
AIALMR_T2	Percent of Responses by Item Category for the Document, Expository, and Narrative Items (Reading Literacy Study)

The data almanacs are provided as PDF files to be read with Adobe Acrobat Reader 4.0 or higher, as ASCII formatted files (LST), and as MS Word format files (DOC) to be read and edited with any text editor or word processing software that can read files of their size.

The files display student-weighted summary statistics, by grade, for each participating country on each variable. The almanacs also display the international averages for each variable, with each country weighted equally.

Background Item Data Almanacs

Background Data Almanacs include all student, home, teacher, and school background variables as well as derived variables based on these. The data presented in the almanacs use the student as the unit of analysis, even if the information is based on the teacher or school questionnaire items. Therefore, the weighted percentages in the teacher and school almanacs reflect the percentage of students to whom the data apply.

There are two types of displays in the Background Data Almanacs, depending on whether data are categorical or continuous.

The display for categorical variables includes:

- Sample size (number of students, teachers or schools included in the sample)

- Valid number of cases (counts of students, teachers or schools for whom valid data were obtained)
- Weighted percentages of students corresponding to each valid response option (percentages based only on the students with valid data)
- Weighted percentages of students for whom none of the valid response options were selected (the variable is coded as “Not Administered” or “Omitted”).

In cases where the question was “Not Applicable” based on the response to an explicit filter question, the percentage of students for whom the variable is coded as Not Applicable is also displayed. The percent Not Applicable is based on only the students with valid data. The last line in the almanac displays the international average in each category of response.

The display for continuous variables includes:

- Sample size (number of students, teachers or schools in the sample)
- Number of valid cases (counts of students, teachers or schools with valid data)
- Weighted percentages of students for whom the variable is coded as “Not Administered”, “Not Applicable”, or “Omitted”
- The mean, mode, minimum, maximum, and the 5th, 10th, 25th, 50th, 75th, 90th, and 95th percentiles across students
- The last line in the almanac displays the international averages.

The Background Data Almanacs also display for each variable the question that was asked, the location in the corresponding questionnaire and the variable name in the data file.

Achievement Item Data Almanacs

Achievement Item Data Almanacs include summary information about the response types given by students. They contain, for each country, the percent of students choosing each option for multiple-choice items or the percent of students receiving each score in the scoring guide for the constructed-response items. They also contain the percent of boys and girls obtaining the maximum score on the item and display the international average percentages for each response category, with each country weighted equally.

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 almanacs are defined below:

- **N:** The number of students to whom the item was administered.
- **A, B, C, D and E:** Used for multiple-choice items only, these represent the percentage of students choosing each one of the response options for the item. Not reached items are included in the denominator for these calculations.
- **Other Incorrect:** Used for multiple-choice items, it indicates the percent of students who omitted or gave an invalid response to the item.
- **Scoring Guide Scores (0, 1, 2, 3):** Used for free-response items only, these represent the percentage of student responses assigned each of the scores in the scoring rubric for the test item. Not reached items are excluded from the denominator for these calculations.
- **Diff (Item Difficulty):** The percentage of students that responded correctly to the item. This was used only for the multiple-choice items. When computing this statistic, not reached items were treated as administered and so the values might differ from those presented in the International Item Statistics Almanac.
- **V1, V2, V3:** Used only for the free-response items, these indicate the percentage of students that scored 1 point or better on the item (V1), 2 points or better (V2), and 3 points or better (V3).
- **Invalid:** Used for multiple-choice items only, this indicates the percentage of students that gave an invalid response to the item (e.g., multiple selections).
- **Not Reached:** Indicates the percentage of students that did not reach the test item.
- **Omit:** Indicates the percentage of students that omitted the test item.
- **Girl (% Right), Boy (% Right):** This indicates the percentage of girls and boys that obtained the maximum score on the item.

8.4 Codebook Files

All information related to the structure of the data files as well as the source, format, descriptive labels, and response option codes for all variables are contained in Codebook files. The Codebook files corresponding to each variable are described in

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

The naming convention of the Codebook files is the same used for the Background and Achievement files described earlier except for the following:

- The fourth through sixth characters identify the file as a codebook and it is always “CBK.”
- The three-character file extension used for the Codebook files is always “SDB” which stands for Standard DBase format.

The Codebook files can be read with most database and spreadsheet software.

The fields contained in the codebooks are the following:

Variable Number: The first column (Var. No.) contains a sequential number for each variable in each Codebook file.

Question: The second column contains an abbreviated variable identifier providing descriptive information needed to identify the content of the question and/or the source for each type of variable.

Variable Name: The third column (Variable Name) contains the variable name associated with each variable included in the International Data files. The naming system used for each variable type is described in the previous sections on the contents of data files.

Variable Label: The fourth column (Label) contains an extended textual variable label of up to 40 characters associated with each variable, providing more descriptive information about the content of each variable. For multiple-choice achievement items, the variable label includes the correct response option enclosed in brackets. During data analysis, the variable labels can be printed out to enhance understanding of the results.

Code: The fifth column (Code) contains the codes used for variable responses. For variables where numeric data are supplied in response to open-ended questions, the keyword VALUE is entered in the Code column. For categorical variables, all possible response options are listed. Any missing codes are also included for either numerical or categorical variables. For example, for multiple-choice achievement items, the code options are A, B, C, D, and E while for the free-response achievement items, the code options are the two-digit numerical codes described in Chapter 4.

Option: The sixth column (Option) includes a textual description of each type of response option. For variables containing numeric data, it contains an explanation of the values contained in the variable.

Location/Format: The seventh column presents the location and format of each variable in the raw data files. The location/format indicates the pattern used to write each value of a numeric or categorical variable, with a general structure of XX-YY/<N or C>ZZ.Z.

The numbers preceding the forward slash (/) indicate the location of the variable and refer to its position in the raw data file (starting [XX] – ending [YY] column positions). The <N or C> after the slash identifies the variable as numerical (N) or categorical (C). The numeric codes after the forward slash (ZZ.Z) indicate the total number of digits (including the decimal point) and the number of decimal places associated with each variable (e.g., 2.0 = 2 integer digits, 0 decimal places; 6.2 = six total digits: 3 integer digits, decimal point, and two decimal digits).

8.5 Program Files

Three different types of program files are provided for use in analyses of the PIRLS data files:

- Data Access Control Files
- Jackknife Statistics Program Files
- Scoring Program Files

The Data Access Control files are provided to convert the ASCII-format raw data files into SAS data sets or SPSS system files. The Jackknife Statistics Program files are used to compute the statistics as well as the standard errors associated with these statistics, using the jackknife repeated replication method discussed in Chapter 7.

The Scoring Program files are required to convert achievement item response codes to the score values used in the computation of international scores. For all program files, two versions are provided: one for SAS programs and one for SPSS programs. The file extension (SAS or SPS) is used to identify the respective SAS and SPSS program files.

Each data file has its own control file. Control files are named according to the same convention as their corresponding data and Codebook files, with the following exceptions:

- The fourth through sixth characters identify the file as a control file and it is always “CTR.”

- 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.TXT Item Information for the PIRLS 2001 Test Items

AITINF11.TXT Item 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.

9

Performing Analyses with the PIRLS Data Using SPSS

9.1	Overview	9-3
9.2	Contents of the CD	9-4
9.3	Creating SPSS System Files	9-6
	<i>Exhibit 9.1 Extract from SPSS Control Code for Creating a Student Background SPSS System File</i>	9-7
9.4	Scoring the Items	9-7
	<i>Exhibit 9.2 Extracted Sections of SPSS Macro SCOREIT Used to Convert Cognitive Item Response Codes to Correctness-Score Levels</i>	9-9
9.5	Basic Analyses with the PIRLS Data: Means, Percentages, Regression Coefficients, and Their JRR Standard Errors	9-10
	Computing Sampling and Imputation Variance for Plausible Values Using SPSS (JACKPV.SPS)	9-10
	<i>Exhibit 9.3 SPSS Control Code and Extract of Output File for Using the Macro JACKPV.SPS</i>	9-14
	Computing Sampling Variance for Variables Other Than Plausible Values Using SPSS (JACKGEN.SPS)	9-15
	<i>Exhibit 9.4 SPSS Control Code and Extract of Output File for Using the Macro JACKGEN.SPS</i>	9-19
	Computing Regression Coefficients and Their JRR Standard Errors for Variables Other Than Plausible Values (JACKREG.SPS)	9-20
	<i>Exhibit 9.5 SPSS Control Code and Extract of Output File for Using the Macro JACKREG.SPS</i>	9-23
	Computing Regression Coefficients and Their JRR Standard Errors with Plausible Values (JACKREGP.SPS)	9-24
	<i>Exhibit 9.6 SPSS Control Code and Extract of Output File for Using the Macro JACKREGP.SPS</i>	9-27
9.6	Conducting Analyses with the PIRLS 2001 International Database: Student Level	9-28
	Example Analysis with Student-Level Variables Not Using Plausible Values	9-28
	<i>Exhibit 9.7 SPSS Control Statements for Performing Analyses with Student-Level Variables NOT Using Plausible Values (EXAMPLE1.SPS)</i>	9-30
	<i>Exhibit 9.8 Extract of SPSS Computer Output for Performing Analyses with Student-Level Variables Not Using Plausible Values (EXAMPLE 1)</i>	9-31
	Example Analysis with Student-Level Variables Using Plausible Values	9-31
	<i>Exhibit 9.9 Sample Exhibit for Student-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report</i>	9-32

Exhibit 9.10	SPSS Control Statements for Performing Analyses with Student-Level Variables Involving Plausible Values (EXAMPLE2.SPS)	9-33
Exhibit 9.11	Extract of SPSS Computer Output for Performing Analyses with Student-Level Variables Involving Plausible Values (EXAMPLE 2)	9-34

9.7 Performing Analyses with Teacher-Level Variables 9-35

Exhibit 9.12	Sample Exhibit for Teacher-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report	9-36
Exhibit 9.13	SPSS Control Statement for Performing Teacher Level Analysis Using Plausible Values	9-37
Exhibit 9.14	Extract of SPSS Output for Performing Teacher Level Analysis Using Plausible Values (EXAMPLE 3)	9-38

9.8 Performing Analyses with School-Level Variables 9-40

Exhibit 9.15	Sample Exhibit for School-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report	9-41
Exhibit 9.16	SPSS Control Statements for Performing Student-Weighted Analyses with School-Level Variables (EXAMPLE4.SPS)	9-42
Exhibit 9.17	Extract of SPSS Computer Output for Performing Student-Weighted Analyses with School-Level Variables(EXAMPLE 4)	9-43

9.9 Performing Analyses with Home Background Variables 9-44

Exhibit 9.18	Sample Exhibit for Home Background Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report	9-45
Exhibit 9.19	SPSS Control Statements for Performing Home Background Analyses (EXAMPLE5.SPS)	9-46
Exhibit 9.20	Extract of SPSS Computer Output for Performing Home Background Variables (EXAMPLE 5)	9-47

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, based on student, home, teacher and school data, replicate some of the 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 macro language. If you have some programming experience in this statistical package, then you will be able to make the necessary modifications to the macros to obtain the desired results. Before using these macros, you should create a system file in SPSS that contains the variables necessary for the analysis. As part of this chapter we describe the control files included on the CD that can be used to do this.

9.2 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 SPSS and SAS programs |
| CODEBOOK: | Contains Codebook files with a description of each variable |
| ALMANACS: | Contains Data Almanacs |
| ITEMS: | Contains the Item Information files and PDF versions of the released reading passages and their questions |
| REPORTS: | Contains PDF versions of the <i>PIRLS International Report, Technical Report, Trends in Children's Reading Literacy Achievement 1991-2001</i> , and this User Guide. |

The directory names within each CD and the file names generally follow the DOS naming convention: file names with up to eight characters, followed by a three-character extension (as in FILENAME.EXT). Files with the same names are complementary to each other, and the extension identifies their function or type. The extensions used in the files contained on the CDs are the following:

- | | |
|------|------------------------------------|
| .SAS | SAS Control file or program |
| .SPS | SPSS Control file or program |
| .DAT | ASCII Data file |
| .LST | Almanacs in ASCII format |
| .DOC | Almanacs in MSWord format |
| .PDF | Almanacs and reports in PDF format |
| .DBF | Codebook in Dbase readable format |
| .TXT | Item Information tables. |

The DATA sub-directory contains the PIRLS and Reading Literacy Study data files in ASCII format. The data files that are in this directory are described in Chapter 8 of this guide. Each of these files has two corresponding sets of control files in the PROGRAMS sub-directory. One of these reads the ASCII data and creates a SAS data set, the other reads the ASCII data and creates an SPSS system file. This chapter will focus on the files that can be used with 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.
These programs are included only on the CD in the PROGRAMS sub-directory.

9.3 Creating SPSS System Files

The CD contains SPSS control code to read each one of the ASCII data files and create an SPSS system file. Each of these control files contains information on the location of each variable in the file, its format, a descriptive label, its categories (in the case of categorical variables), and code for handling missing data. The control and data files have been created to facilitate access of the data on a country-by-country basis. The command lines in the control files should be edited to produce programs that will create SPSS system files for any specified country. While most of the program code is functional as provided, you will need to specify the input and output directories. Performing analyses that require the data from more than one country will necessitate merging or appending the respective data files into a larger one. Alternatively, you can access the data and compute the necessary statistics on a country-by-country basis by reading one file at a time, computing the necessary statistics, and then moving on to the next country's data. The method you choose will depend greatly on the storage and processing capacity of the computer system that you use. For the examples that we present in this User Guide we have combined the data files of five countries into one larger data file.

We have also provided SPSS code that will allow you to merge files from several countries into one. This file is called JOIN_R1.SPS and when using it you need to specify the file type you want to merge (ASG, ATG, AST, ACG or ASH), the countries to be merged, and the variables on which to sort the resulting file. The resulting file is named according to the naming convention explained in Chapter 8, but with digits 4 through 6 being the letters "ALL."

When creating a system file using SPSS, you will need to do the following:

1. Open the corresponding control file, for example, ASGCTRR1.SPS.
2. In the line where it reads 'let !datapath = !unquote("<path>")' enter the path where the raw data is located.
3. In the line where it reads 'let !savpath = !unquote("<path>")' enter the path where you want to store the system file that you will be creating.
4. At the bottom of the files in the line where it reads "create country =" select the three character identification code for each country whose data you want to create. If you leave this line unedited the data for all countries listed will be created.

5. Submit the code for processing. After processing is complete you will find the corresponding system file in the location you specified in step 3 above.

For example, Exhibit 9.1 presents an extract from an SPSS control code in which the student background data for Bulgaria (BGR), Canada (CAN), England (ENG), and the United States (USA) will be read from the directory “c:\pirls\data\.” Separate SPSS system files for each country will be stored in the same directory “c:\pirls\data\” under the name ASG<3-character country code>R1. The sections that need to be modified are presented in **bold letters** in this exhibit.

Exhibit 9.1 Extract from SPSS Control Code for Creating a Student Background SPSS System File

```
SET DECIMAL = DOT.

define create (country=!charend('/')).

!let !datpath = !unquote("c:\pirls\data\").
!let !savpath = !unquote("c:\pirls\data\").

!do !cty !in(!country).

!let !indata = !concat('\',!datpath,asg,!cty,'r1.dat').
!let !savdata = !concat('\',!savpath,asg,!cty,'r1.sav').

* NOTE: Assignment of variable names.
* NOTE: Please change <path> and <filename> according to your system.
*       Change this in two of the following lines of this program.

*NOTE: Assignment of variable names.

data list file = !indata
.
.
.
save outfile = !savdata.
!doend
!enddefine

create country = BGR ENG CAN USA.
```

9.4 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 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.

The achievement data files contained on the CD include the score given to each item administered to a student. You might want to work with these item data after they are recoded to the right-wrong format in the case of multiple-choice items, or to the level of correctness in the case of the constructed-response items. To this effect, we have included on the CD a set of programs in SPSS that will allow you to recode the items from the achievement test to their right-wrong or correctness-level format. These programs contain a macro called SCOREIT and the necessary call to this macro so that all the items in the corresponding file are scored. This program will convert the response option codes for multiple-choice items to dichotomous score levels (0 or 1) based on scoring keys. Three files are included to provide control code to perform the recodes of the test items in the achievement test file:

ASASCRR1	Scores the PIRLS 2001 test items
ASASCRT2	Scores the Reading Literacy Study items.

When using these programs, you must first consider the recoding scheme that you want to work with. For example, under certain circumstances you might want to recode the not reached responses (code 6) as incorrect (code 0), whereas under other circumstances you might want to recode these responses as not administered or invalid responses (code 8). In the case of PIRLS and Reading Literacy Study, not reached responses were recoded as not administered (and effectively as omitted

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 used for the analysis. This is done by using the INCLUDE statement in the corresponding program after reading the data file. When using SPSS, the scoring program code should be included after the “GET FILE” statement that reads the system file containing the item responses. Both of these programs recode the items into the existing variables. If you want to preserve the original answers and codes assigned to the items, then the file with the recoded item variables needs to be saved under a different file name. A copy of the macro that scores the items in SPSS is presented in Exhibit 9.2.

Exhibit 9.2 Extracted Sections of SPSS Macro SCOREIT Used to Convert Cognitive Item Response Codes to Correctness-Score Levels

```

SET MPRINT=ON.
DEFINE SCOREIT (Type = !charend('\') /
                Item = !charend('\') /
                RIGHT = !charend('\') /
                nr = !charend('\') /
                na = !charend('\') /
                om = !charend('\') /
                other = !charend('\') ).

!If (!UPCASE(!Type)=MC) !Then
  !Do !I !In(!Item).
  Recode !I      (!RIGHT =1)
                (!nr    =0)
                (!na    =sysmis)
                (!om    =0)
                (!other =0)
                (Else   =0).
  !DoEnd.
!IfEnd.

!If (!UPCASE(!Type)=OE) !Then
  !Do !I !In(!Item).
  Recode !I      (1 =1)
                (2 =2)
                (3 =3)
                (0 =0)
                (!nr    = 0)
                (!na    = sysmis)
                (!om    = 0)
                (!other = 0)
                (Else   = 0).
  !DoEnd.
!IfEnd.
!enddefine.

SCOREIT Type = MC /
        Item = <list items where option A is the correct one> /
        RIGHT = 1 / nr = 6 / na = 8 / om = 9 / other = 7.

SCOREIT Type = MC /

```

```

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 file where it will be used by issuing the following command under SPSS:

```
include '<path>jackpv.sps' .
```

where <path> points to the specific drive and directory where the macro JACKPV.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 invoking 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 as a CVAR, usually IDCNTRY. |
| PVS | These are the plausible values to be used in the analysis. The plausible values need to be specified in the form “Plausible Value 1 to Plausible Value 5” as in “ASRREA01 to ASRREA05.” Although in most cases you will want to use all five available plausible values, the program will also work when fewer are specified. You should always use at least two plausible values. |
| NPV | This is the number of plausible values that will be used for the analysis. Generally you will want to use all five plausible values for the analysis although under some circumstances fewer can be used. You should always use at least two plausible values for any analysis. |
| 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 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\asgallr1.sav" /
  cvar = idcntry itsex /
  pvs = ASRREA01 to ASRREA05 /
  npv = 5 /
  jkz = JKZONE /
  jkr = JKREP /
  njkz = 75 /
  WGT = TOTWGT .
```

it will compute the mean reading achievement and its standard error for boys and girls within each country using five plausible values and the variable TOTWGT as the sampling weight. It will also compute the percentages of boys and girls within the country, and their corresponding standard errors. 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 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 variable is called TOTWGT.

N

Contains the number of cases in the groups defined by each specific combination of categories for the classification variables. In the example this is the number of boys and girls in the sample for each country.

PCT

Contains the percentages of people in the groups for the classification variable listed last, within the specific combination of the categories defined by the classification variables. In the example, it is the percentage of boys and girls within each country.

PCT_SE

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

MNX

Contains the means for the first plausible value for the groups defined by the corresponding combinations of classification variable categories.

MNX_SE

Contains the JRR standard errors of the means for the first plausible value for the groups computed using the JRR method for computing the standard error. This does not include the imputation error that should be computed when using plausible values.

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 /
  cvar = idcntry itsex /
  pvs = ASRREA01 TO ASRREA05 /
  npv = 5 /
  jkz = JKZONE /
  jkr = JKREP /
  njkz = 75 /
  WGT = TOTWGT .

print formats idcntry itsex n (F6.0) totwgt (f10.0)
  mnpv mnpv_se mnx mnx_se pct pct_se (f6.2).

report format=list
/ var = idcntry itsex n totwgt mnpv mnpv_se mnx mnx_se pct pct_se.
```

IDCNTRY	ITSEX	N	TOTWGT	MNPV	MNPV_SE	MNX	MNX_SE	PCT	PCT_SE
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	.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
100	2	1675	46418	538.06	4.67	537.63	4.63	48.53	.91
124	1	4121	109940	552.89	2.62	553.12	2.56	49.53	.67
124	2	4131	112039	535.58	2.55	535.67	2.52	50.47	.67
.
705	1	1468	10551	512.34	2.52	511.39	2.44	50.09	.82
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
752	2	3079	60551	550.31	2.53	550.83	2.45	51.26	.73
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)

In this section, example SPSS code that can be used to compute the JRR standard errors for means and percentages of variables other than plausible values is described. This code is provided in the form of an SPSS macro called JACKGEN.SPS 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 a variable of choice, and the JRR standard errors of these means. Although you can compute percentage and mean estimates using other basic SPSS commands, the macro JACKGEN.SPS also computes the JRR error estimate for these means and percentages.

When using this macro, you need to specify a set of classification variables, one analysis 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 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 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 IDCNTY. |
| 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    = jkzone       /  
  jkr    = jkrep        /  
  wgt    = totwgt       .
```

it will compute the mean of the age variable (ASDAGE) and its standard error, within each country, using the variable TOTWGT as the sampling weight. The data will be read from the system file ASGALLR1.

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 our example above there is one variable in the resulting system file. This is IDCNTRY. There is one unique occurrence for each category for this variable.

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.

PCT

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 by country, and the mean reading achievement scores. This code is included in the file SampleJackGEN.SPS.

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).

report format=list automatic
  / var   = idcntry n totwgt mnx mnx_se pct pct_se.

```

IDCNTRY	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
32	3284	705891	10.18	.03	100.00	.00
84	2853	7255	9.79	.06	100.00	.00
100	3450	95447	10.93	.02	100.00	.00
124	8253	222012	10.02	.01	100.00	.00
.						
.						
.						
705	2952	21066	9.80	.01	100.00	.00
752	6043	118083	10.80	.01	100.00	.00
792	5125	977316	10.23	.02	100.00	.00
807	3711	27365	10.65	.01	100.00	.00
840	3763	3802557	10.22	.02	100.00	.00
926	3155	592507	10.21	.01	100.00	.00
927	2717	64375	9.78	.01	100.00	.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 entry, or line, for each of the values of the variable IDCNTRY. The first column has the country code, the second column has the number of students in each sample after selecting only those cases where ASDGAGE is not missing, and the third column is the total weight this sample represents in the whole population. This is followed by the mean age of the sampled students and the corresponding standard error.

For example, from line one in the above results, we can determine that Argentina sampled 3,284 students from the population. These students represent 705,891 students in the whole population, with a mean age of 10.18 and a standard error of .03. Since there is only one grade in the category, each line represents 100 percent of the students within a country, with a zero standard error.

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 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, as for example 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 IDCNTY.
XVAR	This is a list of independent variables, at least one, that under the regression model will be used as predictors of the dependent variable specified in DVAR. These independent variables can be continuous or categorical, or any other type of coded variable. For example, it could be the variable ITSEX as originally coded in the data files, or dummy coded as 1 or 0.
DVAR	This is the dependent variable that under the regression model is predicted by the variable or variables specified by the XVAR parameter. Only one variable can be listed here. If you want to use the same set of predictor variables to predict two different variables, then the macro needs to be called separately to generate each set of results.
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 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 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    /
        dvar   = asdgstdt  /
        njkz   = 75        /
        jkz    = jkzone    /
        jkr    = jkrep     /
        wgt    = totwgt    .
```

it will compute the regression equation for the variable REGSEX as a predictor of the number of hours spent studying. 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 called REG, and can be found in the default directory. 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.

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 model and their corresponding jackknifed standard errors. The coefficient zero (B00) is the intercept for the model. The other coefficients receive a sequential number starting with 01. This sequential number corresponds to the order of the variables in the list of variables specified in the parameter XVAR.

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 9.5. This code is included in the file SampleJackREG.SPS.

Exhibit 9.5 SPSS Control Code and Extract of Output File for Using the Macro JACKREG.SPS

```

get file = "c:\pirls\data\asgallr1.sav"
  / keep = idcntry idgrader itsex asbgtvdy jkzone jkrep totwgt.

select if not(missing(itsex)).
compute regsex = itsex-1.

save outfile = student.

include "c:\pirls\programs\jackreg.sps".

jackreg infile = student /
  cvar = idcntry idgrader /
  xvar = regsex /
  dvar = asbgtvdy /
  njkz = 75 /
  jkz = jkzone /
  jkr = jkrep /
  wgt = totwgt .

print formats idcntry n (F6.0) Mult_RSQ (f5.3)
  SS_Total SS_Reg SS_Res (F10.0) B00 B00.SE B01 B01.SE (f6.2).

list vars = idcntry n Mult_RSQ SS_Total SS_Reg SS_Res B00 B00.SE B01 B01.SE.

```

IDCNTRY	N	MULT_RSQ	SS_TOTAL	SS_REG	SS_RES	B00	B00.SE	B01	B01.SE
32	3054	.000	1144825	249	1144576	3.21	.05	-.04	.05
84	2695	.004	12998	58	12940	2.51	.07	.18	.07
100	3413	.007	109945	781	109164	3.04	.03	.18	.04
124	8185	.022	300190	6635	293554	2.67	.03	.35	.04
.
705	2939	.037	21578	792	20786	2.68	.04	.39	.04
752	5891	.008	85409	693	84716	2.96	.02	.15	.02
792	5046	.004	951597	3744	947853	2.44	.03	.12	.03
807	3559	.008	27141	230	26911	2.61	.04	.19	.04
840	3740	.014	5817066	83710	5733356	2.83	.05	.30	.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 watching TV or videos daily) by country and by grade. In the listing of the results we can see that there is one entry or line for each value of the variable IDCNTRY. The first column has the country code and the second column has the number of students in each sample. Keep in mind that this number corresponds only to those students who had valid data for the variables in the regression model. Then the multiple R squared is listed for each country, followed by the sums of squares for the model and corresponding regression coefficients. Because of the way in which the variable REGSEX is coded, the constant (B00) is the mean value for the variable ASBGTVDY for the girls and the error of the constant (B00_SE) is standard error

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 (IDCOUNTRY = 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.SPS. This macro 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 this macro, 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 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 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 "JACKREGP" 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>jackregp.sps' .
```

where <path> points to the specific drive and directory where the macro JACKREGP.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, as for example 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.
- XVAR This is a list of independent variables, at least, which under the regression model will be used as predictors of the dependent variable specified by the plausible values. These independent variables can be continuous or categorical, or any other type of coded variable. For example, it could be the variable ITSEX as originally coded in the data files, or dummy coded as 1 or 0.
- ROOTPV This is the prefix used to identify the plausible values for the achievement scale of interest. This corresponds to the first 7 characters of the plausible value variables. For example, the root of the reading plausible value is "ASRREA0."
- NPV This is the number of plausible values that will be used in the analysis. Generally you will want to use all five plausible values for the analysis although under some circumstances fewer can be used. You should always use at least two plausible values for any analysis.
- 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 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 /
          wgt    = totwgt .
```

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.

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 model and their corresponding jackknifed standard errors combined with the imputation error. The coefficient zero (B00) is the intercept for the model. The other coefficients receive a sequential number starting with 01. This sequential number corresponds to the order of the variables in the list of variables specified in the parameter XVAR.

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 9.6. This code is included on the CD in the file SampleJackREGP.SPS.

Exhibit 9.6 SPSS Control Code and Extract of Output File for Using the Macro JACKREGP.SPS

```
get file = "c:\pirls\data\asgallr1.sav" .

select if not(missing(itsex)).
compute regsex = itsex-1.

save outfile = student.

include "c:\pirls\programs\jackregP.sps".

jackregp infile = student /
  cvar      = idcntry idgrader /
  xvar      = regsex /
  rootpv    = asrrea0 /
  npv       = 5 /
  njkz      = 75 /
  jkz       = jkzone /
  jkr       = jkrep /
  wgt       = totwgt .

print formats idcntry n (F6.0) Mult_RSQ (f5.3)
          SS_Total SS_Reg SS_Res (F12.0) B00 B00.SE B01 B01.SE (f6.2).

select if idgrader = 2.
list vars = idcntry N Mult_RSQ SS_Reg SS_Total SS_Res b00 b00.se b01 b01.se .
```

IDCNTRY	N	MULT_RSQ	SS_REG	SS_TOTAL	SS_RES	B00	B00.SE	B01	B01.SE
32	3292	.009	58398795	6439906220	6381507424	428.18	6.24	-18.14	4.74
84	2829	.017	1332607	80587404	79254797	341.73	5.44	-27.17	4.97
100	3458	.022	14038000	650701756	636663756	562.29	3.69	-24.23	3.61
124	8252	.015	16645023	1146374231	1129729208	552.89	2.62	-17.31	2.05

.									
.									
792	5125	.013	91215640	7264514554	7173298914	459.37	3.96	-19.32	3.08
807	3711	.010	2982206	290818644	287836438	452.24	5.14	-20.87	3.57
840	3763	.012	313113311	26311018108	25997904796	551.06	3.80	-18.07	4.11
926	3155	.017	74526387	4435757667	4361231280	563.74	3.87	-22.43	3.34
927	2717	.011	4794722	456110590	451315868	536.54	3.94	-17.26	3.98

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 plausible values of reading achievement by country. In the listing of the results we can see that there is one entry or line for each value of the variable IDCNTRY after selecting only those cases with IDGRADER=2. The first column has the country code and the second column has the number of students in each sample. Because of the way in which REGSEX is coded, the constant (B00) is the mean reading achievement for the girls, the error of the constant (B00.SE) is the standard error of this estimate. The first regression coefficient (B01) is the difference in the average value in reading achievement between the boys and the girls, and the standard error of this coefficient (B01.SE) is the combined jackknifed sampling error and imputation error of this difference.

For example, from the first line in the output we can say that in Argentina (IDCNTRY = 32) data is available for 3,292 cases. The average of the squared of the multiple correlation between gender and reading achievement is .009 with girls achieving on average 428.18 and boys scoring on average 18.14 points lower than girls. This difference is statistically significant ($18.14 / 4.74 = 3.83$). Overall, we can see that girls outperform boys in reading achievement in all countries in the table.

9.6 Conducting Analyses with the PIRLS 2001 International Database: Student Level

Many analyses of the PIRLS data can be undertaken using student background data as outcome variable. We have already presented some examples in the previous sections when explaining how to use the macros provided with the data files. We will now proceed to work through additional examples of actual analyses where all the steps are undertaken, including the calling of the corresponding SPSS macro.

Example Analysis with Student-Level Variables Not Using Plausible Values

In the first example, we want to find out the amount of television that students, on average, watch every day. From the student background questionnaire we see that question 6 asks “About how much time do you spend watching television or videos outside of school on a normal day?” The response options for this question are: “No time,” “Up to one hour,” “From 1 hour up to 3 hours,” “From 3 hours up to 5

hours,” and “5 hours or more.” Since we want to give a numeric value to the number of hours that students report watching television, we recode the response options to a new variable called TVWATCH that will have values of 0.0, 0.5, 2.0, 4.0, and 7.0, respectively.¹

Since we want to find the average number of hours that students watch television, we need to use the macro JACKGEN. In order to find out these results we need to undertake several steps. After reviewing the codebooks and the questionnaire information, we find that the question about the number of hours students spend watching television or videos corresponds to the variable ASBGTVDY, and that this variable is found in the Student Background data file. Our next step is to review the documentation of national adaptations to the questionnaires to ensure that there were no deviations listed for this variable (see the supplemental material on the accompanying CD). Since no changes were made, we can continue with our analysis without any modifications.

We then proceed to read the SPSS system files that contain this variable. The SPSS code is presented in Exhibit 9.7 and is included on the CD under the name EXAMPLE1.SPS. Extracts from the results obtained from this program are displayed in Exhibit 9.8. We have included as part of the program the corresponding value labels and format statements so that the different categories or groups are labeled appropriately.

Notice that one of the steps in this program is to select only those students in the fourth grade who have non-missing data in the variable. In general, this type of analysis is quite feasible with a powerful desktop computer; however, you need to keep in mind that computing and storage requirements for this type of analysis are quite demanding and increase with the number of countries and variables.

In general, to compute means and percentages using the Student Background data files, you need to do the following:

- Identify the variable, or variables, of interest in the student file and find out about any specific national adaptations to the variable.
- Retrieve the relevant variables from the data files, including the sampling weights, JRR replication information, and any other variables used in the selection of cases.

■ ■ ■

¹ We have used this particular recoding for this variable for illustration purposes only. Other recodings might result in a more accurate representation of actual students' behavior.

- 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\asgbgrr1.sav"
 / file = "c:\pirls\data\asgsver1.sav"
 / file = "c:\pirls\data\asgturrl.sav"
 / file = "c:\pirls\data\asgusar1.sav".

select if not(missing(asbgtvdy)).

value labels
 idcntry 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 idcntry .

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 .

```

Exhibit 9.8 Extract of SPSS Computer Output for Performing Analyses with Student-Level Variables Not Using Plausible Values (EXAMPLE 1)

Example 1: PIRLS 2001 User Guide						
PAGE 1						
COUNTRY ID	N	TOTWGT	MNX	MNX_SE	PCT	PCT_SE
Argentina	3059	643973	2.98	.07	11.51	.54
Bulgaria	3413	94245	2.62	.06	1.68	.05
Sweden	5891	115776	2.30	.03	2.07	.06
Turkey	5046	964738	1.52	.05	17.24	.65
United States	3740	3776946	2.48	.09	67.50	.97

In this example, each country’s mean value for TVWATCH is reported for the fourth grade. The results are presented by country. The country is in the first column. The second column has the number of students sampled in the country (with data available for the variable TVWATCH). The third column corresponds to the weights given these sampled students in the whole population, followed by their mean for TVWATCH and the corresponding standard errors. The last two columns represent the percentages of students sampled responding and the corresponding standard error.

For example, from the first line of the report, we can say that data for Argentina is available for 3,059 cases. These students represent a population of approximately 643,973 students. Argentinean students spend, on average, 2.98 hours watching television or videos outside of school. The standard error of this mean is .07.

Example Analysis with Student-Level Variables Using Plausible Values

In this example, we want to replicate the results presented in Exhibit 4.8 of the international report (replicated results are presented in Exhibit 9.9). We are interested in looking at the fourth graders’ reports on the number of books in their home and their reading achievement. Since the results in this exhibit are based on plausible values, we need to use the macro JACKPV.

Exhibit 9.9 Sample Exhibit for Student-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report

Countries	More than 200 Books		101-200 Books		26-100 Books		11-25 Books		0-10 Books	
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement
Argentina	6 (0.7)	474 (12.6)	7 (0.7)	484 (11.1)	21 (1.6)	466 (9.1)	24 (1.2)	446 (6.8)	42 (2.5)	416 (5.5)
Belize	8 (0.9)	363 (12.0)	7 (0.8)	348 (11.4)	17 (1.1)	357 (8.3)	27 (1.1)	343 (6.5)	41 (2.0)	304 (5.0)
Bulgaria	27 (1.4)	583 (3.8)	16 (0.9)	579 (4.9)	24 (1.1)	565 (3.8)	13 (1.0)	537 (5.7)	20 (1.8)	498 (8.1)
Canada (O,Q)	22 (0.9)	562 (3.4)	22 (0.8)	565 (3.3)	35 (0.7)	546 (2.4)	15 (0.8)	519 (3.3)	6 (0.5)	492 (4.7)
Colombia	5 (0.5)	451 (9.9)	7 (0.5)	455 (9.0)	18 (1.1)	457 (7.3)	24 (1.0)	438 (4.7)	46 (1.8)	401 (4.8)
Cyprus	12 (0.8)	492 (5.6)	13 (0.8)	520 (5.7)	35 (1.1)	510 (3.8)	27 (1.1)	490 (3.6)	12 (0.9)	466 (6.1)
Czech Republic	20 (1.3)	557 (3.7)	21 (1.0)	552 (3.1)	41 (1.4)	538 (2.8)	15 (1.0)	512 (3.1)	4 (0.4)	479 (7.9)
England	20 (1.0)	577 (5.0)	23 (1.2)	579 (4.1)	34 (1.2)	555 (3.5)	16 (1.0)	518 (3.8)	7 (0.6)	477 (6.7)
France	19 (1.0)	555 (4.1)	19 (0.9)	545 (3.1)	36 (1.0)	526 (2.6)	19 (0.9)	506 (3.6)	8 (0.5)	472 (6.1)
Germany	16 (0.7)	570 (2.6)	16 (0.7)	568 (2.9)	35 (0.8)	546 (2.1)	24 (0.8)	523 (2.4)	9 (0.6)	486 (4.2)
Greece	15 (1.1)	544 (6.0)	15 (0.9)	548 (4.3)	35 (1.4)	536 (5.7)	25 (1.3)	514 (4.9)	9 (0.9)	488 (7.1)
Hong Kong, SAR	9 (0.6)	533 (5.0)	10 (0.5)	536 (4.0)	30 (0.8)	538 (3.0)	26 (0.7)	529 (3.4)	26 (1.1)	515 (4.4)
Hungary	28 (1.3)	568 (3.1)	19 (0.7)	559 (2.8)	29 (1.0)	544 (2.9)	16 (0.8)	522 (2.8)	7 (0.8)	481 (5.0)
Iceland	23 (0.7)	528 (3.3)	25 (0.6)	532 (2.1)	35 (0.7)	515 (2.3)	13 (0.5)	490 (3.7)	4 (0.3)	455 (6.1)
Iran, Islamic Rep. of	4 (0.4)	466 (8.0)	5 (0.4)	469 (9.0)	12 (0.9)	473 (6.1)	22 (1.3)	440 (4.6)	56 (2.5)	390 (3.4)
Israel	19 (1.1)	527 (5.7)	16 (0.8)	536 (4.7)	33 (1.1)	532 (3.2)	21 (1.0)	506 (4.5)	10 (0.9)	472 (7.5)
Italy	13 (0.6)	551 (4.2)	14 (0.8)	560 (3.9)	30 (0.8)	553 (3.6)	29 (1.0)	538 (2.9)	14 (0.8)	499 (4.2)
Kuwait	12 (0.6)	418 (6.0)	11 (0.6)	415 (6.7)	24 (0.8)	411 (5.4)	24 (1.0)	405 (5.6)	29 (1.5)	386 (6.5)
Latvia	24 (1.3)	555 (3.1)	24 (1.1)	557 (3.6)	34 (0.9)	546 (2.5)	13 (0.8)	525 (4.3)	5 (0.6)	513 (6.4)
Lithuania	9 (0.6)	562 (5.0)	16 (1.0)	568 (4.0)	39 (1.1)	552 (2.6)	25 (1.1)	529 (3.5)	10 (1.0)	505 (5.1)
Macedonia, Rep. of	8 (0.8)	465 (10.6)	6 (0.5)	487 (7.9)	27 (1.5)	491 (4.8)	34 (1.6)	454 (5.6)	24 (1.8)	415 (7.4)
Moldova, Rep. of	6 (0.9)	542 (8.6)	8 (0.8)	538 (5.4)	21 (1.1)	515 (4.4)	27 (1.3)	488 (4.4)	37 (2.0)	468 (5.1)
Morocco	2 (0.5)	--	3 (0.4)	331 (17.2)	9 (1.2)	372 (13.9)	18 (1.5)	351 (8.5)	68 (2.4)	362 (12.4)
Netherlands	12 (0.8)	570 (3.8)	16 (0.7)	569 (3.0)	36 (0.9)	560 (2.6)	25 (1.0)	547 (2.7)	11 (0.8)	523 (5.3)
New Zealand	23 (1.4)	560 (5.3)	23 (1.1)	556 (4.0)	33 (1.3)	533 (4.5)	14 (0.9)	493 (5.9)	8 (0.7)	464 (7.2)
Norway	27 (1.1)	518 (3.6)	25 (0.9)	517 (3.9)	33 (1.2)	495 (3.6)	10 (0.9)	463 (6.5)	4 (0.5)	435 (7.5)
Romania	9 (0.9)	565 (5.2)	9 (0.8)	558 (5.7)	25 (1.1)	534 (4.2)	25 (1.3)	508 (5.0)	32 (1.8)	482 (9.5)
Russian Federation	17 (1.0)	541 (4.7)	16 (0.8)	543 (4.4)	35 (1.1)	535 (4.0)	21 (1.1)	519 (5.8)	11 (1.0)	496 (9.3)
Scotland	19 (0.9)	553 (5.9)	18 (1.1)	554 (6.1)	34 (1.0)	539 (3.9)	18 (0.8)	506 (4.1)	11 (0.8)	466 (6.6)
Singapore	16 (0.7)	569 (5.7)	19 (0.6)	560 (4.9)	37 (1.0)	540 (4.1)	19 (0.9)	500 (5.3)	10 (0.7)	442 (7.8)
Slovak Republic	15 (1.0)	542 (4.5)	22 (1.0)	535 (4.1)	40 (1.4)	524 (3.0)	17 (1.1)	496 (3.8)	7 (1.2)	450 (15.5)
Slovenia	15 (1.0)	513 (4.1)	19 (0.9)	523 (3.2)	37 (1.1)	511 (2.0)	21 (1.1)	484 (3.4)	8 (0.7)	454 (6.4)
Sweden	31 (1.5)	577 (2.7)	26 (0.7)	572 (2.5)	31 (1.1)	552 (2.3)	9 (0.7)	530 (4.0)	3 (0.5)	503 (6.9)
Turkey	5 (0.6)	494 (10.8)	5 (0.5)	499 (8.3)	20 (1.1)	478 (4.3)	30 (1.1)	455 (3.7)	40 (2.0)	422 (3.8)
United States	21 (1.2)	556 (5.8)	22 (1.1)	568 (4.3)	31 (1.1)	551 (3.8)	17 (1.0)	522 (4.2)	9 (1.0)	477 (7.0)
International Avg.	15 (0.2)	529 (1.3)	15 (0.1)	525 (1.1)	30 (0.2)	514 (0.8)	21 (0.2)	490 (0.8)	19 (0.2)	459 (1.2)

Background data provided by students.
 () Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.
 An "r" indicates data are available for 70-84% of the students. An "s" indicates data are available for 50-69% of the students. An "x" indicates data are available for less than 50% of the students.
 A dash (-) indicates data are not available. A tilde (~) indicates insufficient data to report achievement.

To replicate the results in this exhibit, we need to undertake several steps. After reviewing the codebooks and the questionnaire information we find that the question about the number of books in the home (see the CD for a copy of the student questionnaires) corresponds to the variable ASBGBOOK, and this variable is found in the Student Background data file. Our next step is to review the documentation of national adaptations to the questionnaires to ensure that there were no deviations listed for this variable (also on the CD). If no changes were made we can continue with our analysis without any modifications.

We then proceed to read from the SPSS system file that contains this variable. The SPSS code is presented in Exhibit 9.10. Extracts from the results obtained from this program are displayed in Exhibit 9.11. We have included as part of the program the

corresponding value labels and format statements so that the different categories or groups are labeled appropriately.

Notice that one of the steps in this program is to select only those students in the fourth grade who chose one of the five options presented in the question.

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

- Identify the variable or variables of interest in the student file and find out about any specific national adaptations to the variable.
- Retrieve the relevant variables from the data files, including the achievement score, sampling weights, JRR replication information, and any other variables used in the selection of cases.
- Use the macro JACKPV with the corresponding arguments and parameters.
- Print out the result file.

Exhibit 9.10 SPSS Control Statements for Performing Analyses with Student-Level Variables Involving Plausible Values (EXAMPLE2.SPS)

```
title "Example 2: PIRLS 2001 User Guide".

add files
 / file = "c:\pirls\data\asgargr1.sav"
 / file = "c:\pirls\data\asgbgrr1.sav"
 / file = "c:\pirls\data\asgsver1.sav"
 / file = "c:\pirls\data\asgturrl.sav"
 / file = "c:\pirls\data\asgusarl.sav" .

select if not(missing(asbgbook)) and idgrader = 2.

value labels
 idcntry 32 'Argentina' 100 'Bulgaria' 752 'Sweden' 792 'Turkey' 840 'United
 States' .
 asbgbook 1 '0 -10 BOOKS' 2 '11-25 BOOKS' 3 '26-100 BOOKS' 4 '101-200 BOOKS'
 5 'MORE THAN 200 BOOKS' .

save outfile = student.

* Now use the macro to get the results.
include "c:\pirls\programs\jackpv.sps".

jackpv infile= student /
 cvar = idcntry asbgbook /
 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   = asbgbook (label) n totwgt mnpv mnpv_se pct pct_se
  / break = idcntry .

```

Exhibit 9.11 Extract of SPSS Computer Output for Performing Analyses with Student-Level Variables Involving Plausible Values (EXAMPLE 2)

Example 2: PIRLS 2001 User Guide
PAGE 1

COUNTRY ID	GEN/HOW MANY BOOKS AT HOME	N	TOTWGT	MNPV	MNPV_SE	PCT	PCT_SE
Argentina	0-10 BOOKS	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	499	107115	465.93	9.13	21.18	1.61
	101-200 BOOKS	181	33400	483.59	11.11	6.60	.67
	MORE THAN 200 BOOKS	144	29995	473.98	12.58	5.93	.73
Bulgaria	0-10 BOOKS	583	18440	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	.86
	MORE THAN 200 BOOKS	962	25005	582.60	3.79	27.20	1.44
Sweden	0-10 BOOKS	163	3794	503.47	6.92	3.29	.51
	11-25 BOOKS	505	10812	530.15	4.05	9.39	.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	.68
	MORE THAN 200 BOOKS	1970	35198	576.76	2.68	30.56	1.46
Turkey	0-10 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	.54
	MORE THAN 200 BOOKS	223	45775	494.30	10.81	4.83	.64
United States	0-10 BOOKS	314	334096	476.66	6.98	8.93	.95
	11-25 BOOKS	654	653582	521.51	4.23	17.47	.96
	26-100 BOOKS	1167	1171921	551.13	3.84	31.33	1.06
	101-200 BOOKS	799	804711	568.38	4.29	21.51	1.10
	MORE THAN 200 BOOKS	767	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.

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

Countries	Every Day		3-4 Days a Week		Fewer than 3 Days a Week	
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement
Argentina	28 (4.1)	415 (11.4)	53 (4.8)	416 (9.2)	19 (3.6)	432 (13.6)
Belize	71 (5.2)	324 (5.6)	24 (4.5)	325 (17.1)	5 (2.1)	314 (19.7)
Bulgaria	31 (4.4)	551 (8.2)	60 (4.5)	554 (5.0)	9 (2.3)	543 (10.4)
Canada (O,Q)	67 (3.5)	548 (2.7)	24 (3.3)	540 (4.3)	8 (1.8)	530 (7.8)
Colombia	58 (4.3)	426 (5.0)	28 (3.9)	408 (10.0)	13 (3.1)	442 (18.8)
Cyprus	79 (4.6)	494 (3.5)	17 (4.2)	489 (6.8)	4 (1.9)	500 (14.6)
Czech Republic	44 (4.0)	536 (4.0)	42 (4.3)	535 (3.5)	13 (3.1)	547 (5.9)
England	74 (3.6)	551 (4.2)	23 (3.6)	564 (8.8)	3 (1.7)	570 (8.8)
France	66 (3.9)	524 (2.8)	28 (3.7)	527 (5.2)	7 (1.9)	531 (7.3)
Germany	51 (3.0)	537 (2.5)	32 (2.8)	545 (3.4)	16 (2.3)	539 (3.8)
Greece	81 (3.5)	522 (3.6)	13 (3.5)	535 (12.6)	7 (2.9)	550 (20.4)
Hong Kong, SAR	17 (3.4)	525 (7.4)	38 (4.6)	534 (5.2)	46 (4.4)	525 (4.4)
Hungary	40 (4.2)	543 (3.6)	60 (4.2)	545 (3.0)	0 (0.0)	~ ~
Iceland	37 (0.4)	511 (1.7)	41 (0.4)	513 (2.4)	22 (0.3)	516 (2.6)
Iran, Islamic Rep. of	32 (4.0)	412 (10.4)	52 (4.3)	421 (6.7)	17 (3.4)	404 (8.8)
Israel	55 (4.3)	505 (7.1)	35 (4.2)	512 (7.4)	11 (2.7)	515 (10.9)
Italy	36 (3.6)	535 (4.7)	49 (3.9)	544 (3.6)	15 (2.6)	549 (5.9)
Kuwait	82 (3.0)	399 (5.7)	12 (2.5)	398 (8.3)	6 (1.6)	412 (7.5)
Latvia	66 (4.3)	540 (3.1)	32 (4.3)	555 (4.2)	2 (1.1)	~ ~
Lithuania	67 (4.2)	543 (3.6)	29 (4.0)	542 (5.4)	4 (1.6)	561 (2.6)
Macedonia, Rep. of	41 (3.8)	459 (9.2)	39 (4.4)	431 (8.3)	20 (3.7)	429 (14.6)
Moldova, Rep. of	32 (3.8)	503 (7.1)	63 (3.9)	489 (5.3)	5 (1.8)	455 (12.4)
Morocco	35 (5.3)	340 (11.4)	59 (5.5)	344 (11.7)	7 (2.4)	351 (31.6)
Netherlands	55 (4.3)	557 (3.2)	39 (4.2)	549 (4.9)	6 (1.8)	550 (7.3)
New Zealand	68 (3.7)	522 (4.2)	30 (3.8)	543 (8.2)	2 (1.0)	~ ~
Norway	58 (4.5)	503 (4.6)	32 (3.9)	493 (4.4)	10 (3.1)	496 (8.5)
Romania	71 (4.1)	521 (5.4)	24 (3.7)	495 (12.2)	4 (2.0)	482 (26.1)
Russian Federation	87 (2.6)	529 (4.7)	12 (2.5)	522 (11.0)	1 (0.6)	~ ~
Scotland	45 (4.9)	526 (5.4)	45 (4.9)	528 (5.7)	11 (2.9)	547 (13.0)
Singapore	45 (4.0)	522 (8.2)	34 (4.0)	530 (9.6)	22 (3.3)	532 (11.9)
Slovak Republic	48 (3.9)	514 (4.5)	52 (3.9)	521 (3.7)	0 (0.0)	~ ~
Slovenia	36 (4.1)	499 (3.6)	34 (4.0)	498 (3.5)	30 (4.2)	509 (4.1)
Sweden	57 (3.6)	562 (3.3)	30 (3.4)	556 (3.5)	13 (2.2)	568 (3.4)
Turkey	54 (5.0)	454 (5.2)	37 (4.5)	439 (6.7)	10 (2.7)	459 (9.9)
United States	93 (2.1)	541 (4.2)	6 (2.0)	555 (7.5)	1 (0.5)	~ ~
International Avg.	54 (0.7)	500 (1.0)	35 (0.7)	500 (1.3)	10 (0.4)	495 (2.2)

Background data provided by teachers.
 () Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.
 An "r" indicates data are available for 70-84% of the students. An "s" indicates data are available for 50-69% of the students. An "x" indicates data are available for less than 50% of the students.
 A dash (-) indicates data are not available. A tilde (~) indicates insufficient data to report achievement.

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\astbgr1.sav"
 / file = "c:\pirls\data\astswr1.sav"
 / file = "c:\pirls\data\asttur1.sav"
 / file = "c:\pirls\data\astusar1.sav" .

sort cases by idcntry idteach idlink.
save outfile = studteac.

add files
 / file = "c:\pirls\data\atgargr1.sav"
 / file = "c:\pirls\data\atgbgr1.sav"
 / file = "c:\pirls\data\atgswr1.sav"
 / file = "c:\pirls\data\atgtur1.sav"
 / file = "c:\pirls\data\atgusar1.sav".

sort cases by idcntry idteach idlink atbgract.

save outfile = teacher.

* Now merge the two files.
match files
 / file=studteac
 / table=teacher
 / by idcntry idteach idlink.

select if not(missing(atbgract)) and inprl01=1.

* Define the format for the variables used.
value labels
```

```

idcntry 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 = idcntry atbgract /
pvs = asrrea01 to asrrea05 /
npv = 5 /
jkz = JKZONE /
jkr = JKREP /
njkz = 75 /
WGT = TCHWGT .

sort cases by idcntry.

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

COUNTRY ID	GEN/RDG INSTR WITH STD/HOW OFTEN	N	TCHWGT	MNPV	MNPV_SE	PCT	PCT_SE
Argentina	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	471	123414	431.83	13.63	19.10	3.64
Bulgaria	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
Sweden	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
Turkey	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	388	86759	459.33	9.93	9.53	2.70
United States	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	.50	.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.

9.8 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 always preferable to analyze school-level variables as attributes of students, rather than as elements in their own right.

Although the school samples were not designed primarily to report on school characteristics, it is possible to use school weights to provide reasonable estimates of percentages and averages for primary or middle schools in each country. The following examples, however, describe only analyses based on student-weighted data.

For student-weighted analyses, the school-level data are analyzed to make statements about the number of students attending schools with one characteristic or another rather than the number of schools with certain characteristics. When school-level variables are analyzed, we recommend that you merge the selected school-level variables with the student-level file, and then use the sampling and weight 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 SPSS.

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 located” (ACBGCOMM). The response options for this question were “Urban,” “Suburban,” and “Rural.” These results can be found in Exhibit 7.1 of the *PIRLS International Report* (replicated results are presented in Exhibit 9.15).

Exhibit 9.15 Sample Exhibit for School-Level Analysis Involving Plausible Values Taken From the *PIRLS 2001 International Report*

Countries	Urban		Suburban		Rural	
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement
Argentina	75 (4.5)	427 (6.5)	19 (4.0)	395 (15.2)	5 (2.0)	412 (15.8)
Belize	38 (6.3)	361 (9.4)	15 (6.5)	318 (28.0)	48 (6.2)	296 (7.3)
Bulgaria	64 (2.8)	566 (4.2)	12 (2.6)	542 (17.1)	24 (2.3)	516 (9.1)
Canada (O,Q)	40 (3.7)	540 (3.7)	39 (3.5)	552 (3.4)	21 (2.9)	536 (4.3)
Colombia	70 (3.0)	435 (4.8)	2 (1.1)	~ ~	29 (3.0)	388 (7.3)
Cyprus	r 44 (3.1)	502 (5.3)	25 (3.2)	497 (7.5)	30 (2.7)	482 (4.9)
Czech Republic	55 (3.4)	540 (3.0)	17 (3.3)	537 (4.8)	28 (3.2)	529 (6.0)
England	48 (5.0)	542 (4.2)	32 (4.6)	557 (6.9)	21 (3.7)	574 (5.3)
France	36 (4.1)	525 (5.4)	29 (4.2)	523 (4.6)	35 (3.6)	527 (3.3)
Germany	33 (2.8)	523 (3.2)	22 (2.9)	546 (4.3)	45 (3.8)	547 (2.0)
Greece	--	--	--	--	--	--
Hong Kong, SAR	52 (3.0)	533 (3.8)	46 (2.7)	526 (5.2)	2 (1.1)	~ ~
Hungary	28 (2.5)	566 (4.3)	36 (2.3)	546 (4.9)	36 (1.7)	526 (2.8)
Iceland	r 36 (0.4)	519 (2.3)	43 (0.4)	511 (2.0)	21 (0.3)	499 (3.5)
Iran, Islamic Rep. of	48 (3.7)	444 (6.3)	9 (2.4)	413 (14.5)	43 (3.1)	375 (4.5)
Israel	52 (3.8)	525 (5.6)	22 (3.2)	531 (7.6)	27 (3.3)	461 (11.4)
Italy	76 (3.2)	543 (2.5)	14 (2.5)	537 (6.4)	10 (2.2)	529 (7.6)
Kuwait	18 (2.7)	399 (10.2)	79 (2.7)	399 (5.5)	3 (0.5)	426 (20.0)
Latvia	44 (3.8)	559 (3.8)	18 (3.9)	551 (4.6)	37 (3.0)	523 (3.2)
Lithuania	71 (2.8)	552 (3.2)	6 (2.1)	542 (11.8)	23 (2.4)	515 (4.9)
Macedonia, Rep. of	57 (3.3)	472 (5.3)	15 (2.8)	462 (15.3)	28 (3.1)	387 (9.6)
Moldova, Rep. of	28 (3.3)	515 (6.8)	14 (3.8)	494 (12.2)	59 (3.7)	480 (5.0)
Morocco	41 (4.4)	354 (9.1)	21 (4.4)	323 (10.8)	39 (4.1)	352 (24.9)
Netherlands	31 (3.9)	541 (6.6)	22 (3.7)	562 (3.9)	47 (3.8)	558 (3.2)
New Zealand	38 (4.0)	533 (6.1)	40 (3.6)	522 (6.4)	23 (2.9)	540 (8.9)
Norway	19 (3.5)	505 (7.7)	26 (3.9)	511 (6.2)	56 (3.5)	492 (3.5)
Romania	50 (2.5)	526 (5.1)	4 (2.0)	473 (15.9)	45 (2.8)	499 (8.2)
Russian Federation	55 (2.5)	539 (4.8)	3 (1.9)	553 (6.8)	43 (2.4)	512 (5.8)
Scotland	35 (4.3)	517 (6.3)	39 (5.3)	538 (6.6)	26 (4.5)	534 (6.9)
Singapore	100 (0.0)	528 (5.2)	0 (0.0)	~ ~	0 (0.0)	~ ~
Slovak Republic	50 (3.7)	531 (3.6)	10 (2.5)	522 (7.8)	40 (3.1)	501 (4.5)
Slovenia	39 (3.3)	509 (3.3)	27 (3.7)	499 (4.2)	34 (3.0)	495 (3.6)
Sweden	15 (3.0)	567 (5.4)	66 (3.8)	560 (2.8)	18 (3.5)	560 (5.4)
Turkey	36 (4.2)	464 (6.0)	39 (4.1)	455 (6.1)	25 (3.5)	423 (8.2)
United States	32 (3.5)	519 (6.6)	35 (4.6)	572 (4.6)	33 (3.2)	537 (7.3)
International Avg.	46 (0.6)	507 (0.9)	25 (0.6)	502 (1.8)	29 (0.5)	485 (1.5)

Background data provided by schools.
 () Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.
 An "r" indicates data are available for 70-84% of the students. An "s" indicates data are available for 50-69% of the students. An "x" indicates data are available for less than 50% of the students.
 A dash (-) indicates data are not available. A tilde (~) indicates insufficient data to report achievement.

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 ACBGCOMM in the School Background file, and the student weights and plausible values in the Student Background file. We then proceed to read the School Background file that has the variables that are relevant to our analysis. We then read the variables of interest from the student data file. We then proceed to merge the school information with the student information using the variables IDCNTRY and IDSCHOOL as merge variables, and then use the macro JACKPV to obtain the corresponding percentages of students within each group, and their mean achievement scores in reading. The computer code used to run this analysis in SPSS can be found in Exhibit 9.16 and an extract of the results is shown in Exhibit 9.17.

Exhibit 9.16 SPSS Control Statements for Performing Student-Weighted Analyses with School-Level Variables (EXAMPLE4.SPS)

```

title "Example 4: PIRLS 2001 User Guide".

add files
 / file = "c:\pirls\data\asgargr1.sav"
 / file = "c:\pirls\data\asgbgrr1.sav"
 / file = "c:\pirls\data\asgswr1.sav"
 / file = "c:\pirls\data\asgturr1.sav"
 / file = "c:\pirls\data\asgusar1.sav".

sort cases by idcntry idschool.

save outfile = student
 / keep = idcntry idschool idstud asrrea01 to asrrea05 inprl01 totwgt jkzone
 jkrep.

add files
 / file = "c:\pirls\data\acgargr1.sav"
 / file = "c:\pirls\data\acgbgrr1.sav"
 / file = "c:\pirls\data\acgswr1.sav"
 / file = "c:\pirls\data\acgturr1.sav"
 / file = "c:\pirls\data\acgusar1.sav".

sort cases by idcntry idschool.

save outfile = school
 / keep = idcntry idschool acbgcomm.

* Now merge the two files.
match files
 / file=student
 / table=school
 / by idcntry idschool.

select if not(missing(acbgcomm)) and inprl01=1.

* Define the format for the variables used.
value labels
 idcntry 32 'Argentina' 100 'Bulgaria' 752 'Sweden' 792 'Turkey' 840 'United
 States' /
 acbgcomm 1 'Urban' 2 'Suburban' 3 'Rural'.

save outfile = merged.

* Now use the macro to get the results.
include "c:\pirls\programs\jackpv.sps".

```



```

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: PIRLS 2001 User Guide
PAGE 1

COUNTRY ID	GEN/CHARACTER OF SCH LOCATION AREA	N	TOTWGT	MNPV	MNPV_SE	PCT	PCT_SE
Argentina	Urban	2498	508261	427.13	6.46	75.42	4.50
	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	1881	363850	454.53	6.09	38.53	4.09
	Rural	1191	238830	422.52	8.17	25.29	3.49
United States	Urban	1498	1174714	518.55	6.56	32.10	3.48
	Suburban	1383	1274273	572.29	4.55	34.82	4.62
	Rural	776	1210399	537.42	7.35	33.08	3.21

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 school locations 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, 2,498 students who represent 508,261 students in the population attend schools that are located in an urban area. These students have a mean reading achievement of 427.13 with a standard error of 6.46, and represent 75.42 percent of the sampled students. This percentage has a standard error of 4.50.

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

- Identify the variable or variables of interest in the student file and find out about any specific national adaptations to the variable.
- Retrieve the relevant variables from the student files, including the achievement score, sampling weights, JRR replication information, and any other variables used in the selection of cases.
- Retrieve the relevant classification variable or variables from the school database.
- 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.

9.9 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 9.18, 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."

Exhibit 9.18 Sample Exhibit for Home Background Analysis Involving Plausible Values Taken From the *PIRLS 2001 International Report*

Countries	High PATR		Medium PATR		Low PATR		Percentage of Students at High Level of PATR
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Hungary	74 (1.0)	556 (2.2)	23 (0.9)	518 (2.5)	3 (0.4)	509 (8.3)○
Norway	73 (1.1)	512 (3.2)	23 (1.0)	481 (4.5)	4 (0.5)	470 (10.8)○
Sweden	71 (0.8)	571 (2.1)	24 (0.9)	547 (3.4)	5 (0.4)	536 (4.5)○
Scotland	s 69 (1.3)	554 (3.9)	24 (1.2)	520 (6.3)	7 (0.6)	511 (9.3)○
England	s 69 (1.3)	586 (3.5)	26 (1.2)	543 (5.0)	6 (0.6)	535 (8.2)○
Iceland	r 67 (0.8)	522 (1.6)	30 (0.7)	510 (2.7)	3 (0.4)	489 (8.1)○
Slovak Republic	65 (1.2)	533 (3.0)	32 (1.2)	503 (3.8)	3 (0.4)	483 (10.8)○
New Zealand	r 65 (1.2)	559 (3.8)	30 (1.2)	501 (4.0)	5 (0.6)	488 (10.1)○
Czech Republic	r 64 (1.1)	551 (2.7)	32 (0.9)	528 (3.0)	4 (0.5)	515 (9.8)○
Canada (O,Q)	r 64 (1.0)	561 (2.4)	32 (1.0)	534 (2.7)	4 (0.3)	525 (5.4)○
Greece	64 (1.4)	539 (3.5)	31 (1.3)	505 (4.9)	6 (0.6)	505 (9.6)○
Netherlands	s 62 (1.3)	571 (2.6)	33 (1.1)	553 (3.0)	5 (0.6)	537 (7.4)○
Slovenia	61 (1.0)	516 (2.3)	37 (1.0)	487 (2.7)	2 (0.3)	~ ~○
Cyprus	s 57 (1.4)	507 (3.5)	38 (1.5)	477 (4.6)	5 (0.6)	466 (6.9)○
Italy	56 (1.2)	560 (2.1)	37 (1.1)	527 (2.9)	7 (0.5)	523 (5.2)○
Germany	r 55 (1.1)	559 (2.0)	35 (1.0)	528 (2.5)	10 (0.6)	519 (3.9)○
Bulgaria	54 (1.8)	574 (3.7)	37 (1.4)	538 (5.0)	9 (1.2)	497 (11.6)○
France	53 (1.1)	545 (2.5)	43 (1.1)	512 (3.2)	4 (0.4)	497 (6.1)○
Russian Federation	51 (1.3)	536 (4.6)	45 (1.2)	522 (5.2)	5 (0.5)	516 (6.6)○
Macedonia, Rep. of	r 49 (1.9)	479 (4.8)	48 (1.8)	426 (6.8)	3 (0.4)	420 (12.2)○
Latvia	48 (1.3)	560 (2.9)	50 (1.4)	537 (2.7)	3 (0.4)	532 (13.8)○
Iran, Islamic Rep. of	43 (1.4)	438 (4.8)	52 (1.2)	402 (4.8)	5 (0.8)	373 (16.9)○
Lithuania	42 (1.2)	558 (2.9)	52 (1.2)	537 (3.0)	6 (0.5)	528 (6.1)○
Belize	s 41 (1.8)	358 (6.1)	52 (1.7)	309 (6.7)	7 (1.0)	315 (16.8)○
Colombia	r 41 (1.4)	443 (5.4)	52 (1.3)	414 (5.0)	7 (0.6)	428 (7.4)○
Singapore	40 (0.9)	553 (4.6)	54 (0.9)	517 (5.5)	6 (0.3)	517 (6.9)○
Romania	34 (1.6)	542 (4.2)	55 (1.5)	501 (5.5)	11 (1.2)	492 (10.6)○
Kuwait	r 34 (0.6)	413 (4.2)	61 (0.6)	396 (5.1)	5 (0.3)	388 (7.2)○
Argentina	s 33 (1.6)	462 (7.1)	62 (1.7)	417 (6.8)	5 (0.5)	410 (11.0)○
Turkey	29 (1.4)	483 (5.0)	63 (1.3)	440 (3.5)	7 (0.8)	436 (6.8)○
Hong Kong, SAR	29 (0.9)	539 (3.1)	65 (0.8)	527 (3.1)	6 (0.4)	526 (6.2)○
Moldova, Rep. of	29 (1.5)	516 (5.1)	63 (1.6)	487 (4.2)	8 (0.7)	463 (7.5)○
Israel	x x	x x	x x	x x	x x	x x○
Morocco	--	--	--	--	--	--○
United States	--	--	--	--	--	--○
International Avg.	53 (0.2)	524 (0.7)	42 (0.2)	492 (0.8)	5 (0.1)	482 (1.6)○

Index of Parents' Attitudes Toward Reading (PATR)

Based on parents' agreement with the following: I read only if I have to; I like talking about books with other people; I like to spend my spare time reading; I read only if I need information; and Reading is an important activity in my home. Average is computed across the 5 items based on a 4-point scale: Disagree a lot = 1, Disagree a little = 2, Agree a little = 3, and Agree a lot = 4. Responses for negative statements were reverse-coded. High level indicates an average of 2 through 3, and Low level indicates an average of 1 to less than 2.

() Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent. An "r" indicates data are available for 70-84% of the students. An "s" indicates data are available for 50-69% of the students. An "x" indicates data are available for less than 50% of the students. A dash (-) indicates data are not available. A tilde (~) indicates insufficient data to report achievement.

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.

Exhibit 9.19 SPSS Control Statements for Performing Home Background Analyses (EXAMPLE5.SPS)

```

title "Example 5: PIRLS 2001 User Guide".

add files
 / file = "c:\pirls\data\asgargr1.sav"
 / file = "c:\pirls\data\asgbgr1.sav"
 / file = "c:\pirls\data\asgsver1.sav"
 / file = "c:\pirls\data\asgturr1.sav" .

sort cases by idcntry idstud.

save outfile = student
 / keep = idcntry idschool idstud asrrea01 to asrrea05 inprl01 totwgt jkzone
jkrep.

add files
 / file = "c:\pirls\data\ashargr1.sav"
 / file = "c:\pirls\data\ashbgr1.sav"
 / file = "c:\pirls\data\ashsver1.sav"
 / file = "c:\pirls\data\ashturr1.sav" .

sort cases by idcntry idstud.

save outfile = home
 / keep = idcntry idstud asdhpatr.

* Now merge the two files.
match files
 / file = student
 / file = home
 / by idcntry idstud.

select if not(missing(asdhpatr)) and inprl01=1.

* Define the format for the variables used.
value labels
 idcntry 32 'Argentina' 100 'Bulgaria' 752 'Sweden' 792 'Turkey' 840 'United
States' /
 asdhpatr 1 'High' 2 'Medium' 3 'Low'.

save outfile = merged.

* Now use the macro to get the results.
include "c:\pirls\programs\jackpv.sps".
jackpv infile= merged /
 cvar = idcntry asdhpatr /
 pvs = asrrea01 to asrrea05 /

```

```

npv = 5 /
jkg = JKZONE /
jkr = JKREP /
njkg = 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 = asdhpatt (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: PIRLS 2001 User Guide
PAGE 1

COUNTRY ID	INDEX/PARENTS ATTITUDE TO READING (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 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, 689 students who represent 138,668 students in the population have parents or guardians with a high positive attitude towards reading, and have a mean reading achievement score of 462.12 with a standard error of 7.11. These students represent 32.87 percent of the sampled students and this percentage has a standard error of 1.59.

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

- Identify the variable or variables of interest in the student file and find out about any specific national adaptations to the variable. Note that in our example we have not included data for the United States since the Home Questionnaire was not administered there.
- Retrieve the relevant variables from the student files, including the achievement score, sampling weights, JRR replication information, and any other variables used in the selection of cases.
- Retrieve the relevant classification variable or variables from the home background file.
- Merge the variables from the school database onto the student database using the variables IDCNTRY and IDSTUD.
- Use the macro JACKGEN or JACKPV with the corresponding arguments and parameters.
- Print out the result file.

10

Performing Analyses with the PIRLS Data Using SAS

10.1 Overview	10-3
10.2 Contents of the CD	10-4
10.3 Creating SAS Datasets	10-6
<i>Exhibit 10.1 Extract from SAS Control Code for Creating a Student Background SAS Dataset</i>	10-7
10.4 Scoring the Items	10-8
<i>Exhibit 10.2 Extracted Sections of SAS Macro SCOREIT Used to Convert Cognitive Item Response Codes to Correctness-Score Levels</i>	10-9
10.5 Basic Analyses with the PIRLS Data: Means, Percentages, Regression Coefficients and Their JRR Standard Errors	10-10
Computing Sampling and Imputation Variance for Plausible Values Using SAS (JACKPV.SAS)	10-10
<i>Exhibit 10.3 SAS Control Code and Extract of Output File for Using the Macro JACKPV.SAS</i>	10-14
Computing Sampling Variance for Variables Other Than Plausible Values Using SAS (JACKGEN.SAS)	10-15
<i>Exhibit 10.4 SAS Control Code and Extract of Output File for Using the Macro JACKGEN.SAS</i>	10-19
Computing Regression Coefficients and Their JRR Standard Errors for Variables Other Than Plausible Values (JACKREG.SAS)	10-19
<i>Exhibit 10.5 SAS Control Code and Extract of Output File for Using the Macro JACKREG.SAS</i>	10-23
Computing Regression Coefficients and Their JRR Standard Errors with Plausible Values (JACKREGP.SAS)	10-24
<i>Exhibit 10.6 SAS Control Code and Extract of Output File for Using the Macro JACKREGP.SAS</i>	10-27
10.6 Conducting Analyses with the PIRLS 2001 International Database: Student Level	10-28
Example Analysis with Student-Level Variables Not Using Plausible Values	10-28
<i>Exhibit 10.7 SAS Control Statements for Performing Analyses with Student-Level Variables NOT Using Plausible Values (EXAMPLE1.SAS)</i>	10-30
<i>Exhibit 10.8 Extract of SAS Computer Output for Performing Analyses with Student-Level Variables Not Using Plausible Values (EXAMPLE 1)</i>	10-30
Example Analysis with Student-Level Variables Using Plausible Values	10-31
<i>Exhibit 10.9 Sample Exhibit for Student-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report</i>	10-32

<i>Exhibit 10.10</i>	<i>SAS Control Statements for Performing Analyses with Student-Level Variables Involving Plausible Values (EXAMPLE2.SAS)</i>	10-33
<i>Exhibit 10.11</i>	<i>Extract of SAS Computer Output for Performing Analyses with Student-Level Variables Involving Plausible Values (EXAMPLE 2)</i>	10-34
10.7	Performing Analyses with Teacher-Level Variables	10-35
<i>Exhibit 10.12</i>	<i>Sample Exhibit for Teacher-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report</i>	10-36
<i>Exhibit 10.13</i>	<i>SAS Control Statement for Performing Teacher-Level Analysis Using Plausible Values</i>	10-37
<i>Exhibit 10.14</i>	<i>Extract of SAS Output for Performing Teacher-Level Analysis Using Plausible Values (EXAMPLE 3)</i>	10-38
10.8	Performing Analyses with School-Level Variables	10-39
<i>Exhibit 10.15</i>	<i>Sample Exhibit for School-Level Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report</i>	10-41
<i>Exhibit 10.16</i>	<i>SAS Control Statements for Performing Student-Weighted Analyses with School-Level Variables (EXAMPLE4.SAS)</i>	10-42
<i>Exhibit 10.17</i>	<i>Extract of SAS Computer Output for Performing Student-Weighted Analyses with School-Level Variables(EXAMPLE 4)</i>	10-42
10.9	Performing Analyses with Home Background Variables	10-44
<i>Exhibit 10.18</i>	<i>Sample Exhibit for Home Background Analysis Involving Plausible Values Taken From the PIRLS 2001 International Report</i>	10-45
<i>Exhibit 10.19</i>	<i>SAS Control Statements for Performing Home Background Analyses (EXAMPLE5.SAS)</i>	10-46
<i>Exhibit 10.20</i>	<i>Extract of SAS Computer Output for Performing Home Background Variables (EXAMPLE 5)</i>	10-47

10.1 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 SAS 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 percentages of students in specified subgroups, the mean reading achievement for those groups, and the corresponding standard errors for the percentages and mean statistics. Additionally, some examples of regression statistics are presented. These analyses, based on student, home, teacher and school data, replicate some of the 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 SAS 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 SAS macro language. If you have some programming experience in this statistical package, then you will be able to make the necessary modifications to the macros to obtain the desired results. Before using these macros, you should create a system file in SAS that contains the variables necessary for the analysis. As part of this chapter we describe the control files included on the CD that can be used to do this.

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 variable
ALMANACS:	Contains data almanacs
ITEMS:	Contains the Item Information files and PDF versions of the released reading passages and their questions
REPORTS:	Contains PDF versions of <i>the PIRLS International Report, Trends in Children's Reading Literacy Achievement 1991-2001, Technical Report</i> , and this User Guide.

The directory names within each CD and the file names generally follow the DOS naming convention: file names with up to eight characters, followed by a three-character extension (as in FILENAME.EXT). Files with the same names are complementary to each other, and the extension identifies their function or type. The extensions used in the files contained in the CDs are the following:

.SAS	SAS Control file or program
.SPSS	SPSS Control file or program
.DAT	ASCII Data file
.LST	Almanacs in ASCII format
.DOC	Almanacs in MSWord format
.PDF	Almanacs and reports in PDF format
.DBF	Codebook in Dbase readable format
.TXT	Item Information tables.

The DATA sub-directory contains the PIRLS and Reading Literacy Study data files in ASCII format. The data files that are in this directory are described in Chapter 8 of this guide. Each of these files has two corresponding sets of control files in the

PROGRAMS sub-directory. One of these reads the ASCII data and creates an SPSS system file, the other reads the ASCII data and creates a SAS dataset. This chapter will focus on the files that can be used with SAS.

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 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.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 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.SAS.

JACKREG.SAS

This macro program in SAS 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.SAS

This macro program in SAS 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 macros and prints out the results. These programs are discussed later in the chapter.

EXAMPLE1.SAS, EXAMPLE2.SAS, EXAMPLE3.SAS, EXAMPLE4.SAS, EXAMPLE5.SAS

These are the programs used in the samples presented later in this chapter.

These programs are included only in the CD in the PROGRAMS sub-directory.

10.3 Creating SAS Datasets

The CD contains SAS control code to read each one of the ASCII data files and create a SAS dataset. Each of these control files contains information on the location of each variable in the file, its format, a descriptive label, its categories (in the case of categorical variables), and code for handling missing data. The control and data files have been created to facilitate access of the data on a country-by-country basis. The command lines in the control files should be edited to produce programs that will create SAS datasets for any specified country. While most of the program code is functional as provided, you will need to specify the input and output directories. Performing analyses that require the data from more than one country will necessitate merging or appending the respective data files into a larger one. Alternatively, you can access the data and compute the necessary statistics on a country-by-country basis by reading one file at a time, computing the necessary statistics, and then moving on to the next country's data. The method you choose will depend greatly on the storage and processing capacity of the computer system that you use. For the examples that we present in this User Guide we have combined the data files of five countries into one larger data file.

We have also provided SAS code that will allow you to merge files from several countries into one. This file is called JOIN_R1.SAS and when using it you need to specify the file type you want to merge (ASG, ATG, AST, ACG or ASH), the countries to be merged, and the variables on which to sort the resulting file. The resulting file is named according to the naming convention explained in Chapter 8, but with digits 4 through 6 being the letters "ALL."

When creating a dataset using SAS, you will need to do the following:

1. Open the corresponding control file, for example, ASGCTRR1.SAS .
2. In the line where it reads "%LET CTY = ..." select the three character identification code for each country whose data you want to create. If you leave this line unedited the data for all countries listed will be created.
3. In the line where it reads "LIBNAME..." you will need to enter the path where you want to store the datasets that you will be creating.
4. In the line where it reads "FILENAME rawinp..." enter the path where the raw data is located.

5. Submit the code for processing. After processing is complete you will find the corresponding datasets in the location you specified in step 3 above.

For example, Exhibit 10.1 presents an extract from SAS code in which the student background data for Bulgaria (BGR), Canada (CAN), England (ENG) and the United States (USA) will be read from the directory "c:\pirls\data\." Separate SAS datasets for each country will be stored in the same directory "c:\pirls\data\" under the name ASG<3-character country code>R1. The sections that need to be modified are presented in **bold letters** in this exhibit.

Exhibit 10.1 Extract from SAS Control Code for Creating a Student Background SAS Dataset

```
%let cty = BGR CAN ENG USA;

Options nosource nocenter;

LIBNAME libdat "c:\pirls\data" ;
LIBNAME library "c:\pirls\data";

%macro doit;
  %let i = 1;
  %do %while(%length(%scan(&cty,&i)));
    %let ccode = %lowercase(%scan(&cty,&i));

    %let fname= asg&ccode.r1 ;

    FILENAME rawinp "c:\pirls\data\&fname..dat" ;

    %PUT INFO: data set definition ;
    %PUT INFO: data set definitions may need to be modified ;

    *Output data set ;
    data libdat.&fname;

    RUN ;

    %let i = %eval(&i + 1);
  %end;
%mend doit;

%doit;
```

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 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.

The achievement data files contained on the CD include the score given to each item administered to a student. You might want to work with these item data after they are recoded to the right-wrong format in the case of multiple-choice items, or to the level of correctness in the case of the constructed-response items. To this effect, we have included in the CD a set of programs in SAS that will allow you to recode the items from the achievement test to their right-wrong or correctness-level format. These programs contain a macro called SCOREIT and the necessary call to

this macro so that all the items in the corresponding file are scored. This program will convert the response option codes for multiple-choice items to dichotomous score levels (0 or 1) based on scoring keys. Three files are included to provide control code to perform the recodes of the test items in the achievement test file:

ASASCRR1	Scores the PIRLS 2001 test items
ASASCRT2	Scores the Reading Literacy Study items.

When using these programs, you must first consider the recoding scheme that you want to work with. For example, under certain circumstances you might want to recode the not reached responses (code 6) as incorrect (code 0), whereas under other circumstances you might want to recode these responses as not administered or invalid responses (code 8). In the case of PIRLS and Reading Literacy Study, not reached responses were recoded as not administered (and effectively as omitted 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 SAS programs used for the analysis. This is done by using the INCLUDE statement in the corresponding program after reading the data file. When using SAS, the scoring program code should be included after the "DATA" statement that reads the dataset containing the item responses. Both of these programs recode the items into the existing variables, so if you want to preserve the original answers and codes assigned to the items, then the file with the recoded item variables needs to be saved under a different file name. A copy of the macro that scores the items in SAS is presented in Exhibit 10.2.

Exhibit 10.2 Extracted Sections of SAS Macro SCOREIT Used to Convert Cognitive Item Response Codes to Correctness-Score Levels

```
%MACRO SCOREIT (ITEM, TYPE, RIGHT, NR, NA, OM, OTHER) ;
%IF %upcase(&TYPE) = "MC" %THEN %DO ;
  SCORE = 0 ;
  IF &ITEM = &RIGHT THEN SCORE = 1 ;
  IF &ITEM = &NR THEN SCORE = 0 ;
  IF &ITEM = &NA THEN SCORE = . ;
  IF &ITEM = &OM THEN SCORE = 0 ;
  IF &ITEM = &OTHER THEN SCORE = 0 ;
  &ITEM = SCORE;
%END;
%IF %upcase(&TYPE) = "OE" %THEN %DO;
  SCORE = 0 ;
  IF &ITEM = 3 THEN SCORE = 3 ;
  IF &ITEM = 2 THEN SCORE = 2 ;
  IF &ITEM = 1 THEN SCORE = 1 ;
  IF &ITEM = 0 THEN SCORE = 0 ;
  IF &ITEM = &NR THEN SCORE = 0 ;
  IF &ITEM = &NA THEN SCORE = . ;
```

```

        IF &ITEM = &OM THEN SCORE = 0 ;
        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 the form of an SAS 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 SAS macro language in order to use JACKPV.SAS. 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 SAS 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 string “%JACKPV” should not be used within that session 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>jackpv.sas'.
```

where <path> points to the specific drive and directory where the macro JACKPV.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 invoking 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 as a CVAR, usually IDCNTY.
ROOTPV	This is the prefix used to identify the plausible values for the achievement scale of interest. This corresponds to the first 7 characters of the plausible value variables. For example, the root of the overall reading plausible value is “ASRREA0.”
NPV	This is the number of plausible values that will be used for the analysis. Generally you will want to use all five plausible values for the analysis although under some circumstances fewer can be used. You should always use at least two plausible values for any analysis.

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 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 SAS 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.sas".
%jackpv(totwgt , jkzone, jkrep, 75, idcntry itsex, asrrea0, 5, asgallm2);
```

it will compute the mean reading achievement and its standard error for boys and girls within each country using five plausible values and the variable TOTWGT as the sampling weight. It will also compute the percentages of boys and girls within

the country, and their corresponding standard errors. 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 FINAL 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 the above example there are two classification variables in the FINAL data set, IDCNTY 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 variable is called TOTWGT.

N

Contains the number of cases in the groups defined by each specific combination of categories for the classification variables. In the example this is the number of boys and girls in the sample for each country.

PCT

Contains the percentages of people in the groups for the classification variable listed last, within the specific combination of the categories defined by the classification variables. In the example, it is the percentage of boys and girls within each country.

PCT_SE

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

MNX

Contains the means for the first plausible value for the groups defined by the corresponding combinations of classification variable categories.

MNX_SE

Contains the JRR standard errors of the means for the first plausible value for the groups computed using the JRR method for computing the standard error. This does not include the imputation error that should be computed when using plausible values.

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 ar1 "c:\pirls\data";

data student;
  set ar1.asgallr1;
  where itsex in(1,2);

%include "c:\pirls\programs\jackpv.sas";
%jackpv (totwgt , jkzone, jkrep, 75, idcntry itsex, asrrea0, 5, student);

proc print data=final noobs;
  var idcntry itsex N totwgt mnpv mnpv_se mnx mnx_se pct pct_se;
  format idcntry itsex N 6.0 totwgt 10.0 mnpv mnpv_se mnx mnx_se pct pct_se
  6.2;
run;
```

IDCNTRY	ITSEX	N	TOTWGT	mnpv	mnpv_se	mnx	mnx_se	pct	pct_se
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
124	2	4131	112039	535.58	2.55	535.67	2.52	50.47	0.67
.
705	1	1468	10551	512.34	2.52	511.39	2.44	50.09	0.82
705	2	1484	10515	490.65	2.38	490.97	2.28	49.91	0.82
752	1	2965	57583	572.27	2.62	572.06	2.53	48.74	0.73
752	2	3079	60551	550.31	2.53	550.83	2.45	51.26	0.73
792	1	2426	470383	459.37	3.96	458.57	3.90	48.13	0.92
792	2	2699	506933	440.06	3.73	440.15	3.69	51.87	0.92
807	1	1814	13392	452.24	5.14	452.15	5.06	48.94	0.88
807	2	1897	13973	431.37	4.76	431.15	4.73	51.06	0.88
840	1	1911	1927005	551.06	3.80	551.97	3.73	50.68	0.82
840	2	1852	1875552	532.99	4.89	531.22	4.66	49.32	0.82

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 SAS (JACKGEN.SAS)

In this section, example SAS code that can be used to compute the JRR standard errors for means and percentages of variables other than plausible values is described. This code is provided in the form of an SAS macro called JACKGEN.SAS 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 a variable of choice, and the JRR standard errors of these means. Although you can compute percent and mean estimates using other basic SAS commands, the macro JACKGEN.SAS also computes the JRR error estimate for these means and percentages.

When using this macro, you need to specify a set of classification variables, one analysis 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 to be processed.

You need to know some basic SAS 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 SAS 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 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.
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 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 the following code:

```
%include "c:\pirls\programs\jackgen.sas".  
%jackgen(totwgt , jkzone, jkrep, 75, idcntry, asdgage, asgallr1);
```

it will compute the mean of the age variable (ASDGAGE) and its standard error, within each country, using the variable TOTWGT as the sampling weight. The data will be read from the dataset ASGALLR1.

The file that contains these results is then called FINAL 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 is one variable in the resulting dataset. This is IDCNTRY. There is one unique occurrence for each category for this variable.

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.

PCT

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 SAS procedure of choice. An example call to this macro, and a subset of the resulting file, is presented in Exhibit 10.4. In this example, the macro will compute the percentages by country, and the mean reading achievement scores. This code is included in the file SampleJackGEN.SAS.

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

```

libname ar1 "c:\pirls\data";

data student;
  set ar1.asgallr1;
  where inpr101=1 and asdgage > 0;

%include "c:\pirls\programs\jackgen.sas";
%jackgen (totwgt , jkzone, jkrep, 75, idcntry idgrader, asdgage, student);

proc print data=final noobs;
  var idcntry N totwgt mnx mnx_se pct pct_se;
  format idcntry n 6.0 totwgt 10.0 mnx mnx_se pct pct_se 6.2;

run;

```

	IDCNTRY	N	TOTWGT	mnx	mnx_se	pct	pct_se
	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 entry, or line, for each of the values of the variable IDCNTRY. The first column has the country code, the second column has the number of students in each sample after selecting only those cases where ASDGAGE is not missing, and the third column is the total weight this sample represents in the whole population. This is followed by the mean age of the sampled students and the corresponding standard error.

For example, from line one in the above results, we can determine that Argentina sampled 3,284 students from the population. These students represent 705,891 students in the whole population, with a mean age of 10.18 and a standard error of .03. Since there is only one grade in the category, each line represents 100 percent of the students within a country, with a zero standard error.

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

In this section, example SAS 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 SAS macro program called

JACKREG.SAS. 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.SAS, 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 SAS 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 SAS 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 string “%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 SAS:

```
%include '<path>jackreg.sas'.
```

where <path> points to the specific drive and directory where the macro JACKREG.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, as for example 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.

- XVAR This is a list of independent variables, at least one, that under the regression model will be used as predictors of the dependent variable specified in DVAR. These independent variables can be continuous or categorical, or any other type of coded variable. For example, it could be the variable ITSEX as originally coded in the data files, or dummy coded as 1 or 0.
- DVAR This is the dependent variable that under the regression model is predicted by the variable or variables specified by the XVAR parameter. Only one variable can be listed here. If you want to use the same set of predictor variables to predict two different variables, then the macro needs to be called separately to generate each set of results.
- 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 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 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:

```
%jackreg(totwgt , jkzone, jkrep, 75, idcntry idgrader, regsex, asdgstdt, asgallr1);
```

it will compute the regression equation for the variable REGSEX as a predictor of the number of hours spent studying. 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 called REG, and can be found in the default directory. 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 model and their corresponding jackknifed standard errors. The coefficient zero (B00) is the intercept for the model. The other coefficients receive a sequential number starting with 01. This sequential number corresponds to the order of the variables in the list of variables specified in the parameter XVAR.

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 ar1 "c:\pirls\data";

data student;
  set ar1.asgallr1;
  where itsex in(1,2) and inprl01=1;
  regsex = itsex - 1;

%include "c:\pirls\programs\jackreg.sas";
%jackreg (totwgt , jkzone, jkrep, 75, idcntry, regsex, asbgtdy, student);

proc print data=reg noobs;
  var idcntry N Mult_RSQ SS_Res SS_Reg SS_Total b00 b00_se b01 b01_se;
  format idcntry n 6.0 Mult_RSQ 5.3
         SS_Total SS_Reg SS_Res 10.0 B00 B00_SE B01 B01_SE 6.2;
run;
```

IDCNTRY	N	Mult_RSQ	SS_Res	SS_Reg	SS_Total	b00	b00_se	b01	b01_se
32	3054	0.000	1144576	249	1144825	3.21	0.05	-0.04	0.05
84	2695	0.004	12940	58	12998	2.51	0.07	0.18	0.07
100	3413	0.007	109164	781	109945	3.04	0.03	0.18	0.04
124	8185	0.022	293554	6635	300190	2.67	0.03	0.35	0.04
.
705	2939	0.037	20786	792	21578	2.68	0.04	0.39	0.04
752	5891	0.008	84716	693	85409	2.96	0.02	0.15	0.02
792	5046	0.004	947853	3744	951597	2.44	0.03	0.12	0.03
807	3559	0.008	26911	230	27141	2.61	0.04	0.19	0.04
840	3740	0.014	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 watching TV or video daily) by country and by grade. In the listing of the results we can see that there is one entry or line for each value of the variable IDCNTRY. The first column has the country code and the second column has the number of students in each sample. Keep in mind that this number corresponds only to those students who had valid data for the variables in the regression model. Then the multiple R squared is listed for each country, followed by the sums of squares for the model and corresponding regression coefficients. Because of the way in which the variable REGSEX is coded, the constant (B00) is the mean value for the variable ASBGTVDY for the girls and the error of the constant (B00_SE) is standard error 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.SAS)

In this section example SAS 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 SAS macro called JACKREGP.SAS. This macro 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 this macro, 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 SAS 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 SAS interactively, then the macro needs to be called once at the beginning of the session, and 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 string “%JACKREGP” 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>jackregp.sas'.
```

where <path> points to the specific drive and directory where the macro JACKREGP.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, as for example 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.
XVAR	This is a list of independent variables, at least, which under the regression model will be used as predictors of the dependent variable specified by the plausible values. These independent variables can be continuous or categorical, or any other type of coded variable. For example, it could be the variable ITSEX as originally coded in the data files, or dummy coded as 1 or 0.
ROOTPV	This is the prefix used to identify the plausible values for the achievement scale of interest. This corresponds to the first 7 characters of the plausible value variables. For example, the root of the reading plausible value is "ASRREA0."
NPV	This is the number of plausible values that will be used in the analysis. Generally you will want to use all five plausible values for the analysis although under some circumstances fewer can be used. You should always use at least two plausible values for any analysis.
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 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 model and their corresponding jackknifed standard errors combined with the imputation error. The coefficient zero (B00) is the intercept for the model. The other coefficients receive a sequential number starting with 01. This sequential number corresponds to the order of the variables in the list of variables specified in the parameter XVAR.

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 10.6. This code is included on the CD in the file SampleJackREGP.SAS.

Exhibit 10.6 SAS Control Code and Extract of Output File for Using the Macro JACKREGP.SAS

```
libname ar1 "c:\pirls\data";

data student;
  set ar1.asgallr1;
  where itsex in(1,2) and inprl01=1;
  regsex = itsex - 1;

%include "c:\pirls\programs\jackregP.sas";
%jackregP (totwgt , jkzone, jkrep, 75, idcntry, regsex , asrrea0, 5, student);

proc print data=reg noobs;
  var idcntry N Mult_RSQ SS_Res SS_Reg SS_Total b00 b00_se b01 b01_se;
  format idcntry n 6.0 Mult_RSQ 5.3
         SS_Total SS_Reg SS_Res 12.0 B00 B00_SE B01 B01_SE 6.2;
run;
```

IDCNTRY	n	Mult_RSQ	SS_Res	SS_Reg	SS_Total	b00	b00_se	b01	b01_se
32	3292	0.009	6381507424	58398795	6439906220	428.18	6.24	-18.14	4.74
84	2829	0.017	79254797	1332607	80587404	341.73	5.44	-27.17	4.97
100	3458	0.022	636663756	14038000	650701756	562.29	3.69	-24.23	3.61
124	8252	0.015	1129729208	16645023	1146374231	552.89	2.62	-17.31	2.05
.
705	2952	0.023	105914973	2481133	108396106	512.34	2.52	-21.69	2.85
752	6044	0.028	497158256	14240985	511399241	572.27	2.62	-21.96	2.61
792	5125	0.013	7173298914	91215640	7264514554	459.37	3.96	-19.32	3.08
807	3711	0.010	287836438	2982206	290818644	452.24	5.14	-20.87	3.57
840	3763	0.012	25997904796	313113311	26311018108	551.06	3.80	-18.07	4.11

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 plausible values of reading achievement by country. In the listing of the results we can see that there is one entry or line for each value of the variable IDCNTRY after selecting only those cases with IDGRADER=2. The first column has the country code and the second column has the number of students in each sample. Because of the way in which REGSEX is coded, the constant (B00) is the mean reading achievement for the girls, the error of the constant (B00.SE) is the standard error of this estimate. The first regression coefficient (B01) is the difference in the average value in reading achievement between the boys and the girls, and the standard error of this coefficient (B01.SE) is the combined jackknifed sampling error and imputation error of this difference.

For example, from the first line in the output we can say that in Argentina (IDCNTRY = 32) data is available for 3,292 cases. The average of the squared of the multiple correlation between gender and reading achievement is .009 with girls achieving on average 428.18 and boys scoring on average 18.14 points lower than girls. This difference is statistically significant ($18.14 / 4.74 = 3.83$). Overall, we can see that girls outperform boys in reading achievement in all countries in the table.

10.6 Conducting Analyses with the PIRLS 2001 International Database: Student Level

Many analyses of the PIRLS data can be undertaken using student background data as outcome variable. We have already presented some examples in the previous sections when explaining how to use the macros provided with the data files. We will now proceed to work through additional examples of actual analyses where all the steps are undertaken, including the calling of the corresponding SAS macro.

Example Analysis with Student-Level Variables Not Using Plausible Values

In the first example, we want to find out the amount of television that students, on average, watch every day. From the student background questionnaire we see that question 6 asks “About how much time do you spend watching television or videos outside of school on a normal day?” The response options for this question are: “No time,” “Up to one hour,” “From 1 hour up to 3 hours,” “From 3 hours up to 5 hours,” and “5 hours or more.” Since we want to give a numeric value to the number of hours that students report watching television, we recode the response

options to a new variable called TVWATCH that will have values of 0.0, 0.5, 2.0, 4.0, and 7.0, respectively.¹

Since we want to find the average number of hours that students watch television, we need to use the macro JACKGEN. In order to find out these results we need to undertake several steps. After reviewing the codebooks and the questionnaire information, we find that the question about the number of hours students spend watching television or videos corresponds to the variable ASBGTVDY and this variable is found in the Student Background data file. Our next step is to review the documentation of national adaptations to the questionnaires to ensure that there were no deviations listed for this variable (see the supplemental material on the accompanying CD). Since no changes were made, we can continue with our analysis without any modifications.

We then proceed to read the SAS datasets that contain this variable. The SAS code is presented in Exhibit 10.7 and is included in the CD under the name EXAMPLE1.SAS. Extracts from the results obtained from this program are displayed in Exhibit 10.8. We have included as part of the program the corresponding value labels and format statements so that the different categories or groups are labeled appropriately.

Notice that one of the steps in this program is to select only those students in the fourth grade who have non-missing data in the variable. In general, this type of analysis is quite feasible with a powerful desktop computer; however, you need to keep in mind that computing and storage requirements for this type of analysis are quite demanding and increase with the number of countries and variables.

In general, to compute means and percentages using the Student Background data files, you need to do the following:

- Identify the variable, or variables, of interest in the student file and find out about any specific national adaptations to the variable.
- Retrieve the relevant variables from the data files, including the sampling weights, JRR replication information, and any other variables used in the selection of cases.
- Use the macro JACKGEN with the corresponding arguments and parameters.
- Print out the result file.

□□□□□□

¹ We have used this particular recoding for this variable for illustration purpose only. Other recodings might result in a more accurate representation of actual students' behavior.

Exhibit 10.7 SAS Control Statements for Performing Analyses with Student-Level Variables NOT Using Plausible Values (EXAMPLE1.SAS)

```

libname ar1 "c:\pirls\data";
%include "c:\pirls\Programs\jackgen.sas";

title "Example 1: PIRLS 2001 User Guide";

data student;
set ar1.asgargr1
    ar1.asgbgrr1
    ar1.asgswer1
    ar1.asgturrl
    ar1.asgusar1;
where nmiss(asbgtvdy) = 0 and inprl01 = 1 and idgrader = 2;
select(asbgtvdy);
    when(1) tvwatch = 0.0;
    when(2) tvwatch = 0.5;
    when(3) tvwatch = 2.0;
    when(4) tvwatch = 4.0;
    when(5) tvwatch = 7.0;
    otherwise;
end;

* Define the format for the variables used;
proc format library=work;
value country
    32="Argentina" 100="Bulgaria" 752="Sweden" 792="Turkey" 840="United
States";

%jackgen (totwgt , jkzone, jkrep, 75, idcntry, tvwatch, student);

proc print data=final noobs;
var idcntry N totwgt mnx mnx_se pct pct_se;
format idcntry country.
    n 6.0 totwgt 10.0 mnx pct mnx_se pct_se 6.2;

run;

```

Exhibit 10.8 Extract of SAS Computer Output for Performing Analyses with Student-Level Variables Not Using Plausible Values (EXAMPLE 1)

Example 1: PIRLS 2001 User Guide						
IDCNTRY	N	TOTWGT	mnx	mnx_se	pct	pct_se
Argentina	3059	643973	2.98	0.07	11.51	0.54
Bulgaria	3413	94245	2.62	0.06	1.68	0.05
Sweden	5891	115776	2.30	0.03	2.07	0.06
Turkey	5046	964738	1.52	0.05	17.24	0.65
United States	3740	3776946	2.48	0.09	67.50	0.97

In this example, each country’s mean value for TVWATCH is reported for the fourth grade. The results are presented by country. The country is in the first column. The second column has the number of students sampled in the country (with data available for the variable TVWATCH). The third column corresponds to the weights given these sampled students in the whole population,

followed by their mean for TVWATCH and the corresponding standard errors. The last two columns represent the percentages of students sampled responding and the corresponding standard error.

For example, from the first line of the report, we can say that data for Argentina is available for 3,059 cases. These students represent a population of approximately 643,973 students. Argentinian students spend, on average, 2.98 hours watching television or videos outside of school. The standard error of this mean is .07.

Example Analysis with Student-Level Variables Using Plausible Values

In this example, we want to replicate the results presented in Exhibit 4.8 of the international report (replicated results are presented in Exhibit 10.9). We are interested in looking at the fourth graders' reports on the number of books in their home and their reading achievement. Since the results in this exhibit are based on plausible values, we need to use the macro JACKPV.

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

Countries	More than 200 Books		101-200 Books		26-100 Books		11-25 Books		0-10 Books	
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement
Argentina	6 (0.7)	474 (12.6)	7 (0.7)	484 (11.1)	21 (1.6)	466 (9.1)	24 (1.2)	446 (6.8)	42 (2.5)	416 (5.5)
Belize	8 (0.9)	363 (12.0)	7 (0.8)	348 (11.4)	17 (1.1)	357 (8.3)	27 (1.1)	343 (6.5)	41 (2.0)	304 (5.0)
Bulgaria	27 (1.4)	583 (3.8)	16 (0.9)	579 (4.9)	24 (1.1)	565 (3.8)	13 (1.0)	537 (5.7)	20 (1.8)	498 (8.1)
Canada (O,Q)	22 (0.9)	562 (3.4)	22 (0.8)	565 (3.3)	35 (0.7)	546 (2.4)	15 (0.8)	519 (3.3)	6 (0.5)	492 (4.7)
Colombia	5 (0.5)	451 (9.9)	7 (0.5)	455 (9.0)	18 (1.1)	457 (7.3)	24 (1.0)	438 (4.7)	46 (1.8)	401 (4.8)
Cyprus	12 (0.8)	492 (5.6)	13 (0.8)	520 (5.7)	35 (1.1)	510 (3.8)	27 (1.1)	490 (3.6)	12 (0.9)	466 (6.1)
Czech Republic	20 (1.3)	557 (3.7)	21 (1.0)	552 (3.1)	41 (1.4)	538 (2.8)	15 (1.0)	512 (3.1)	4 (0.4)	479 (7.9)
England	20 (1.0)	577 (5.0)	23 (1.2)	579 (4.1)	34 (1.2)	555 (3.5)	16 (1.0)	518 (3.8)	7 (0.6)	477 (6.7)
France	19 (1.0)	555 (4.1)	19 (0.9)	545 (3.1)	36 (1.0)	526 (2.6)	19 (0.9)	506 (3.6)	8 (0.5)	472 (6.1)
Germany	16 (0.7)	570 (2.6)	16 (0.7)	568 (2.9)	35 (0.8)	546 (2.1)	24 (0.8)	523 (2.4)	9 (0.6)	486 (4.2)
Greece	15 (1.1)	544 (6.0)	15 (0.9)	548 (4.3)	35 (1.4)	536 (5.7)	25 (1.3)	514 (4.9)	9 (0.9)	488 (7.1)
Hong Kong, SAR	9 (0.6)	533 (5.0)	10 (0.5)	536 (4.0)	30 (0.8)	538 (3.0)	26 (0.7)	529 (3.4)	26 (1.1)	515 (4.4)
Hungary	28 (1.3)	568 (3.1)	19 (0.7)	559 (2.8)	29 (1.0)	544 (2.9)	16 (0.8)	522 (2.8)	7 (0.8)	481 (5.0)
Iceland	23 (0.7)	528 (3.3)	25 (0.6)	532 (2.1)	35 (0.7)	515 (2.3)	13 (0.5)	490 (3.7)	4 (0.3)	455 (6.1)
Iran, Islamic Rep. of	4 (0.4)	466 (8.0)	5 (0.4)	469 (9.0)	12 (0.9)	473 (6.1)	22 (1.3)	440 (4.6)	56 (2.5)	390 (3.4)
Israel	19 (1.1)	527 (5.7)	16 (0.8)	536 (4.7)	33 (1.1)	532 (3.2)	21 (1.0)	506 (4.5)	10 (0.9)	472 (7.5)
Italy	13 (0.6)	551 (4.2)	14 (0.8)	560 (3.9)	30 (0.8)	553 (3.6)	29 (1.0)	538 (2.9)	14 (0.8)	499 (4.2)
Kuwait	12 (0.6)	418 (6.0)	11 (0.6)	415 (6.7)	24 (0.8)	411 (5.4)	24 (1.0)	405 (5.6)	29 (1.5)	386 (6.5)
Latvia	24 (1.3)	555 (3.1)	24 (1.1)	557 (3.6)	34 (0.9)	546 (2.5)	13 (0.8)	525 (4.3)	5 (0.6)	513 (6.4)
Lithuania	9 (0.6)	562 (5.0)	16 (1.0)	568 (4.0)	39 (1.1)	552 (2.6)	25 (1.1)	529 (3.5)	10 (1.0)	505 (5.1)
Macedonia, Rep. of	8 (0.8)	465 (10.6)	6 (0.5)	487 (7.9)	27 (1.5)	491 (4.8)	34 (1.6)	454 (5.6)	24 (1.8)	415 (7.4)
Moldova, Rep. of	6 (0.9)	542 (8.6)	8 (0.8)	538 (5.4)	21 (1.1)	515 (4.4)	27 (1.3)	488 (4.4)	37 (2.0)	468 (5.1)
Morocco	2 (0.5)	--	3 (0.4)	331 (17.2)	9 (1.2)	372 (13.9)	18 (1.5)	351 (8.5)	68 (2.4)	362 (12.4)
Netherlands	12 (0.8)	570 (3.8)	16 (0.7)	569 (3.0)	36 (0.9)	560 (2.6)	25 (1.0)	547 (2.7)	11 (0.8)	523 (5.3)
New Zealand	23 (1.4)	560 (5.3)	23 (1.1)	556 (4.0)	33 (1.3)	533 (4.5)	14 (0.9)	493 (5.9)	8 (0.7)	464 (7.2)
Norway	27 (1.1)	518 (3.6)	25 (0.9)	517 (3.9)	33 (1.2)	495 (3.6)	10 (0.9)	463 (6.5)	4 (0.5)	435 (7.5)
Romania	9 (0.9)	565 (5.2)	9 (0.8)	558 (5.7)	25 (1.1)	534 (4.2)	25 (1.3)	508 (5.0)	32 (1.8)	482 (9.5)
Russian Federation	17 (1.0)	541 (4.7)	16 (0.8)	543 (4.4)	35 (1.1)	535 (4.0)	21 (1.1)	519 (5.8)	11 (1.0)	496 (9.3)
Scotland	19 (0.9)	553 (5.9)	18 (1.1)	554 (6.1)	34 (1.0)	539 (3.9)	18 (0.8)	506 (4.1)	11 (0.8)	466 (6.6)
Singapore	16 (0.7)	569 (5.7)	19 (0.6)	560 (4.9)	37 (1.0)	540 (4.1)	19 (0.9)	500 (5.3)	10 (0.7)	442 (7.8)
Slovak Republic	15 (1.0)	542 (4.5)	22 (1.0)	535 (4.1)	40 (1.4)	524 (3.0)	17 (1.1)	496 (3.8)	7 (1.2)	450 (15.5)
Slovenia	15 (1.0)	513 (4.1)	19 (0.9)	523 (3.2)	37 (1.1)	511 (2.0)	21 (1.1)	484 (3.4)	8 (0.7)	454 (6.4)
Sweden	31 (1.5)	577 (2.7)	26 (0.7)	572 (2.5)	31 (1.1)	552 (2.3)	9 (0.7)	530 (4.0)	3 (0.5)	503 (6.9)
Turkey	5 (0.6)	494 (10.8)	5 (0.5)	499 (8.3)	20 (1.1)	478 (4.3)	30 (1.1)	455 (3.7)	40 (2.0)	422 (3.8)
United States	21 (1.2)	556 (5.8)	22 (1.1)	568 (4.3)	31 (1.1)	551 (3.8)	17 (1.0)	522 (4.2)	9 (1.0)	477 (7.0)
International Avg.	15 (0.2)	529 (1.3)	15 (0.1)	525 (1.1)	30 (0.2)	514 (0.8)	21 (0.2)	490 (0.8)	19 (0.2)	459 (1.2)

Background data provided by students.
 (.) Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.
 An "r" indicates data are available for 70-84% of the students. An "s" indicates data are available for 50-69% of the students. An "x" indicates data are available for less than 50% of the students.
 A dash (-) indicates data are not available. A tilde (~) indicates insufficient data to report achievement.

To replicate the results in this exhibit, we need to undertake several steps. After reviewing the codebooks and the questionnaire information we find that the question about the number of books in the home (see the CD for a copy of the student questionnaires) corresponds to the variable ASBGBOOK, and this variable is found in the Student Background data file. Our next step is to review the documentation of national adaptations to the questionnaires to ensure that there were no deviations listed for this variable (also on the CD). If no changes were made we can continue with our analysis without any modifications.

We then proceed to read from the SAS dataset that contains this variable. The SAS code is presented in Exhibit 10.10. Extracts from the results obtained from this program are displayed in Exhibit 10.11. We have included as part of the program the corresponding value labels and format statements so that the different categories or groups are labeled appropriately.

Notice that one of the steps in this program is to select only those students in the fourth grade who chose one of the five options presented in the question.

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

- Identify the variable or variables of interest in the student file and find out about any specific national adaptations to the variable.
- Retrieve the relevant variables from the data files, including the achievement score, sampling weights, JRR replication information, and any other variables used in the selection of cases.
- Use the macro JACKPV with the corresponding arguments and parameters.
- Print out the result file.

Exhibit 10.10 SAS Control Statements for Performing Analyses with Student-Level Variables Involving Plausible Values (EXAMPLE2.SAS)

```

libname ar1 "c:\pirls\data";
%include "c:\pirls\Programs\jackpv.sas";

title "Example 2: PIRLS 2001 User Guide";

data student;
  set ar1.asgargr1
      ar1.asgbgrr1
      ar1.asgswer1
      ar1.asgturrl
      ar1.asgusar1;
  where nmiss(asbgbook) = 0 and inpr101 = 1 and idgrader = 2;

  * Define the format for the variables used;
proc format library=work;
value books
  1='0 -10 BOOKS' 2='11-25 BOOKS' 3='26-100 BOOKS' 4='101-200 BOOKS'
  5='MORE THAN 200 BOOKS';
value country
  32="Argentina" 100="Bulgaria" 752="Sweden" 792="Turkey" 840="United
  States";

%jackpv (totwgt , jkzone, jkrep, 75, idcntry asbgbook, asrrea0, 5, student);
proc print data=final noobs;
  by idcntry ;
  var asbgbook N totwgt mnpv mnpv_se pct pct_se;
  format idcntry country. asbgbook books.
         N 6.0 totwgt 10.0 mnpv mnpv_se pct pct_se 6.2;

run;

```

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*=Argentina

ASBGBOOK          N          TOTWGT          mnpv          mnpv_se          pct          pct_se
0 -10 BOOKS          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          499          107115          465.93          9.13          21.18          1.61
101-200 BOOKS          181          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          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          TOTWGT          mnpv          mnpv_se          pct          pct_se
0 -10 BOOKS          163          3794          503.47          6.92          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          mnpv          mnpv_se          pct          pct_se
0 -10 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 States

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          653582          521.51          4.23          17.47          0.96
26-100 BOOKS          1167          1171921          551.13          3.84          31.33          1.06
101-200 BOOKS          799          804711          568.38          4.29          21.51          1.10
MORE THAN 200 BOOKS          767          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 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 *PIRLS International Report* (replicated results are presented in Exhibit 10.12).

Exhibit 10.12 Sample Exhibit for Teacher-Level Analysis Involving Plausible Values Taken From the *PIRLS 2001 International Report*

Countries	Every Day		3-4 Days a Week		Fewer than 3 Days a Week	
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement
Argentina	28 (4.1)	415 (11.4)	53 (4.8)	416 (9.2)	19 (3.6)	432 (13.6)
Belize	71 (5.2)	324 (5.6)	24 (4.5)	325 (17.1)	5 (2.1)	314 (19.7)
Bulgaria	31 (4.4)	551 (8.2)	60 (4.5)	554 (5.0)	9 (2.3)	543 (10.4)
Canada (O,Q)	67 (3.5)	548 (2.7)	24 (3.3)	540 (4.3)	8 (1.8)	530 (7.8)
Colombia	58 (4.3)	426 (5.0)	28 (3.9)	408 (10.0)	13 (3.1)	442 (18.8)
Cyprus	79 (4.6)	494 (3.5)	17 (4.2)	489 (6.8)	4 (1.9)	500 (14.6)
Czech Republic	44 (4.0)	536 (4.0)	42 (4.3)	535 (3.5)	13 (3.1)	547 (5.9)
England	74 (3.6)	551 (4.2)	23 (3.6)	564 (8.8)	3 (1.7)	570 (8.8)
France	66 (3.9)	524 (2.8)	28 (3.7)	527 (5.2)	7 (1.9)	531 (7.3)
Germany	51 (3.0)	537 (2.5)	32 (2.8)	545 (3.4)	16 (2.3)	539 (3.8)
Greece	81 (3.5)	522 (3.6)	13 (3.5)	535 (12.6)	7 (2.9)	550 (20.4)
Hong Kong, SAR	17 (3.4)	525 (7.4)	38 (4.6)	534 (5.2)	46 (4.4)	525 (4.4)
Hungary	40 (4.2)	543 (3.6)	60 (4.2)	545 (3.0)	0 (0.0)	--
Iceland	37 (0.4)	511 (1.7)	41 (0.4)	513 (2.4)	22 (0.3)	516 (2.6)
Iran, Islamic Rep. of	32 (4.0)	412 (10.4)	52 (4.3)	421 (6.7)	17 (3.4)	404 (8.8)
Israel	55 (4.3)	505 (7.1)	35 (4.2)	512 (7.4)	11 (2.7)	515 (10.9)
Italy	36 (3.6)	535 (4.7)	49 (3.9)	544 (3.6)	15 (2.6)	549 (5.9)
Kuwait	82 (3.0)	399 (5.7)	12 (2.5)	398 (8.3)	6 (1.6)	412 (7.5)
Latvia	66 (4.3)	540 (3.1)	32 (4.3)	555 (4.2)	2 (1.1)	--
Lithuania	67 (4.2)	543 (3.6)	29 (4.0)	542 (5.4)	4 (1.6)	561 (2.6)
Macedonia, Rep. of	41 (3.8)	459 (9.2)	39 (4.4)	431 (8.3)	20 (3.7)	429 (14.6)
Moldova, Rep. of	32 (3.8)	503 (7.1)	63 (3.9)	489 (5.3)	5 (1.8)	455 (12.4)
Morocco	35 (5.3)	340 (11.4)	59 (5.5)	344 (11.7)	7 (2.4)	351 (31.6)
Netherlands	55 (4.3)	557 (3.2)	39 (4.2)	549 (4.9)	6 (1.8)	550 (7.3)
New Zealand	68 (3.7)	522 (4.2)	30 (3.8)	543 (8.2)	2 (1.0)	--
Norway	58 (4.5)	503 (4.6)	32 (3.9)	493 (4.4)	10 (3.1)	496 (8.5)
Romania	71 (4.1)	521 (5.4)	24 (3.7)	495 (12.2)	4 (2.0)	482 (26.1)
Russian Federation	87 (2.6)	529 (4.7)	12 (2.5)	522 (11.0)	1 (0.6)	--
Scotland	45 (4.9)	526 (5.4)	45 (4.9)	528 (5.7)	11 (2.9)	547 (13.0)
Singapore	45 (4.0)	522 (8.2)	34 (4.0)	530 (9.6)	22 (3.3)	532 (11.9)
Slovak Republic	48 (3.9)	514 (4.5)	52 (3.9)	521 (3.7)	0 (0.0)	--
Slovenia	36 (4.1)	499 (3.6)	34 (4.0)	498 (3.5)	30 (4.2)	509 (4.1)
Sweden	57 (3.6)	562 (3.3)	30 (3.4)	556 (3.5)	13 (2.2)	568 (3.4)
Turkey	54 (5.0)	454 (5.2)	37 (4.5)	439 (6.7)	10 (2.7)	459 (9.9)
United States	93 (2.1)	541 (4.2)	6 (2.0)	555 (7.5)	1 (0.5)	--
International Avg.	54 (0.7)	500 (1.0)	35 (0.7)	500 (1.3)	10 (0.4)	495 (2.2)

Background data provided by teachers.
 () Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.
 An "r" indicates data are available for 70-84% of the students. An "s" indicates data are available for 50-69% of the students. An "x" indicates data are available for less than 50% of the students.
 A dash (–) indicates data are not available. A tilde (~) indicates insufficient data to report achievement.

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 IDCNTY, 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 ar1.astargr1
      ar1.astbgrr1
      ar1.astswer1
      ar1.astturrl
      ar1.astusar1;

proc sort data=studteac;
  by idcnty idteach idlink;

data teacher;
  set ar1.atgargr1
      ar1.atgbgrr1
      ar1.atgswer1
      ar1.atgturrl
      ar1.atgusar1;

proc sort data=teacher;
  by idcnty idteach idlink;

* Now merge the two files;
data merged;
  merge studteac teacher;
  by idcnty idteach idlink;
  if nmiss(atbgract) = 0 and inpr101 = 1;

* Define the format for the variables used;
proc format library=work;
  value country
    32="Argentina" 100="Bulgaria" 752="Sweden" 792="Turkey" 840="United
States";
  value ract 1='Every Day' 2='3-4 Days per week' 3='Fewer than 3 days per
week';

%jackpv(tchwggt , jkzzone, jkrep, 75, idcnty atbgract, asrrea0, 5, merged);

proc print data=final noobs;
```

```

by idcntry ;
var atbgract N tchwgt mnpv mnpv_se pct pct_se;
format idcntry 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)

Example 3: PIRLS 2001 User Guide

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	471	123414	431.83	13.63	19.10	3.64
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
COUNTRY ID=Turkey						
ATBGRACT	N	TCHWGT	mnpv	mnpv_se	pct	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	388	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 (IDCOUNTRY, IDTEACHER, 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 IDCOUNTRY, IDTEACHER, 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 always preferable to analyze school-level variables as attributes of students, rather than as elements in their own right.

Although the school samples were not designed primarily to report on school characteristics, it is possible to use school weights to provide reasonable estimates of percentages and averages for primary or middle schools in each country. The following example, however, describe only analyses based on student-weighted data.

For student-weighted analyses, the school-level data are analyzed to make statements about the number of students attending schools with one characteristic or another rather than the number of schools with certain characteristics. When school-level variables are analyzed, we recommend that you merge the selected school-level variables with the student-level file, and then use the sampling and weight

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 located” (ACBGCOMM). The response options for this question were “Urban,” “Suburban,” and “Rural.” These results can be found in Exhibit 7.1 of the *PIRLS International Report* (replicated results are presented in Exhibit 10.15).

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 ACBGCOMM in the School Background file, and the student weights and plausible values in the Student Background file. We then proceed to read the School Background file that has the variables that are relevant to our analysis. We then read the variables of interest from the student data file. We then proceed to merge the school information with the student information using the variables IDCNTRY and IDSCHOOL as merge variables, and then use the macro JACKPV to obtain the corresponding percentages of students within each group, and their mean achievement scores in reading. The computer code used to run this analysis in SAS can be found in Exhibit 10.16 and an extract of the results is shown in Exhibit 10.17.

Exhibit 10.15 Sample Exhibit for School-Level Analysis Involving Plausible Values Taken From the *PIRLS 2001 International Report*

Countries	Urban		Suburban		Rural	
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement
Argentina	75 (4.5)	427 (6.5)	19 (4.0)	395 (15.2)	5 (2.0)	412 (15.8)
Belize	38 (6.3)	361 (9.4)	15 (6.5)	318 (28.0)	48 (6.2)	296 (7.3)
Bulgaria	64 (2.8)	566 (4.2)	12 (2.6)	542 (17.1)	24 (2.3)	516 (9.1)
Canada (O,Q)	40 (3.7)	540 (3.7)	39 (3.5)	552 (3.4)	21 (2.9)	536 (4.3)
Colombia	70 (3.0)	435 (4.8)	2 (1.1)	--	29 (3.0)	388 (7.3)
Cyprus	44 (3.1)	502 (5.3)	25 (3.2)	497 (7.5)	30 (2.7)	482 (4.9)
Czech Republic	55 (3.4)	540 (3.0)	17 (3.3)	537 (4.8)	28 (3.2)	529 (6.0)
England	48 (5.0)	542 (4.2)	32 (4.6)	557 (6.9)	21 (3.7)	574 (5.3)
France	36 (4.1)	525 (5.4)	29 (4.2)	523 (4.6)	35 (3.6)	527 (3.3)
Germany	33 (2.8)	523 (3.2)	22 (2.9)	546 (4.3)	45 (3.8)	547 (2.0)
Greece	--	--	--	--	--	--
Hong Kong, SAR	52 (3.0)	533 (3.8)	46 (2.7)	526 (5.2)	2 (1.1)	--
Hungary	28 (2.5)	566 (4.3)	36 (2.3)	546 (4.9)	36 (1.7)	526 (2.8)
Iceland	36 (0.4)	519 (2.3)	43 (0.4)	511 (2.0)	21 (0.3)	499 (3.5)
Iran, Islamic Rep. of	48 (3.7)	444 (6.3)	9 (2.4)	413 (14.5)	43 (3.1)	375 (4.5)
Israel	52 (3.8)	525 (5.6)	22 (3.2)	531 (7.6)	27 (3.3)	461 (11.4)
Italy	76 (3.2)	543 (2.5)	14 (2.5)	537 (6.4)	10 (2.2)	529 (7.6)
Kuwait	18 (2.7)	399 (10.2)	79 (2.7)	399 (5.5)	3 (0.5)	426 (20.0)
Latvia	44 (3.8)	559 (3.8)	18 (3.9)	551 (4.6)	37 (3.0)	523 (3.2)
Lithuania	71 (2.8)	552 (3.2)	6 (2.1)	542 (11.8)	23 (2.4)	515 (4.9)
Macedonia, Rep. of	57 (3.3)	472 (5.3)	15 (2.8)	462 (15.3)	28 (3.1)	387 (9.6)
Moldova, Rep. of	28 (3.3)	515 (6.8)	14 (3.8)	494 (12.2)	59 (3.7)	480 (5.0)
Morocco	41 (4.4)	354 (9.1)	21 (4.4)	323 (10.8)	39 (4.1)	352 (24.9)
Netherlands	31 (3.9)	541 (6.6)	22 (3.7)	562 (3.9)	47 (3.8)	558 (3.2)
New Zealand	38 (4.0)	533 (6.1)	40 (3.6)	522 (6.4)	23 (2.9)	540 (8.9)
Norway	19 (3.5)	505 (7.7)	26 (3.9)	511 (6.2)	56 (3.5)	492 (3.5)
Romania	50 (2.5)	526 (5.1)	4 (2.0)	473 (15.9)	45 (2.8)	499 (8.2)
Russian Federation	55 (2.5)	539 (4.8)	3 (1.9)	553 (6.8)	43 (2.4)	512 (5.8)
Scotland	35 (4.3)	517 (6.3)	39 (5.3)	538 (6.6)	26 (4.5)	534 (6.9)
Singapore	100 (0.0)	528 (5.2)	0 (0.0)	--	0 (0.0)	--
Slovak Republic	50 (3.7)	531 (3.6)	10 (2.5)	522 (7.8)	40 (3.1)	501 (4.5)
Slovenia	39 (3.3)	509 (3.3)	27 (3.7)	499 (4.2)	34 (3.0)	495 (3.6)
Sweden	15 (3.0)	567 (5.4)	66 (3.8)	560 (2.8)	18 (3.5)	560 (5.4)
Turkey	36 (4.2)	464 (6.0)	39 (4.1)	455 (6.1)	25 (3.5)	423 (8.2)
United States	32 (3.5)	519 (6.6)	35 (4.6)	572 (4.6)	33 (3.2)	537 (7.3)
International Avg.	46 (0.6)	507 (0.9)	25 (0.6)	502 (1.8)	29 (0.5)	485 (1.5)

Background data provided by schools.
 () Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent.
 An "r" indicates data are available for 70-84% of the students. An "s" indicates data are available for 50-69% of the students. An "x" indicates data are available for less than 50% of the students.
 A dash (-) indicates data are not available. A tilde (~) indicates insufficient data to report achievement.

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

```

title "Example 4: PIRLS 2001 User Guide";

libname ar1 "c:\pirls\data";
%include "c:\pirls\Programs\jackpv.sas";

data student;
  set ar1.asgargr1
      ar1.asgbgrr1
      ar1.asgswer1
      ar1.asgturr1
      ar1.asgusar1;

proc sort data=student;
  by idcntry idschool;

data school;
  set ar1.acgargr1
      ar1.acgbgrr1
      ar1.acgswer1
      ar1.acgturr1
      ar1.acgusar1;

proc sort data=school;
  by idcntry idschool;

* Now merge the two files;
data merged;
  merge student school;
  by idcntry idschool;
  if nmiss(acbgcomm) = 0 and inpr101 = 1;

* Define the format for the variables used;
proc format library=work;
  value country
    32="Argentina" 100="Bulgaria" 752="Sweden" 792="Turkey" 840="United
States";
  value comm 1='Urban' 2='Suburban' 3='Rural';

%jackpv(totwgt , jkzone, jkrep, 75, idcntry acbgcomm, asrrea0, 5, merged);

proc print data=final noobs;
  by idcntry ;
  var acbgcomm N totwgt mnpv mnpv_se pct pct_se;
  format idcntry country. acbgcomm comm.
         N 6.0 totwgt 10.0 mnpv mnpv_se pct pct_se 6.2;

run;

```

Exhibit 10.17 Extract of SAS Computer Output for Performing Student-Weighted Analyses with School-Level Variables(EXAMPLE 4)

```

Example 4: PIRLS 2001 User Guide

*COUNTRY ID*=Argentina
ACBGCOMM      N      TOTWGT      mnpv      mnpv_se      pct      pct_se
Urban          2498      508261      427.13      6.46      75.42      4.50
Suburban       472      129481      395.02      15.18      19.21      4.04
Rural          156      36134      412.00      15.75      5.36      2.03

*COUNTRY ID*=Bulgaria
ACBGCOMM      N      TOTWGT      mnpv      mnpv_se      pct      pct_se

```


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
COUNTRY ID=Sweden						
ACBGCOMM	N	TOTWGT	mnpv	mnpv_se	pct	pct_se
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
COUNTRY ID=Turkey						
ACBGCOMM	N	TOTWGT	mnpv	mnpv_se	pct	pct_se
Urban	1887	341684	464.38	5.96	36.18	4.24
Suburban	1881	363850	454.53	6.09	38.53	4.09
Rural	1191	238830	422.52	8.17	25.29	3.49
COUNTRY ID=United States						
ACBGCOMM	N	TOTWGT	mnpv	mnpv_se	pct	pct_se
Urban	1498	1174714	518.55	6.56	32.10	3.48
Suburban	1383	1274273	572.29	4.55	34.82	4.62
Rural	776	1210399	537.42	7.35	33.08	3.21

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 school locations 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, 2,498 students who represent 508,261 students in the population attend schools that are located in an urban area. These students have a mean reading achievement of 427.13 with a standard error of 6.46., and represent 75.42 percent of the sampled students. This percentage has a standard error of 4.50.

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

- Identify the variable or variables of interest in the student file and find out about any specific national adaptations to the variable.
- Retrieve the relevant variables from the student files, including the achievement score, sampling weights, JRR replication information, and any other variables used in the selection of cases.
- Retrieve the relevant classification variable or variables from the school database.

- 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."

Exhibit 10.18 Sample Exhibit for Home Background Analysis Involving Plausible Values Taken From the *PIRLS 2001 International Report*

Countries	High PATR		Medium PATR		Low PATR		Percentage of Students at High Level of PATR
	Percent of Students	Average Achievement	Percent of Students	Average Achievement	Percent of Students	Average Achievement	
Hungary	74 (1.0)	556 (2.2)	23 (0.9)	518 (2.5)	3 (0.4)	509 (8.3)	
Norway	73 (1.1)	512 (3.2)	23 (1.0)	481 (4.5)	4 (0.5)	470 (10.8)	
Sweden	71 (0.8)	571 (2.1)	24 (0.9)	547 (3.4)	5 (0.4)	536 (4.5)	
Scotland	s 69 (1.3)	554 (3.9)	24 (1.2)	520 (6.3)	7 (0.6)	511 (9.3)	
England	s 69 (1.3)	586 (3.5)	26 (1.2)	543 (5.0)	6 (0.6)	535 (8.2)	
Iceland	r 67 (0.8)	522 (1.6)	30 (0.7)	510 (2.7)	3 (0.4)	489 (8.1)	
Slovak Republic	65 (1.2)	533 (3.0)	32 (1.2)	503 (3.8)	3 (0.4)	483 (10.8)	
New Zealand	r 65 (1.2)	559 (3.8)	30 (1.2)	501 (4.0)	5 (0.6)	488 (10.1)	
Czech Republic	r 64 (1.1)	551 (2.7)	32 (0.9)	528 (3.0)	4 (0.5)	515 (9.8)	
Canada (O,Q)	r 64 (1.0)	561 (2.4)	32 (1.0)	534 (2.7)	4 (0.3)	525 (5.4)	
Greece	64 (1.4)	539 (3.5)	31 (1.3)	505 (4.9)	6 (0.6)	505 (9.6)	
Netherlands	s 62 (1.3)	571 (2.6)	33 (1.1)	553 (3.0)	5 (0.6)	537 (7.4)	
Slovenia	61 (1.0)	516 (2.3)	37 (1.0)	487 (2.7)	2 (0.3)	~ ~	
Cyprus	s 57 (1.4)	507 (3.5)	38 (1.5)	477 (4.6)	5 (0.6)	466 (6.9)	
Italy	56 (1.2)	560 (2.1)	37 (1.1)	527 (2.9)	7 (0.5)	523 (5.2)	
Germany	r 55 (1.1)	559 (2.0)	35 (1.0)	528 (2.5)	10 (0.6)	519 (3.9)	
Bulgaria	54 (1.8)	574 (3.7)	37 (1.4)	538 (5.0)	9 (1.2)	497 (11.6)	
France	53 (1.1)	545 (2.5)	43 (1.1)	512 (3.2)	4 (0.4)	497 (6.1)	
Russian Federation	51 (1.3)	536 (4.6)	45 (1.2)	522 (5.2)	5 (0.5)	516 (6.6)	
Macedonia, Rep. of	r 49 (1.9)	479 (4.8)	48 (1.8)	426 (6.8)	3 (0.4)	420 (12.2)	
Latvia	48 (1.3)	560 (2.9)	50 (1.4)	537 (2.7)	3 (0.4)	532 (13.8)	
Iran, Islamic Rep. of	43 (1.4)	438 (4.8)	52 (1.2)	402 (4.8)	5 (0.8)	373 (16.9)	
Lithuania	42 (1.2)	558 (2.9)	52 (1.2)	537 (3.0)	6 (0.5)	528 (6.1)	
Belize	s 41 (1.8)	358 (6.1)	52 (1.7)	309 (6.7)	7 (1.0)	315 (16.8)	
Colombia	r 41 (1.4)	443 (5.4)	52 (1.3)	414 (5.0)	7 (0.6)	428 (7.4)	
Singapore	40 (0.9)	553 (4.6)	54 (0.9)	517 (5.5)	6 (0.3)	517 (6.9)	
Romania	34 (1.6)	542 (4.2)	55 (1.5)	501 (5.5)	11 (1.2)	492 (10.6)	
Kuwait	r 34 (0.6)	413 (4.2)	61 (0.6)	396 (5.1)	5 (0.3)	388 (7.2)	
Argentina	s 33 (1.6)	462 (7.1)	62 (1.7)	417 (6.8)	5 (0.5)	410 (11.0)	
Turkey	29 (1.4)	483 (5.0)	63 (1.3)	440 (3.5)	7 (0.8)	436 (6.8)	
Hong Kong, SAR	29 (0.9)	539 (3.1)	65 (0.8)	527 (3.1)	6 (0.4)	526 (6.2)	
Moldova, Rep. of	29 (1.5)	516 (5.1)	63 (1.6)	487 (4.2)	8 (0.7)	463 (7.5)	
Israel	x x	x x	x x	x x	x x	x x	
Morocco	--	--	--	--	--	--	
United States	--	--	--	--	--	--	
International Avg.	53 (0.2)	524 (0.7)	42 (0.2)	492 (0.8)	5 (0.1)	482 (1.6)	

Index of Parents' Attitudes Toward Reading (PATR)

Based on parents' agreement with the following: I read only if I have to; I like talking about books with other people; I like to spend my spare time reading; I read only if I need information; and Reading is an important activity in my home. Average is computed across the 5 items based on a 4-point scale: Disagree a lot = 1, Disagree a little = 2, Agree a little = 3, and Agree a lot = 4. Responses for negative statements were reverse-coded. High level indicates an average of 2 through 3, and Low level indicates an average of 1 to less than 2.

() Standard errors appear in parentheses. Because results are rounded to the nearest whole number, some totals may appear inconsistent. An "r" indicates data are available for 70-84% of the students. An "s" indicates data are available for 50-69% of the students. An "x" indicates data are available for less than 50% of the students. A dash (-) indicates data are not available. A tilde (~) indicates insufficient data to report achievement.

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 IDCNTY 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";

libname arl "c:\pirls\data";
%include "c:\pirls\Programs\jackpv.sas";

data student;
  set arl.asgargr1
      arl.asgbgrr1
      arl.asgswer1
      arl.asgturrl;

proc sort data=student;
  by idcntry idstud;

data home;
  set arl.ashargr1
      arl.ashbgrr1
      arl.ashswer1
      arl.ashturrl;

proc sort data=home;
  by idcntry idstud;

* Now merge the two files;
data merged;
  merge student home;
  by idcntry idstud;
  if nmiss(asdhpatr) = 0 and inpr101 = 1;

* Define the format for the variables used;
proc format library=work;
  value country
    32="Argentina" 100="Bulgaria" 752="Sweden" 792="Turkey" 840="United
States";
  value patr 1='Low' 2='Medium' 3='High';

%jackpv(totwgt , jkzone, jkrep, 75, idcntry asdhpatr, asrrea0, 5, merged);

proc print data=final noobs;
  by idcntry ;
  var asdhpatr N totwgt mnpv mnpv_se pct pct_se;
  format idcntry country. asdhpatr patr.
         N 6.0 totwgt 10.0 mnpv mnpv_se pct pct_se 6.2;

run;

```

Exhibit 10.20 Extract of SAS Computer Output for Performing Home Background Variables (EXAMPLE 5)

Example 5: PIRLS 2001 User Guide						
COUNTRY ID=Argentina						
ASDHPATR	N	TOTWGT	mpv	mpv_se	pct	pct_se
Low	689	138668	462.12	7.11	32.87	1.59
Medium	1234	263009	416.76	6.81	62.35	1.66
High	99	20151	409.70	11.05	4.78	0.50
COUNTRY ID=Bulgaria						
ASDHPATR	N	TOTWGT	mpv	mpv_se	pct	pct_se
Low	1776	47355	574.10	3.73	54.34	1.77
Medium	1136	32371	538.07	4.99	37.15	1.37
High	230	7417	496.82	11.58	8.51	1.24
COUNTRY ID=Sweden						
ASDHPATR	N	TOTWGT	mpv	mpv_se	pct	pct_se
Low	4022	76513	571.35	2.13	71.24	0.83
Medium	1204	25579	546.77	3.40	23.82	0.85
High	237	5303	536.14	4.53	4.94	0.40
COUNTRY ID=Turkey						
ASDHPATR	N	TOTWGT	mpv	mpv_se	pct	pct_se
Low	1351	256311	482.51	4.99	29.34	1.42
Medium	2854	552066	439.75	3.48	63.20	1.26
High	323	65102	436.43	6.84	7.45	0.76

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 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, 689 students who represent 138,668 students in the population have parents or guardians with a high positive attitude towards reading, and have a mean reading achievement score of 462.12 with a standard error of 7.11. These students represent 32.87 percent of the sampled students and this percentage has a standard error of 1.59.

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

- Identify the variable or variables of interest in the student file and find out about any specific national adaptations to the variable. Note that in our example we have not included data for the United States since the Home Questionnaire was not administered there.

- Retrieve the relevant variables from the student files, including the achievement score, sampling weights, JRR replication information, and any other variables used in the selection of cases.
- Retrieve the relevant classification variable or variables from the home background file.
- Merge the variables from the school database onto the student database using the variables IDCNTRY and IDSTUD.
- Use the macro JACKGEN or JACKPV with the corresponding arguments and parameters.
- Print out the result file.

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 Center at Boston College

Ina V.S. Mullis, Co-Director

Michael O. Martin, Co-Director

Eugenio J. Gonzalez, Director of Operations and Data Analysis

Ann Kennedy, PIRLS Project Coordinator

Cheryl Flaherty, PIRLS Research Associate

Teresa A. Smith, TIMSS Science Coordinator

Robert A. Garden, TIMSS Mathematics Coordinator

Steven J. Chrostowski, TIMSS Project Coordinator

Ebru Erberber, TIMSS Research Associate

Monica Guidi, TIMSS Research Assistant

María José Ramirez, Graduate Assistant

Alka Arora, Graduate Assistant

Joseph Galia, Statistician Programmer

Isaac Li, Statistician Programmer

José R. Nieto, Publications Manager

Mario Pita, Data Graphics Specialist

Betty Hugh, Data Graphics Specialist

Susan Messner, Data Graphics Specialist

Marcie Petras, Manager of Office Administration

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

A.3 PIRLS Advisory Committees

The PIRLS Reading Development Group (RDG) contributed their invaluable expertise to developing and reviewing the framework and reading test. The RDG also worked with the PIRLS Reading Coordinator, NFER, and ISC staff members to develop the descriptions of achievement at international benchmarks. The PIRLS Questionnaire Development Group (QDG) helped develop the PIRLS questionnaires, including writing items and reviewing drafts of all questionnaires.

PIRLS Reading Development Group

Marilyn Binkley

National Center for Education Statistics

United States

Karl Bluemel

Vienna School Board

Austria

Sue Horner

Qualifications and Curriculum Authority

England

Pirjo Linnakylä

University of Jyväskylä

Finland

Martine Rémond

Institut National de la Recherche Pédagogique

France

William Tunmer

Massey University

New Zealand

Tan See Keen

Ministry of Education

Singapore

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
University of Hamburg

Greece

Georgia Kontogiannopoulou-Polydorides
Costas Basbas
University of Athens

Hong Kong, SAR

Tse Shek-Kam
The University of Hong Kong

Hungary

Péter Vari
Emese Felvégi
National Institute of Public Education
Centre for Evaluation Studies

Iceland

Einar Gudmundsson
University of Iceland

Islamic Republic of Iran

Abdol'azim Karimi
Ministry of Education

Israel

Elite Olshtain
Hebrew University

Ruth Zuzovsky
Tel Aviv University

Italy

Gabriella Pavan de Gregorio
National Institute for the Evaluation
of the Educational System (CEDE)

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)

