

# PIRLS Sampling Design

Pierre Foy

Marc Joncas

## 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 – Main Survey* and *School Coordinator Manual – Main Survey* (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 primary school level.

### 5.2.1 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 – not part of the target population.

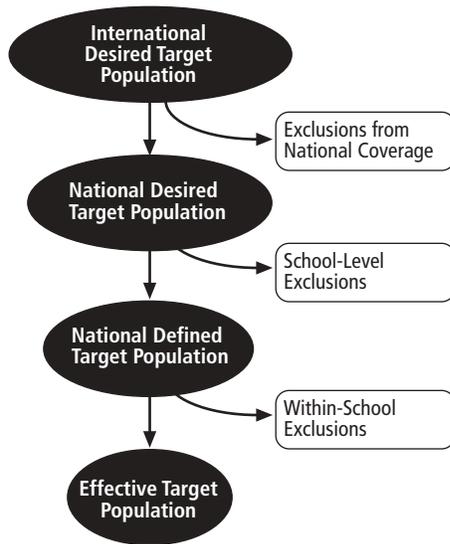
Although countries were expected to do everything possible to maximize coverage of the population by the sampling plan, schools could be excluded, where necessary, from the sampling frame for the following reasons:

- They were in geographically remote regions.
- They were of extremely small 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

**Exhibit 5.1:** Relationship Between the Desired Populations and Exclusions

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. Exhibit 5.1 illustrates the relationship between the desired populations and the excluded populations. 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.

### 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,<sup>1</sup> which may be stratified; the second stage consists of a sample of one or more classrooms from the target grade in sampled schools.

1 In some very large countries, it was necessary to include an extra preliminary stage, where school districts were sampled first, and then schools.

### 5.3.1 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.

### 5.3.2 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.

**Exhibit 5.2: PIRLS Sample-Design Table**

Minimum Cluster Size		Intraclass Correlations								
		0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
5	a	150	169	201	233	265	297	329	361	393
	n	750	845	1 005	1 165	1 325	1 485	1 645	1 805	1 965
10	a	150	150	161	197	233	269	305	341	377
	n	1 500	1 500	1 610	1 970	2 330	2 690	3 050	3 410	3 770
15	a	150	150	150	184	222	259	296	334	371
	n	2 250	2 250	2 250	2 760	3 330	3 885	4 440	5 010	5 565
20	a	150	150	150	178	216	254	292	330	368
	n	3 000	3 000	3 000	3 560	4 320	5 080	5 840	6 660	7 360
25	a	150	150	150	175	213	251	291	328	367
	n	3 750	3 750	3 570	4 375	5 325	6 275	7 250	8 220	9 175
30	a	150	150	150	172	211	250	288	327	366
	n	4 500	4 500	4 500	5 160	6 330	7 500	8 640	9 810	10 980
35	a	150	150	150	170	209	248	287	326	365
	n	5 250	5 250	5 250	2 950	7 315	8 680	10 045	11 410	12 775
40	a	150	150	150	169	208	247	286	325	364
	n	6 000	6 000	6 000	6 760	8 320	9 880	11 440	13 000	14 960
45	a	150	150	150	168	207	246	285	325	364
	n	6 750	6 750	6 750	7 560	9 315	11 070	12 825	14 625	16 380
50	a	150	150	150	167	207	246	285	324	363
	n	7 500	7 500	7 500	8 350	10 350	12 300	14 250	16 200	18 150

a = Number of sampled schools

n = Number of sampled students in target grade

Note: Minimum Cluster Size is number of students selected in each sampled school (generally the average classroom size)

Sample-design tables, such as the one in Exhibit 5.2, 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. For example, on the basis of Exhibit 5.2, a participant whose intra-class correlation was expected to be 0.6, with an average classroom size of 30, would need to sample a minimum of 250 schools. 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 section 5.4.1).

### 5.3.3 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.

### 5.3.4 Replacement Schools

Although PIRLS participants were expected to make great efforts to secure the participation of sampled schools, it was anticipated that a 100 percent participation rate would not be possible in all countries. To avoid sample-size losses, a mechanism was instituted to identify, *a priori*, replacement schools for each sampled school. For each sampled school, the next school on the ordered school sampling frame was identified as its replacement – and the one after that as a second replacement, should it be needed (see Exhibit 5.3 for an example).

The use of implicit stratification variables and the subsequent ordering of the school sampling frame by size ensured that any sampled school's replacement would have similar characteristics. Although this approach does not guarantee avoiding response bias, it tends to minimize the potential for bias, and was deemed more acceptable than over-sampling to accommodate a low response rate.

## 5.4 First Sampling Stage

The sample-selection method used for the first sampling stage in PIRLS made use of a systematic probability-proportional-to-size (PPS) technique. In order to use this method, it was necessary to have some measure of the size (MOS) of the sampling units. Ideally, this was the number of sampling elements within the unit (e.g., the

number of students in the school in the target grade). If this was unavailable, some other highly correlated measure, such as total school enrollment, was used.

The schools in each explicit stratum were listed in order of the implicit stratification variables – together with the MOS for each school. Schools were further sorted by MOS within implicit stratification variables. The measures of sizes were accumulated from school to school, and the running total (the cumulative MOS) was listed next to each

**Exhibit 5.3:** Application of the PPS Systematic Sampling Method to PIRLS

Total MOS: 392 154      Sampling Interval: 2 614.3600  
School Sample: 150      Random Start: 1 135.1551

School Code	School MOS	Cumulative MOS	Sample
939438	532	532	
26825	517	1049	
277618	487	1536	–
228882	461	1997	R1
833389	459	2456	R2
386017	437	2893	
986694	406	3299	
41733	385	3684	
56595	350	4034	–
945801	341	4375	R1
865982	328	4703	R2
700089	311	5014	
656616	299	5313	
647690	275	5588	
381836	266	5854	
510529	247	6101	
729813	215	6316	
294281	195	6511	–
16174	174	6685	R1
292526	152	6837	R2
541397	133	6970	
502014	121	7091	
662598	107	7198	
821732	103	7301	
436600	97	7398	

– = Sampled School

R1, R2 = Replacement Schools

school (see Exhibit 5.3). The cumulative MOS was a measure of the size of the population of sampling elements; dividing it by the number of schools to be sampled gave the sampling interval.

The first school was sampled by choosing a random number in the range between 1 and the sampling interval. The school whose cumulative MOS contained the random number was the sampled school. By adding the sampling interval to that first random number, a second school was identified. This process of consistently adding the sampling interval to the previous selection number resulted in a PPS sample of the required size.

Among the many benefits of this sample-selection method are that it was easy to implement, and that it was easy to verify that it was implemented properly. The latter is critical, since one of PIRLS's main objectives was to ensure that a sound sampling methodology had been used.

Exhibit 5.3 illustrates the PPS systematic sampling method applied to a fictitious sampling frame. The first three sampled schools are shown, as well as their pre-selected replacement schools – should the originally selected schools not participate.

#### 5.4.1 Small Schools

Small schools tend to be problematic in PPS samples because students sampled from them get very large sampling weights, which can increase sampling variance. Also, when the school size falls below the minimum cluster size, it reduces the overall student sample size. In PIRLS, a school was

deemed to be small if the number of students in the target grade was less than the minimum cluster size. For example, if the minimum cluster size was set at 20, then a school with fewer than 20 students in the target grade was considered a small school.

The PIRLS approach for dealing with small schools consisted of two steps:

- **Identifying extremely small schools.** Extremely small schools were defined as schools with fewer students than a quarter of the minimum cluster size. For example, if the minimum cluster size was set at 20, then schools with fewer than five students in the target grade were considered extremely small schools. If student enrollment in these schools was less than 2 percent of the eligible population, then these schools could be excluded – provided the overall inclusion rate met the 95 percent criterion (see section 5.2.1).
- **Creating an explicit stratum of small schools.** If fewer than 10 percent of eligible students were enrolled in small schools, then no additional actions were required. If, however, more than 10 percent of eligible students were enrolled in small schools, then it was necessary to create an explicit stratum for small schools. The number of schools to be sampled from this explicit stratum would remain proportional to the stratum size, but all schools would have an equal probability of selection. This action would ensure greater stability in the resulting sampling weights.

### 5.4.2 Optional Preliminary Sampling Stage

Very large countries had an opportunity to introduce a preliminary sampling stage before sampling schools. The Russian Federation and the United States availed themselves of this option. In these countries, the first step was to draw a sample of geographic regions using PPS sampling. Then a sample of schools was drawn from each sampled region. This design was used mostly as a cost-reduction measure, where the construction of a comprehensive list of schools would have been either impossible or prohibitively expensive. Also, the additional sampling stage reduced the dispersion of the school sample, thereby potentially reducing travel costs. Sampling guidelines were put in place to ensure that an adequate number of units would be sampled from this preliminary stage. The sampling frame had to consist of at least 80 primary sampling units, of which at least 40 had to be sampled at this stage.

## 5.5 Second Sampling Stage

The second sampling stage consisted of selecting classrooms within sampled schools. As a rule, one classroom per school was sampled, although some participants opted to sample two classrooms. All classrooms were selected with equal probabilities for all countries.

### 5.5.1 Small Classrooms

Generally, classrooms in an education system tend to be of roughly equal size. Occasionally, however, small classrooms are devoted to special situations, such as remedial or accelerated programs. These classrooms can become problematic – since they

can lead to a shortfall in sample size – and thus introduce some instability in the resulting sampling weights, when classrooms are selected with PPS.

In order to avoid these problems, it was suggested that any classroom smaller than half the specified minimum cluster size be combined with another classroom from the same grade and school. For example, if the minimum cluster size was set at 30, then any classroom with fewer than 15 students should be combined with another. The resulting pseudo-classroom would then constitute a sampling unit.

## 5.6 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 subsample of the PIRLS sample of schools, most of the required sampling tasks were carried out during the PIRLS school sampling.

**Exhibit 5.4:** Countries Participating in the Trends in IEA's Reading Literacy Study

Country	Primary School Target Grade
Greece	4
Hungary	3
Iceland	4
Italy	4
New Zealand	4
Singapore	3
Slovenia	3
Sweden	3
United States	4

The nine countries that took part in the Trends in IEA's Reading Literacy Study and their target grades are presented in Exhibit 5.4.

### 5.6.1 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 in Hungary, Singapore, and Sweden.

### 5.6.2 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.

### 5.6.3 Replacement Schools

Because schools sampled for the trend study were also sampled for PIRLS, first and second replacement schools for this study are the same ones identified for PIRLS.

There were, however, three exceptions to this rule:

- **A sampled school had only one classroom and agreed to participate in the study.** In this case, the only available classroom in the sampled school was assigned to PIRLS and the PIRLS first replacement school became the trend study sampled school. This left the PIRLS second replacement school as the only 10-year trend study replacement school.

- **A sampled school refused to participate, but the corresponding PIRLS first replacement school agreed to participate and had only one classroom.** In this case, the PIRLS first replacement was used for PIRLS and the PIRLS second replacement school became the trend study first replacement school. In this scenario, there is no second replacement school for the trend study.
- **Both the sampled school and the PIRLS first replacement school refused to participate, but the PIRLS second replacement school agreed to participate and had only one classroom.** In this case, there was no trend study replacement school and the sampled school had a non-participation status.
- Percentages:  $p \pm 7$  percent (where  $p$  is a student-level percentage estimate)
- Correlations:  $r \pm 0.14$  (where  $r$  is a student-level correlation estimate).

The Trends in IEA's Reading Literacy Study focuses mainly on student achievement, but can also report results from schools and classrooms. Based on a minimum sample size of 75 schools, such results should have 95 percent confidence limits for means and percentages in the range  $\pm 23$  percent of their standard deviations.

## 5.7 Sampling Precision and Sample Size

With a single classroom sampled from only half of the 150 schools sampled for PIRLS, the number of students sampled for Trends in IEA's Reading Literacy Study should be roughly half the number of students sampled for PIRLS. This translates into a loss of sampling precision when compared with PIRLS. To get an idea of the resulting standard of sampling precision, the 95 percent confidence limits given earlier in Section 5.3.2 are simply multiplied by  $\sqrt{2}$ . This gives the following 95 percent confidence limits for sample estimates of population means, percentages and correlation coefficients:

- Means:  $m \pm 0.14s$  (where  $m$  is a student mean estimate and  $s$  is its estimated standard deviation for students)

## References

---

IEA. (2001). *W3S: Within-School Sampling Software*. Hamburg: IEA Data Processing Center.

Foy, P., & Joncas, M. (2000). TIMSS Sample Design. In M.O. Martin, K.D. Gregory, & S.E. Stemler. *TIMSS 1999 Technical Report*. Chestnut Hill, MA: Boston College.

Progress in International Reading Literacy Study (PIRLS). (1999). *School Sampling Manual – Version 2* (Doc. Ref.: PIRLS 99-0019). Prepared by Pierre Foy & Marc Joncas, Statistics Canada. Chestnut Hill, MA: Boston College.

Progress in International Reading Literacy Study (PIRLS). (2000). *School Sampling Guide for the 10-Year Trend Study – Version 2* (Doc. Ref.: PIRLS 00-0001). Prepared by Pierre Foy & Marc Joncas, Statistics Canada. Chestnut Hill, MA: Boston College.

Progress in International Reading Literacy Study (PIRLS). (2001a). *School Coordinator Manual – Main Survey* (Doc. Ref.: PIRLS 01-0003). Prepared by the International Study Center. Chestnut Hill, MA: Boston College.

Progress in International Reading Literacy Study (PIRLS). (2001b). *Survey Operations Manual – Main Survey* (Doc. Ref.: PIRLS 01-0001). Prepared by the International Study Center. Chestnut Hill, MA: Boston College.

