where *Y* denotes the set of responding original-sample schools; *N* denotes the set of eligible non-responding original sample schools; *Wi* denotes the base weight for school *i;* *Wi* = 1*/Pi*, where *Pi* denotes the school selection probability for school *i;* and *Ei* denotes the enrollment size of age-eligible students, as indicated on the sampling frame.

In addition to the 349 participating schools from the original sample, 21 substitute schools participated for a total of 370 participating schools at grade 4 in the United States (see table A-2). This gives a weighted (and unweighted) school participation rate after substitution of 85 percent (see table A-1).[10]

Classroom sample. Schools agreeing to participate in PIRLS were asked to list their 4th-grade classes as the basis for sampling at the classroom level. At this time, schools were given the opportunity to identify special classes—classes in which all or most of the students had intellectual or functional disabilities or were non-native-language speakers. While these classes were regarded as eligible, the students as a group were treated as "excluded" since, in the opinion of the school, their disabilities or language capabilities would render meaningless their performance on the assessment. Fifty 4th-grade schools excluded classes and 669 students were excluded from participation in PIRLS as a result.

Prior to sampling, classes with fewer than 15 students were collapsed with other classes into what are called pseudo-classrooms. Creating pseudo-classrooms in this way ensured that all eligible classrooms in a school had at least 20 students. Up to four eligible classrooms were selected, with classes being randomly assigned to TIMSS or PIRLS. In schools with only one classroom, this classroom was selected with certainty and randomly assigned to TIMSS or PIRLS. Some 1,257 classrooms were selected as a result of this process. All selected classrooms participated in PIRLS yielding a classroom response rate of 100 percent (Mullis, et al. 2012, exhibit C.8).

Student sample. Schools were asked to list the students in each of the classrooms. A total of 14,253 students were listed as a result, and 12,726 4th-grade students participated in PIRLS 2011. These students are identified by IEA as "sampled students in participating schools" (see table A-2).

This pool of students is reduced by within-school exclusions and withdrawals. At the time schools listed the students in the sampled classrooms, they had the opportunity to identify particular students who were not suited to take the test because of physical or intellectual disabilities (i.e., students with disabilities who had been mainstreamed) or because they were non-English-language speakers. Schools identified a total of 830 students they wished to have excluded from the assessment; also by the time of the assessment a further 169 of the listed students had withdrawn from the school or classroom. In total, the pool of 14,253 sampled students was reduced by 999 students (830 excluded and 169 withdrawn) to yield 13,254 "eligible" students. The number of eligible students is used as the base for calculating student response rates (Mullis, et al. 2012, exhibit C.6).

The number of eligible students was further reduced on assessment day by 528 student absences, leaving 12,726 "assessed students" identified as having completed a PIRLS 2011 assessment booklet (see table A-2). IEA defines the student response rate as the number of students assessed as a percentage of the number of eligible students which, in this case, yields a weighted (and unweighted) student response rate of 96 percent (see table A-1).

Note that the 669 students excluded because whole classes were excluded do not figure in the calculation of student response rates. They do, however, figure in the calculation of the coverage of the International Target Population. Together, these 669 students excluded prior to classroom sampling, plus the 830 within-class exclusions, resulted in an overall student exclusion rate of 7 percent (see table A-1 and Mullis, et.al. 2012, exhibit C.3). The reported coverage of the International Target Population, then, is 93 percent (see Mullis, et. al. 2012, exhibit C.3).

Combined participation rates. For the results for an education system to be included in the PIRLS international report without a response rate annotation, the IEA requires a "combined" or overall response rate—expressed as the product of (a) the (unrounded) weighted school response rate without substitute schools and (b) the (unrounded) weighted student response rate—of at least 75 percent (after rounding to the nearest whole percent). The overall response rate for the United States, 76.6 percent without substitute schools, meets this requirement. However, the United States did include substitute schools because its school-level response rate was less than 85 percent, and, absent advance knowledge of the student-level response rate, introducing substitute schools was a prudent approach to take. For the results of an education system to be included in the PIRLS international report without a student inclusion annotation, the IEA requires a student inclusion rate of at least 95 percent. Because 7 percent of the 4th-grade student population was excluded in the United States, the overall U.S. student inclusion rate was 93 percent. For this reason, the U.S. 4th-grade results in the PIRLS international report carry a coverage annotation indicating that coverage of the defined student population was less than the IEA standard of 95 percent.

Tables A-1 and A-2 are extracts from the international report exhibits noted above and are designed to summarize information on school and student responses rates and coverage of the target populations in each nation.

[10]Substitute schools are matched pairs and do not have an independent probability of selection. NCES standards (Standard 1-3-8) indicate that, in these circumstances, response rates should be calculated without including substitute schools (National Center for Education Statistics 2002). PIRLS response rates denoted as "before replacement" conform to this standard. PIRLS response rates denoted as "after replacement" are not consistent with NCES standards since, in the calculation of these rates, substitute schools are treated as the equivalent of sampled schools.

Table A-1. Coverage of target populations, school participation rates, and student response rates, by education system: 2011

Education system	Percentage of international desired population coverage	National desired population overall exclusion rate	Weighted school participation rate before substitution	Weighted school participation rate after substitution	Weighted student response rate	Combined weighted school participation and student response rate
Australia	100	4	96	98	95	93
Austria	100	5	100	100	98	98
Azerbaijan	100	7	84	100	100	100
Belgium (French)-BEL	100	6	77	85	97	82
Bulgaria	100	3	97	100	95	95
Canada	100	10	98	98	96	94
Chinese Taipei-CHN	100	1	100	100	99	99
Colombia	100	2	89	99	97	95
Croatia	100	8	99	100	95	95
Czech Republic	100	5	90	99	94	94
Denmark	100	7	87	98	97	95
England-GBR	100	2	73	87	94	82
Finland	100	3	97	99	96	95
France	100	5	98	100	98	97
Georgia	92	5	97	98	98	96
Germany	100	2	96	99	96	95
Hong Kong-CHN	100	12	86	88	94	83
Hungary	100	4	98	99	97	96
Indonesia	100	3	100	100	97	97
Iran, Islamic Rep. Of	100	5	100	100	99	99
Ireland	100	3	98	100	95	95
Israel	100	25	98	99	94	93
Italy	100	4	81	98	96	95
Lithuania	93	6	94	100	94	94
Malta	100	4	100	100	95	95
Morocco	100	2	99	99	96	95
Netherlands	100	4	68	92	97	89
New Zealand	100	3	93	99	94	93
Northern Ireland-GBR	100	4	62	85	93	79
Norway	100	4	57	83	86	71
Oman	100	2	98	98	98	96
Poland	100	4	100	100	96	96
Portugal	100	3	87	99	95	93
Qatar	100	6	100	100	99	99
Romania	100	4	99	100	97	97
Russian Federation	100	5	100	100	98	98
Saudi Arabia	100	2	95	100	98	98
Singapore	100	6	100	100	96	96
Slovak Republic	100	5	95	99	97	96
Slovenia	100	3	96	97	97	95
Spain	100	5	96	99	97	96
Sweden	100	4	97	99	92	91
Trinidad and Tobago	100	1	99	99	96	95
United Arab Emirates	100	3	100	100	97	97
United States	100	7	80	85	96	81

See notes at end of table.

Table A-1. Coverage of target populations, school participation rates, and student response rates, by education system: 2011—Continued

Benchmarking education systems	Percentage of international desired population coverage	National desired population overall exclusion rate	Weighted school participation rate before substitution	Weighted school participation rate after substitution	Weighted student response rate	Combined weighted school participation and student response rate
Alberta-CAN	100	7	97	99	95	94
Ontario-CAN	100	8	99	99	96	95
Quebec-CAN	100	4	95	96	96	92
Maltese-MLT	100	4	100	100	94	94
Andalusia-ESP	100	5	99	99	97	96
Abu Dhabi-UAE	100	3	99	99	97	96
Dubai-UAE	100	5	99	99	96	94
Florida-USA	89	13	96	96	95	91

NOTE: Education systems in the Southern hemisphere administered PIRLS 2011 in the fall of 2010 while those in the Northern hemisphere administered the assessment in the spring of 2011. Italics indicate participants identified and counted in this report as an education system and not as a separate country. The international desired population refers to the sample and not the responding schools, classes, and students.
SOURCE: International Association for the Evaluation of Educational Achievement (IEA), Progress in International Reading Literacy Study (PIRLS), 2011.

Table A-2. Total number of schools and students, by education system: 2011

Education system	Schools in original sample	Eligible schools in original sample	Schools in original sample that participated	Substitute schools	Total schools that participated	Sampled students in participating schools	Students assessed
Australia	290	284	275	5	280	6,709	6,126
Austria	160	158	158	0	158	4,976	4,670
Azerbaijan	170	169	142	27	169	5,098	4,881
Belgium (French)-BEL	150	150	115	12	127	3,910	3,727
Bulgaria	150	147	142	5	147	5,725	5,261
Canada	1,142	1,125	1,106	5	1,111	25,707	23,206
Chinese Taipei-CHN	150	150	150	0	150	4,376	4,293
Colombia	157	152	131	19	150	4,309	3,966
Croatia	152	152	150	2	152	5,097	4,587
Czech Republic	180	178	161	16	177	4,895	4,556
Denmark	240	236	207	25	232	4,994	4,594
England-GBR	150	148	109	20	129	4,243	3,927
Finland	150	146	141	4	145	4,914	4,640
France	175	175	170	4	174	4,638	4,438
Georgia	180	177	172	1	173	4,958	4,796
Germany	200	199	190	7	197	4,229	4,000
Hong Kong-CHN	154	150	130	2	132	4,189	3,875
Hungary	150	150	146	3	149	5,488	5,204
Indonesia	158	158	158	0	158	5,049	4,791
Iran, Islamic Rep. Of	250	244	244	0	244	5,932	5,758
Ireland	152	151	148	3	151	4,849	4,524
Israel	153	153	150	2	152	4,579	4,186
Italy	205	205	166	36	202	4,529	4,189
Lithuania	160	154	145	9	154	5,140	4,661
Malta	99	96	96	0	96	3,958	3,598
Morocco	289	287	284	0	284	8,381	7,805
Netherlands	151	151	97	41	138	4,179	3,995
New Zealand	201	199	180	12	192	6,192	5,644
Northern Ireland-GBR	160	160	100	36	136	3,942	3,586
Norway	150	145	85	35	120	3,921	3,190
Oman	338	333	327	0	327	10,840	10,394
Poland	150	150	150	0	150	5,316	5,005
Portugal	150	150	133	15	148	4,428	4,085
Qatar	175	167	166	0	166	4,394	4,120
Romania	150	148	147	1	148	4,879	4,665
Russian Federation	202	202	202	0	202	4,693	4,461
Saudi Arabia	175	171	163	8	171	4,625	4,507
Singapore	176	176	176	0	176	6,687	6,367
Slovak Republic	200	198	187	10	197	5,933	5,630
Slovenia	202	201	193	2	195	4,674	4,512
Spain	314	314	308	4	312	9,223	8,580
Sweden	161	153	148	4	152	5,209	4,622
Trinidad and Tobago	150	150	149	0	149	4,190	3,948
United Arab Emirates	478	460	458	0	458	15,372	14,618
United States	450	437	349	21	370	14,253	12,726

See notes at end of table.

Table A-2. Total number of schools and students, by education system: 2011—Continued

Benchmarking education systems	Schools in original sample	Eligible schools in original sample	Schools in original sample that participated	Substitute schools	Total schools that participated	Sampled students in participating schools	Students assessed
Alberta-CAN	150	147	143	2	145	4,292	3,789
Ontario-CAN	200	191	188	1	189	4,932	4,561
Quebec-CAN	200	197	189	1	190	4,529	4,244
Maltese-MLT	99	95	95	0	95	3,942	3,548
Andalusia-ESP	150	150	149	0	149	4,652	4,333
Abu Dhabi-UAE	168	165	164	0	164	4,308	4,146
Dubai-UAE	152	139	138	0	138	6,497	6,061
Florida-USA	81	80	77	0	77	3,052	2,598

NOTE: Education systems in the Southern hemisphere administered PIRLS 2011 in the fall of 2010, while those in the Northern hemisphere administered the assessment in the spring of 2011. Italics indicate participants identified and counted in this report as an education system and not as a separate country.
SOURCE: International Association for the Evaluation of Educational Achievement (IEA), Progress in International Reading Literacy Study (PIRLS), 2011.

Nonresponse bias in the U.S. PIRLS sample

NCES standards require a nonresponse bias analysis when the response rate of any sampled unit falls below 85 percent (Standard 2-2-2, NCES Statistical Standards, 2002). Because the response rate for U.S. schools was below 85 percent, a nonresponse bias analysis was conducted. It took a form similar to that adopted for TIMSS 2003 (Ferraro and Van de Kerckhove 2006). A full report of this study will be included in a technical report to be released with the U.S. national PIRLS dataset. The response rate in Florida was sufficiently high so that a nonresponse bias analysis was not required.

Three methods were chosen to perform this analysis. The first method focused exclusively on the *sampled schools* and ignored substitute schools. The schools were weighted by their school base weights, excluding any nonresponse adjustment factor. The second method focused on *sampled schools plus substitute schools*, treating as nonrespondents those schools from which a final response was not received from the original or substitute school. Again, schools were weighted by their base weights, with the base weight for each substitute school set to the base weight of the original school that it replaced. The third method repeated the analyses from the second method using nonresponse adjusted weights.[11]

In order to compare PIRLS respondents and nonrespondents, it was necessary to match the sample of schools back to the sample frame to identify as many characteristics as possible that might provide information about the presence of nonresponse bias.[12] The characteristics available for analysis in the sampling frame were taken from the CCD for public schools, and from the PSS for private schools. For categorical variables, the distribution of the characteristics for respondents was compared with the distribution for all schools. The hypothesis of independence between a given school characteristic and the response status (whether or not the school participated) was tested using a Rao-Scott modified chi-square statistic. For continuous variables, summary means were calculated and the difference between means was tested using a *t* test. Note that this procedure took account of the fact that the two samples in question were not independent samples, but in fact the responding sample was a subsample of the full sample. This effect was accounted for in calculating the standard error of the difference. Note also that in those cases where both samples were weighted using just the base weights, the test is exactly equivalent to testing that the mean of the respondents was equal to the mean of the nonrespondents.

In addition, multivariate logistic regression models were set up to identify whether any of the school characteristics were significant in predicting response status when the effects of all potential influences were considered simultaneously.

Public and private schools were modeled together using the following variables:[13] community level (central city, urban fringe/large town, rural/small town); control of school (public or private); census region (Northeast, Southeast, Central, West); poverty level (percentage of students in school eligible for free or reduced-price lunch);[14] number of students enrolled in grade 4; total number of students; and percentage minority students.[15]

Results for the original sample of schools. In the analyses for the original sample of schools, all substituted schools were treated as nonresponding schools. The results of these analyses follow.

In the investigation into nonresponse bias at the school level for PIRLS 4th-grade schools, comparisons between schools in the eligible sample and participating schools showed that there was no relationship between response status for eight of the twelve school characteristics available for analysis. In the original sample, a separate variable-by-variable bivariate analyses identified four variables that were found to be statistically significantly predictors of response status related to: school control, community level, 4th-grade enrollment, and students eligible for free or reduced-price lunch. When all school-level factors were considered simultaneously in a regression analysis, four variables were found to be statistically significant predictors of response status: private schools, high poverty, total school enrollment, and 4th-grade enrollment. The second method focused on sampled schools plus substitute schools, treating as nonrespondents those schools from which a final response was not received from the original or substitute school. This model (using as a predictor percent minority rather than percent in various race/ethnicity

[11]A detailed treatment of the meaning and calculation of sampling weights, including the nonresponse adjustment factors, is provided in the *TIMSS and PIRLS Methods and Procedures* (Martin and Mullis 2011).

[12]Comparing characteristics for respondents and nonrespondents is not always a good measure of nonresponse bias if the characteristics are either unrelated or weakly related to more substantive items in the survey. Nevertheless, this is often the only approach available.

[13]NAEP region and community level were dummy coded for the purposes of these analyses. In the case of NAEP region, "West" was used as the reference group. For community level, "urban fringe/large town" was chosen as the reference group.

[14]The measure of school poverty is based on the proportion of students in a school eligible for the free or reduced-price lunch (FRPL) program, a federally assisted meal program that provides nutritionally balanced, low-cost or free lunches to eligible children each school day. For the purposes of the nonresponse bias analyses, schools were classified as "low poverty" if less than 50 percent of the students were eligible for FRPL, and "high poverty" if 50 percent or more of students were eligible. Since the nonresponse bias analyses involve both participating and nonparticipating schools, they are based, out of necessity, on data from the sampling frame. PIRLS data are not available for nonparticipating schools. The school frame data are derived from the CCD and PSS. The CCD data provide information on the percentage of students in each school who are eligible for free or reduced-price lunch, but are limited to public schools. The PSS data do not provide the same information for private schools. In the interest of retaining all of the schools and students in these analyses, private schools were assumed to be low-poverty schools–that is, they were assumed to be schools in which less than 50 percent of students were eligible for FRPL.

[15]Two forms of this school attribute were used in the analyses. In the bivariate analyses the percentage of each race/ethnic group was related separately to participation status. In the logistic regression analyses a single measure was used to characterize each school, namely, "percentage of minority students."