Age distribution and reading achievement configurations among fourth-grade students in PIRLS 2006

Michael O. Martin, Ina V. S. Mullis, and Pierre Foy
TIMSS & PIRLS International Study Center, Boston College, Chestnut Hill, Massachusetts, USA

This study used the Progress in International Reading Literacy Study (PIRLS) 2006 reading achievement results for fourth-grade students in 36 countries to show how variations in countries' policies on age of school entry and promotion and retention make it difficult to determine a straightforward relationship between age-within-grade and achievement across countries. Although, in some countries, older Grade 4 students had higher achievement than younger students, older students did not necessarily perform better. Many countries had a substantial proportion of older students in the fourth grade (students whose age suggests they belong in a higher grade), and these older students had significantly lower achievement than their younger classmates. The different configurations of age-within-grade across countries make it problematic to statistically adjust countries' average reading achievement for differences in age. Thus, comparisons across countries in how chronological age and amount of schooling influence achievement must be made with care.

IERI Monograph Series: Issues and Methodologies in Large-Scale Assessments Volume 4

INTRODUCTION

The International Association for the Evaluation of Educational Achievement (IEA)'s assessment of student achievement in mathematics and science (TIMSS) and assessment of reading achievement (PIRLS) are grade-based assessments, which means that the target population is all the students at a particular grade level—four and eight for TIMSS and four for PIRLS. IEA's focus on grade as the basis for its target populations contrasts with the Organisation for Economic Co-operation and Development (OECD)'s Programme for International Student Assessment (PISA) survey, which assesses mathematics literacy, science literacy, and reading literacy among 15-year-olds (OECD, 2001).

IEA considers that, for its major purpose of improving educational achievement, amount of schooling provides a more useful basis than age for interpreting and using the results of international comparisons. Although the PISA approach of assessing students by age provides an absolutely clear basis for defining the samples of students to be assessed, TIMSS and PIRLS assess student achievement by grade because this maximizes the opportunity to link educational achievement to policies, curriculum, and instructional practices. Since most educational initiatives are implemented at particular grade levels, TIMSS and PIRLS data can readily be used as a basis for implementing and monitoring educational reform and improving student achievement. Also, because mathematics, science, and (perhaps to a lesser extent) reading are curriculum dependent and learned largely in school, IEA views amount of instruction as the major determinant of educational achievement.

A number of studies during the past several decades have investigated the effects of chronological age and amount of schooling on achievement, with most showing that the effects of schooling are greater than the effects of age. For example, Cliffordson and Gustafsson (2007) used regression analyses of simultaneous variation in age and length of schooling to investigate the effects of age and amount of schooling on Swedish military enlistees' test scores. After reviewing the enlistment scores of an entire cohort of 18-year-old males, the authors concluded that both age and schooling increased performance on two aspects of student ability—crystallized intelligence and general visualization—and that the effects of amount of schooling were considerably stronger than the effects of age.

Recently, interest in using TIMSS and PIRLS data to research the effects of age and amount of schooling has increased. However, using grade (determined by years of formal schooling) as the basis for comparison makes variation in the age of the students an important consideration. We discuss, in the next section, several attempts to model variation in age or to adjust for it, but caution that if analytical approaches such as these are to be more widely used, there needs to be a clear understanding of how policies and practices influencing assignment to grade relate to age and how this association varies from country to country. Within a single grade level, the TIMSS and PIRLS demographic data show that average student age can vary across countries by as much as one year (Mullis, Martin, Foy, et al., 2008; Mullis, Martin, Kennedy, & Foy, 2007). In general, this is because students' age of entry to primary school varies (typically, children enter school at six to seven years of age), and because promotion and retention policies also vary. Also, the practices of parents can be flexible; sometimes they accord with the formal policies and sometimes not. Because it is difficult to assemble information describing variations in patterns of age-within-grade across countries as well as providing explanations for the differences, such information unfortunately is not readily available.

The central purpose of this paper is to describe the distribution of student age within the fourth grade in the countries that participated in PIRLS 2006 and to examine how that distribution related to achievement. We used the PIRLS 2006 international database (Foy & Kennedy, 2008) containing reading achievement results for fourth-grade students in 40 countries[1] in this study in order to examine the distributions of fourth-graders by birth month. We considered age distributions in the light of the participating countries' reports of their policies and practices regarding ages of entry to primary school and policies on promotion and retention.[2] We augmented this information with more recent material from the *TIMSS 2007 Encyclopedia* (Mullis, Martin, Olson, et al., 2008). Finally, we looked at the relationships between various patterns of distribution by birth month and reading achievement in several of the PIRLS 2006 countries.

RESEARCH USING INTERNATIONAL DATA TO INVESTIGATE THE EFFECTS OF AGE AND AMOUNT OF SCHOOLING

In one notable attempt to examine the interrelationships among age, grade, and achievement, Luyten (2006) applied a regression discontinuity approach in a multilevel modeling context using data from TIMSS 1995. The regression discontinuity approach is a quasi-experimental analytic technique that can be applied to achievement data from adjacent grades to estimate the effect of schooling.[3] Luyten explained that this approach, sometimes referred to as a cutting-point design, had been applied previously to estimate the independent effects of age and schooling on achievement in a 1973 study of primary education in Israel. In the initial study, Cahan and Cohen (1989) found that the effect of one year of schooling was substantial, having as much as twice the effect of one year of age.

The drawback of the regression discontinuity approach is that it requires achievement data from countries where admission to school depends primarily on a student's date of birth. If the country has a strict policy of admitting students to school on a particular date when they are at a particular age (e.g., all children who have turned six by January 1), the data from two consecutive grades can be used to assess the absolute

1 As described in a later section, this paper is restricted to data from 36 countries, with the Flemish and French parts of Belgium seen as separate entities.

2 These reports were assembled in the *PIRLS 2006 Encyclopedia* (Kennedy, Mullis, Martin, & Trong, 2007).

3 The present paper includes an application of the regression discontinuity design to data from two countries, Iceland and Norway, both of which participated in PIRLS 2006 with two adjacent grades.

contribution of schooling to students' achievement. While the regular increase in age by birth month for the students can be related to achievement for each grade, we can anticipate a discontinuity between the oldest students in the lower grade and the youngest students in the higher grade. That is, the students beginning the next higher grade will be not be much older (born in January rather than in the previous December), but they will have had one more year of school. Based on the 26 countries participating in TIMSS 1995 at the primary-school level, Luyten (2006) found eight countries that met his criteria of having a sharp cutoff date, namely, Cyprus, Greece, England, Iceland, Japan, Norway, Scotland, and Singapore. The analysis of the TIMSS 1995 mathematics and science data from Grades 3 and 4 in these countries showed strong schooling effects on achievement.

In another attempt to control for age, Luyten, Peschar, and Coe (2008) used data from PISA 2000 to analyze the effects of age and schooling on the reading literacy of 15-year-olds in England. Capitalizing on the fact that children in England begin school strictly in accordance with their date of birth and that approximately half the national sample of 15-year-olds was in Grade 10 and half in Grade 11, these authors employed a regression discontinuity approach within a multilevel modeling framework to examine the effect of an extra grade of schooling while controlling for students' age. Luyten and his colleagues found a modest positive effect of schooling on reading performance, even though the PISA reading assessment was not designed to reflect a specific school curriculum. Also, we might expect that students typically would have mastered their reading skills by Grade 10 and thus have less room for growth than would have been the case earlier in their schooling careers. Yet, the authors noted, only one third of the extra grade effect could be accounted for by age differences, a finding which implies that "most of the advantage for the students in the upper grade must be due to the effect of schooling" (Luyten et al., 2008, p. 336).

Recognizing that policies on age of school entry and promotion and retention result in variation in student age, Cliffordson (2008) also employed a between-grade regression discontinuity design to analyze TIMSS 1995 data from Sweden; her focus was on Grades 6, 7, and 8. Although this design provides a powerful technique for analyzing adjacent grade data, such as those in TIMSS 1995, the author acknowledged that when the design is used to separate the effects of age and grade, it "relies on the assumptions that there is a sharp age-based decision rule for grade assignment, and that the regression of performance on age is linear" (p. 3). Accordingly, one of Cliffordson's major purposes for conducting the study was to assess the extent to which these assumptions were tenable for the Swedish TIMSS data. Cliffordson's results for her analyses comparing the achievement of sixth- and seventh-grade students in mathematics and science were broadly in line with the studies described above, with the effect of schooling about twice as strong as the effect of age. Analyses based on the data for the seventh- and eighth-grade students showed a weaker schooling effect, particularly for science, although it was still stronger than the age effect.

Cliffordson's (2008) analyses also revealed that the bias in estimating age and school effects due to departures from a sharp age-based decision rule for grade assignment was relatively small in the Swedish data. However, she noted that this outcome was probably because the percentage of "normal-aged" students (students in the appropriate grade for their age) was very high in the Swedish sample. Departures from the grade assignment rule affected only about 3.5% of students, and so the potential for bias was limited. She cautioned that her results did not imply that bias would remain low if the proportion of students not of normal age for their cohort was excessive.

The studies described above show that it is possible to use quasi-experimental techniques such as regression discontinuity analysis with adjacent grade data to estimate the effect of age-within-grade on student achievement, provided that students are assigned to grades solely on the basis of age and that no other factors related to achievement are involved. For example, Van Damme, Vanhee, and Pustjens (2008), generalizing from previous work, statistically adjusted average reading achievement to control for differences in average age among the countries that participated in PIRLS 2006. They applied the same correction factor to all countries, regardless of the relationship between age and reading achievement evident in each. However, it is fundamental to recognize that students' age-within-grade can be manipulated by policy decisions and, in particular, that policies on age of entry to school and policies on promotion/retention can influence not just the average age but also the age distribution within a grade.

Conscious of the need for an analytic approach to estimating the effects of schooling that makes less restrictive assumptions about the role of age, Luyten and Veldkamp (2008) applied a two-step procedure suggested by Heckman (1979) to reanalyze the TIMSS 1995 data for Grades 3 and 4. This procedure, as a first step, explicitly models the role of age in assigning students to a grade and then, as a second step, includes a correction factor derived from the model in a regression equation. The intention behind including the correction factor is to control for any selection bias caused by factors other than age influencing the grade assignment process.

Luyten and Veldkamp (2008) found it possible to model the grade-assignment process adequately in 15 of the 26 countries that participated in TIMSS 1995 at the primary level, and to apply the Heckman procedure successfully in these countries. On average, across the countries, the authors estimated that schooling accounted for 51% of the mathematics achievement difference between third and fourth grades and for 45% of the science difference. The effect sizes varied widely across countries, from 29% in England to 72% in the Netherlands for mathematics and from 14% in Canada to 69% in Korea for science.

Both Cliffordson (2008) and Luyten and Veldkamp (2008) acknowledged, when reporting the results of their studies, that anyone wanting to use adjacent grades to estimate the effect of schooling across countries needs to understand how students' ages are used to assign students to grades. Taking the above methodology a step further, Cliffordson and Gustafsson (2010) reanalyzed the Swedish TIMSS 1995 data using an instrumental variables regression approach that relaxed the assumption that there is a sharp age-based decision rule for grade assignment. Finding the results for this one country encouraging, the authors recommended that the method be further investigated with data from other countries. However, because the instrumental variables approach is also based on creating an instrumental variable that adequately models the age−grade assignment process, it too requires an understanding of how policies and practice with regard to age of entry to school and grade promotion/retention influence the distribution of age-within-grade.

Because the 40 countries that participated in PIRLS 2006 represented a wide range of policy approaches, we considered that an analysis of the age distribution within grade in terms of the different entry and promotion and retention policies of these countries could be used to investigate the relationship between these policies and the distribution of students' age-within-grade. Educational researchers can use this information to inform attempts to model the grade assignment process more effectively, while educational decisionmakers can use it to understand the effects on achievement of students starting school at younger or older ages, as well as the likely effects on achievement of various promotion and retention policies.

DISTRIBUTION OF STUDENT AGES IN THE PIRLS 2006 TARGET GRADE

As described in the *PIRLS 2006 Encyclopedia* (Kennedy et al., 2007), countries' admission policies generally require children to begin primary school when they are six or seven years old. As already discussed, a number of the studies about the effects of age and schooling on achievement have used the regression discontinuity model, which requires data from a country with a strict adherence to a nationwide cutoff point that determines if a student is in a higher or a lower grade. If we assume a uniform birth rate across the months of a year,[4] an age-of-entry policy based on students' dates of birth would result in an approximately equal proportion of students in each birth month. Thus, a full cohort of students would contain approximately equal percentages of students born in January, February, March, and so on. Because there are 12 months in a year, approximately eight percent of students would be evident in each birth month. We could expect to see this pattern in countries with strict policies on age of entry and automatic promotion from grade to grade.

How countries implement policies about age of entry into school varies from country to country, thereby contributing to the variety of age distributions across those countries. Countries with more flexible practices with respect to age of entry can have different age distributions within a grade, as can countries with promotion policies based on examination results, or retention or acceleration practices based on school or teacher recommendations.

4 Although the odds of being born in a particular month vary somewhat across the 12 months of the year, fluctuations from month to month are relatively small and the odds are sufficiently uniform for the purposes of this argument.

Figure 1 shows the distribution of students' ages, on average, as well as by birth month across the PIRLS 2006 countries that assessed students in their fourth year of formal schooling and that collected data according to the Northern Hemisphere schedule of April to May, 2005.[5] The figure presents the countries in decreasing order by predominant age cohort; countries where students were older, on average, are at the top of the figure, and those where students were younger are at the bottom. The bar chart accompanying each country depicts the percentage of students at each month of birth; the 12 adjacent months that include the greatest percentage of students are highlighted to identify the predominant age cohort for the grade (i.e., the 12 adjacent months that define the greatest percentage of students). This approach not only groups countries with similar age distributions but also shows the cumulative effects of school admission policies and promotion and retention practices through the grades on the distribution of student ages at the fourth grade. In general, few countries had all of their students within this 12-month interval; most had more widespread distributions, in some cases spanning several years.

The countries in the first grouping in Figure 1 require students to begin primary school in the calendar year in which they turn seven. Thus, the students assessed in PIRLS 2006 predominantly were born in the 1995 calendar year. The countries included in this first group were Denmark, Bulgaria, Latvia, Lithuania, Moldova, Romania, the Russian Federation, and Sweden. Students in these countries were among the oldest in the PIRLS 2006 assessment—approximately 10.8 years of age, on average.

Among this group of countries, Sweden was the country where the predominant age cohort included practically all students at fourth grade; there were very few younger or older students. This pattern indicates a relatively strict adherence to the nationwide cutoff dates, as discussed by the researchers who used the Swedish data from TIMSS 1995 (see above). Also, Sweden has a policy of automatic promotion from grade to grade. However, the other countries in this group had noticeable percentages of older and, in some cases, younger students. Among these, Latvia, Lithuania, Romania, and the Russian Federation reported that promotion from grade to grade depended on students' progress; retention was thus a contributing factor to having older students.

The set of countries in Figure 1 in which the predominant cohort was born in 1995 is followed by several countries where the predominant cohort was not confined to a single calendar year but spanned two years. In Hungary, where children normally should be six years old by May 31 in order to begin school in September, the predominant cohort extended from March 1995 through February 1996, as also was the case for Macedonia. Both countries also had substantial percentages of older and younger students in the fourth grade, although only Hungary reported basing its promotion policy on students' progress.

5 PIRLS 2006 also included two countries that tested at Grade 5 (South Africa and Luxembourg), three countries where the school year is conducted on a Southern Hemisphere schedule (Singapore and New Zealand, in addition to South Africa), and five Canadian provinces. However, to simplify the basis for comparing age distributions across countries, we have not included these participants in Figures 1, 2, and 3.

Figure 1: PIRLS 2006 percentage of students by month of birth

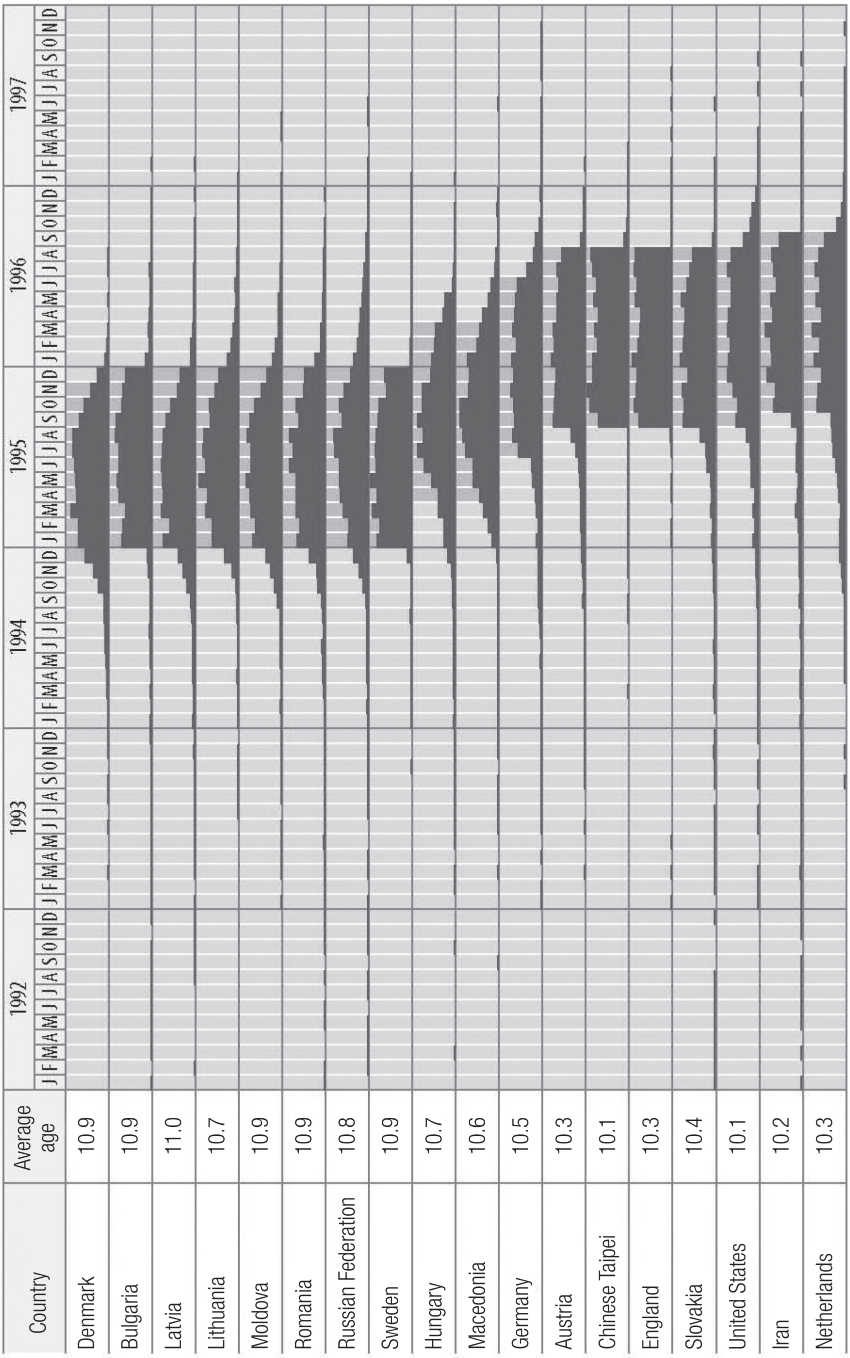
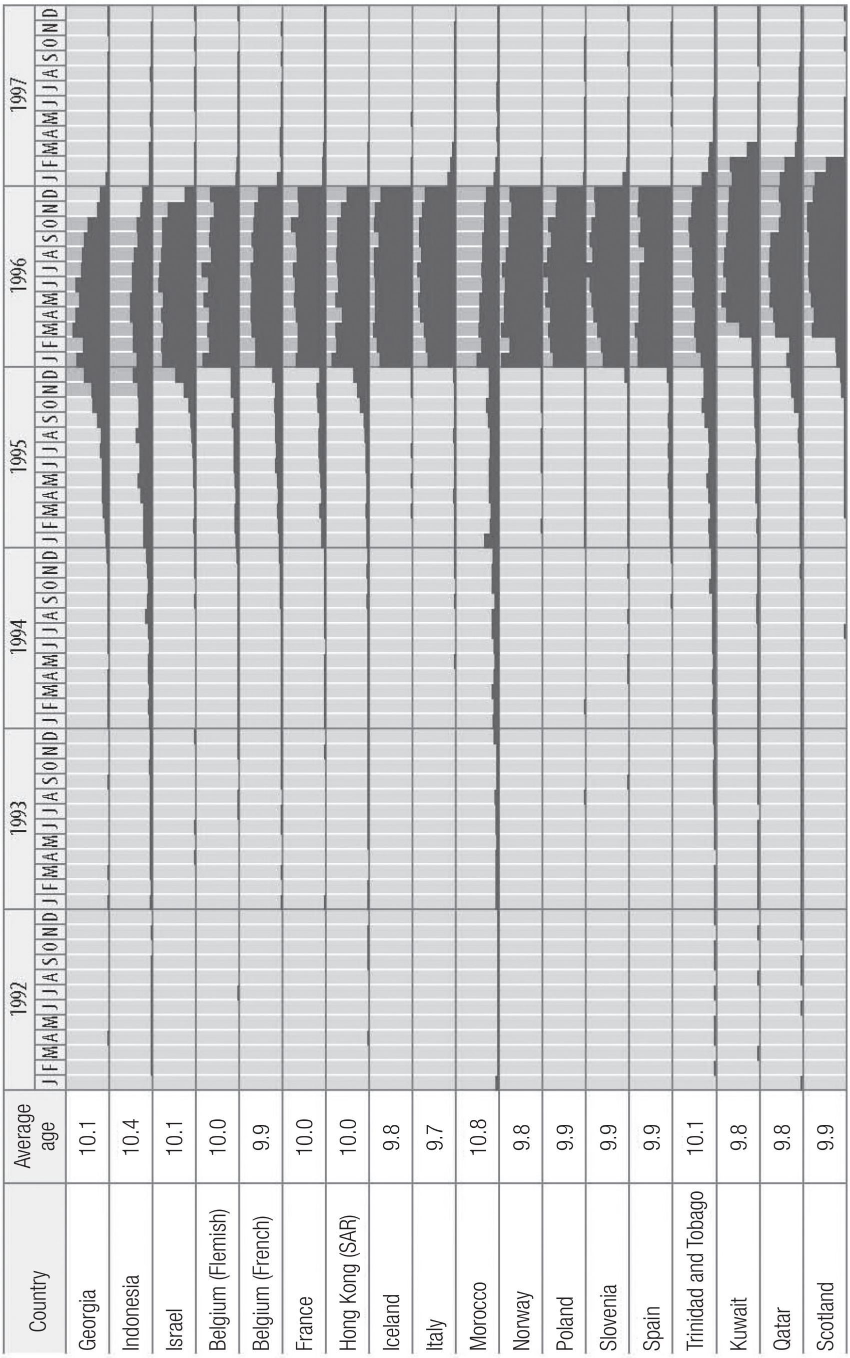


Figure 1: PIRLS 2006 percentage of students by month of birth (Contd.)



Children in Germany must be six years old by the end of June 30 to begin school in September or, upon special request, by December 31 of that year. Germany's predominant age cohort reflected this policy, spanning July 1995 through June 1996, but it also had some older and younger students in Grade 4. In Germany, policies about promotion and retention vary by *Länder* (federal state).

Next, in Figure 1, comes a group of countries where the practice is to admit children who have reached their sixth birthday by the beginning of September, when the school year begins. In these countries, which include Austria, Chinese Taipei, England, the Slovak Republic, and the United States, the predominant age cohort extended from September 1995 through to August 1996. In Austria, Chinese Taipei, and the Slovak Republic, children must be six years old on September 1 to begin primary school. Children in England must begin school at the start of the term following their fifth birthday. Age-of-entry policy and practice in the United States varies from state to state, but judging by the country's age distribution, it is common practice to admit children to school on the basis of their age at the beginning of September. Almost all of the Grade 4 students in Chinese Taipei and England were included in the predominant age cohort, reflecting strict adherence to the nationwide entrance policies and automatic promotion from grade to grade. This was less the case in Austria, the Slovak Republic, and the United States, which have, to varying degrees, promotion policies that depend on students' progress and can thus result in relatively large percentages of older students.

Figure 1 next shows Iran, where children must be six years old by September 20 to begin school, and the Netherlands, where children begin school at age six. Although their predominant age cohorts extended from October 1995 through September 1996, Iran and the Netherlands had many older students. We could have anticipated this pattern for Iran because, for each grade, students must pass an examination to be promoted. In the Netherlands, the decision is left to the schools. In Georgia and Indonesia, the predominant age cohort spanned November 1995 through October 1996, although there were also many older and younger students. The policy in Georgia is that students must have turned six by the end of December in order to enter school, while in Indonesia students can begin school at six years of age, but *must* enter by seven years of age. In Israel, where students should be age six by the beginning of the school year (which varies somewhat), the predominant age cohort extended from December 1995 through November 1996, and contained some older students and a few younger ones.

For the PIRLS 2006 countries, the most common age-of-entry practice was to admit children to primary school in the calendar year in which they turned six; for these 11 countries, the predominant age cohort coincided with the 1996 calendar year. The 11 countries included both the Flemish and French parts of Belgium, as well as France, Hong Kong SAR, Iceland, Italy, Morocco, Norway, Poland, Slovenia, and Spain. There were some slight deviations from the calendar-year policy in Hong Kong SAR, where students must be 5.8 years old by September, and in Italy, which also has an examination for early admission (perhaps explaining its somewhat younger students, on average).

Six countries in this group—Iceland, Italy, Norway, Poland, Slovenia, and Spain—had almost all of their Grade 4 students in the predominant age group. These countries reported automatic promotion, except in Slovenia, where retention can begin in the fourth grade, and in Spain, where students need to demonstrate basic competencies to be promoted. In Morocco, the distribution of students spanned a very wide range of birth months, and students were considerably older, on average, than in the other countries in this grouping. Although Morocco reported a policy of automatic promotion, it is a country striving to increase school enrollment at all levels, despite facing many economic challenges. With the exception of Morocco, and taking into account the pattern whereby fourth-graders in the countries where students enter school during the calendar year in which they turn six are typically one year younger when they enter school than are the Grade 4 students in the first group of countries shown in Figure 1 (i.e., students entering school in the calendar year in which they turn seven), it follows that students in this group of countries were among the youngest students assessed in PIRLS 2006. They were, on average, 9.9 years of age.

With respect to the remaining several countries in Figure 1, students in Trinidad and Tobago can start school when they are five years old, but the PIRLS data for this country showed quite a wide distribution of ages by birth month. Trinidad and Tobago also has a struggling economy, making it difficult to enforce educational policies. Students in Kuwait can start school if they are 5.5 years old by September, and those in Scotland can begin school between the ages of 4.5 and 6 depending on where their birth month falls in relation to the beginning of the school year in August. The students in Kuwait, Qatar, and Scotland, where the predominant age cohort extended from March 1996 through February 1997, were among the youngest of the PIRLS 2006 participants.

The information presented in Figure 1 makes apparent the considerable variation across countries in how they assign students to grades. This variation differentially influences the relationship between age and grade from country to country, making the task of comparing the effects of age and amount of schooling on achievement an extremely complicated one. In the majority of the PIRLS 2006 countries, substantial percentages of the fourth-grade students were not what Cliffordson (2007) referred to as "normal-aged" for their grade. According to Luyten (2006), application of the regression discontinuity approach presents no significant problems as long as the percentage of older and the percentage of younger students do not exceed five. However, that criterion was not met by many of the PIRLS 2006 countries.

Figure 2 presents, for each PIRLS 2006 country, the percentage of fourth-grade students included in the predominant age cohort, together with the percentage of older and younger students. The figure is organized in decreasing order by percentage of students in the predominant age cohort. The PIRLS 2006 countries in the upper part of the figure have all or practically all of their students included in a single

12-month period (e.g., from January to December or September to August), while those in the lower part have substantial percentages of older students and also some younger students.

As shown in Figure 2, eight PIRLS 2006 countries had almost all of their fourth-grade students in the predominant age cohort—95% or more in accordance with Luyten's aforementioned criterion. These countries included Iceland (100%), Norway, and England (99%), Poland (98%), Chinese Taipei (97%), Slovenia (96%), and Sweden (95%). In each case, the practice for age of entry was to admit an entire age cohort (i.e., students born in a single 12-month period) and to have automatic promotion from grade to grade, at least in the grades preceding fourth grade. Interestingly, England, Iceland, and Norway were three of the countries that Luyten (2006) studied during his analysis of the TIMSS 1995 data.

As is also evident in Figure 2, many of the participating countries had distributions that did not fit this pattern of adhering to a strict cutoff date. They instead had age distributions that featured a long tail of older students as well as some younger students. In more than half of the PIRLS 2006 countries, more than 10% of students in the grade tested were older than the students in the predominant age cohort. The most extreme examples in Figure 2 are Trinidad and Tobago, Indonesia, and Morocco, where, in particular, between one third and one half of the students were older than those in the predominant age cohort.

However, in these countries, the lack of economic development is hampering the implementation of educational policies. As noted in the *PIRLS 2006 International Report* (Mullis, Martin, Kennedy, & Foy, 2007), the majority of the countries had a Human Development Index (HDI) of 0.9 or higher (indicating high levels of school enrollment and a good standard of living). In comparison, the relatively low HDI values for these three countries (e.g., 0.64 for Morocco) highlight the impact of economic development on students' educational achievement and, indeed, on a country's ability to implement strict policies about age of entry to school and (more generally) to increase school enrollments.

READING ACHIEVEMENT AND STUDENT AGE

If, as seems intuitive from a perspective based solely on maturation, older students within a grade would have higher reading achievement than younger students at that grade, then we could expect that countries with older students should have, on average, higher achievement than countries with younger students. As a result of the cross-national variations in age of entry and promotion and retention practices described above, the average age of students in the PIRLS 2006 countries that tested at the fourth grade ranged from 9.7 years (Italy) to 11.0 (Latvia). If being older of itself conveys an advantage in reading literacy, then the PIRLS data should show higher average reading achievement for the countries with the older students and lower average achievement for countries with younger students. However, this pattern did not appear to always be the case.

Figure 2: PIRLS 2006 average student age and age cohorts

Country	Percentage of students					
	Older than the predominant age cohort		In the predominant age cohort		Younger than the predominant age cohort	
Iceland	0	(0.1)	100	(0.1)	0	(0.1)
Norway	0	(0.1)	99	(0.2)	0	(0.1)
England	1	(0.1)	99	(0.4)	1	(0.4)
Poland	1	(0.2)	98	(0.3)	1	(0.1)
Chinese Taipei	1	(0.2)	97	(0.3)	2	(0.2)
Slovenia	2	(0.2)	96	(0.3)	2	(0.2)
Sweden	3	(0.4)	95	(0.5)	2	(0.3)
Italy	3	(0.4)	93	(0.6)	4	(0.6)
Scotland	7	(1.0)	93	(1.0)	0	(0.1)
Spain	8	(0.7)	91	(0.8)	1	(0.2)
Israel	8	(0.6)	89	(0.6)	3	(0.2)
Bulgaria	7	(0.9)	88	(1.1)	5	(0.6)
Hong Kong (SAR)	15	(1.1)	85	(1.1)	0	(0.0)
Belgium (Flemish)	14	(0.9)	85	(0.9)	2	(0.2)
Kuwait	13	(0.7)	84	(0.7)	3	(0.3)
Denmark	12	(0.6)	84	(0.7)	4	(0.4)
Lithuania	7	(0.5)	84	(0.7)	9	(0.5)
Slovakia	15	(0.8)	84	(0.8)	1	(0.1)
Austria	16	(0.7)	82	(0.7)	2	(0.3)
Iran	18	(1.2)	82	(1.2)	0	(0.1)
France	16	(1.0)	81	(1.0)	2	(0.3)
Macedonia	9	(0.7)	80	(0.9)	10	(0.7)
Moldova	13	(0.8)	80	(0.9)	7	(0.7)
Belgium (French)	18	(0.9)	80	(1.0)	2	(0.3)
Romania	14	(0.9)	79	(1.2)	7	(1.1)
Latvia	16	(0.8)	79	(1.2)	5	(1.0)
Hungary	17	(1.0)	77	(1.0)	6	(0.5)
United States	15	(1.3)	77	(1.2)	8	(0.7)
Germany	15	(1.3)	77	(1.2)	8	(0.5)
Georgia	18	(1.1)	76	(1.0)	6	(0.6)
Netherlands	21	(1.5)	75	(1.4)	4	(0.4)
Qatar	21	(0.5)	74	(0.5)	6	(0.3)
Russian Federation	11	(0.6)	73	(1.1)	16	(1.0)
Trinidad and Tobago	33	(1.5)	60	(1.4)	8	(0.7)
Indonesia	42	(1.9)	50	(1.7)	8	(0.5)
Morocco	50	(1.4)	46	(1.3)	4	(0.4)

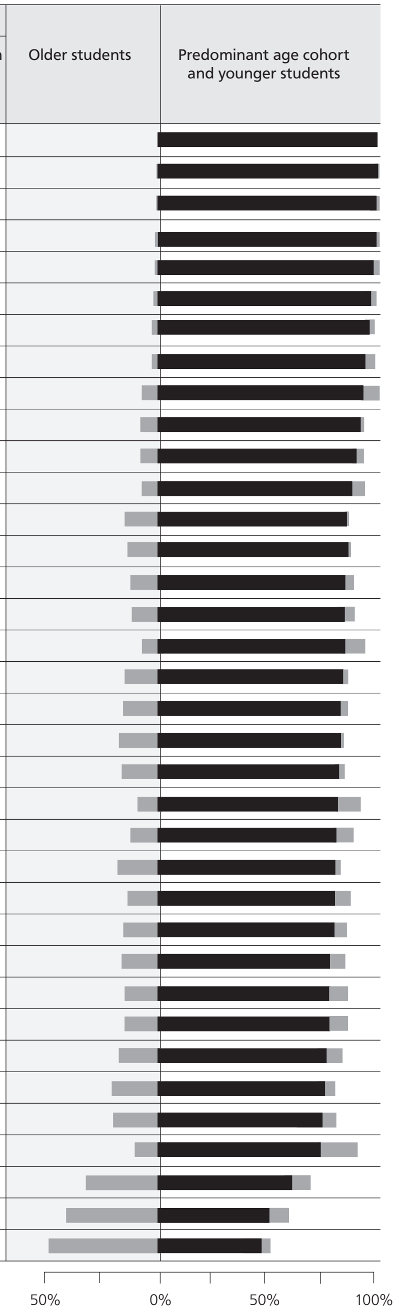


Figure 3 plots, for each PIRLS 2006 country shown in Figures 1 and 2, average reading achievement against the average age of the PIRLS 2006 students. The Pearson correlation between age and achievement for these 36 countries is 0.15, which is only slightly positive. But what is important to observe here is that the countries with developing economies tended to have older students, as evidenced by Morocco. As another example (not included in Figure 3), South Africa (with a low HDI value of 0.65) participated in PIRLS 2006 at Grade 5 and had the oldest students, on average, at 11.9 years old. Yet, despite its students being a year older than even the oldest group of fourth-grade participants and having experienced one more year of school, South Africa had the lowest average achievement across all the countries. Thus, it appears that degree of economic development can have a much larger impact on educational achievement than can age-within-grade or even amount of schooling.

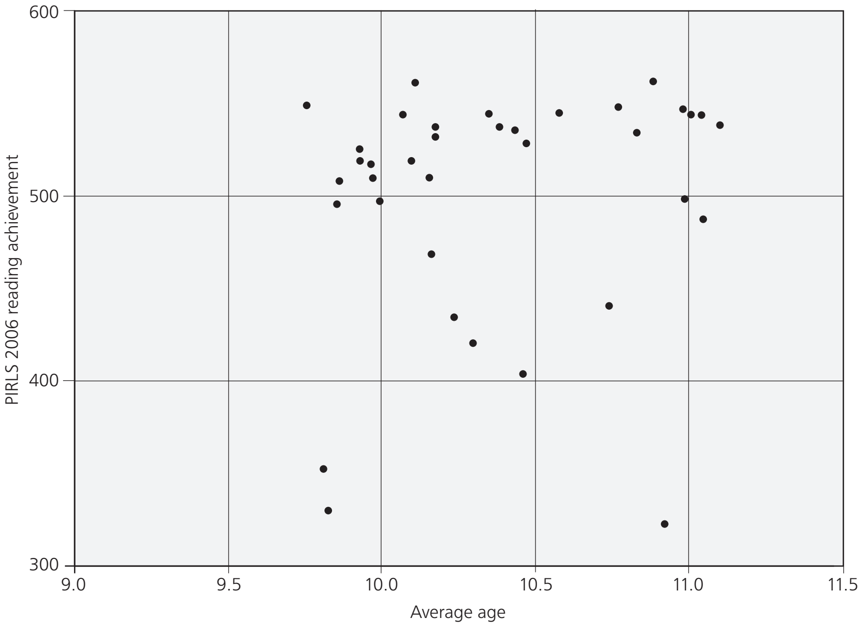
AGE-WITHIN-GRADE AND STUDENT READING ACHIEVEMENT

Many factors other than average age can influence a country's average reading achievement. As such, there are many possible explanations for why a simple plot of countries' average achievement against average age might not show an expected positive relationship. For example, and as we have already noted, economic development and the investment in education that this makes possible may have a stronger effect on literacy levels than average age of students at a grade level. However, even if there is an underlying maturational effect, that is, older students being more able than younger, it is possible that the combined effect of a country's policy and practice on age of entry to primary school together with its promotion and retention policies could disrupt this expected positive relationship between age and achievement.

Figures 1 and 2 above showed many examples of countries with sizeable percentages of PIRLS 2006 fourth graders who were older than the predominant age cohort. In general, these countries have policies that result in students being retained and repeating a grade. In a country where students have to demonstrate a certain level of progress in order to be promoted to the next grade, the weaker students who have to repeat a grade will be older, on average, than the students who are in that grade for the first time. Similarly, in a country where children are not obliged to begin school strictly on the basis of age, there may be a tendency for parents to hold back for another year those children they consider not quite ready for the rigors of schooling. Some of these students also may turn out to be among the less able of their age cohort. These factors, operating separately or in combination, can create a situation in which the older students in a grade level are also among the academically less able. If sufficiently large numbers of such students are involved, this factor can affect the relationship between age and achievement within the grade.

The first part of this study has shown that the distribution of students' ages within fourth grade varied very widely across the PIRLS 2006 countries, with many societal and systemic factors coming into play. Despite this variation, we identified a number of countries where the policy appeared to be to admit children to school solely on the

Figure 3: PIRLS 2006 Grade 4 average reading achievement by average age

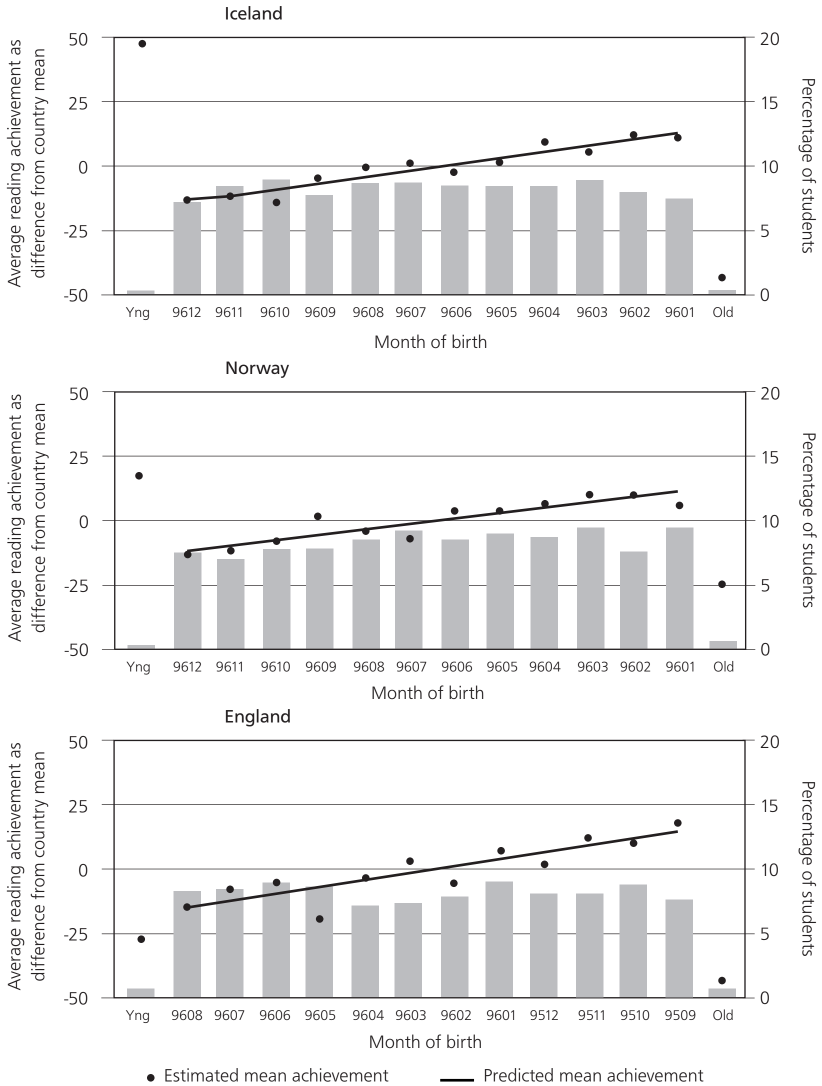


basis of their chronological age, and to automatically promote this cohort through the grades, at least as far as Grade 4. Based on the findings of previous research, we could assume that any underlying "natural" or maturational relationship between age and achievement would be most apparent in such countries—those where all the students in the grade come from a single one-year age cohort and where there has been no grade repetition for weaker students or promotion for high-performing students. We selected, from the eight PIRLS 2006 countries with age distributions meeting these criteria (see Figure 4), Iceland, Norway, and England, in order to explore this issue. These three countries had the greatest percentages of target-grade students (essentially all students) in the predominant age cohort, and all three have the policy of promoting primary-school students automatically from one grade to the next.

For Iceland, Norway, and England, Figure 4 presents the percentage of students at each birth-month of the predominant age cohort, together with average reading achievement for each month. When reading across this figure for each country, we can see that the younger students are to the left and the older students to the right. The bars representing the percentages of students at each month of birth are labeled accordingly (e.g., "9612" refers to the year 1996 and the month of December, or December 1996). Students younger than those in the predominant age cohort are represented by the bar labeled "Yng" and those older by the bar labeled "Old."

As is evident in Figure 4, each of the three countries has an approximately uniform distribution of students across the 12 months of the predominant age cohort and very few younger or older students. Accompanying the percentage of students

Figure 4: Example countries with all or almost all students in the predominant age cohort: Iceland, Norway, England



at each birth month is the average reading achievement for the students at each birth month, expressed as a difference from the overall mean for the country. The left side of the display for each country presents a scale ranging from -50 to +50, with 0 indicating the country's overall mean reading achievement. The black dot for each month shows the degree to which students with that birth-month had average achievement below or above the overall mean for the country. For all three countries, the achievement pattern is the same: average reading achievement is lowest among the younger students and gradually increases as students get older. Therefore, as shown by the regression line, there is a positive relationship between age and reading achievement. The very small percentages of students who were younger or older than the predominant age cohort have little impact on the age–achievement relationship, even though the very youngest students have the highest achievement and the very oldest have the lowest achievement.

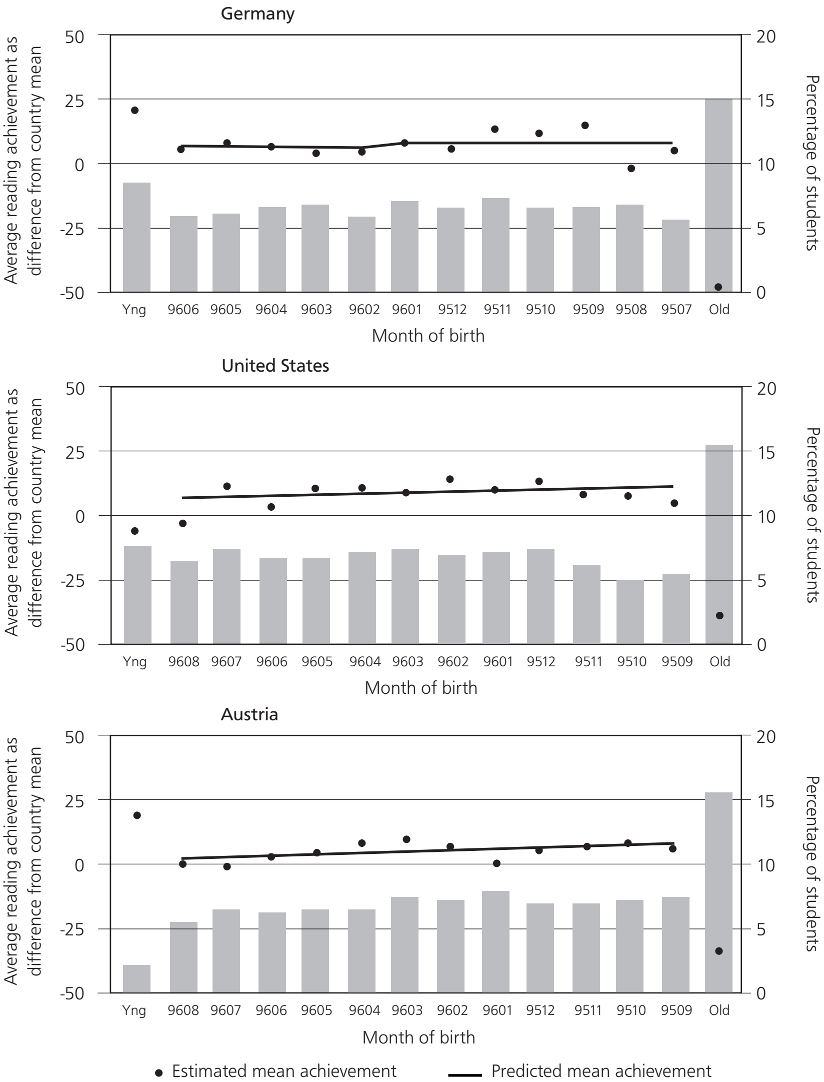
The display for Germany, the United States, and Austria in Figure 5 provides a contrasting example. These countries are three of the many PIRLS 2006 countries that had a relatively large percentage of students older than the predominant age cohort. In Austria, Germany, and the United States, approximately 15% of the Grade 4 students assessed were older than the predominant age cohort (see Figure 2). Additionally, in Germany and the United States, eight percent of the students were younger than the predominant cohort, and in Austria, two percent. In all three countries, the older students had average achievement well below the students in the predominant age cohort, and the younger students had achievement above them, at least in Austria and Germany. Furthermore, the achievement pattern evident in Figure 4, that is, higher achievement among the older students in the predominant age cohort, is not evident in Figure 5. It appears that, in countries such as these, the combined effect of entry and promotion and retention practices is that of producing a predominant age cohort with uniform reading achievement, together with an older group with lower achievement and a (smaller) younger group with higher achievement. Note that the regression lines in Figure 5 are based on students in the predominant age cohort only. Including the younger and older students would have made the age–achievement relationships even less positive, and somewhat distorted.

Consideration of the information presented in Figures 4 and 5 makes clear that no simple, consistent relationship within a grade level between student age and reading achievement applied to all countries at fourth grade. This finding implies that any attempt to make a global statistical adjustment to countries' average achievement in order to account for differences in average age is likely to be misleading.

GROWTH IN READING ACHIEVEMENT FROM FOURTH TO FIFTH GRADE

In this last section of this paper, we provide an example of how the regression discontinuity approach can be applied to the PIRLS 2006 data. As shown in Figure 4, Iceland and Norway were two of the countries with strict adherence to an age-of-entry cutoff and automatic promotion through the grades. In addition, Iceland and Norway were the two countries that administered the PIRLS 2006 assessment at the fifth grade as well as at the fourth grade. Although the fifth-grade samples were somewhat smaller than the fourth-grade samples reported internationally for PIRLS 2006, the fifth-grade samples were nationally representative and so provided us with an opportunity to examine the relationship between age, grade, and reading achievement (for information on the fifth-grade samples, see Martin, Mullis, & Kennedy, 2007).

Figure 5: Example countries with students older or younger that in the predominant age cohort: Germany, United States, Austria



Figures 6, 7, and 8 summarize the results of the series of regression analyses that we conducted in order to model the relationship between age-within-grade (represented by birth-month) and reading achievement for Iceland and Norway. Model 1 in Figure 6 shows the regression coefficients resulting from fitting separate equations to the fourth- and fifth-grade data in each country. The slopes of the regression lines in this model for the fourth and fifth grades are approximately parallel (2.5 and 2.2, respectively, for Iceland, and 2.1 and 2.9, respectively, for Norway). Accordingly, for this exploratory analysis, we decided to fit parallel slopes for the next step, as shown in Figure 6, Model 2. The parallel slopes for the two regression lines imply that the relationship between age and reading achievement is the same in the fourth and fifth grades in each country.

The fitted lines with parallel slopes are presented graphically in Figure 6, which also shows average reading achievement by birth month for fourth and fifth grade for both Iceland and Norway. Fitting separate lines with parallel slopes to fourth and fifth grade (2.3 for Iceland and 2.5 for Norway) provides estimates of the average difference in reading achievement between the two grades. These are 38.3 score points for Iceland and 42.3 points for Norway (see Figure 6, Model 2). According to the parallel-slopes model, the fifth-grade students in Iceland scored 38.3 points higher, on average, than the fourth-grade students on the PIRLS 2006 assessment, while the fifth-graders in Norway scored 42.3 points higher than the fourth-graders. These estimates are reasonably close to the actual differences between the grades: 40 points for Iceland and 37.4 points for Norway (Figure 6, Model 1).

The approximately 40-points difference in the reading achievement of the fourth- and fifth-grade students is an estimate of the average growth in achievement that could be expected in these two countries as students move from the fourth through to the fifth grade. This increase is less than that found by Luyten (2006) with respect to the TIMSS mathematics and science data. However, we could anticipate that the increase in reading achievement would be less because students have essentially learned the basics of reading by fourth grade but would still be developing their mathematics skills between third and fourth grades.

Figure 6: PIRLS 2006 reading achievement as a function of birth month within grade

Model 1: Reading achievement by birth-month, separate models by grade

Country	Grade	Intercept	Slope
Iceland	4	496.8 (2.09)	2.5 (0.33)
	5	536.8 (5.21)	2.2 (0.59)
Norway	4	486.2 (3.59)	2.1 (0.43)
	5	523.6 (7.48)	2.9 (1.32)

Model 2: Reading achievement by birth month, parallel regression lines

Country	Intercept	Grade	Slope
Iceland	497.6 (2.06)	38.3 (3.69)	2.3 (0.33)
Norway	483.7 (4.69)	42.3 (4.23)	2.5 (0.71)

Model 3: Reading achievement by birth month, extended regression lines

Country	Intercept	Grade	Slope
Iceland	497.6 (2.06)	10.3 (5.81)	2.3 (0.33)
Norway	483.7 (4.69)	12.1 (9.10)	2.5 (0.71)

This growth in reading achievement reflects not only the effects of an extra year of schooling but also other learning experiences in the home and the community as well as cognitive maturation as the students became a year older. These factors are inextricably interwoven, because all the students attended school for a year and, at the same time, all the students grew older and more mature. Therefore, it is not possible from data such as these to say with any certainty how much of the 40-points growth from fourth grade to fifth grade in Iceland and Norway can be attributed to the effects of a year of schooling and how much can be attributed to other, incidental maturational factors. However, the regression discontinuity technique makes it possible to capitalize on the common relationship between age and reading achievement in these two grades to estimate the effect of schooling in these countries.

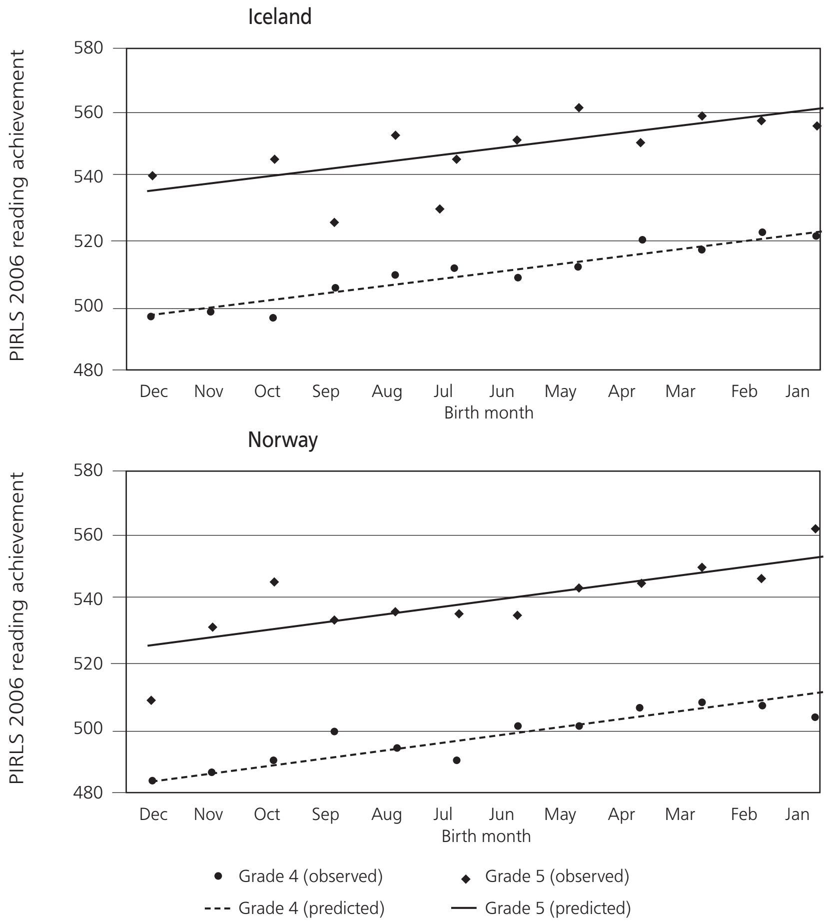
The predominant age cohort for the PIRLS 2006 fourth-grade students in Iceland and Norway was the cohort born between January 1 and December 31, 1996; the corresponding cohort for the fifth-grade students was those students born between January 1 and December 31, 1995. As Figure 7 shows, there was a steady increase in average reading achievement at fourth grade from the youngest students (those born in December 1996) to the oldest students (those born in January 1996) and also at fifth grade from the youngest students (those born in December 1995) to the oldest students (those born in January 1995). Because the regression lines for the two grades in the figure are parallel within each country, the increases from month to month at fourth grade are the same and the increases from month to month at fifth grade are the same.

Figure 7 depicts average achievement for the two grade levels along a continuous two-year birth-month continuum, beginning on the left with those born in December 1996 (the youngest fourth-grade students), and extending on the right to those born in January 1995 (the oldest fifth-grade students). The cutoff point for the two grades is between January 1996 and December 1995. A discontinuity in the regression lines at this point would indicate an achievement difference associated with a change from fourth to fifth grade, whereas a continuous straight line would imply no discernable grade effect.

For both Iceland and Norway, Figure 8 shows a discontinuity in the fitted regression line, with higher achievement associated with fifth grade in both countries. For Iceland, the grade effect, adjusted for birth month, was 10.3 score points; for Norway, it was 12.1 points (see Figure 6, Model 3). One interpretation of this finding with respect to Iceland is that we could expect the extra year of schooling to give a hypothetical fifth-grade student born in a particular birth month a 10.3-point advantage over a hypothetical fourth-grade student born in the same birth month. Similarly, in Norway, we could expect that a hypothetical fifth-grade student would have a 12.1-point advantage over a hypothetical fourth-grade student born in the same month.

Essentially, the foregoing analysis implies that the 38.3-point difference between the average reading achievement of the fourth- and fifth-grade students in Iceland could be separated into a 10.3-point grade effect and a 28-point age-within-grade effect,

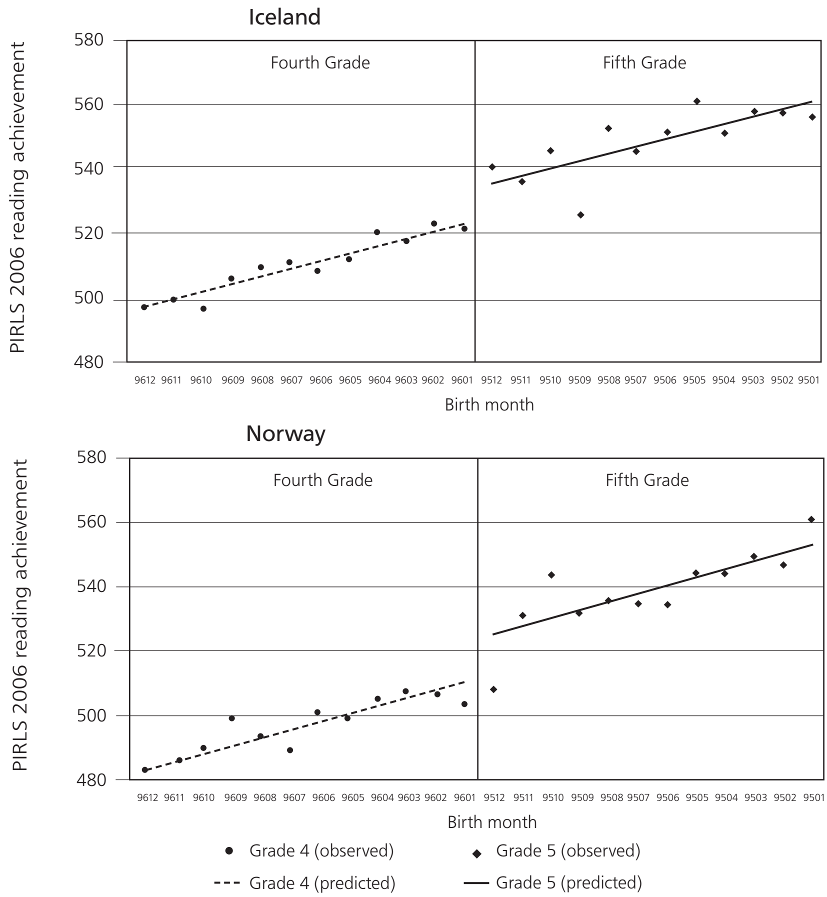
Figure 7: PIRLS 2006 reading achievement as a function of birth month at Grades 4 and 5: Iceland and Norway



while the 42.3-point difference in Norway could be separated into a 12.1-points grade effect and a 30.2-point age-within-grade effect.

When considering this finding, it is important to recognize that the relative sizes of the grade effect and the age-within-grade effect are determined by the slopes of the age-within-grade regression lines. The steeper the slopes of the regression lines for each grade, the more the lines will align across the grades and the smaller will be the estimate of the grade effect. However, as we demonstrated earlier, the slopes of the age-within-grade regression lines are greatly influenced by the factors that determine the distribution of age-within-grade. And, as we also observed earlier, Iceland and Norway are two of the countries where age alone is the criterion for beginning school. In both countries, the relationship between age and achievement within grade was positive.

Figure 8: PIRLS 2006 reading achievement as a function of birth month across Grades 4 and 5 consecutively: Iceland and Norway



DISCUSSION AND CONCLUSION

Analyses by Cliffordson (2008), Cliffordson and Gustafsson (2010), and Luyten and Valdecamp (2008) have demonstrated the promise of new, powerful analytic techniques for disentangling the effects of age and grade in large-scale cross-sectional student achievement data. However, these methods can depend on restrictive assumptions about the relationship between age and grade, as is the case for the regression discontinuity technique used in several important studies (Cahan & Cohen, 1987; Cliffordson & Gustafsson 2007; Luyten, 2006). Even when recent investigations that relax these assumptions, such as the Heckman approach applied by Luyten and Valdecamp, or the instrumental variables regression approach of Cliffordson and Gustafsson, are taken into account, these approaches require use of adjacent grade data from countries where it is possible to explicitly model the relationship between age and grade. The present study has shown that although the relationship between age and grade may be straightforward and amenable to analysis in some countries, in general the complexity of the age–grade relationship remains a barrier to the application of any straightforward adjustment for the effects of age differences to the grade-level achievement results reported by TIMSS and PIRLS.

We used PIRLS 2006 reading achievement data for fourth-grade students to conduct the age-by-birth-month analysis for 36 countries. The analysis showed that variation in policies of age of entry to school resulted in different average ages for the fourth-grade students from country to country. In line with the three most prevalent policies across the participating PIRLS countries, the oldest fourth-grade students in the assessment were from countries that admit students in the calendar year in which they turn seven, the next oldest from countries that admit children who have reached their sixth birthday by September, and the youngest from countries where students begin primary school in the calendar year in which they turn six. Promotion and retention policies also vary across the PIRLS countries to the extent that some of the PIRLS 2006 countries had a relatively large percentage of "older" students by fourth grade. Thus, students who were retained and repeated grades were older than the predominant age cohort for their grade. Finally, students in several of the countries facing challenges associated with economic development were older, on average, than the average age of their international peers.

With so many factors influencing the within-grade age distribution, the finding that the relationship at fourth grade between students' age and their reading achievement differed across the PIRLS 2006 countries is understandable. In the relatively few countries with strict age cutoffs for entering school and policies of automatic promotion, students were equally distributed by birth month across the fourth grade. In these countries, the older fourth-graders had systematically higher average reading achievement than the younger students. However, in the many countries with relatively large percentages of older students, these students had lower reading achievement than the younger students. Thus, across the PIRLS 2006 countries at fourth grade, there was only a slightly positive correlation (0.15) between age and reading achievement. Also, the PIRLS 2006 countries with the fewest economic resources tended to have the oldest students and the lowest average reading achievement, a finding that highlights the dominant importance of economic development for students' educational achievement.

TIMSS' and PIRLS' choice of grade as the basis for monitoring and reporting student achievement makes student age an important contextual variable to consider when interpreting achievement differences across countries. This choice also means, however, that participating countries can monitor the relationship between changes in policies related to students' ages within grades and improvements (or declines) in average achievement. For example, both the Russian Federation and Slovenia have lowered the age of entry into school to provide an extra year of primary school (now four years of schooling rather than three) for their students. In both cases, the increase in schooling has coincided with an increase in average reading achievement (Mullis, Martin, Kennedy, & Foy, 2007).

Because Iceland and Norway are two countries that admit children to school strictly on the basis of age and automatically promote them through the grades, these two countries, by assessing two grades in PIRLS 2006 (fifth as well as fourth grade), were able to examine the relationship between age-within-grade and reading achievement using the regression discontinuity approach. They were also able to examine these results in the light of existing curricular emphases and prevalent instructional approaches to reading.

This study has shown that the interrelationships among achievement, grade, and age vary from country to country and often are extremely complicated. These complications are rooted in countries' policies on age of school entry, on promotion and retention, and on how each country actually implements—in accordance with its economic, social, and cultural contexts—these policies in practice. Some countries adopt a strict policy of admitting students to primary school solely on the basis of age and maintain the resulting age distribution by way of automatic promotion through the grades. Others allow parents some latitude in choosing when to have their child begin school, and may permit or require students who are struggling to repeat a year. The study also has shown that the policies that countries implement with regard to these issues are reflected in the distributions of students' ages within grade and in the relationship between age and reading achievement at fourth grade.

References

Cahan, S., & Cohen, N. (1989). Age versus schooling effects on intelligence development. *Child Development*, *60*(5), 1239–1249.

Cliffordson, C. (2008, September). *Effects of schooling and age on performance in mathematics and science: A between-grade regression discontinuity design applied to Swedish TIMSS 1995 data*. Paper presented at the third IEA International Research Conference (IRC-2008), Taipei, Taiwan.

Cliffordson, C., & Gustafsson, J.-E. (2007). Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling. *Intelligence*, *36*(2), 143–152.

Cliffordson, C., & Gustafsson, J.-E. (2010). *Effects of schooling and age on performance in mathematics and science: A between-grade regression discontinuity design with instrumental variables applied to Swedish TIMSS 1995 data*. Paper presented at the fourth IEA International Research Conference (IRC-2010), Gothenburg, Sweden.

Foy, P., & Kennedy, A. M. (Eds.). (2008). *PIRLS 2006 user guid*e. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica*, *51*, 153–161.

Kennedy, A. M., Mullis, I. V. S., Martin, M. O., & Trong, K. L. (Eds.). (2007). *PIRLS 2006 encyclopedia*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Luyten, H. (2006). An empirical assessment of the absolute effect of schooling: Regression-discontinuity applied to TIMSS 1995. *Oxford Review of Education*, *32*, 397–429.

Luyten, H., Peschar, J., & Coe, R. (2008). Effects of schooling on reading performance, reading engagement, and reading activities of 15-year-olds in England. *American Educational Research Journal*, *45*(2), 319–342.

Luyten, H., & Veldkamp, B. (2008, September). *Assessing the effect of schooling with cross-sectional data: Between grade differences addressed as a selection-bias problem*. Paper presented at the third IEA International Research Conference (IRC-2008), Taipei, Taiwan.

Martin, M. O., Mullis, I. V. S., & Kennedy, A. M. (Eds.). (2007). *PIRLS 2006 technical report*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., & Foy, P., with Olson, J. F., Preuschoff, A. C., Erberber, E., ... Galia, J. (2008). *TIMSS 2007 international mathematics report: Findings from IEA's Trends in International Mathematics and Science Study at the fourth and eighth grades*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Kennedy, A. M., & Foy, P. (2007). *PIRLS 2006 international report: IEA's Progress in International Reading Literacy Study in primary schools in 40 countries*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Martin, M. O., Olson, J. F., Berger, D. R., Milne, D., & Stanco, G. M. (Eds.). (2008). *TIMSS 2007 encyclopedia: A guide to mathematics and science education around the world*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Organisation for Economic Co-operation and Development (OECD). (2001). *Knowledge and skills for life: First results from the OECD Programme for International Student Assessment (PISA) 2000*. Paris, France: Author.

Van Damme, J., Vanhee, L., & Pustjens, H. (2008, September). *Explaining reading achievement in PIRLS by age and SES*. Paper presented at the third IEA International Research Conference (IRC-2008), Taipei, Taiwan.