



What makes a good reader? Worldwide insights from PIRLS 2016

João Marôco¹ 

© Springer Nature B.V. 2020

Abstract

Using hierarchical linear models, this study probes into student, family, teacher, and schools' variables that can explain the variation in Progress in International Reading Literacy Study (PIRLS) 2016 results. Students' confidence in reading, early literacy tasks, and parents' expectations are the strongest explanatory variables of reading literacy. Teachers' perception of class instruction being limited by students' needs is the strongest explanatory variable of PIRLS achievement, although this was not consistently verified among all countries. No teaching strategies or other related variables emerged consistently as explanatory variables in every country. A similar result was observed in schools where the percentage of economic disadvantage students was the most consistent explanatory variable of PIRLS results. The present analysis shows that although student variables are the most consistent explanatory variables among participating countries, a general conclusion of what makes a good reader worldwide must consider all student, teacher, and school variables conjointly, acknowledging the existence of between-country variation.

Keywords Reading literacy · PIRLS 2016 · Student, families, teachers, and schools' explanatory variables

Introduction

In today's driven information world, being able to read and comprehend what was read is a key skill for full citizenship. Becoming a proficient reader at an early age is fundamental not only for the development and intellectual maturation of children, but also to advance successfully in school, in the workplace, and contemporary societies.

✉ João Marôco
jpmaroco@gmail.com

¹ William James Center for Research, ISPA-Instituto Universitário de Ciências Psicológicas, Sociais e da Vida, Rua Jardim do Tabaco, 34, 1149-041 Lisbon, Portugal

Aware of the importance of reading in modern-day societies, the International Association for the Evaluation of Educational Achievement (IEA)—a cooperative of institutes and agencies, both government and non-governmental, that develop research in education—has been promoting the *Progress in International Reading Literacy Study* (PIRLS) every 5 years since 2001. PIRLS is a collaborative research project between the participating countries and the IEA, under the direction of the TIMSS & PIRLS International Study Center at Boston College. The study evaluates reading literacy in children with 4 years of formal schooling (excluding the kindergarten or pre-primary education years) by sampling a large number of students in each participant.

The IEA defines reading literacy as “the ability to understand and use those written language forms required by society and/or valued by the individual. Readers can construct meaning from texts in a variety of forms. They read to learn, to participate in communities of readers in school and everyday life, and for enjoyment” (Mullis, Martin, & Sainsbury, 2015, p. 12). The PIRLS framework emphasizes reading for two main purposes: (1) literary experience and (2) acquisition and use of information (Mullis et al., 2015). The age of the students evaluated by the PIRLS assessments—9–10 years—reinforces the importance of reading literacy in modern educational systems: until this age, children learn to read; from then on, they read to learn (Martin, Mullis, & Foy, 2017). It is therefore fundamental that children are competent readers by the time they leave primary school.

Data gathered by PIRLS provide comparative information on how well a child reads as assessed by a comprehensive test of reading literacy. The test focuses on four broad-based reading comprehension processes employed by fourth-graders: (1) focusing on and retrieval of explicitly stated information; (2) making straightforward inferences; (3) interpreting and integrating ideas and information; and (4) evaluating and criticizing content and textual elements (Mullis et al., 2015). PIRLS also collects considerable background information through a series of questionnaires aimed at the participating countries’ educational systems, schools, teachers, students, and students’ families (Hooper, Mullis, & Martin, 2015). This information allows for the characterization of reading opportunities, strategies, and contexts, as well as for the identification of variables that can influence learning. In assessing reading skills and the contexts in which they develop, PIRLS opens the way for the diagnosis of educational, socioeconomic, and cultural areas where education systems can invest to make reading an effective tool for acquiring knowledge and active citizenship.

The development of the PIRLS framework, tests, and questionnaires is the result of collaborative work between groups of IEA specialists and the PIRLS national coordinators. National coordinators compile data on the educational systems of each of the participants for the Encyclopedia of PIRLS to characterize and frame the educational systems of their countries (Mullis, Martin, Goh, & Prendergast, 2017).

The collaborative nature of PIRLS, together with a range of technical validations imposed on the process, ensures that the results in each edition of the study present transcultural validity, concurrent validity, and reliability. The technical validations, that cover different areas of the assessment, include procedures for translation, adaptation, and delivery of the tests and questionnaires, sampling, and literacy estimation methods that are anchored in the results of previous PIRLS editions.

In 2016, fifty countries, 12,000 schools, 16,000 teachers, 310,000 parents, and 319,000 students participated in the fourth edition of PIRLS (Mullis, Martin, Foy, & Hooper, 2017a). Fourth graders from Russia ($M=581$, $SE=2.2$) and Singapore ($M=576$, $SE=3.2$) had the highest mean reading achievement. On the scale opposite, 4th graders from Morocco ($M=358$, $SE=3.9$) and South-Africa ($M=320$, $SE=4.4$) had the lowest country mean achievements. The gap between top and low-performers countries is more than two PIRLS standard-deviation (200 points), an equivalent to a four-year¹ gap in the reading literacy of 10-year old students from high versus low performing countries. More girls than boys are good readers, and good readers have an early start in learning and home environments that support learning. Good readers attend schools that put a high priority in reading instruction, are academically oriented, and are well resourced (Mullis, Martin, Foy, & Hooper, 2017c). Although the TIMSS and PIRLS International Study Center has released an overall outlook on what makes a good reader, the subject of explanatory variables of reading attainment within participating countries has been conspicuously absent from most of the recent studies undertaken by the educational research community. Many models of reading comprehension have been proposed (see, e.g., Harker, 1972; Joshi & Aaron, 2000) but those are limited in the number of explanatory variables that can explain reading literacy. For example, in Joshi & Aaron (2000) model, reading comprehension is simply estimated as the product of Decoding and Listening Comprehension plus the Speed of Processing. For more recent models, as the Component Model of Reading (Aaron, Joshi, Gooden, & Bentum, 2008), reading literacy of children is explained not only by cognitive factors (e.g., word recognition and reading fluency) but also by psychological (e.g., reading self-motivation and enjoyment; teachers' expectations) and environmental factors (e.g., home environment and parents' engagement; schools' resources for reading). The PIRLS framework follows closely the Component Model of Reading wide and comprehensive network of possible explanatory variables of reading literacy including students (cognitive and psychological domains), and families, teachers, and schools' (environmental domain). Initially developed for the 2001 edition of PIRLS, but updated for each subsequent assessment cycle, the PIRLS framework emphasis has been shifting from "demonstrating fluency and basic comprehension to demonstrating the ability to apply what is read to new situations or projects" (Mullis et al., 2015, p. 11). The framework explicit recognizing self (e.g., self-confidence in reading), families (e.g., home resources for learning or parents education), teachers (e.g., strategies used for teaching reading) and schools (e.g., reading resources like libraries or schools' emphasis in academic success) as complex explanatory variables of reading achievement at early ages.

There is a long-standing debate around the role of student and family background and school quality in shaping learning. Most efforts to address this question using international large scale student assessments (ILSA) are based on TIMSS (mathematics and science literacies) and PISA (mathematics, reading and science literacies) data on a country or regional levels (see e.g., Karakolidis, Pitsia, & Emvalotis,

¹ Within a country, a mean difference between 1/3 and 1/2 standard deviation in the PIRLS scale was estimated to approximate one school year (Schwippert & Goy, 2007, p. 27).

2016; Thien & Ong, 2015). A few PIRLS reports on reading literacy have surfaced, but once again limited to a country or a few countries analysis (Araújo & Costa, 2015; Costa & Araújo, 2017; Marôco, 2018; Park, 2008; Tse, Lam, Lam, Loh, & Westwood, 2005). Furthermore, a 2015 literature review on PIRLS data usage found out that despite PIRLS being a reading literacy study, its data is hardly used for reading research (Lenkeit, Chan, Hopfenbeck, & Baird, 2015). Using data from 25 countries who took part in the 2001 PIRLS edition, Park (2008) uncovered that home literacy resources were consistent explanatory variables of reading achievement, although effects sizes varied substantially across countries. A small number of generalized multi-country analysis have been produced recently, and yet again, focusing mainly on mathematics and science and its correlates measured at a single level (either student/family or school) (see e.g., Guo, Marsh, Parker, & Dicke, 2018; Schmidt, Burroughs, Zoido, & Houang, 2015). It is not clear how student/family level variables versus school level variables relate to student achievement overall or how these patterns replicate across different countries. Using data from the 1970s gathered by IEA on mathematics and science literacy from a series of countries with different gross national products, Stephen Heyneman and William Loxley proposed that in developing nations, school variables are more important than family socioeconomic status, in determining academic achievement (Heyneman & Loxley, 1982). However, the ‘Heyneman-Loxley’ hypothesis did not hold when data from TIMSS on the relationship between family background and mathematics and science achievement gathered in the mid-1990 was reassessed. Baker, Goesling, and Letendre (2002) found out that the relationship between family and school variables and student achievement was similar across nations, regardless of national income. These authors suggested that the spread of mass schooling has reduced the ‘Heyneman-Loxley’ effect. A 2008 study of PIRLS 2001 literacy done by Park (2008) uncovered evidence that partially supports the presence of the Heyneman-Loxley effect in some countries, but not in others. The effects of early home literacy activities, parental attitudes towards learning, and home resources for learning varied between countries according to their level of economic development (Park, 2008).

Research questions

In this paper, looking at PIRLS data collected in the 48 countries that took part in the 2016 edition of the study and have a complete dataset, I inquire onto the three following research questions:

(Q1): How does between-schools’ variation relate to country performance in PIRLS?

(Q2): Which school, teacher, family, and student variables, as set by the cognitive, psychological, and environmental domains of the Component of Reading Model, can explain the within country’s variation on reading achievement?

(Q3): Are the PIRLS literacy explanatory variables for top-achievers substantially different from those of low-achievers?

Within-school variation may correlate negatively with country mean scores, i.e., countries where schools are more heterogeneous have lower PIRLS scores than countries where schools are more homogeneous. If this is the case, then student/family level variables will correlate less with student achievement than school-related variables, thus confirming the Heyneman-Loxley' effect. If not, then I expect that the stronger explanatory variables of student reading achievement will not differ between low and top-performer countries as demonstrated by Baker et al. (2002) with data from the mid-1990 TIMSS editions.

An exploratory hierarchical joint analysis of literacy regressed on student, family, teacher, and school-level organizational, socioeconomic and cultural variables, taking into account the complex sample design of PIRLS and the countries' within schools' variation, can provide useful insights for policy recommendations and education practices aimed at improving reading literacy.

Methods

PIRLS 2016 aims at assessing the reading achievement of fourth graders in 50 participating countries using a set of standardized reading tests forms and context questionnaires to characterize education systems, schools, teachers, students, and their families. These are briefly described in the following sections. A detailed description of the PIRLS methods and experimental design can be found in the PIRLS 2016 Methods and Procedures manual edited by Martin, Mullis, and Hooper (2017).

Participants

Three hundred and seventeen thousand students of both sexes (mean age = 10.2, $SD = 0.4$) attending the 4th year of schooling during the 2015/2016 academic year in 50 countries were enrolled in PIRLS 2016. The 11 benchmark regions were not included in this study, as well as the two countries that did not apply the Student's family questionnaire were removed from the overall analysis. A minimum sample size of 4000 students per country (mean = 5874, $SD = 3175$) was selected by a multi-stage probabilistic sampling procedure as defined by the IEA-Boston College consortium responsible for 2016 PIRLS (LaRoche, Joncas, & Foy, 2017). In the first sampling stage, countries were stratified into regions defined by a set of local stratum variables. In the second stage, around 200 schools per country (mean = 208, $SD = 138$) were selected by systematic sampling proportional to size. Finally, in each selected school, one or two grade four (from primary education) classes were randomly sampled according to the number and size of the classes in the selected schools. All students in the selected classes whose participation was authorized by their parents and who met the eligibility criteria for PIRLS (did not have special educational needs, and were native speakers of the test language or students whose mother tongue was not the test language but had more than 1 year of language learning) were assessed (overall weighted mean participation rate was 95%, $SD = 3%$). School, class, and student sampling weights were derived from the multi-stage

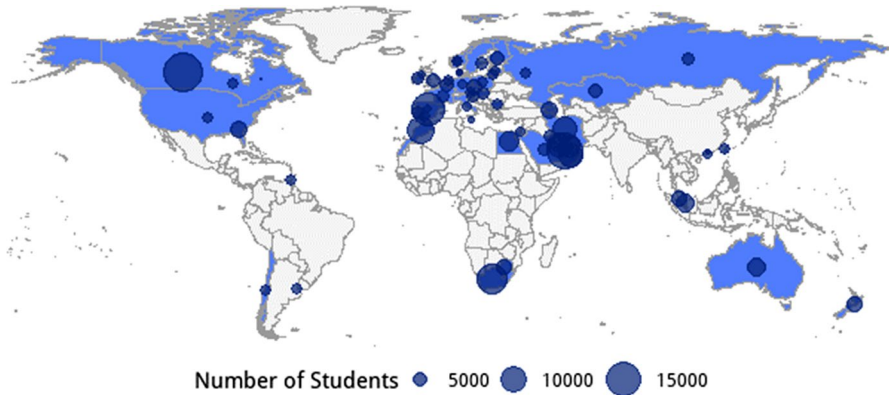


Fig. 1 Countries and benchmark entities, and students per participant taking part in PIRLS 2016

sampling accounting for any disproportional sampling of sub-groups and non-participation (LaRoche et al., 2017). Figure 1 shows the geographical distribution of PIRLS 2016 participants.

The PIRLS test

The PIRLS 2016 test was composed of a set of 16 test forms (booklets). Each booklet comprised a literary text and an informative text followed by a set of multiple-choice and constructed-response items that evaluated reading literacy and its dimensions (reading purposes and comprehension processes) (Mullis & Prendergast, 2017). Each student responded to one of the 16 test forms, according to a matrix-sampling booklet design with planned missing by design items, and to a sociodemographic questionnaire (Martin et al., 2017). The translation and adaptation of the tests, from the original English version, was carried out by the National PIRLS research centers. The validity of the translations and adaptations was verified by an independent translation agency subcontracted by IEA (Ebbs & Wry, 2017).

The PIRLS questionnaires

In parallel to the tests, the PIRLS study deploys a series of questionnaires to school principals, teachers of the selected classes, parents, and students. Based on the questionnaire responses, the IEA-Boston College consortium produced indexes and psychometric scales that allow for the characterization of the educational stakeholders' opinions and perceptions of the educational, professional, and sociodemographic contexts of the school community (Hooper & Fishbein, 2017). Amongst the several scales and indices reported by PIRLS, and following the Component of Reading Model three domains predictors of reading literacy, the ones with a larger explaining power in the HLM models are described briefly in the following section (see Hooper and Fishbein, 2017 and Martin et al., 2017 for full descriptions).

Student and family-level variables

Home resources for Learning (Home Res for Learning) is a scale based on students' and parents' responses regarding the home possessions that may facilitate reading literacy (e.g., books at home; the highest level of education of parents; highest parent's occupational status). Higher scores indicate higher availability of resources. Reported Cronbach's α for this scale (see Martin et al., 2017) ranged from as low as .51 (Saudi Arabia) to .81 (Hungary) with most countries displaying values around .6.

Student Bullying (Stud Bullying) is a scale based on students' responses to how often they experienced bullying behaviors (e.g., making fun or calling names; spread lies; threatening). Higher scores of the scale indicate a lower frequency of bullying. Reported Cronbach's α (see Martin et al., 2017) ranged from .77 (Kuwait) to .86 (Australia) with most countries displaying values greater than .7.

Students Confident in Reading (Stud Conf Reading) is a scale based on students' degree of agreement with statements about their confidence and liking about reading (e.g., I usually do well in reading; reading is easy for me; reading is harder for me than any other subjects). Higher scores indicate higher confidence. Reported Cronbach's α for this scale (see Martin et al., 2017) ranged from as low as .53 (Saudi Arabia) to .83 (Belgium(Flemish)) with most countries displaying values above .7.

Students like reading (Stud Like Read) is a scale based on students' responses to items that evaluate the enjoyment and liking about reading (e.g., I enjoy reading; I like talking about what I read with other people; reading is boring). Higher scores indicate higher liking. Reported Cronbach's α for this scale (see Martin et al., 2017) ranged from as low as .71 (Iran) to .90 (Poland) with most countries displaying values above .8.

Students Sense of School Belonging (Stud Sense of Sch Belong) is a scale based on students' degree of agreement with statements like "I like being in school"; "I feel safe when I am at school"; or "Teachers are fare to me". Higher scores indicate a higher sense of belonging. Reported Cronbach's α for this scale (see Martin et al., 2017) ranged from as low as .59 (Morocco) to .82 (Australia or Qatar) with most countries displaying values above .7.

Expected level of education of Child (Exp Level Ed Child) is an index based on the parents' response to the level of education that they expect their child to reach.

Early Literacy Activities Before Beginning Primary School (Early Lit Actv Befor Sch) is a scale based on parents' report on the frequency of the children doing literacy activities before entering school (e.g.; read books; tell stories; write letters or words). Higher scale scores indicate higher activities. Reported Cronbach's α for this scale (see Martin et al., 2017) ranged from as low as .70 (Kazakhstan, Oman,...) to .85 (Bulgaria and Egypt) with most countries displaying values above .7.

Could Do Early Literacy Tasks When Beginning Primary School (Early Lit Tasks) is a scale derived from parents' responses to how well their children could do tasks like "Read some Words"; "Write some letters"; "Write sentences" before entering primary school. Higher scores indicate higher ability. Reported Cronbach's α for this scale (see Martin et al., 2017) ranged from .80 (e.g. Macao SAR) to .94 (Bulgaria) with most countries displaying values around .9.

Parents' Perceptions of Their Child's School (Parents Perce Sch) is a scale derived from the parents' responses about the school of their children (e.g., My child

school promotes high academic standards; My child school cares about my child progress). Higher scale scores indicate better perceptions about the school. Reported Cronbach's α for this scale (see Martin et al., 2017) ranged from .75 (Malta) to .90 (Sweden) with most countries displaying values around .8.

Parents Like Reading (Parents Like Read) is a scale that measures how students' parents feel about reading as derived from their agreement with sentences like "I read only if I have to"; "I enjoy reading" or "Reading is one of my favorite hobbies". Higher scores indicate higher liking. Reported Cronbach's α for this scale (see Martin et al., 2017) ranged from .72 (Kazakhstan) to .90 (e.g. Austria or Hungary) with most countries displaying values above .8.

Teachers and schools-level variables

Classroom Instruction Limited by Student Attributes (Class Instr Lim by Stud) is a scale derived from the teachers' reports on the extent to which their classroom instruction in reading was limited by students' preparedness and readiness to learn (e.g., lacking skills; sleep-deprived; poor nutrition; disruptive students or with learning impairments). Higher scale scores indicate lower limitations on instruction. Reported Cronbach's α for this scale (see Martin et al., 2017) ranged from .61 (Italy) to .83 (e.g. Australia) with most countries displaying values above .7.

Safe and Orderly School (Safe and Ord Sch) is a scale derived from teachers' degree of agreement with statements such as "I feel safe at this school"; "This school has clear rules about student conduct" or "The students are respectful to teachers". Higher scores on the scale indicate higher safety and orderly schools. Reported Cronbach's α for this scale (see Martin et al., 2017) ranged from .61 (Georgia) to .90 (United States of America) with most countries displaying values above .8.

Time spent by teacher in Reading Instruction (Time Spent Read Instr) is a measure derived from principals' reports of total instruction time per year and teachers' reports on language instruction and reading times.

Teacher asks students to Reading Silently (Read Silently) is a measure of how often teachers ask their students to read silently (from "never or almost never" to "every day or almost every day").

Teachers Teaching Students strategies for Decoding Words and Sounds (Decoding Words) is a measure of how often teachers teach strategies for decoding words and sounds (from "never or almost never" to "every day or almost every day").

Teachers Teaching Students how to Summarize the main ideas of a Text (Summarize Main Ideas) is a measure of how often teachers teach students how to summarize the main ideas of the text (from "never or almost never" to "every day or almost every day").

Teachers asking Students to Locate Information in the Text (Locate Info) is a measure of how often teachers ask students how to locate the main ideas of the text (from "never or almost never" to "every day or almost every day").

Percentage of Economic Disadvantage Students (Economic Disad) a report by Principals' on the percentage of students coming from economically disadvantage homes (from "0 to 10%" to "more than 50%").

Shortage of Instructional Materials (Short Instruct Materials) a report by principals on how much the school capacity to provide instruction is affected by the shortage of instructional materials (from “not at all” to “a lot”).

Percentage of Students Entering with Literacy Skills (Stud Enter Lit Skills) is a measure derived from principals’ answers to 6 questions regarding the percentage of students that enter schools with reading literacy skills, e.g., “read some words”, “recognize most of the letters of the alphabet”, “read sentences” or “write letters of the alphabet” using a ranking from “less than 25%” to “more than 75%”.

Instruction Affected by Reading Materials Shortage (Instr Aff by Read Short) is a scale derived by principals’ reports how reading instruction is affected by 12 school and classroom resources (e.g., “instructional material (e.g., textbooks)”, “Teachers with a specialization in reading”, or “Computer technology for teaching and learning” reported on a rating scale from “not at all” to “a lot”. Higher scale scores indicate less shortage of resources. Reported Cronbach’s α for this scale (see Martin et al., 2017) ranged from .77 (The Netherlands) to .96 (United Arab Emirates) with most countries displaying values above .8.

Parents Expectations (Parent Expect) a measure by principals on their perception about parents’ expectations for student achievement (from “very high” to “very low”).

School Discipline (Sch Discipl) is a scale derived from principals’ answers to ten potential school discipline problems, e.g., “Classroom disturbance”, “Cheating” or “Vandalism” reported on a scale from “Not a problem” to “Serious problem”. Higher scale scores indicate more disciplinary problems. Reported Cronbach’s α for this scale (see Martin et al., 2017) ranged from .73 (Macao SAR) to .97 (Georgia) with most countries displaying values above .8.

School Emphasis in Academic Success (Sch Enph Acad Success) is a scale derived from principals’ answers to 13 questions regarding aspects of school emphasis on academic success, e.g., “Teacher’s understanding the school’s curricular goals”, “Parental expectations for student achievement” or “Students’ desire to do well in school” with higher scores indicating higher emphasis. Reported Cronbach’s α for this scale (see Martin et al., 2017) ranged from .84 (Germany) to .93 (Australia) with most countries displaying values around .85.

Testing and coding procedures

The administration of PIRLS tests and questionnaires in the participating countries followed a standardized procedure defined by IEA. The duration of the test was 80 min, with an interval of no more than 30 min at the end of the first 40 min. A student sociodemographic questionnaire (20 min) was applied at the end of the test (Johansone, 2017). The school, teacher, and family surveys were conducted before the testing sessions with no time limit to answer the questionnaires. Answers to multiple-choice and constructed-response test items were coded by a panel of national coders previously trained by the National Centers according to the coding scripts produced by the IEA. The inter-coder reliability was controlled by the IEA.

Estimation of students' reading literacy

Student reading literacy (proficiency) scores were estimated by the psychometrics team of the IEA-Boston College consortium using Item Response Theory and Imputation of missing values by regression on latent (conditioning) variables methods (Foy & Yin, 2017). The score estimates, presented in the form of five plausible values, are stored in the PIRLS database that was made available to the public in December 2017 (*TIMSS & PIRLS—PIRLS 2016 International Database*, n.d.). The plausible values are estimates of the score that a student is expected to obtain in reading, considering his/her latent aptitude in the evaluated reading purposes/processes and context variables that characterize the student. Standardized student performance estimates were then converted into the PIRLS scale, established in the first edition of the study, to vary between 0 and 1000 points, with an average of 500 points and a standard deviation of 100 points (Foy & Yin, 2017).

Statistical analysis

The identification of variables, scales, and indexes that were able to explain the variation of the reading literacy scores was made through a hierarchical linear regression of the five plausible values for reading literacy on student and family (level 1) and teachers and schools (level 2) explanatory variables. The level 1 (within) model for student i within school j is:

$$Y_{ij} = \beta_{0j} + \beta_{1j} \mathbf{W}_{ij} + \varepsilon_{ij} \quad (1)$$

where Y_{ij} is the reading literacy imputed from the five plausible values, β_{0j} are school random effects associated with the model intercept, β_{1j} are the fixed effects for the \mathbf{W}_{ij} level 1 (within) explanatory variables ($\beta_{1j} = \gamma_{10}$) and ε_{ij} are the random errors for student i within school j . The level 2 (between) model is:

$$\beta_{0j} = \gamma_{00} + \gamma_{01} \mathbf{B}_j + \mathbf{u}_{0j} \quad (2)$$

where γ_{00} is the common intercept for schools, γ_{01} are the fixed effects for the \mathbf{B}_j level 2 (between) explanatory variables and \mathbf{u}_{0j} are schools' random effects associated with the model intercept. The final mixed model, obtained by combining Eqs. (1) and (2), is:

$$Y_{ij} = \gamma_{00} + \gamma_{10} \mathbf{W}_{ij} + \gamma_{01} \mathbf{B}_j + \mathbf{u}_{0j} + \varepsilon_{ij} \quad (3)$$

where $\varepsilon_{ij} \sim N(0, \sigma)$. Only explanatory variables that could be predictors of reading literacy according to the Component Model of Reading (Aaron et al., 2008) were included in the models. Furthermore, variables displaying strong multicollinearity ($VIF > 5$) were removed before the HLM Analysis (see e.g., Montgomery & Peck, 1982). Thus, the models presented in this study do not include all the variables, indexes and psychometric scales (more than 100) that can be found in the international report (for a full description of all the scales and indices see Hooper et al., 2015) (Table 1 lists the variables present in the HLM models for each country).

After pairwise missing data deletion, the proportion of non-missing data in all variables was larger than 0.99 for all countries.

In the first step of the analysis, the intraclass correlation coefficient (ICC) for schools within countries was estimated from a basal model with no explanatory variables ($Y_{ij} = \gamma_{00} + \mathbf{u}_{0j} + \varepsilon_{ij}$) as $ICC = \sigma_{u_{0j}}^2 / (\sigma_{u_{0j}}^2 + \sigma_{\varepsilon_{ij}}^2)$ with corresponding design effect estimated as $1 + (n - 1) \times ICC$ where n is the average cluster size (Stapleton, 2013). ICC greater than or equal to 0.1 and design effect larger than 2 were considered an indicator of cluster effects that must be considered to obtain efficient estimates with the HLM model (Muthén & Satorra, 1995; Stapleton, 2013). In the second stage of the analysis, hierarchical linear models with random intercepts and homogeneous slopes between-schools were analyzed for each of the PIRLS 2016 participating countries. Level 1 explanatory variables (students and family) were centered by the grand mean of the country while level 2 explanatory variables (teachers and schools) were centered by the group mean. To account for the 2-stage sampling designs (PPS and cluster sampling), students (level 1) weights and teachers/schools (level 2) weights were recalculated from the information in the PIRLS 2016 international database according to procedures implemented in the IEA Data Analysis center (Stein-Planck, 2017, pers. comm.). Teacher variables were aggregated by the mean, for interval scaled variables, or the median, for ordinal scaled variables within schools. The validation of the model assumptions, namely the normal distribution of the residues, was safeguarded by the large sample size and consequent application of the central limit theorem. All HLM analyses were performed on the five plausible values for reading literacy using the imputation procedure implemented in the *Mplus* software (v 7.2, Muthén & Muthén, Los Angeles) with robust maximum likelihood estimation (MLR). A diagonal covariance structure for the null model and unstructured for the 2-level model were assumed. No other covariance structures were tested. Model fit was evaluated from the model's R^2 for level 1 and level 2 variables estimated as $R^2 = 1 - \sigma_{\varepsilon_{ij}}^2 / (\sigma_{u_{0j}}^2 + \sigma_{\varepsilon_{ij}}^2)$ which applies to both within- and between levels (Liu, Zumbo, & Wu, 2014). *Mplus* input files and results of the analyses retrieval were optimized through the R software (R Core Team, 2017) package *MplusAutomation* (Hallquist & Wiley, 2018). No missing data imputation was performed during the HLM analysis (pairwise deletion). The HLM models were fitted in three stages to account for model complexity and possible algorithmic convergence issues. In the first, only **W** variables were added, and the statistically significant ones selected; in the second, **B** variables were probed, selecting the statistically significant ones. Finally, in the third stage, the full two-level HLM model was fitted using the level 1 and level 2 selected variables in the two previous steps. Summary measures were obtained with the *skimr* package for R (McNamara et al., 2018). Student's t-tests were used to probe the significance of the overall mean standardized regression coefficients. Effects with $p < .05$ were considered statistically significant. Standardized regression coefficients below 0.10 ($\beta < 0.10$), even if statistically significant, were considered irrelevant effect sizes at this level of analysis.

Results

In the PIRLS 2016 edition, the Russian Federation ($M=581$, $SE=2.2$) followed by Singapore ($M=576$, $SE=3.2$), Hong Kong SAR ($M=569$, $SE=2.7$), Ireland ($M=567$, $SE=2.5$), and Finland ($M=566$, $SE=1.8$) were the top-5 performers (scores at or greater than the scale mean + $0.5SD$; green-colored in Fig. 2). Israel ($M=530$, $SE=2.5$), Portugal ($M=528$, $SE=2.3$), New Zealand ($M=523$, $SE=523$), France ($M=511$, $SE=2.2$), and Chile ($M=494$, $SE=2.5$) are mid performers with average proficiency around the mid-point of the PIRLS scale (yellow-colored in Fig. 2). Malta, ($M=452$, $SE=1.8$), UAE ($M=450$, $SE=3.2$), Morocco ($M=358$, $SE=3.9$), and South-Africa ($M=320$, $SE=4.4$) are examples of low-performers (scores at or below the Mean - $0.5SD$; red-colored in Fig. 2) (see Mullis et al., 2017a for the full description of the results). Figure 2 illustrates the average achievement of the PIRLS 2016 participants.

Looking at the PIRLS average achievement for participating countries, a strong linear relationship emerges between the mean PIRLS score and the between-schools' variation as probed by the Intraclass Correlation Coefficient (ICC) for schools within countries ($r=-.624$, $p<.001$). Countries with lower PIRLS scores show a larger variation between schools than countries with higher scores who display more homogenous schools (see Fig. 3a). The support for the 'Heyneman-Loxley' effect is, overall, weak. Only 19% of the ICC variation can be explained by the GDP per capita ($r=-0.436$; $p=.012$) (Fig. 3b). However, there are quite striking differences between countries suggesting that specific, within-countries, effects must be in place.

Student and family-level variables

An HLM regression of PIRLS plausible values on students and families (level 1), teachers and schools (level 2) with random intercepts and fixed slopes allowed for

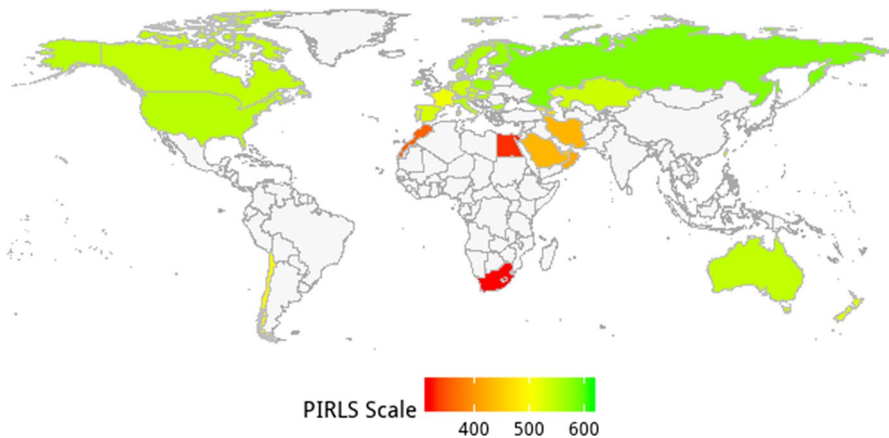


Fig. 2 PIRLS 2016 mean achievement per participating country. (Color figure online)

Table 1 Standardized HLM regression coefficients (β) and standard errors (S.E.) for PIRLS 2016 reading literacy regressed on student and family (Level 1), and teachers and schools (Level 2) predictors, models' R^2 , ICC and design effects (DE)

Level	Predictor	ARE (450)		AUS (544)		AUT (541)		AZE (472)		BGR (552)		BHR (446)		CAN (543)		CHL (494)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Student and Family	Sex (ASBG01) (0-girl, 1-boy)	-0.072	0.015	-0.082	0.028	-0.011	0.022	-0.057	0.031	-0.063	0.025	-0.077	0.026	-0.061	0.02	-0.004	0.033
	Often Speak Lang Test (ASBG03)	-0.023	0.015	-0.007	0.032	-0.145	0.025	-0.025	0.025	-0.038	0.032	-0.036	0.016	0.021	0.019	-0.086	0.028
	Home Res for Learn (ASB-GHRL)	0.052	0.019	0.065	0.033	0.179	0.028	0.085	0.034	0.078	0.036	0.008	0.019	0.116	0.022	-0.026	0.03
	Stud Bullying (ASBGSB)	0.049	0.016	0.01	0.029	0.033	0.023	0.048	0.032	0.097	0.023	0.047	0.018	0.049	0.017	0.102	0.029
	Stud Conf Reading (ASBGSCR)	0.316	0.016	0.386	0.033	0.304	0.02	0.204	0.028	0.26	0.026	0.314	0.019	0.35	0.017	0.273	0.029
	Stud Like Read (ASBGSLR)	0.011	0.019	-0.038	0.035	-0.034	0.025	0.05	0.032	-0.029	0.027	-0.004	0.018	-0.045	0.018	-0.016	0.03
	Stud Sense of Sch Belong (ASB-GSSB)	0.003	0.016	0.028	0.032	-0.005	0.021	0.084	0.034	-0.08	0.023	-0.011	0.019	0.043	0.02	0.033	0.034
	Level Ed Father (ASBH18A)	0.082	0.015	0.026	0.036	0	0.026	0.036	0.028	0.064	0.021	0.107	0.019	0.019	0.018	0.064	0.03
	Level Ed Mather (ASBH18B)	0.033	0.018	0.035	0.042	-0.008	0.024	-0.005	0.035	0.04	0.03	0.085	0.022	0.055	0.017	0.092	0.029
	Exp Level Ed Child (ASBH19)	0.059	0.016	0.086	0.033	0.165	0.021	0.118	0.036	0.119	0.036	0.147	0.019	0.076	0.022	0.129	0.031
	Early Lit Actv Befor Sch (ASB-HELA)	0.043	0.021	-0.033	0.04	0.052	0.034	-0.043	0.049	0.048	0.033	0.022	0.025	0.045	0.026	-0.022	0.044

Table 1 (continued)

Level	Predictor	ARE (450)		AUS (544)		AUT (541)		AZE (472)		BGR (552)		BHR (446)		CAN (543)		CHL (494)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
	Early Literacy Tasks (ASB-HELT)	0.191	0.019	0.176	0.03	0.019	0.02	0.082	0.043	0.12	0.024	0.222	0.016	0.14	0.019	0.207	0.032
	Early Num Actv Befor Sch (ASB-HENA)	-0.015	0.021	0.014	0.039	-0.058	0.029	0.019	0.041	-0.009	0.033	0.023	0.024	-0.041	0.028	0.019	0.037
	Parents Perce Sch (ASBHPCS)	0.006	0.016	-0.029	0.035	-0.012	0.027	0.006	0.031	-0.019	0.019	0.005	0.02	-0.038	0.016	-0.067	0.032
	Parents Like Read (ASBHPLR)	0.035	0.015	0.117	0.038	0.073	0.023	0.003	0.034	0.017	0.026	0.003	0.017	0.045	0.021	0.055	0.031
R ²	(Within)	0.261	0.012	0.326	0.029	0.3	0.02	0.159	0.022	0.271	0.02	0.328	0.016	0.264	0.059	0.244	0.024
Teacher	Class Instr Lim by Stud (ATBGSOSG)	0.171	0.045	0.284	0.082	0.038	0.096	0.244	0.128	0.164	0.113	0.039	0.073	0.172	0.073	0.235	0.089
	Safe and Ord Sch (ATBGSOSG)	0.157	0.069	0.114	0.1	0.198	0.128	-0.033	0.111	0.086	0.102	0.309	0.083	0.094	0.076	-0.049	0.057
	Time Spent Read Instr (ATBR07G)	0.014	0.039	0.119	0.089	0.029	0.069	0.093	0.107	-0.064	0.079	-0.023	0.05	0.013	0.086	0.01	0.077
	Read Silently (ATBR10CG)	0.094	0.047	0.124	0.088	-0.055	0.106	0.011	0.143	-0.001	0.109	-0.033	0.085	0.046	0.047	-0.048	0.105
	Decoding Words (ATBR10DG)	0.099	0.051	0.12	0.082	0.188	0.095	-0.067	0.105	-0.152	0.083	0.084	0.07	0.003	0.062	-0.003	0.079
	Summarize Main Ideas (ATBR10FG)	0.064	0.054	0.098	0.106	-0.007	0.106	-0.122	0.117	-0.254	0.094	0.076	0.072	-0.054	0.065	0.142	0.076
	Locate Info (ATBR12AG)	0.035	0.056	-0.239	0.093	0.244	0.083	0.331	0.134	-0.096	0.098	-0.022	0.068	-0.004	0.057	0.095	0.064

Table 1 (continued)

Level	Predictor	ARE (450)		AUS (544)		AUT (541)		AZE (472)		BGR (552)		BHR (446)		CAN (543)		CHL (494)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
School	Economic Disadv (ACBG03A)	-0.06	0.059	-0.14	0.118	-0.369	0.11	0.034	0.091	-0.002	0.088	-0.053	0.075	-0.124	0.068	-0.286	0.079
	Short Instruct Materials (ACB-G12BD)	-0.041	0.086	-0.108	0.129	-0.003	0.091	-0.029	0.135	0.159	0.093	-0.02	0.112	0.025	0.093	0.139	0.136
	Parent Expect (ACBG13H)	-0.121	0.065	-0.143	0.141	-0.106	0.112	-0.112	0.126	-0.011	0.112	-0.067	0.091	-0.164	0.079	0.043	0.141
	Sch Discipl (ACB-GDAS)	0.143	0.056	0.112	0.086	0.009	0.11	-0.053	0.094	-0.086	0.07	0.195	0.064	0.068	0.067	-0.047	0.109
	Sch Enph Acad Success (ACB-GEAS)	0.031	0.077	0.088	0.159	0.116	0.137	-0.336	0.168	0.359	0.109	0.101	0.109	0.082	0.098	0.453	0.147
	Stud enter Lit Skills (ACB-GELS)	0.297	0.056	0.048	0.102	0.111	0.086	0.13	0.109	0.18	0.1	0.092	0.074	0.041	0.062	0.29	0.073
	Instr Aff by Read Short (ACB-GRRS)	0.153	0.097	-0.167	0.125	-0.078	0.146	-0.263	0.104	-0.064	0.133	0.165	0.135	-0.089	0.077	-0.028	0.108
	Size School Library (ACDG09)	-0.073	0.051	-0.007	0.081	-0.012	0.091	-0.211	0.101	-0.141	0.071	-0.23	0.066	-0.014	0.058	-0.182	0.067
R ²	(Between)	0.514	0.043	0.594	0.073	0.499	0.097	0.314	0.098	0.595	0.063	0.511	0.067	0.276	0.013	0.717	0.073
ICC		0.50		0.19		0.15		0.39		0.41		0.29		0.21		0.27	
DE		18		5		5		14		12		9		5		8	

Table 1 (continued)

Level	Predictor	CZE (543)		DEU (537)		DNK (547)		ENG (559)		ESP (528)		FIN (566)		FRA (511)		GEO (488)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Student and Family	Sex (ASBG01) (0-girl, 1-boy)	-0.034	0.017	-0.012	0.025	-0.068	0.027	-0.06	0.019	-0.035	0.021	-0.09	0.017	-0.06	0.023	-0.058	0.033
	Often Speak Lang Test (ASBG03)	0.017	0.017	-0.107	0.025	-0.053	0.025	0.009	0.017	0.005	0.017	-0.068	0.02	-0.043	0.019	0.029	0.029
	Home Res for Learn (ASB-GHRL)	0.14	0.023	0.148	0.03	0.113	0.037	0.223	0.016	0.099	0.031	0.134	0.023	0.111	0.033	0.054	0.033
	Stud Bullying (ASBGSB)	0.066	0.015	0.012	0.027	-0.013	0.037	0.08	0.021	0.05	0.02	0.03	0.018	0.052	0.027	0.038	0.029
	Stud Conf Reading (ASBGSCR)	0.294	0.02	0.247	0.031	0.364	0.027	0.44	0.017	0.276	0.02	0.241	0.02	0.29	0.019	0.26	0.029
	Stud Like Read (ASBGSLR)	-0.03	0.019	0.025	0.036	0.057	0.034	-0.041	0.017	-0.023	0.021	0.055	0.02	0.001	0.022	-0.012	0.037
	Stud Sense of Sch Belong (ASBGSSB)	0.005	0.022	0.043	0.03	0.035	0.038	0.03	0.022	0.007	0.026	0.013	0.021	-0.023	0.023	0.006	0.03
	Level Ed Father (ASBH18A)	0.011	0.022	0.031	0.032	0.013	0.03	-	-	0.068	0.021	0.029	0.023	0.049	0.023	0.066	0.033
	Level Ed Mather (ASBH18B)	-0.015	0.02	0.018	0.029	0.054	0.035	-	-	0.031	0.024	0.023	0.026	-0.02	0.027	0.052	0.026
	Exp Level Ed Child (ASBH19)	0.224	0.021	0.227	0.033	0.04	0.031	-	-	0.135	0.019	0.103	0.02	0.245	0.026	0.177	0.031
Early Lit Actv Befor Sch (ASBELA)	0.015	0.022	0.079	0.037	0.022	0.054	-	-	-0.005	0.022	-0.02	0.024	-0.019	0.034	-0.032	0.033	

Table 1 (continued)

Level	Predictor	CZE (543)		DEU (537)		DNK (547)		ENG (559)		ESP (528)		FIN (566)		FRA (511)		GEO (488)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
	Early Literacy Tasks (ASB-HELT)	0.098	0.019	0.042	0.028	0.153	0.035	–	–	0.185	0.018	0.23	0.021	0.115	0.027	0.082	0.033
	Early Num Actv Befor Sch (ASBHENA)	–0.027	0.022	–0.096	0.035	–0.055	0.037	–	–	0.003	0.024	–0.023	0.027	0.021	0.028	–0.005	0.04
	Parents Perce Sch (ASBH-PCS)	–0.075	0.017	–0.016	0.022	0.011	0.025	–	–	–0.016	0.021	–0.018	0.018	–0.059	0.021	0.019	0.026
	Parents Like Read (ASBH-PLR)	0.078	0.021	0.062	0.042	0.026	0.032	–	–	0.025	0.021	0.074	0.02	0.049	0.026	0.041	0.027
R ²	(Within)	0.314	0.015	0.315	0.022	0.304	0.022	0.312	0.015	0.259	0.016	0.316	0.11	0.314	0.022	0.217	0.026
Teacher	Class Instr Lim by Stud (ATBGSLIG)	0.149	0.099	0.243	0.089	–0.006	0.132	0.312	0.015	0.158	0.063	0.177	0.138	0.262	0.109	0.057	0.137
	Safe and Ord Sch (ATBG-SOSG)	0.035	0.121	0.323	0.093	0.221	0.117	0.482	0.092	0.037	0.072	0.036	0.156	–0.01	0.114	0.238	0.103
	Time Spent Read Instr (ATBR07G)	–0.005	0.096	–0.086	0.098	0.233	0.104	0.027	0.13	–0.094	0.061	–0.009	0.12	0.201	0.094	–0.003	0.08
	Read Silently (ATBR10CG)	–0.116	0.113	–0.047	0.074	0.056	0.106	–0.067	0.09	0.121	0.056	–0.066	0.108	–0.044	0.153	0.094	0.102
	Decoding Words (ATBR10DG)	0.051	0.103	–0.133	0.088	0.196	0.109	0.026	0.093	0.07	0.063	0.012	0.106	–0.027	0.13	–0.11	0.087

Table 1 (continued)

Level	Predictor	CZE (543)		DEU (537)		DNK (547)		ENG (559)		ESP (528)		FIN (566)		FRA (511)		GEO (488)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
School	Summarize Main Ideas (ATBR10FG)	0.023	0.096	0.043	0.09	0.049	0.094	0.08	0.095	-0.106	0.064	0.265	0.136	-0.045	0.128	-0.217	0.135
	Locate Info (ATBR12AG)	0.182	0.112	-0.066	0.092	-0.223	0.111	0.077	0.095	0.049	0.077	-0.185	0.142	0.195	0.11	-0.106	0.132
	Economic Disadv (ACBG03A)	-0.417	0.095	-0.121	0.099	-0.214	0.097	-0.225	0.112	-0.005	0.089	-0.007	0.137	-0.448	0.088	-0.001	0.107
	Short Instruct Materials (ACBG12BD)	-0.079	0.155	0.114	0.105	-0.058	0.11	0.187	0.119	-0.09	0.085	-0.109	0.116	0.14	0.115	0.032	0.131
	Parent Expect (ACBG13H)	-0.029	0.144	0.024	0.09	-0.025	0.137	-0.235	0.145	-0.242	0.107	-0.216	0.138	-0.16	0.132	-0.148	0.112
	Sch Discipl (ACBGDAS)	0.055	0.094	-0.006	0.095	-0.024	0.108	0.209	0.099	0.176	0.085	0.038	0.143	0.034	0.104	0.105	0.125
	Sch Enph Acad Success (ACB-GEAS)	0.077	0.124	0.242	0.124	0.227	0.159	0.153	0.146	-0.136	0.092	0.144	0.142	-0.257	0.159	-0.018	0.123
	Stud enter Lit Skills (ACB-GELS)	0.29	0.096	-0.029	0.077	0.144	0.084	0.038	0.082	0.094	0.06	0.061	0.145	-0.049	0.107	-0.108	0.08
	Instr Aff by Read Short (ACBGRRS)	-0.181	0.108	0.125	0.127	0.071	0.131	0.137	0.127	0.285	0.111	-0.051	0.162	0.163	0.102	0.249	0.132
	Size School Library (ACDG09)	0.095	0.099	0.001	0.087	-0.326	0.112	-0.076	0.106	-0.203	0.056	0.043	0.109	0.191	0.101	-0.096	0.107
R ² (Between)		0.337	0.092	0.538	0.081	0.466	0.105	0.482	0.092	0.463	0.092	0.33	0.016	0.395	0.086	0.33	0.089
ICC		0.20		0.42		0.12		0.11		0.21		0.09		0.14		0.25	
DE		8		9		3		4		6		4		5		8	

Table 1 (continued)

Level	Predictor	HNK (569)		HUN (554)		IRL (567)		IRN (428)		ISR (530)		ITA (548)		KAZ (536)		LTU (548)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Student and Family	Sex (ASBG01) (0-girl, 1-boy)	-0.024	0.024	-0.065	0.026	-0.063	0.026	-0.173	0.053	-0.038	0.024	0.003	0.022	-0.062	0.027	-0.033	0.032
	Often Speak Lang Test (ASBG03)	0.074	0.029	-0.039	0.022	0.032	0.025	-0.016	0.034	0.019	0.027	-0.094	0.031	0.003	0.024	0.011	0.025
	Home Res for Learn (ASB-GHRL)	0.036	0.036	0.116	0.032	0.134	0.032	0.038	0.039	0.141	0.031	0.099	0.034	0.044	0.033	0.086	0.03
	Stud Bullying (ASBGSB)	0.006	0.027	0.029	0.023	0.049	0.021	0.005	0.028	-	-	-0.002	0.026	0.055	0.022	0.088	0.033
	Stud Conf Reading (ASBGSCR)	0.279	0.025	0.331	0.023	0.309	0.022	0.34	0.028	0.364	0.023	0.225	0.025	0.173	0.027	0.285	0.025
	Stud Like Read (ASBGSLR)	-0.018	0.036	-0.04	0.027	-0.005	0.031	0.048	0.028	0.002	0.022	-0.036	0.024	0.029	0.03	-0.028	0.034
	Stud Sense of Sch Belong (ASBGSSB)	0.085	0.025	0.03	0.02	0.001	0.029	-0.042	0.031	0.009	0.023	0.078	0.025	0.02	0.034	-0.035	0.027
	Level Ed Father (ASBH18A)	-0.022	0.029	0.026	0.019	0.001	0.029	0.054	0.034	0.06	0.022	0.05	0.024	0.006	0.029	0.03	0.024
	Level Ed Mather (ASBH18B)	0.016	0.03	-0.008	0.027	0.039	0.025	-0.021	0.044	-0.001	0.023	0.057	0.03	-0.008	0.032	0	0.026

Table 1 (continued)

Level	Predictor	HNK (569)		HUN (554)		IRL (567)		IRN (428)		ISR (530)		ITA (548)		KAZ (536)		LTU (548)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
	Exp Level Ed Child (ASBH19)	0.071	0.033	0.274	0.022	0.123	0.033	0.097	0.04	0.192	0.024	0.131	0.031	0.073	0.028	0.128	0.03
	Early Lit Actv Befor Sch (ASBHELA)	-0.016	0.04	-0.016	0.035	0.001	0.04	-0.042	0.043	0.007	0.031	0.011	0.033	-0.036	0.036	-0.072	0.026
	Early Literacy Tasks (ASBHELT)	0.193	0.029	0.036	0.029	0.178	0.026	0.038	0.039	0.095	0.02	0.099	0.028	0.11	0.029	0.266	0.023
	Early Num Actv Befor Sch (ASBHENA)	-0.068	0.037	-0.02	0.027	0.003	0.044	0.063	0.036	-0.003	0.029	-0.016	0.032	0.023	0.031	0.024	0.031
	Parents Perce Sch (ASBHPCS)	0.017	0.025	0.018	0.021	-0.007	0.021	-0.013	0.032	-0.002	0.022	0.021	0.023	-0.036	0.03	0.018	0.024
	Parents Like Read (ASBHPLR)	-0.018	0.028	0.039	0.025	0.065	0.024	0.088	0.032	-0.012	0.024	0.072	0.026	0.045	0.025	0.044	0.024
R ²	(Within)	0.174	0.017	0.335	0.018	0.327	0.027	0.228	0.025	0.298	0.019	0.22	0.024	0.097	0.015	0.347	0.022
Teacher	Class Instr Lim by Stud (ATBGSLIG)	0.302	0.095	0.185	0.078	0.262	0.104	0.197	0.096	0.196	0.078	0.109	0.098	-0.068	0.106	0.127	0.076
	Safe and Ord Sch (ATBGSOSG)	0.082	0.103	0.17	0.109	0.163	0.115	-0.19	0.105	-0.002	0.088	0.152	0.127	0.123	0.096	0.085	0.079
	Time Spent Read Instr (ATBR07G)	-0.031	0.088	0.063	0.08	-0.095	0.105	-0.219	0.071	0.121	0.059	-0.166	0.101	0.143	0.087	-0.036	0.099
	Read Silently (ATBR10CG)	0.092	0.087	-0.098	0.07	-0.083	0.104	-0.106	0.115	-0.004	0.071	0.118	0.095	-0.184	0.082	-0.009	0.069

Table 1 (continued)

Level	Predictor	HNK (569)		HUN (554)		IRL (567)		IRN (428)		ISR (530)		ITA (548)		KAZ (536)		LTU (548)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
School	Decoding Words (ATBR10DG)	-0.005	0.114	0.053	0.065	0.043	0.118	0.006	0.113	0.063	0.065	-0.061	0.083	-0.106	0.113	-0.028	0.108
	Summarize Main Ideas (ATBR10FG)	-0.009	0.136	0.13	0.064	-0.113	0.103	-0.084	0.109	0.108	0.077	0.235	0.092	0.137	0.064	-0.076	0.099
	Locate Info (ATBR12AG)	-0.027	0.11	0.159	0.059	-0.115	0.119	-0.055	0.082	0.044	0.071	-0.124	0.078	-0.001	0.124	-0.061	0.09
	Economic Disadv (ACBG03A)	0.024	0.12	-0.247	0.092	-0.1	0.14	-0.322	0.103	-0.275	0.096	-0.169	0.11	-0.276	0.105	-0.51	0.079
	Short Instruct Materials (ACBG12BD)	-0.044	0.138	0.108	0.091	0.176	0.122	-0.018	0.119	-0.218	0.132	-0.13	0.137	0.273	0.167	-0.212	0.102
	Parent Expect (ACBG13H)	-0.137	0.143	-0.216	0.132	-0.446	0.188	-0.259	0.129	-0.037	0.086	-0.214	0.123	-0.349	0.141	0.103	0.089
	Sch Discipl (ACBGDAS)	0.299	0.1	0.205	0.106	-0.016	0.118	0.053	0.086	-	-	0.258	0.094	-0.001	0.097	-0.072	0.09
	Sch Enph Acad Success (ACB-GEAS)	0.053	0.136	-0.049	0.128	-0.109	0.179	-0.026	0.125	-0.003	0.086	-0.168	0.181	-0.233	0.15	0.313	0.118
	Stud enter Lit Skills (ACB-GELS)	-0.098	0.125	0.032	0.06	0.145	0.148	-0.23	0.107	0.055	0.093	0.221	0.101	0.136	0.092	-0.013	0.078
	Instr Aff by Read Short (ACBGRRS)	-0.053	0.169	-0.048	0.105	0.069	0.134	0.075	0.113	0.198	0.116	-0.221	0.164	0.281	0.158	-0.065	0.087
Size School Library (ACDG09)	0.034	0.07	-0.158	0.074	0.145	0.081	-0.228	0.092	0.105	0.07	-0.222	0.124	0.005	0.102	-0.054	0.098	

Table 1 (continued)

Level	Predictor	HNK (569)		HUN (554)		IRL (567)		IRN (428)		ISR (530)		ITA (548)		KAZ (536)		LTU (548)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
R ²	(Between)	0.269	0.088	0.609	0.06	0.674	0.084	0.487	0.082	0.576	0.061	0.372	0.096	0.311	0.085	0.548	0.069
ICC		0.25		0.32		0.13		0.39		0.34		0.15		0.27		0.32	
DE		7		11		5		7		9		5		9		8	
Level	Predictor	LVA (558)		MAC (546)		MAR (358)		MLT (452)		NIR (565)		NLD (545)		NOR (559)		NZL (523)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Student and Family	Sex (ASBG01) (0-girl, 1-boy)	-0.049	0.022	-0.046	0.023	-0.088	0.036	-0.04	0.025	-0.087	0.036	-0.061	0.028	-0.098	0.023	-0.1	0.032
	Often Speak Lang Test (ASBG03)	-0.049	0.018	-0.067	0.02	0.031	0.031	-0.093	0.021	-0.012	0.036	-0.065	0.032	-0.067	0.019	-	-
	Home Res for Learn (ASB-GHRL)	0.117	0.026	0.103	0.025	-0.002	0.042	0.068	0.028	0.046	0.053	0.071	0.036	0.099	0.021	0.137	0.04
	Stud Bullying (ASBGSSB)	0.081	0.02	0.072	0.024	0.095	0.027	0.059	0.019	0.045	0.039	0.042	0.034	-0.018	0.028	0.06	0.036
	Stud Conf Reading (ASBGSCR)	0.293	0.024	0.287	0.022	0.252	0.032	0.332	0.021	0.335	0.045	0.26	0.023	0.334	0.022	0.375	0.028
	Stud Like Read (ASBGSLR)	-0.063	0.021	0.04	0.028	0.074	0.033	-0.027	0.024	0.001	0.043	0.079	0.028	0.051	0.024	-0.044	0.029
	Stud Sense of Sch Belong (ASBGSSB)	-0.007	0.027	0.031	0.021	0.021	0.029	0.077	0.027	-0.03	0.036	0.04	0.044	-0.003	0.026	0.031	0.037
	Level Ed Father (ASBH18A)	0.007	0.025	0.019	0.024	0.045	0.038	0.018	0.021	0.021	0.046	0.023	0.036	0.032	0.024	0.015	0.035

Table 1 (continued)

Level	Predictor	LVA (558)		MAC (546)		MAR (358)		MLT (452)		NIR (565)		NLD (545)		NOR (559)		NZL (523)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
	Level Ed Mather (ASBH18B)	0.007	0.027	-0.034	0.024	-0.008	0.037	0.001	0.023	0.063	0.047	-0.014	0.036	0.058	0.03	0.034	0.036
	Exp Level Ed Child (ASBH19)	0.118	0.029	0.074	0.021	0.116	0.041	0.237	0.021	0.194	0.04	0.26	0.03	0.088	0.024	0.067	0.032
	Early Lit Actv Befor Sch (ASBHELA)	-0.014	0.026	-0.043	0.023	-0.007	0.047	0.051	0.03	-0.015	0.052	0.041	0.04	0.012	0.026	0.058	0.045
	Early Literacy Tasks (ASBHELT)	0.222	0.022	0.189	0.026	0.17	0.04	0.154	0.023	-	-	0.061	0.032	0.17	0.021	0.058	0.033
	Early Num Actv Befor Sch (ASBHENA)	-0.049	0.031	-0.007	0.029	0.013	0.04	-0.024	0.034	0.056	0.055	-0.036	0.037	-0.048	0.024	0.028	0.04
	Parents Perce Sch (ASBHPCS)	-0.012	0.024	0	0.019	0.027	0.028	0.003	0.018	0.055	0.035	0.059	0.033	0.026	0.019	0.009	0.032
	Parents Like Read (ASBHPLR)	0.049	0.024	-0.008	0.024	0.007	0.039	-0.014	0.021	0.045	0.037	0.062	0.039	0.043	0.021	0.064	0.031
R ²	(Within)	0.292	0.017	0.22	0.016	0.214	0.027	0.321	0.107	0.292	0.028	0.294	0.022	0.312	0.019	0.618	0.08
Teacher	Class Instr Lim by Stud (ATBGSLIG)	0.004	0.079	0.164	0.161	0.199	0.094	0.196	0.144	0.484	0.19	0.251	0.157	0.252	0.117	-0.001	0.09
	Safe and Ord Sch (ATBGSOSG)	0.104	0.09	0.165	0.172	0.069	0.107	0.103	0.138	-0.23	0.143	0.078	0.168	0.195	0.108	0.028	0.113

Table 1 (continued)

Level	Predictor	LVA (558)		MAC (546)		MAR (358)		MLT (452)		NIR (565)		NLD (545)		NOR (559)		NZL (523)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
	Time Spent Read Instr (ATBR07G)	-0.059	0.074	-0.047	0.15	0.042	0.082	-0.017	0.162	0.149	0.164	0.006	0.138	-0.037	0.084	0.151	0.114
	Read Silently (ATBR10CG)	-0.021	0.1	-0.31	0.158	-0.055	0.07	-0.121	0.162	-0.083	0.184	-0.287	0.119	-0.064	0.124	-0.049	0.122
	Decoding Words (ATBR10DG)	-0.122	0.091	0.074	0.169	0.049	0.078	0.139	0.144	0.017	0.145	-0.343	0.135	0.133	0.106	0.22	0.081
	Summarize Main Ideas (ATBR10FG)	0.086	0.088	0.01	0.163	0.037	0.107	0.078	0.138	-0.111	0.182	0.203	0.11	-0.147	0.114	-0.12	0.099
	Locate Info (ATBR12AG)	-0.019	0.091	-0.332	0.159	-0.155	0.087	-0.011	0.142	0.007	0.146	0.039	0.119	-0.161	0.103	0.015	0.099
School	Economic Disadv (ACBG03A)	-0.312	0.12	0.043	0.137	-0.119	0.073	-0.058	0.148	-0.492	0.215	-0.217	0.174	0.035	0.096	-0.243	0.124
	Short Instruct Materials (ACBG12BD)	-0.233	0.136	-0.088	0.242	-0.094	0.112	0.194	0.185	-0.097	0.246	-0.15	0.15	-0.013	0.128	-0.153	0.165
	Parent Expect (ACBG13H)	-0.102	0.113	-0.072	0.215	-0.144	0.118	0.002	0.186	0.185	0.209	-0.192	0.191	-0.222	0.141	-0.359	0.149
	Sch Discipl (ACBGDAS)	0.112	0.096	-0.244	0.139	0.081	0.088	0.314	0.191	0.043	0.191	0.13	0.156	0.011	0.13	0.279	0.129
	Sch Enph Acad Success (ACB- GEAS)	0.269	0.116	0.16	0.137	0.086	0.125	0.084	0.194	0.1	0.172	0.015	0.2	0.198	0.14	-0.225	0.187
	Stud enter Lit Skills (ACB- GELS)	0.008	0.093	0.027	0.16	0.229	0.104	0.018	0.097	0.012	0.146	0.01	0.106	-0.111	0.092	0.166	0.104

Table 1 (continued)

Level	Predictor	LVA (558)		MAC (546)		MAR (358)		MLT (452)		NIR (565)		NLD (545)		NOR (559)		NZL (523)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
	Instr Aff by Read Short (ACBGRRS)	-0.231	0.16	-0.355	0.248	-0.045	0.137	0.049	0.178	-0.085	0.182	-0.19	0.161	-0.094	0.185	-0.109	0.155
	Size School Library (ACDG09)	-0.165	0.095	-0.191	0.19	-0.038	0.084	-0.001	0.142	-0.016	0.157	-0.014	0.108	-0.092	0.11	0.239	0.102
R ²	(Between)	0.54	0.083	0.55	0.127	0.395	0.079	0.345	0.018	0.491	0.133	0.518	0.115	0.42	0.091	0.304	0.026
ICC		0.15		0.27		0.45		0.13		0.13		0.12		0.09		0.32	
DE		5		20		7		6		5		5		3		10	
Level	Predictor	OMN (418)		POL (565)		PRT (528)		QAT (442)		RUS (581)		SAU (430)		SGP (576)		SVK (535)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Student and Family	Sex (ASBG01) (0-girl, 1-boy)	-0.115	0.018	-0.061	0.024	0.035	0.021	-0.101	0.019	-0.018	0.026	-0.229	0.079	-0.035	0.016	-0.01	0.015
	Often Speak Lang Test (ASBG03)	0.017	0.02	0.072	0.017	-0.01	0.02	-0.01	0.014	-0.022	0.024	0.141	0.024	0.007	0.017	-0.068	0.018
	Home Res for Learn (ASB-GHRL)	0.031	0.029	0.153	0.032	0.15	0.038	0.034	0.017	0.037	0.035	0.012	0.035	0.16	0.017	0.224	0.027
	Stud Bullying (ASBGSB)	0.003	0.017	0.094	0.022	0.024	0.019	0.067	0.018	0.065	0.03	0.081	0.028	0.074	0.013	0.057	0.017
	Stud Conf Reading (ASBG-SCR)	0.288	0.021	0.288	0.022	0.298	0.018	0.306	0.017	0.328	0.026	0.168	0.024	0.291	0.016	0.271	0.02
	Stud Like Read (ASBGSLR)	0.039	0.02	0.002	0.031	-0.087	0.024	-0.024	0.017	-0.055	0.024	0.104	0.027	0.004	0.016	-0.026	0.024

Table 1 (continued)

Level	Predictor	OMN (418)		POL (565)		PRT (528)		QAT (442)		RUS (581)		SAU (430)		SGP (576)		SVK (535)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
	Stud Sense of Sch Belong (ASBGSSB)	0.03	0.02	-0.07	0.026	0.104	0.024	0.046	0.016	-0.012	0.032	0.009	0.032	0.008	0.015	-0.056	0.026
	Level Ed Father (ASBH18A)	0.095	0.024	0.047	0.019	0.024	0.026	0.117	0.017	0.02	0.025	0.048	0.032	0.067	0.017	-0.03	0.017
	Level Ed Mather (ASBH18B)	0.089	0.019	0.017	0.025	0.024	0.036	0.068	0.015	0.055	0.029	0.081	0.027	0.028	0.017	-0.023	0.02
	Exp Level Ed Child (ASBH19)	0.141	0.019	0.173	0.028	0.135	0.021	0.08	0.016	0.183	0.038	0.109	0.028	0.105	0.015	0.254	0.018
	Early Lit Actv Befor Sch (ASBHELA)	0.02	0.023	-0.019	0.035	0.06	0.027	-0.01	0.019	-0.011	0.031	0.021	0.033	-0.02	0.02	0.019	0.027
	Early Literacy Tasks (ASBHELT)	0.222	0.021	0.182	0.02	0.097	0.021	0.185	0.015	0.134	0.032	0.122	0.031	0.202	0.014	0.044	0.021
	Early Num Actv Befor Sch (ASBHENA)	-0.016	0.021	-0.063	0.034	-0.051	0.029	0.002	0.021	-0.035	0.033	-0.017	0.032	-0.031	0.019	-0.016	0.025
	Parents Perce Sch (ASBHPCS)	-0.015	0.02	-0.037	0.024	-0.008	0.019	0.037	0.015	0	0.022	-0.019	0.033	0.003	0.016	-0.009	0.017
	Parents Like Read (ASBHPLR)	0.029	0.019	0.023	0.023	0.062	0.025	0.046	0.014	0.013	0.028	0	0.025	-0.011	0.014	0.051	0.018
R ²	(Within)	0.315	0.07	0.336	0.018	0.25	0.015	0.282	0.012	0.275	0.02	0.205	0.035	0.33	0.013	0.332	0.018

Table 1 (continued)

Level	Predictor	OMN (418)		POL (565)		PRT (528)		QAT (442)		RUS (581)		SAU (430)		SGP (576)		SVK (535)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Teacher	Class Instr Lim by Stud (ATBG-SLIG)	0.073	0.089	0.223	0.092	0.118	0.08	0.177	0.067	0.051	0.097	0.149	0.114	0.251	0.071	0.048	0.082
	Safe and Ord Sch (ATBGSOSG)	0.16	0.1	0.179	0.106	0.2	0.074	0.038	0.065	0.045	0.108	-0.016	0.111	0.207	0.072	0.095	0.087
	Time Spent Read Instr (ATBR07G)	-0.083	0.073	0.004	0.114	0.116	0.067	0.043	0.067	-0.167	0.097	0.095	0.067	-0.117	0.056	-0.067	0.063
	Read Silently (ATBR10CG)	-0.051	0.099	-0.027	0.088	-0.021	0.069	-0.162	0.076	-0.26	0.082	0.033	0.078	-0.068	0.082	0.021	0.069
	Decoding Words (ATBR10DG)	0.122	0.086	0.293	0.118	-0.038	0.073	0.081	0.078	-0.024	0.083	-0.034	0.097	0.076	0.088	0.155	0.063
	Summarize Main Ideas (ATBR10FG)	0.219	0.098	0.024	0.13	0.068	0.063	0.091	0.084	0.044	0.081	0.126	0.097	0.001	0.079	-0.031	0.074
	Locate Info (ATBR12AG)	-0.119	0.104	0.033	0.1	0.096	0.048	0.004	0.075	0.12	0.095	-0.075	0.087	-0.014	0.069	-0.031	0.072
School	Economic Disadv (ACBG03A)	-0.043	0.099	-0.348	0.116	-0.083	0.109	0.009	0.062	-0.071	0.1	-0.002	0.079	-0.287	0.084	-0.374	0.089
	Short Instruct Materials (ACBG12BD)	0.001	0.112	0.125	0.159	-0.072	0.092	0.094	0.131	-0.222	0.121	0.214	0.104	0.048	0.091	0.062	0.08
	Parent Expect (ACBG13H)	0.05	0.108	0.037	0.132	-0.152	0.12	-0.081	0.068	0.025	0.128	0.199	0.135	-0.01	0.116	-0.086	0.101
	Sch Discipl (ACBGDAS)	-0.034	0.089	0.098	0.11	0.099	0.085	0.027	0.068	0.067	0.1	0.46	0.078	0.088	0.065	-0.009	0.084
	Sch Enph Acad Success (ACB-GEAS)	0.45	0.123	0.242	0.159	0.313	0.151	0.069	0.084	0.172	0.108	0.294	0.152	0.128	0.106	0.273	0.09

Table 1 (continued)

Level	Predictor	OMN (418)		POL (565)		PRT (528)		QAT (442)		RUS (581)		SAU (430)		SGP (576)		SVK (535)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
	Stud enter Lit Skills (ACB-GELS)	0.043	0.111	-0.093	0.124	0.028	0.059	0.192	0.065	0.28	0.112	-0.073	0.104	-0.086	0.059	0.105	0.074
	Instr Aff by Read Short (ACB-GRRS)	-0.064	0.117	0.205	0.14	-0.048	0.137	0.432	0.128	-0.07	0.12	0.115	0.112	-0.056	0.095	-0.006	0.085
	Size School Library (ACDG09)	0.08	0.083	-0.013	0.099	-0.01	0.069	-0.058	0.064	-0.252	0.088	0.056	0.09	-0.13	0.067	-0.082	0.076
R ²	(Between)	0.35	0.016	0.528	0.09	0.508	0.079	0.365	0.069	0.421	0.074	0.366	0.074	0.467	0.062	0.536	0.067
ICC		0.34		0.13		0.19		0.43		0.31		0.48		0.25		0.31	
DE		11		5		5		18		8		12		10		8	
Level	Predictor	SVN (542)		SWE (555)		TTO (479)		TWN (559)		USA (549)		VLG (525)		WAL (497)		ZAF (320)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Student and Family	Sex (ASBG01)	-0.055	0.019	-0.045	0.023	-0.081	0.028	0.023	0.034	-0.033	0.018	-0.043	0.022	-0.056	0.027	-0.083	0.046
	Often Speak Lang Test (ASBG03)	-0.07	0.017	-0.066	0.028	-0.019	0.028	0.011	0.022	-0.063	0.021	-0.13	0.025	-0.05	0.022	-0.015	0.052
	Home Res for Learn (ASB-GHRL)	0.145	0.029	0.082	0.032	0.096	0.03	0.084	0.038	0.115	0.027	0.108	0.027	0.145	0.034	0.054	0.052
	Stud Bullying (ASBGSB)	0.024	0.02	0.023	0.026	0.023	0.026	0.026	0.021	0.379	0.029	0.017	0.023	0.031	0.024	0.142	0.046
	Stud Conf Reading (ASBGSCR)	0.284	0.023	0.27	0.022	0.397	0.026	0.245	0.032	-0.047	0.024	0.262	0.019	0.362	0.021	0.339	0.058

Table 1 (continued)

Level	Predictor	SVN (542)		SWE (555)		TTO (479)		TWN (559)		USA (549)		VLG (525)		WAL (497)		ZAF (320)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
	Stud Like Read (ASBGSLR)	0.027	0.023	0.046	0.031	-0.039	0.023	-0.024	0.042	0.093	0.022	0.007	0.022	0.001	0.024	0.029	0.057
	Stud Sense of Sch Belong (ASBGSSB)	-0.045	0.023	0.015	0.032	0.023	0.032	0.034	0.023	-0.125	0.105	0.08	0.026	0.016	0.027	0.09	0.056
	Level Ed Father (ASBH18A)	0.019	0.021	0.057	0.035	0	0.033	0.059	0.051	0.141	0.083	0.055	0.021	0.049	0.032	0.072	0.038
	Level Ed Mather (ASBH18B)	0.055	0.022	0.034	0.033	0.025	0.029	-0.007	0.04	0.006	0.066	-0.009	0.023	0.01	0.032	0.012	0.055
	Exp Level Ed Child (ASBH19)	0.199	0.026	0.057	0.03	0.157	0.024	0.173	0.033	-0.04	0.059	0.216	0.022	0.163	0.026	0.051	0.072
	Early Lit Actv Befor Sch (ASBHELA)	0.046	0.031	-0.077	0.034	0.012	0.032	0.044	0.038	0.094	0.068	0.105	0.027	0.021	0.024	-0.063	0.101
	Early Literacy Tasks (ASBHELT)	0.122	0.025	0.277	0.026	0.128	0.022	0.27	0.036	-0.176	0.083	0.008	0.02	0.092	0.02	0.122	0.048
	Early Num Actv Befor Sch (ASBHENA)	-0.046	0.03	0.008	0.032	-0.033	0.03	-0.02	0.034	0.14	0.092	-0.064	0.022	-0.048	0.023	0.073	0.077
	Parents Perce Sch (ASBHPCS)	-0.049	0.02	-0.013	0.023	0.044	0.026	-0.086	0.03	-	-	-0.035	0.02	-0.01	0.021	-0.068	0.061
	Parents Like Read (ASBHPLR)	0.067	0.021	0.072	0.03	0.021	0.028	-0.001	0.025	-	-	0.041	0.02	0.087	0.023	0.009	0.06
	R2 (Within)	0.331	0.135	0.317	0.022	0.314	0.021	0.31	0.023	0.19	0.024	0.277	0.018	0.337	0.023	0.27	0.038

Table 1 (continued)

Level	Predictor	SVN (542)		SWE (555)		TTO (479)		TWN (559)		USA (549)		VLG (525)		WAL (497)		ZAF (320)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
Teacher	Class Instr Lim by Stud (ATBGSLIG)	0.035	0.106	-0.014	0.098	0.115	0.081	0.095	0.074	-	-	0.312	0.095	0.211	0.104	0.026	0.102
	Safe and Ord Sch (ATBG- SOSG)	0.06	0.145	0.107	0.121	0.061	0.117	0.062	0.09	-	-	0.122	0.083	0.127	0.093	-0.218	0.099
	Time Spent Read Instr (ATBR07G)	0.061	0.126	0.016	0.082	-0.11	0.074	0.235	0.132	-	-	-0.074	0.074	-0.129	0.082	0.237	0.106
	Read Silently (ATBR10CG)	0.023	0.105	-0.243	0.131	0.086	0.101	0.103	0.1	-	-	-0.113	0.092	0.005	0.078	-0.229	0.103
	Decoding Words (ATBR10DG)	0.085	0.133	0.138	0.109	0.095	0.097	0.056	0.121	-	-	-0.049	0.092	0.053	0.094	0.177	0.162
	Summarize Main Ideas (ATBR10FG)	-0.205	0.128	-0.093	0.113	-0.053	0.101	0.07	0.101	-	-	0.039	0.08	0.103	0.101	0.059	0.125
	Locate Info (ATBR12AG)	0.07	0.138	0.094	0.121	0.064	0.099	0.069	0.1	-	-	0.078	0.077	-0.05	0.087	-0.116	0.104
School	Economic Disadv (ACBG03A)	-0.144	0.097	-0.162	0.163	-0.425	0.093	-0.627	0.105	-0.581	0.098	-0.443	0.114	-0.413	0.084	-0.68	0.16
	Short Instruct Materials (ACBG12BD)	-0.131	0.14	-0.06	0.142	-0.034	0.114	0.115	0.189	-0.087	0.104	-0.046	0.134	0.019	0.078	-0.025	0.104
	Parent Expect (ACBG13H)	-0.099	0.112	-0.142	0.2	0.029	0.141	0.366	0.111	-0.056	0.112	-0.108	0.112	-0.118	0.106	-0.219	0.14
	Sch Discipl (ACBGDAS)	-0.134	0.113	0.041	0.134	-0.209	0.093	-0.052	0.132	0.303	0.075	-0.005	0.088	-0.044	0.072	0.146	0.13

Table 1 (continued)

Level	Predictor	SVN (542)		SWE (555)		TTO (479)		TWN (559)		USA (549)		VLG (525)		WAL (497)		ZAF (320)	
		β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.	β	S.E.
	Sch Enph Acad Success (ACB-GEAS)	0.227	0.141	0.055	0.202	0.152	0.124	0.277	0.131	-0.237	0.12	0.035	0.107	0.078	0.137	-0.05	0.201
	Stud enter Lit Skills (ACB-GELS)	0.195	0.113	0.248	0.136	0.263	0.091	0.052	0.157	0.2	0.088	0.091	0.072	0.099	0.07	-0.147	0.113
	Instr Aff by Read Short (ACBGRRS)	-0.003	0.152	0.126	0.159	0.051	0.103	0.173	0.161	-0.146	0.104	-0.058	0.098	0.069	0.093	-0.104	0.115
	Size School Library (ACDG09)	-0.066	0.207	0.027	0.118	0.021	0.087	0.181	0.13	-0.126	0.094	0.001	0.083	-0.112	0.077	-0.238	0.138
	R2 (Between)	0.366	0.017	0.433	0.096	0.573	0.088	0.607	0.094	0.639	0.064	0.642	0.064	0.685	0.081	0.777	0.071
ICC		0.05		0.14		0.28		0.11		0.20		0.16		0.20		0.46	
DE		2		5		8		4		6		7		7		9	

Values between parenthesis are the countries reading literacy mean for PIRLS 2016. Average cluster size for all countries was 29 (SD=8)

ARE—The United Arab Emirates; AUS—Australia; AUT—Austria, AZE—Azerbaijan; BGR—Bulgaria; BHR—Bahrain, CAN—Canada, CHL—Chile; CZE—Czech Republic; DEU- Germany, DNK—Denmark, ENG—England, ESP—Spain; FIN—Finland; FRA—France, GEO—Georgia; HNK—Hong Kong SAR; HUN—Hungary; IRL—Ireland; IRN—Republic Iran; ISR—Israel; ITA—Italy; KAZ—Kazakhstan; LTU—Lithuania; LVA—Latvia; MAC—Macao SAR; MAR—Morocco; MLT—Malta; NIR—Northern Ireland; NLD—The Netherlands; NOR—Norway; NZL—New Zealand; OMN—Oman; POL—Poland; PRT—Portugal; QAT—Qatar; RUS—Russian Federation; SAU—Saudi Arabia; SGP—Singapore; SVK—Slovakia; SVN—Slovenia; SWE—Sweden; TTO—Trinidad and Tobago; TWN—Taiwan; USA—United States of America; VLG—Belgium (Flemish); WAL—Belgium (French); ZAF—South Africa

The Predictors' codes between parenthesis are the PIRLS 2016 indices and variables names: “-” indicates variables not measured or indices not derived; standardized coefficients (β) in bold-face are statistically significant at $p < .05$

the identification of key explanatory variables that, at each level within-countries, have considerable effect sizes on accounting for the variation of within-country PIRLS 2016 results. Table 1 summarizes the regression coefficients for the selected level 1 and level 2 explanatory variables.

For top-performer countries like the Russian Federation or Singapore (see RUS and SGP in Table 1), the key explanatory variables of student reading literacy are students' confidence in reading (a scale reflecting how much a student is confident on what he/she reads and how good he/she feels at it; see Martin et al., 2017), the parents' expectations on the level of education their child will achieve, and children early literacy tasks when beginning primary school (a scale that evaluates how well the child could do tasks like recognizing most letters of the alphabet, read some words, write letters, etc.). Home resources for learning (a scale derived from home possessions, home study support for learning, and parents' highest level of education and occupation) was a relevant explanatory variable for Singapore, but not for the Russian Federation or Hong Kong SAR, another top-performer. For mid-performer countries with scores around the 500 points PIRLS mean, like Portugal, New Zealand, France, or Chile, the same family and student-level variables (students' confidence in reading, early literacy tasks, and home resources for learning) emerged as the most relevant explanatory variables. For low-performing countries like Morocco or South Africa, students' confidence in reading, students' liking to read, and early literacy tasks were important explanatory variables, but not home resources for learning—in contrast to what was observed for the most top and mid-performing countries. Student bullying (a scale created from the frequency of bullying behaviors, from about weekly to almost never, with higher scores indicating the lower frequency of bullying), was also an explanatory variable for some low-performers (e.g., South Africa), as well as mid-performers (e.g., Chile) or top-performers (e.g., Bulgaria).

Teachers and schools-level variables

At the teacher level, there was a large variation between countries in teachers' strategies and attitudes associated with better PIRLS results. For the Russian Federation, giving students time to read silently (varying from every day to never or almost never) was the teaching strategy associated with better student's results. However, the amount of time spent on reading instruction was the strongest, and a negative explanatory variable of Singapore's students' performance in PIRLS. In some mid-performing countries, like Portugal, locating information was the strongest explanatory variable of PIRLS results, while for others, like Slovakia, decoding words strategies was the strongest explanatory variable. In some low-performers, like South Africa or the United Arab Emirates, time spent by students reading silently and decoding words were the teachers' strategies more strongly associated with better PIRLS results. Problems with discipline and students' difficulties limiting classroom instructions (an inverted scale derived from students lacking prerequisite knowledge or skills, disruptive students, uninterested student, etc.) were common explanatory variables in low-performers (e.g., the United Arab Emirates

Table 2 Mean, SE, Percentile 25 (P25), Percentile (P50) and Percentile 75 (P75), and histograms of standardized coefficients (β) obtained from HLM regression of PIRLS 2016 plausible values on student and family (Level1) and teachers and schools (Level2) predictors for all the countries participating in PIRLS 2016 who dad measured the study variables/indices

Level	Variable	n	Mean	S.E.	P25	P50	P75	Histogram	t (df)	p-value
Student and Family	Sex (0-Female, 1-Male)	48	-0.058	0.006	-0.078	-0.059	-0.035		-9.667(47)	<.001
	Often Speak Lang Test	47	-0.025	0.008	-0.066	-0.022	0.010		-3.125(46)	.003
	Home Res for Learn	46	0.091	0.008	0.052	0.098	0.130		11.375(45)	<.001
	Stud Bullying	46	0.047	0.005	0.024	0.048	0.067		9.4(45)	<.001
	Stud Conf Reading	48	0.300	0.008	0.270	0.290	0.330		37.5(47)	<.001
	Stud Like Read	48	-0.001	0.006	-0.031	-0.005	0.029		-0.167(47)	.868
	Stud Sense of Sch Belong	48	0.018	0.006	-0.004	0.018	0.036		3(47)	.004
	Level Ed Father	46	0.038	0.005	0.019	0.032	0.058		7.6(45)	<.001
	Level Ed Mather	46	0.024	0.005	-0.007	0.024	0.053		4.8(45)	<.001
	Exp Level Ed Child	46	0.140	0.009	0.090	0.130	0.180		15.556(45)	<.001
	Early Lit Actv Befor Sch	46	0.005	0.006	-0.020	0.004	0.036		0.833(45)	.409
	Early Lit Tasks	45	0.140	0.010	0.095	0.130	0.190		14(44)	<.001
	Early Num Actv Befor Sch	46	-0.014	0.005	-0.040	-0.016	0.012		-2.8(45)	.008
	Parents Perce Sch	46	-0.008	0.005	-0.019	-0.009	0.010		-1.6(45)	.117
	Parents Like Read	46	0.039	0.005	0.014	0.043	0.062		7.8(45)	<.001
R ² (Within)	48	0.28	0.008	0.26	0.3	0.33		35(47)	<.001	
Teacher	Class Instr Lim by Stud	48	0.150	0.016	0.050	0.160	0.230		9.375(47)	<.001
	Safe and Ord Sch	48	0.086	0.017	0.037	0.094	0.160		5.059(47)	<.001
	Time Spent Read Instr	48	0.010	0.016	-0.065	0.005	0.071		0.625(47)	.535
	Read Silently	48	-0.045	0.016	-0.087	-0.045	0.021		-2.813(47)	.007
	Decoding Words	48	0.037	0.016	-0.027	0.052	0.096		2.313(47)	.025
	Summarize Main Ideas	48	0.014	0.017	-0.078	0.038	0.087		0.824(47)	.414
	Locate Info	48	-0.004	0.019	-0.080	-0.003	0.072		-0.211(47)	.834
School	Economic Disadv	48	-0.220	0.027	-0.350	-0.190	-0.057		-8.148(47)	<.001
	Short Instruct Materials	48	-0.009	0.017	-0.091	-0.027	0.098		-0.529(47)	.599
	Parent Expect	48	-0.094	0.022	-0.170	-0.110	-0.011		-4.273(47)	<.001
	Sch Discipl	48	0.069	0.020	-0.011	0.054	0.130		3.45(47)	.001
	Sch Emph Acad Sucess	48	0.089	0.026	-0.007	0.087	0.230		3.423(47)	.001
	Stud enter Lit Skills	48	0.076	0.019	0.010	0.058	0.170		4(47)	<.001
	Instr Aff by Read Short	48	0.005	0.023	-0.086	-0.046	0.130		0.217(47)	.829
	Size School Library	48	-0.054	0.019	-0.150	-0.046	0.009		-2.842(47)	.007
R ² (Sch + Teach: Between)	48	0.480	0.019	0.370	0.490	0.570		25.263(47)	<.001	

Bold predictors have a mean value greater than or equal to 0.10 (lower limit for a medium effect size)

and Morocco), mid-performers (e.g., Poland), and top-performers (e.g., Singapore, Northern Ireland) alike.

As was the case for teacher-level factors, considerable differences between countries in terms of school-level reading attainment explanatory variables were also observed. A higher percentage of economically disadvantaged students was associated with lower performance in PIRLS in countries like South Africa (low-performer), Flemish Belgium (mid-performer) or Singapore, and Northern Ireland (top-performers). School emphasis in academic success was identified as an important explanatory variable in low-achieving countries like Oman, mid achieving like Portugal, and also top achievers like Taiwan.

An exploration of Table 1 reveals a common pattern for student and family-level variables associated with PIRLS results, but a strong variation in teacher and school-related variables. A summary of all the explanatory variables in all PIRLS countries with data available for all variables studied is given in Table 2.

The analysis of Table 2 reveals that confidence in reading, parents' higher expectations about their child education, and early literacy tasks before entering primary education were the strongest explanatory variables of PIRLS 2016 reading literacy scores at the students and family level ($p < .001$). However, the effect on average PIRLS attainment described by these explanatory variables varied considerably between countries. On average, student and family variables accounted for 28% of the within-school PIRLS variation. At the teacher level, classroom instruction limited by students' needs was the most common, and negative, explanatory variable of PIRLS results across countries ($p < .001$), with no effective teaching strategies proving to be common to most countries. Finally, and at the school level, the percentage of economically disadvantaged students was the negative explanatory variable of PIRLS results that emerged consistently among participating countries ($p < .001$). School emphasis on academic success emerged as an important and strong explanatory variable in some countries, both low, mid, and top-performers, but the overall common effect across the 48 countries analyzed was quite weak (mean standardized effect of .09) although statistically significant ($p = .001$). On average, teacher and school variables accounted for 48% of the variation of PIRLS results between schools. It is worthwhile to note that a country's poor performance in PIRLS was negatively associated with the between-schools variation ($p < .05$).

Discussion

PIRLS 2016 differences between top-performers and low-achievers were around 200 points or two standard-deviations on the PIRLS scale (see Fig. 2). This means that 4th-grade students in low-performing countries were about four school years behind students from top-performing countries in the same grade as far as reading literacy is concerned. Analysis of within versus between-schools' variation reveals that low-performers have higher variation due to school-level factors when compared to top-performers. For example, in Morocco, ca. 45% of the variation in PIRLS results is attributable to schools, while for top performers like Finland less than 10% of the variation in PIRLS results is due to school differences. A strong negative linear relationship emerges from poor performance in PIRLS and larger school variation within countries ($r = -.62$; $p < .05$).

Analysis of student and family-level variables, at the within school level, revealed some variation but a common set of explanatory variables common to low-performers, mid performers, and top performers countries emerged. Student confidence in reading (how much the students thinks he/she is a good reader), parents' expectation on their children future education, and early literacy task before entering primary education were consistent explanatory variables of PIRLS scores. More than 50% of the participating countries in the PIRLS 2016

edition showed that the student confidence in reading had a standardized effect of at least .29 ($P50 = .29$). That is a change in one standard deviation in the student confidence in reading results in at least 29 points increases in the PIRLS score scale for half of the participating countries. A related construct, reading self-concept, was identified as a significant explanatory variable of reading literacy in PIRLS 2001 for Hong-Kong SAR students (Tse et al., 2005), Australian students, and other 14 countries that participated in PIRLS 2011 (Guo et al., 2018). Similar effects have been observed in several European countries taking part in previous editions of PIRLS (Araújo & Costa, 2015; Costa & Araújo, 2017; Netten, Voeten, Droop, & Verhoeven, 2014). The effect of parents' expectations about their children's future education had a more modest effect across countries. An increase in one standard deviation on the level of expectation resulted, across countries, in at least 13 points increase in the PIRLS score for half of the participating countries ($P50 = .13$). The same result was seen in the early literacy tasks scale. Similar results, as well as expectations and practices of early literacy activities before entering primary education of Canadian children, were reported by Martini and Sénéchal (2012) in their home literacy model, and for a few countries in Park's 2008 study with 25 countries who took part in the PIRLS 2001 study.

In 27 out of the 48 countries that took the full PIRLS test in 2016, girls significantly outperformed boys (see Table 1). However, the average gender effect was around 6 points in the PIRLS scale (see Table 2). When considering the effect of other student and family level on the students' PIRLS scores, the gender effect was not statistically significant for the other 21 participating countries. This contrasts with the PIRLS consortium published information for the gender gap that reported girls scoring significantly higher than boys in 48 countries out of the 50 PIRLS and PIRLS Literacy participants that took the test in 2016 (Mullis et al., 2017a). The results presented in this paper, obtained from the analysis of gender differences after considering other student and family variables, contradicts the previous emphasis on the reading gender gap. Nonetheless, when considering single country studies, there are countries where the gender gap is inexistent (e.g., Portugal and Macao SAR, see Marôco, 2018), mid-sized (e.g. Ireland, see Eivers, Gilleece, & Delaney, 2017) or relatively large (e.g., Saudi Arabia or South Africa, see Spaull, 2017).

Another result that is at odds with previously published research on both PIRLS and other International Large Scale Student Assessments (ILSA) like TIMSS (see, e.g., Ólafsson et al., 2014) or PISA (see, e.g., Karakolidis et al., 2016) is the positive effect of family socioeconomic status on literacy, although with a strong heterogeneity among students and countries (Lagravinese, Liberati, & Resce, 2019; Park, 2008). For example, the PISA's Economic and Socio-Cultural Status (ESCS) was a strong explanatory variable of mathematics achievement in Malaysia but not Singapore (Thien & Ong, 2015). In this study, when taken together with other student and family-level variables like early literacy tasks, socio-economic related variables like home resources for learning did not show a consistent effect across countries. Indeed, the mean effect for this explanatory variable over the 48 analyzed countries was .091 ($P50 = .098$). For about half of the participating countries, an increase in one unit of the home resources for learning scale results in a modest increase of nine points in the PIRLS scale. On the other hand, this study confirms the results of an

earlier analysis of PIRLS 2011 data from four European countries where student early literacy skills, home literacy practices, resources, and behaviors were strongly associated with PIRLS scores (Costa & Araújo, 2017). The association of home resources for learning and early literacy tasks (a toddler can only read or play with a book if a book is available at home, or preschool...) may result in possible suppression effects, although this was not a testable hypothesis in this study. Consistently with this hypothesis, Park (2008) observed a U-shape effect of the number of books at home with the country's economic level on reading literacy.

When considering the influence of teacher-related variables on students' PIRLS achievement, no teacher-related variables (e.g., teaching strategies or teachers professional experience) emerged consistently on all the 48 countries analyzed other than students' difficulties hindering learning. Analysis from other PIRLS editions has pointed out the importance of learning-oriented teaching strategies. Cheung et al. (2017) noted that the results of Hong Kong students in the PIRLS 2006 were significantly correlated with teaching strategies and activities that promoted silent reading. However, the positive effect of teacher strategies on student literacy does not appear to be universal. For example, when analyzing PIRLS 2006 data Shiel and Eivers (2009) found contradictory relationships between student performance and different types of strategies and resources used by teachers from different countries. In this study, in only but a few countries, either low-performers like South Africa, mid-performers like Slovakia and Spain, or top-performers like the Russian Federation, did time spent in reading silently explained PIRLS results. Overall, the mean standardized effect was .045 ($P50 = .045$). There is no common denominator for all countries and education systems that can easily identify the resources and strategies that teachers can use to improve reading literacy.

At the level of school-related variables, school emphasis on the students' academic success has been pinpointed as a relevant explanatory variable in the PIRLS 2016 edition in several individual countries. This was the case in mid-performers like Portugal (Marôco, 2018) and several low performers like Oman, and high performers like Taiwan (Mullis et al., 2017a). But again, this effect was not consistent across countries. Furthermore, some countries displayed effects contrary to the expectation, like what happened with the USA's negative effect of the emphasis on academic success on PIRLS results (see the USA in Table 1). The most consistent school-level explanatory variable of (negative) PIRLS achievement across countries was the percentage of disadvantaged students. Schools with a higher proportion of economically disadvantaged students had consistently lower scores somehow masking the effect of the type of school governance (public or state-funded vs. private owned). This effect was observed in low-performing (e.g. South Africa's $\beta = -0.68$, see Table 1), mid-performing (e.g. Poland's $\beta = -0.348$, see Table 1), and top-performing countries (e.g. Taiwan's $\beta = -0.68$, see Table 1) alike.

Overall, these results on reading literacy using PIRLS 2016 data, do not support a comprehensive Heyneman-Loxley' effect. Student and family-related variables that emerged consistently among the top, mid and low performing countries were quite similar, while school-related variables, like percentage of economically disadvantaged students, although accounting for a larger fraction of the within-country achievement variance ($r^2 = .48$ vs. $r^2 = .28$), were not stronger explanatory variables

at the low-achieving countries than at mid or top-performing countries. Taking the percentage of disadvantaged economic students at school as a proxy for a country's economic health, no stronger and consistent effect of this variable was observed between countries with the higher gross national product (see Fig. 3) like Hong Kong SAR ($\beta=0.024$), Singapore ($\beta=-0.287$) or Taiwan ($\beta=-0.627$) versus low-performing and less economically advantaged countries like Oman ($\beta=-0.043$), Morocco ($\beta=-0.312$) or South Africa ($\beta=-0.68$). These observations are in line with the previous conclusions from Baker et al. (2002) on the attenuation of the socio-economic and national economic development effects on student achievement across developing and developed countries.

Conclusions and Recommendations

Results from this study (see Table 2) demonstrate that models for reading literacy achievement are complex, and involve different hierarchical variables (students and families, teachers, and schools). Although the PIRLS framework and the Aaron et al. Component Model of Reading (Aaron et al., 2008; Joshi & Aaron, 2000), set the stage for the cognitive, psychological and environmental explanatory variables of reading literacy, PIRLS data show strong variation in the explanatory variables of reading literacy between countries and even within countries. No single, universal set of explanatory variables function homogeneously for all children and countries.

The analysis of 48 countries that took part in the PIRLS 2016 assessment and provided student, family, and teacher-level data, as well as school questionnaires shows that countries with lower PIRLS scores displayed higher between-schools variation, thus confirming the hypothesis set by research question 1. On average, between schools' variation accounted for 25% of the countries' PIRLS scores. The percentage of disadvantaged students surfaced as the most consistent school variable explaining PIRLS results. Regarding research question 2, at the student and family level, student confidence in reading, early literacy tasks, and parent expectations were the strongest explanatory variables of reading literacy in PIRLS 2016. Although profusely reported before, gender and home resources for learning were less important explanatory variables of reading literacy when the effects other student and family-related variables were accounted for. At the school and teacher level, teacher perception of the class instruction limited by students' needs and weaknesses was the strongest explanatory variable of PIRLS achievement. No common teaching strategies or other school-related variables emerged consistently in the analyzed countries, disproving the hypothesis proposed by research question 3. These data also disprove the hypothesis that school-related variables are more important explanatory variables of student achievement than student-related variables for low-performer countries when compared with mid or top-performer countries. No consistent pattern was observed between healthier versus poorer countries, corroborating the hypothesis that the Heyneman-Loxley Effect is no longer in place as suggested by others (See Baker et al. 2002). Although school and teacher-related variables account for a larger amount of student reading literacy variance than student

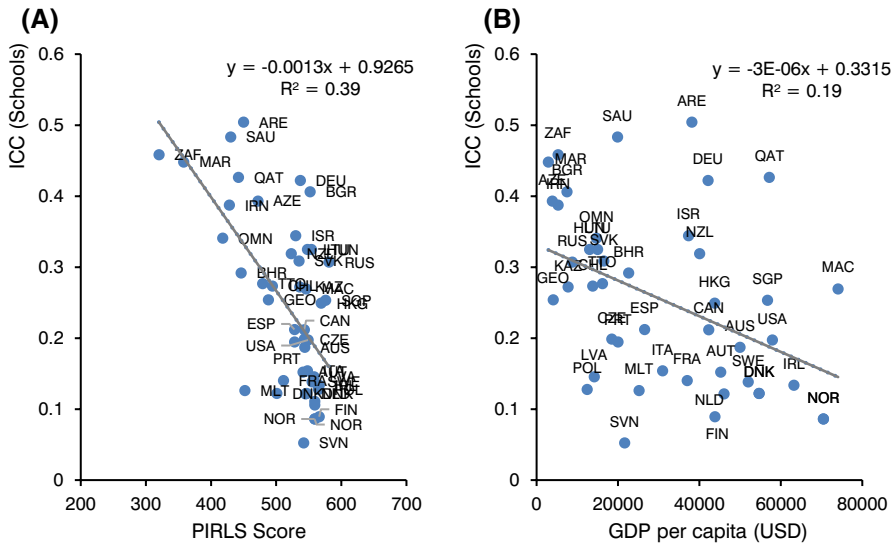


Fig. 3 Relationship between the PIRLS 2016 mean score and the Intraclass Correlation Coefficients (ICC) for schools within countries ($r = -0.62$; $p < .001$) (a); and between GDP per capita and ICC (b). Countries are represented by 3-letter ISO codes. GDP Data was retrieved from the World Bank (<https://data.worldbank.org/indicator/NY.GDP.PCAP.CD>)

and family-related variables, there was no clear pattern for the overall relevance of school and teachers' variables across countries in all levels of country performance.

An increase in the number of formal schooling years across countries, as well as a mass schooling investment, focused on school and teaching quality backed by nationwide uniformization of curricula, learning goals, and teacher training may be the cause of lower effects of school and teacher-related variables on student achievement. On the other hand, the parental expectation on children's education and families' home resources for learning can overcome schools' limitations for learning. This does not imply that within-countries school variation does not occur. As the data shows, low-performing countries do have larger between-schools variation than top-performing countries. However, no common effects of school and teacher-related variables can be generalized to a one-size-fits-all recommendation.

The results presented in this study are significant for both evidence-based policy recommendations and education practices aimed at improving reading literacy. First, the results show that many student and family-related variables are strong explanatory variables of reading literacy variation within countries. A common set of variables to most of the countries include early literacy tasks, parents' expectations, and student confidence in reading. Results presented in this study suggest that investment in the development of pre-school/kindergarten systems can give children a leading advantage for their schooling years. Also, improving parents' expectations on their children's education, through e.g. adult-education programs reinforcing the value of education, skills, and literacy competencies, can lead to better student achievement. Finally, improving student confidence in reading, through e.g. programs to encourage reading both for literary as well as information purposes, can improve

student academic achievement, not only in reading but also in all other disciplines since competence in reading is a must for learning. At the teacher level, policies oriented at reducing the students' limitations hindering learning (e.g., remedial classes, or recovery strategies for students showing learning difficulties) is, according to the PIRLS 2016 data, the most rewarding strategy.

The multi-level analysis considering the complex sampling features of PIRLS—a methodological approach that should be chosen more often with ILSA data (Jerim & Lopez-Agudo, 2017)—in this study draws attention to the fact that individual explanatory variable effects (like gender) may be over or underestimated when considered individually. This corroborates the findings of other studies that analyzed ILSA data other than PIRLS's. Since explanatory variables are not fully independent from each other, their effects should be considered together and not individually as is the frequent cause for the international reports of ILSA and the secondary data analysis research that generally follows. Results gathered in this study show that there is a large between-countries variation in both reading achievement and its explanatory variables recommending against the generalized adoption of policies and practices from top-performing countries as a short-term fix to improve reading literacy in mid or low-performing countries. Even within low-performing countries, there are top-achiever students and schools whose policies and education practices are better suited for the countries' socioeconomic and cultural landscape than possibly those from top-performing countries.

Despite the importance of these conclusions, PIRLS has limitations that must be considered. PIRLS and similar ILSA studies are sample-based studies with non-participation and biases in the selection of students possibly hindering the generalization of conclusions. Cross-cultural differences between-countries and regions may also hinder the transcultural validity of both achievement tests and psychometric scales. Lack of strong measurement invariance, as well as different coverage of the students' abilities by the test items, may render the between-countries mean score comparisons untruthful. Additionally, PIRLS, like all other ILSA, is a correlational study by nature and thus causal inference from significant correlations and regression models' coefficients may not be more than a form of statistical fantasy. Proper care must be exerted before the acceptance of causal effects.

Acknowledgements The author thanks Joana Andrade from the National Foundation for Educational Research in England and Wales (NFER) Center for Statistics for the critical review of the manuscript.

References

- Aaron, P. G., Joshi, R. M., Gooden, R., & Bentum, K. E. (2008). Diagnosis and treatment of reading disabilities based on the component model of reading. *Journal of Learning Disabilities, 41*(1), 67–84. <https://doi.org/10.1177/0022219407310838>.
- Araújo, L., & Costa, P. (2015). Home book reading and reading achievement in EU countries: the Progress in International Reading Literacy Study 2011 (PIRLS). *Educational Research and Evaluation, 21*(5–6), 422–438. <https://doi.org/10.1080/13803611.2015.1111803>.
- Baker, D. P., Goesling, B., & Letendre, G. K. (2002). Socioeconomic status, school quality, and national economic development: A cross-national analysis of the “Heyneman-Loxley effect” on

- mathematics and science achievement. *Comparative Education Review*, 46(3), 291–312. <https://doi.org/10.1086/341159>.
- Cheung, W. M., Lam, J. W. I., Au, D. W. H., So, W. W. Y., Huang, Y., & Tsang, H. W. H. (2017). Explaining student and home variance of Chinese reading achievement of the PIRLS 2011 Hong Kong. *Psychology in the Schools*, 54(9), 889–904. <https://doi.org/10.1002/pits.22041>.
- Costa, P., & Araújo, L. (2017). Skilled students and effective schools: Reading achievement in Denmark, Sweden, and France. *Scandinavian Journal of Educational Research*, 3831, 1–15. <https://doi.org/10.1080/00313831.2017.1307274>.
- Ebbs, D., & Wry, E. (2017). Translation and layout verification for PIRLS 2016. In I. V. S. Mullis, & M. Hooper (Eds.), *Methods and procedures in PIRLS 2016* (pp. 1–15). TIMSS & PIRLS International Study Center. Lynch School of Education, Boston College and International Association of the Evaluation of Educational Achievement. Retrieved July 18, 2018, from <https://timssandpirls.bc.edu/publications/pirls/2016-methods/chapter-7.html>.
- Eivers, E., Gilleece, L., & Delaney, E. (2017). *Reading achievement in PIRLS 2016: Initial report for Ireland*. Retrieved January 12, 2019, from www.erc.ie.
- Foy, P., & Yin, L. (2017). Scaling the PIRLS 2016 achievement data. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and procedures in PIRLS 2016* (pp. 1–38). TIMSS & PIRLS International Study Center. Lynch School of Education, Boston College and International Association of the Evaluation of Educational Achievement. Retrieved July 18, 2018, from <https://timssandpirls.bc.edu/publications/pirls/2016-methods/chapter-12.html>.
- Guo, J., Marsh, H. W., Parker, P. D., & Dicke, T. (2018). Cross-cultural generalizability of social and dimensional comparison effects on reading, math, and science self-concepts for primary school students using the combined PIRLS and TIMSS data. *Learning and Instruction*, 58(January), 210–219. <https://doi.org/10.1016/j.learninstruc.2018.07.007>.
- Hallquist, M. N., & Wiley, J. F. (2018). MplusAutomation: An R package for facilitating large-scale latent variable analyses in Mplus. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(4), 621–638. <https://doi.org/10.1080/10705511.2017.1402334>.
- Harker, W. J. (1972). An evaluative summary of models of reading comprehension. *Journal of Literacy Research*, 5(1), 26–34. <https://doi.org/10.1080/10862967209547019>.
- Heyneman, S. P., & Loxley, W. A. (1982). Influences on academic achievement across high and low income countries: A re-analysis of IEA data. *Sociology of Education*, 55(1), 13. <https://doi.org/10.2307/2112607>.
- Hooper, M., & Fishbein, B. (2017). Developing the PIRLS 2016 context questionnaires. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and procedures in PIRLS 2016* (pp. 1–8). TIMSS & PIRLS International Study Center. Lynch School of Education, Boston College and International Association of the Evaluation of Educational Achievement. Retrieved on July 18, 2018, from <https://doi.org/978-1-88993-44-8>.
- Hooper, M., Mullis, I. V. S., & Martin, M. O. (2015). PIRLS 2016 context questionnaire framework. In I. V. S. Mullis, & M. O. Martin (Eds.), *PIRLS 2016 assessment framework* (2nd Ed., pp. 31–53). TIMSS & PIRLS International Study Center. Lynch School of Education, Boston College and International Association of the Evaluation of Educational Achievement. Retrieved on July 18, 2018, from https://timssandpirls.bc.edu/pirls2016/downloads/P16_FW_Chap2.pdf.
- Jerrim, J., & Lopez-Agudo, L. A. (2017). *What happens when econometrics and psychometrics collide? An example using PISA data*. Department of Quantitative Social Science Working Paper No. 17-04 February 2017.
- Johansone, I. (2017). Survey operations procedures in PIRLS 2016. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and procedures in PIRLS 2016* (pp. 1–26). TIMSS & PIRLS International Study Center. Lynch School of Education, Boston College and International Association of the Evaluation of Educational Achievement. <https://doi.org/978-1-88993-44-8>.
- Joshi, R. M., & Aaron, P. G. (2000). The component model of reading: Simple view of reading made a little more complex. *Reading Psychology*, 21(2), 85–97. <https://doi.org/10.1080/02702710050084428>.
- Karakolidis, A., Pitsia, V., & Emvalotis, A. (2016). Examining students' achievement in mathematics: A multilevel analysis of the Programme for International Student Assessment (PISA) 2012 data for Greece. *International Journal of Educational Research*, 79(1), 106–115.
- Lagravinese, R., Liberati, P., & Resce, G. (2019). The impact of economic, social and cultural conditions on educational attainments. *Journal of Policy Modeling*. <https://doi.org/10.1016/j.jpmlm.2019.03.007>.

- LaRoche, S., Joncas, M., & Foy, P. (2017). Sample design in PIRLS 2016. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and procedures in PIRLS 2016* (pp. 1–34). TIMSS & PIRLS International Study Center. Lynch School of Education, Boston College and International Association of the Evaluation of Educational Achievement. <https://doi.org/978-1-88993-44-8>.
- Lenkeit, J., Chan, J., Hopfenbeck, T. N., & Baird, J. A. (2015). A review of the representation of PIRLS related research in scientific journals. *Educational Research Review*, *16*, 102–115. <https://doi.org/10.1016/j.edurev.2015.10.002>.
- Liu, Y., Zumbo, B. D., & Wu, A. D. (2014). Relative importance of predictors in multilevel modeling. *Journal of Modern Applied Statistical Methods*, *13*, 5–6. <https://doi.org/10.22237/jmasm/1398916860>.
- Marôco, J. P. (2018). O Bom Leitor: Preditores da Literacia de Leitura dos dos Alunos Portugueses no PIRLS 2016. *Revista Portuguesa de Educação*, *31*(2), 37–53.
- Martin, M. O., Mullis, I. V. S., & Foy, P. (2017). Assessment design for PIRLS, PIRLS literacy, and ePIRLS in 2016. In I. V. S. Mullis & M. O. Martin (Eds.), *PIRLS 2016 assessment framework* (2nd Ed., pp. 55–69). TIMSS & PIRLS International Study Center. Lynch School of Education, Boston College and International Association of the Evaluation of Educational Achievement. Retrieved on July 19, 2018, from https://timssandpirls.bc.edu/pirls2016/downloads/P16_FW_Chap3.pdf.
- Martin, M. O., Mullis, I. V. S., & Hooper, M. (Eds.). (2017). *Methods and procedures in PIRLS 2016*. TIMSS & PIRLS International Study Center. Lynch School of Education, Boston College and International Association of the Evaluation of Educational Achievement. Retrieved on July 19, 2018, from <https://timssandpirls.bc.edu/publications/pirls/2016-methods.html>.
- Martin, M. O., Mullis, I. V. S., Hooper, M., Yin, L., Foy, P., Fishbein, B., et al. (2017). Creating and interpreting the PIRLS 2016 context questionnaire scales. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and procedures in PIRLS 2016* (pp. 1–106). TIMSS & PIRLS International Study Center. Lynch School of Education, Boston College and International Association of Educational Achievement. Retrieved on July 19, 2018, from <https://timssandpirls.bc.edu/publications/pirls/2016-methods/chapter-14.html>.
- Martini, F., & Sénéchal, M. (2012). Learning literacy skills at home: Parent teaching, expectations, and child interest. *Canadian Journal of Behavioural Science/Revue Canadienne Des Sciences Du Comportement*, *44*(3), 210–221. <https://doi.org/10.1037/a0026758>.
- McNamara, A., Arino de la Rubia, E., Zhu, H., Ellis, S., & Quinn, M. (2018). *skimr: Compact and flexible summaries of data (R package version 1.0.3) [Computer software]*. <https://cran.r-project.org/package=skimr>.
- Montgomery, D. C., & Peck, E. A. (1982). *Introduction to linear regression analysis*. New York, Brisbane: John Wiley.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017a). *PIRLS 2016 international results in reading*. Retrieved on July 18, 2018, from <http://timssandpirls.bc.edu/pirls2016/international-results/>.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017b). *PIRLS 2016 international results in reading*. TIMSS & PIRLS International Study Center [u.a.]. Retrieved on July 19, 2018, from <http://timssandpirls.bc.edu/pirls2016/international-results/pirls/student-achievement/pirls-achievement-t-results/>.
- Mullis, I. V. S., Martin, M. O., Foy, P., & Hooper, M. (2017c). *What makes a good reader : International findings from PIRLS 2016*. PIRLS 2016 International Results in Reading. TIMSS & PIRLS International Study Center. Retrieved on July 18, 2018, from <http://timssandpirls.bc.edu/pirls2016/international-results/pirls/summary/>.
- Mullis, I. V. S., Martin, M. O., Goh, S., & Prendergast, C. (2017). *PIRLS 2016 Encyclopedia—PIRLS 2016*. PIRLS 2016 Encyclopedia: Education Policy and Curriculum in Reading. Retrieved on July 18, 2018, from <http://timssandpirls.bc.edu/pirls2016/encyclopedia/>.
- Mullis, I. V. S., Martin, M. O., & Sainsbury, M. (2015). PIRLS 2016 reading framework. In I. V. S. Mullis, & M. O. Martin (Eds.), *PIRLS 2016 assessment framework* (2nd Ed., pp. 13–31). TIMSS & PIRLS International Study Center. Lynch School of Education, Boston College and International Association of the Evaluation of Educational Achievement. Retrieved on July 18, 2018, from https://timssandpirls.bc.edu/pirls2016/downloads/P16_FW_Chap1.pdf.
- Mullis, I. V. S., & Prendergast, C. O. (2017). Developing the PIRLS 2016 achievement items. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and procedures in PIRLS 2016* (pp. 1–29). Retrieved on July 19, 2018, from <https://timssandpirls.bc.edu/publications/pirls/2016-methods/chapter-1.html>.

- Muthén, B. O., & Satorra, A. (1995). Complex sample data in structural equation modeling. *Sociological Methodology*, 25, 267. <https://doi.org/10.2307/271070>.
- Netten, A., Voeten, M., Droop, M., & Verhoeven, L. (2014). Sociocultural and educational factors for reading literacy decline in the Netherlands in the past decade. *Learning and Individual Differences*, 32, 9–18. <https://doi.org/10.1016/j.lindif.2014.02.002>.
- Ólafsson, R. F., Grønmo, L. S., Gustafsson, J.-E., Björnsson, J. K., Kupari, P., Mejding, J., et al. (2014). *Northern lights on TIMSS and PIRLS 2011 : Differences and similarities in the Nordic countries*. IMF eLibrary. <https://doi.org/TN2014-528>.
- Park, H. (2008). Home literacy environments and children's reading performance: A comparative study of 25 countries. *Educational Research and Evaluation*, 14(6), 489–505. <https://doi.org/10.1080/13803610802576734>.
- R Core Team. (2017). *R: A language and environment for statistical computing*. (3.3). R Foundation for Statistical Computing. <http://www.r-project.org>.
- Schmidt, W. H., Burroughs, N. A., Zoido, P., & Houang, R. T. (2015). The role of schooling in perpetuating educational inequality: An international perspective. *Educational Researcher*, 44(7), 371–386. <https://doi.org/10.3102/0013189X15603982>.
- Schwippert, K., & Goy, M. (2007). PIRLS 2001 in brief. In W. Bos, M. Krüger-Potratz, J. Henze, S. Hornberg, B. von Kopp, H.-G. Kotthof, et al. (Eds.), *Studies in international comparative and multicultural education* (pp. 21–36). Münster: Waxmann Verlag GmbH.
- Shiel, G., & Eivers, E. (2009). International comparisons of reading literacy: What can they tell us? *Cambridge Journal of Education*, 39(3), 345–360. <https://doi.org/10.1080/03057640903103736>.
- Spaull, N. (2017). *The unfolding reading crisis: The new PIRLS 2016 results*. <http://Nicspaull.com>. <https://nicspaull.com/2017/12/05/the-unfolding-reading-crisis-the-new-pirls-2016-results/>.
- Stapleton, L. M. (2013). Multilevel structural equation modeling with complex sample data. In G. R. Hancock & R. O. Mueller (Eds.), *Quantitative methods in education and the behavioral sciences: Issues, research, and teaching. Structural equation modeling: A second course* (pp. 521–562). IAP Information Age Publishing.
- Thien, L. M., & Ong, M. Y. (2015). Malaysian and Singaporean students' affective characteristics and mathematics performance: evidence from PISA 2012. *SpringerPlus*, 4(1), 563. <https://doi.org/10.1186/s40064-015-1358-z>.
- TIMSS & PIRLS — PIRLS 2016 International Database. (n.d.). Retrieved May 2, 2019, from <https://timssandpirls.bc.edu/pirls2016/international-database/index.html>.
- Tse, S. K., Lam, J., Lam, Y. R., Loh, E., & Westwood, P. (2005). Students' test performance in PIRLS, attitude to reading, and reading self-concept across three ability groups: Data from Hong Kong. *Australian Journal of Learning Disabilities*, 10(1), 9–18.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.