Introduction

TIMSS and PIRLS 2011: Relationships Among Reading, Mathematics, and Science Achievement—Implications for Early Learning

Ina V.S. Mullis and Michael O. Martin
Boston College

Background Descriptions of TIMSS and PIRLS

TIMSS (Trends in International Mathematics and Science Study) is an international assessment of mathematics and science at the fourth and eighth grades that has been conducted every four years since 1995, with the most recent assessment in 2011. In total, more than 600,000 students participated in TIMSS 2011. Countries and regional benchmarking entities could participate in the fourth grade assessment, the eighth grade assessment, or both: fifty-two countries and seven benchmarking entities participated in the fourth grade assessment, and

45 countries and 14 benchmarking entities participated in the eighth grade assessment. Also, several countries, where fourth and eighth grade students were expected to find the TIMSS assessments too difficult, assessed their sixth and ninth grade students.

The TIMSS 2011 achievement results were reported in two companion publications, the *TIMSS 2011 International Results in Mathematics* (Mullis, Martin, Foy, & Arora, 2012) and the *TIMSS 2011 International Results in Science* (Martin, Mullis, Foy, & Stanco, 2012). These reports summarized mathematics and science achievement at the fourth and eighth grades, documented trends in achievement over time for participants in previous TIMSS assessments, and related achievement to the rich array of information about students' characteristics and attitudes as well as their home, school, and classroom contexts for learning.

PIRLS (Progress in International Reading Literacy Study) is an international assessment of reading comprehension at the fourth grade that has been conducted every five years since 2001. In total, approximately 325,000 students participated in PIRLS 2011, including countries assessing students at the sixth as well as the fourth grades, regional participants or language benchmarking efforts, and prePIRLS (an easier version of PIRLS) for students who are still developing their reading skills. Forty-five countries assessed fourth grade students in 2011.

The PIRLS 2011 results were published in *PIRLS 2011 International Results in Reading* (Mullis, Martin, Foy, & Drucker, 2012). This report, which is similar to the TIMSS 2011 volumes for mathematics and science, contains the 2011 reading achievement results for the participating countries and benchmarking entities, shows trends over time for the countries and benchmarking entities that also participated in previous assessments, and relates reading achievement to a number of home, school, and classroom contexts for learning to read. Full details of the methodology underpinning TIMSS and PIRLS in 2011 are presented in *Methods and Procedures in TIMSS and PIRLS 2011* (Martin & Mullis, 2012).

Both TIMSS and PIRLS 2011 continue series of international assessments in mathematics, science, and reading conducted by the International Association for the Evaluation of Educational (IEA). IEA pioneered international comparative assessments of educational achievement in the 1960s to gain a deeper understanding of the effects of policies and practices across countries'

different systems of education. TIMSS and PIRLS are directed by IEA's TIMSS & PIRLS International Study Center at Boston College.

Introduction to the Current Report

In 2011, the TIMSS and PIRLS data collection schedules came into alignment for the first time in the history of these international assessments. This provided countries with the opportunity to assess their fourth grade students in three fundamental curricular areas: mathematics, science, and reading. However, more pertinent to the present report, 34 countries and three benchmarking entities took advantage of this unique opportunity to assess the **same** students in all three subjects. Equally important, because the PIRLS assessment includes a parent questionnaire that provides information describing students' home environments and supports for learning, this home environment information was available for the first time with TIMSS data as well. Taken together, the fourth grade students in these 34 countries and three benchmarking participants have achievement data in the three core academic areas—reading, mathematics, and science—accompanied by an extensive array of background questionnaire data about the home, school, and classroom contexts for learning these three subjects.[1]

Having data on the same students makes it possible to conduct a range of investigations of the important characteristics of home and school influencing early learning, while controlling for extraneous factors. Researchers can apply a variety of modeling techniques to explore these important issues by examining the interrelationships among their underlying components. To facilitate this research, the TIMSS & PIRLS International Study Center created a special international database including only fourth grade students assessed in all three subjects, and achievement scores in reading, mathematics, and science were estimated based on a multidimensional scaling or reading, mathematics, and

1 The TIMSS 2011 fourth grade mathematics and science assessment frameworks were organized around a content dimension (number, geometric shapes and measures, and data display in mathematics; life science, physical science, and earth science in science), and a cognitive dimension (knowing, applying, and reasoning for both mathematics and science). Given the frameworks broad coverage, the assessment item pools were necessarily large—175 items in mathematics, and 217 in science—with about half being multiple choice and half constructed response. TIMSS 2011 also collected extensive information about students' home supports and school environments for learning. The questionnaires given to students, teachers, schools, and parents yielded nearly 20 context questionnaire scales about learning and teaching mathematics and science.

PIRLS assesses two purposes for reading that account for most of the reading done by students in and out of school: for literary experience and to acquire and use information. Within each of these two major purposes four comprehension processes are assessed: retrieving, inferencing, integrating, and evaluating. PIRLS gives students reading passages (texts) approximately 800 words in length and asks them 13–16 questions about each passage. PIRLS 2011 contained 10 passages (5 for each purpose) and 135 questions in total. The PIRLS achievement scale was used to summarize students' performance on the assessment questions. PIRLS also included questionnaires given to students, teachers, schools, and parents that were developed in parallel to those administered with TIMSS. Like TIMSS, the PIRLS background data yielded nearly 20 context questionnaire scales about students' attitudes toward reading as well their supports and instructional experiences in learning to read.

science together (Foy, 2013). The purpose of the special database is to have the most appropriate basis for studying relationships among reading, mathematics, and science teaching and learning. The three separate international reports referenced above, and the separate international databases for TIMSS 2011 mathematics and science (Foy, Arora, & Stanco, 2013) and for PIRLS 2011 reading (Foy & Drucker, 2013) should be used for information about the results in one or another of the three subjects assessed in 2011.

It is anticipated that the primary value of this special TIMSS and PIRLS 2011 data will be realized through in-depth national research, as participating countries use the data for school improvement at the primary level. The intention of this initial book examining relationships among reading, mathematics, and science teaching and learning is to illustrate the potential of the special TIMSS and PIRLS 2011 database and to make some headway in the analysis process. To this end, the book includes four very different analyses as described in the following sections.

Profiles of Achievement in Reading, Mathematics, and Science

In examining the relationships among students' achievement in reading, mathematics, and science, a good starting place is to look at whether primary schools are providing students with a thorough grounding in these core subjects, and establishing a solid foundation for later learning. The first chapter in the book sets the stage for the following three chapters by examining patterns of achievement in reading, mathematics, and science within each of the 34 countries and 3 benchmarking entities.

For each TIMSS and PIRLS 2011 participant, achievement is profiled at the TIMSS and PIRLS High International Benchmark and Low International Benchmark, by providing the percentages of students reaching these benchmarks in all three subjects as well as in each of the three subjects separately.[2] The data also are shown graphically in displays that simultaneously show the results in all three subject areas. These graphics show at a glance which countries are most successful in educating their fourth grade students to high levels, and whether countries are equally successful across all three subjects. Interestingly, most countries are more successful in one or two of the subjects than another, especially when it comes to educating substantial percentages of students to high levels.

2 The Advanced, High, Intermediate, and Low International Benchmarks are specific points on the TIMSS and PIRLS achievement scales. As described in Chapter 1, TIMSS and PIRLS use a scale anchoring procedure to describe what students scoring at these benchmarks know and can do.

Students performing at the High International Benchmarks in all three subjects are very accomplished fourth grade students—able to read relatively complex materials with in-depth understanding, solve a variety of mathematics problems, and show familiarity with a range of scientific information. These students have developed an extremely solid basis for further learning and are well positioned to take advantage of future educational opportunities. However, the TIMSS and PIRLS 2011 data provide evidence that it is a very challenging task to educate students to the level of the high benchmarks by the fourth grade. Only Singapore had more than half of its students reach the high benchmark in all three subjects, and only two more countries, Chinese Taipei and Finland, had at least half of their fourth grade students reach the high benchmark in each subject separately.

More than half the countries, however, were successful in educating 90 percent of more of their students to the Low International Benchmark in all three subjects. These students showed that they can read and comprehend facts, read a variety of simple graphs and tables, know simple mathematics (such as adding, subtracting, and basic geometric figures), and know science facts about health, ecosystems, and animals. Although these students have lower achievement than those at the high level, they do have a well-rounded foundation in core concepts and skills that provides a good basis for further learning. In comparison, students who have not learned the basic fundamentals of reading, mathematics, and science by the end of their fourth year of schooling may be at some risk for future academic success.

The profiles across countries of the percentages of fourth grade students reaching high and basic levels of achievement help to situate countries with respect to their relative performance in reading, mathematics, and science. In addition, these profiles provide a good foundation for considering the results of the more complicated analyses presented in the subsequent three chapters.

Impact of Reading Ability on Mathematics and Science Achievement: An Analysis by Item Reading Demands

The *TIMSS 2011 Assessment Frameworks* (Mullis, Martin, Ruddock, O'Sullivan, & Preuschoff, 2009), developed collaboratively with the participating countries through a series of reviews, describe the mathematics and science content and cognitive processes that were to be assessed. Both the mathematics and science frameworks require assessing rather sophisticated reading demands even at the fourth grade. For example, topics in the mathematics and science

content domains specify that students should be able to solve routine and non-routine problems set in everyday contexts and conduct inquiries about various phenomena. Understanding the descriptions of the situations for these types of problems necessarily involves reading. Moreover, in mathematics, the Data Display content area is based on "reading and interpreting" tables, pictographs, bar graphs, and pie charts as well as creating such data displays. The science framework requires comprehending descriptions of experiments and investigations as well as a variety of models and diagrams. Finally, and perhaps most importantly, both mathematics and science can be generally regarded as specialized languages with their own technical vocabularies.

In developing items to assess student achievement in mathematics and science, TIMSS makes every effort to avoid unnecessary reading so that the language used is no more complex than necessary to frame the question (and responses for multiple choice items). However, inevitably the assessment items in both mathematics and science assessments vary considerably in the reading demands they place on students. Reading requirements can be quite minimal, as in items requiring students to complete a calculation or identify the smallest or largest quantity, and most of these items are short. However, some items can have more substantial reading demands, as in those requiring students to understand a description of a science experiment or phenomenon and then apply their knowledge or explain their reasoning.

The availability of PIRLS 2011 data on reading achievement provided an ideal opportunity to investigate the relationship between reading ability and the reading demands of the TIMSS 2011 fourth grade mathematics and science assessment items. Fourth grade students are likely to be at a disadvantage in learning mathematics and science as well as demonstrating high performance on the TIMSS assessment if they lack reading skills.

Essentially, the study examined the hypothesis that students with high reading ability would not be impacted by the level of reading demands in the TIMSS items, but that poorer readers would score lower on the items with highest reading demands than on the items with the lowest reading demands.

Essentially, a coding scheme was used to categorize the TIMSS mathematics and science items into groups (low, medium, and high) in terms of the reading demands they place on the student. The coding scheme evaluated each item in terms of length, technical vocabulary, and density of graphical displays (pictorial representations, models, tables, and graphs). The basic approach used in the analysis was to examine, for each participating country and benchmarking

participant, the relationship between fourth grade students' reading ability as measured by PIRLS and their performance on TIMSS items with increasing levels of reading demands.

The methods used to evaluate the reading demands of the items provided additional insight into the TIMSS items from the perspective of "mathematics reading" and "science reading." Although the total number of words was the strongest factor, technical vocabulary and complicated diagrams also contribute to reading demands. The most significant contribution, however, was gaining a deeper understanding of the interconnectedness among curriculum coverage, instructional emphasis, cognitive processing, and reading ability and the challenges in trying to disentangle the various roles they have in affecting student achievement. The results varied from country to country and even between mathematics and science within countries, yet there was support for the idea that more reading demands can make the fourth grade TIMSS items more challenging for weaker readers even in the context of variation in curriculum coverage and that assessing more complex cognitive processing often involves more reading.

What are the Characteristics of Effective Schools in Reading, Mathematics, and Science?

In order to address this question, school effectiveness analyses were conducted to study what makes schools successful, beyond having a majority of students in attendance from advantaged socioeconomic backgrounds. From an analytic perspective, school effectiveness studies make use of multilevel modeling in order to analyze the relationship between school factors and achievement after controlling for the influences of students' home backgrounds.

The research in this chapter began with a strong conceptual model of school effectiveness based on the existing body of school effectiveness research and the factors that influenced school quality as documented in the *TIMSS 2011* and *PIRLS 2011 International Reports*. According to the conceptual model, an effective school was safe and orderly, supported academic success, had adequate facilities and equipment, was staffed with well-prepared teachers, had well-resourced classrooms, and provided effective instruction. From the vast amount of contextual background data available in TIMSS and PIRLS 2011, eventually eleven context questionnaire scales were combined into five robust school effectiveness measures that were available in parallel across reading, mathematics, and science: three measures of effective school environment, and

two measures of effective school instruction. The Home Background Control model also included two measures: the Home Resources for Learning scale, and an index of students' ability to do numeracy and literacy tasks when they started school.

Separately for each country—and for reading, mathematics, and science within each country—a series of multilevel regression models were formulated. These models were used to describe how the school explanatory measures were associated with achievement, both before and after controlling for home background at the student and school level.

Although there was variation from country to country, the Home Background Control model was successful in capturing the relationship between home background and students' achievement in reading, mathematics, and science in every country, with the Home Resources for Learning variable the strongest predictor. In fact, 16 of the participants had just one significant predictor after controlling for home background.

The school variables posited by the conceptual model were positively correlated with student achievement in most countries, providing support for the validity of the model. After controlling for home background, of the school environment variables, **Schools Are Safe and Orderly** was related to higher achievement in at least one subject in 15 countries, and **Schools Support Academic Success** in 10 countries. **Students Engaged in Reading, Mathematics, and Science Lessons** was the most powerful school instruction variable, related to higher achievement in at least one subject in 15 countries, again after controlling for home background. All in all, a school that was safe and orderly, promoted academic excellence, and provided engaging instruction, could be considered to have several important characteristics for effectiveness.

It should be realized, however, that countries with little or no differences from school to school in student achievement (including at least seven in this research) provide little scope for an effective school analysis of the type described here. Factors such as the ones considered in this research are still important school factors for supporting high student achievement, but an analysis focused on differences between schools cannot show evidence of their effects.

Home Support for Literacy and Numeracy Achievement

One of the most stable findings in educational research is the impact of students' background on achievement, especially parents' level of education

and occupation or earnings. Also, a great deal of research on child development has highlighted the importance of home environments that stimulate the development of early literacy skills. Consistent with this research, in each PIRLS assessment cycle the PIRLS home background data have shown a strong positive association between student reading achievement at the fourth grade and home educational resources, parents' emphasis on early literacy activities, and children's literacy skills when entering school.

Although there has been less research conducted about early numeracy skills, this is an area of growing interest. Therefore, the PIRLS 2011 Home Background Questionnaire, which was administered to parents of students who participated in both TIMSS and PIRLS, was designed also to collect data on early numeracy activities and children's numeracy skills upon entering school. The literacy and numeracy background data, in association with students' reading, mathematics, and science achievement, provide an excellent opportunity to examine the differential effects of aspects of home environment on student achievement in these essential subjects.

The fourth research study presented in this book, conducted by Jan-Eric Gustafsson, Kajsa Yang Hensen, Monica Rosen from the University of Gothenburg in Sweden, took particular advantage of the information about children's early literacy and numeracy experiences provided by the Home Background Questionnaire. This research adopted a path modeling approach to investigate the extent of the influence of Parental Education and Gender on mathematics, science, and reading achievement at the fourth grade; and the mechanisms through which Parental Education and Gender influence achievement via books in the home, frequency of early literacy and numeracy activities, and the child's ability to carry out literacy and numeracy tasks when starting school. The variables in the model were ordered chronologically and logically. In general, Parental Education and Gender preceded the number of books in the home, which preceded the literacy and numeracy activities with the pre-school child, which preceded the child's early literacy and numeracy skills at the beginning of first grade, which preceded the PIRLS 2011 reading achievement and TIMSS mathematics and science achievement scores at the fourth grade.

In the first step of estimation, a common model was fit based on the pooled data from all 34 countries and three benchmarking entities, after which separate models were fit for each country. In the pooled data, the total effects of Parental Education were substantial for mathematics, science,

and reading (.33, .35, and .35, respectively), and books in the home was an important mediating variable. The common model provided strong support for the hypothesized chain of influence via books, early activities, and ability entering school to achievement. The number of books was related to frequency of activities in the home oriented towards both literacy and numeracy, and these activities influenced the general levels of literacy and numeracy skills the child had developed at the time of entering school. Interestingly, a stronger emphasis on early literacy activities than on numeracy activities influenced both the levels of children's literacy and numeracy skills when entering school as well as their achievement in the fourth grade. Similar results have been found in other studies, perhaps because adequate language skills are a prerequisite for learning mathematics. In comparison, the effects of Gender were much weaker. There were essentially no gender differences in mathematics or science, although the total effect on reading achievement was rather substantial (.12). Also, only a small part of the Gender effect was mediated via the variables in the model, although for girls the early learning activities were oriented more toward literacy than numeracy.

While the overall findings and mechanisms described above were identified in most countries, there were interesting differences across the countries. There is much additional TIMSS 2011 and PIRLS 2011 background data that can be used to expand this research, and the research can be extended in many different directions to investigate further variables and hypothesized mechanisms.

Summary

In summary, the TIMSS and PIRLS 2011 fourth grade combined database provides an important resource for researching the contexts for early in reading, mathematics, and science. The achievement measures are extremely robust and there is a rich array of context questionnaire data. It is hoped that the four chapters of this books will inspire many others to conduct further research and mine this valuable data.

References

Foy, P. (2013). *TIMSS and PIRLS 2011 user guide for the fourth grade combined international database*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved from http://timssandpirls.bc.edu/timsspirls2011/international-database.html

Foy, P., Arora, A., & Stanco, G.M. (2013). *TIMSS 2011 user guide for the international database*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved from http://timssandpirls.bc.edu/timss2011/international-database.html

Foy, P. & Drucker, K.L. (2013). *PIRLS 2011 user guide for the international database*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved from http://timssandpirls.bc.edu/pirls2011/international-database.html

Martin, M.O. & Mullis, I.V.S. (Eds.). (2012). *Methods and procedures in TIMSS and PIRLS 2011*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College. Retrieved from http://timssandpirls.bc.edu/methods/index.html

Martin, M.O., Mullis, I.V.S., Foy, P., & Stanco, G.M. (2012). *TIMSS 2011 international results in science*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., Foy, P., & Arora, A. (2012). *TIMSS 2011 international results in mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., Foy, P., & Drucker, K.T. (2012). *PIRLS 2011 international results in reading*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I.V.S., Martin, M.O., Ruddock, G.J., O'Sullivan, C.Y., & Preuschoff, C. (2009). *TIMSS 2011 assessment frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.