GEM THE CENTRE FOR GLOBAL EDUCATION MONITORING

ASSESSMENT GEMS SERIES

No. 13

May 2020

Trends in International Mathematics and Science Study: Measuring and making international comparisons of student achievement in mathematics and science

Trends in International Mathematics and Science Study (TIMSS) is an ongoing, cyclical international study of student achievement in mathematics and science. It is directed by the International Association for the Evaluation of Educational Achievement (IEA).

Origins and context

The IEA was founded in 1958. It has evolved from a collective of research bodies into a professional organisation with a secretariat based in the Amsterdam (NLD) and a centre devoted to data processing and research based in Hamburg (DEU). Beyond this professional organisation, IEA has over 70 members that are governmental and non-governmental educational research institutions from countries in Africa, Asia, Australasia, Europe, the Middle East and the Americas. Most of the members represent national education systems. IEA also maintains funding and non-funding partnerships (IEA, n.d.-b).[1]

According to IEA's founders, the different education systems across the world together form a kind of educational laboratory, and comparative research into these different systems can reveal important relationships between inputs and outcomes, relationships that would not necessarily be detected if any one system were studied in isolation (IEA, 2014a).

IEA studies seek to understand the processes and products of education by administering cognitive assessments and collecting background data to examine the *intended curriculum*, the *implemented curriculum* and the *attained curriculum* (IEA, n.d.-b). The intended curriculum is concerned with the national, social and educational contexts. It covers what is described in curriculum policies and publications, and how the education system is structured to facilitate the learning that is described in these policies and publications. The implemented curriculum is concerned with the school, teacher and classroom contexts. It covers what is actually taught in the classrooms and how it is taught, including the characteristics of the individuals who are teaching. The attained curriculum is concerned with the learning outcomes and characteristics of students. It covers what students learn, what they think about what they learn, and their backgrounds (Mullis & Martin, 2013).

This three-aspect concept of the curriculum has been used in many of the 30 comparative research studies IEA has conducted since its inception. While TIMSS and the Progress in International Reading Literacy Study (PIRLS) measure performance in basic school subjects, studies have also been conducted in areas such as literature, advanced mathematics and physics, civics and citizenship, and computer and information technology (IEA, 2014b, 2014c).

At the international level, TIMSS is managed by the TIMSS & PIRLS International Study Center at the Lynch School of Education at Boston College. Each participant has a research coordinator team that is responsible for the local implementation of the study.

1 Some of IEA's recent funding partnerships have been with the European Commission, the Ford Foundation, the Inter-American Development Bank, the United Nations Development Program and the World Bank. Its non-funding partnerships include those with the Educational Testing Service (ETS), UNESCO's International Institute for Educational Planning (IIEP UNESCO), Organization of Ibero-American States (OEI), Partnership for Educational Revitalization in the Americas (PREAL), Programme d'Analyse des Systémes Educatifs CONFEMEN (PASEC), Statistics Canada, the Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ), and the West African Examinations Council (WAEC) (IEA, n.d.-b).

Australian Council *for* Educational Research

ACER

ISSN 2203-9406 (Online)

Table 1: Domains in the content and cognitive dimensions in the TIMSS 2019 assessment framework

<table>
<tr><th></th><th colspan="2">MATHS</th><th colspan="2">SCIENCE</th></tr>
<tr><th></th><th>Domains in content dimension</th><th>Domains in cognitive dimension</th><th>Domains in content dimension</th><th>Domains in cognitive dimension</th></tr>
<tr><td>Fourth grade</td><td>• Number
• Measurement and Geometry
• Data</td><td rowspan="2">• Knowing
• Applying
• Reasoning</td><td>• Life Science
• Physical Science
• Earth Science</td><td rowspan="2">• Knowing
• Applying
• Reasoning</td></tr>
<tr><td>Eighth grade</td><td>• Number
• Algebra
• Geometry
• Data and Probability</td><td>• Biology
• Chemistry
• Physics
• Earth Science</td></tr>
</table>

TIMSS has been conducted every four years since 1995. The number of participating countries has grown from 45 in the first cycle (IEA, 2014d) to more than 60 in the seventh cycle (Department of Education USA, 2018). Since the second cycle, some sub-national benchmarking entities (e.g. states, provinces, districts or district consortia) have also participated.

TIMSS is funded by participants and through IEA's funding partnerships.

Purpose

TIMSS measures the mathematics and science performance of students and collects a wide array of contextual information about students, schools, curricula and educational policies and systems. In analysing and reporting on these performance and contextual data, IEA aims to assist participants to make informed decisions to improve mathematics and science teaching and learning (Martin, Mullis, Foy, & Stanco, 2012).

Measurement objectives

Learning domains

TIMSS is an international assessment that is both curriculum-based and standardised. The development of common cognitive tests that adequately represent the curricula of all participants involves extensive research, consultation and consensus-building.

Before the first cycle of TIMSS, IEA established frameworks that provided an organising structure for discussing, classifying and categorising all elements of mathematics and science curricula, from test items to textbook content to statements in official documents.

These curriculum frameworks were used in an initial in-depth analysis of the mathematics and science curricula of first-cycle participants.[2] The findings of the analysis informed the selection of the final items for this first cycle of the assessment (Martin, 1996).

The curriculum frameworks then guided test specifications for the second cycle of TIMSS, and later formed the basis of the TIMSS assessment frameworks that were developed for the third cycle in 2003.

Since 2003, mathematics and science have been considered in the assessment frameworks in two dimensions: the content dimension and the cognitive dimension. The content dimension specifies the subject matter to be assessed, and the cognitive dimension specifies the thinking processes to be assessed. Table 1 below shows the domains in the content and cognitive dimensions for both the fourth grade and the eighth grade (the two grades TIMSS assesses) as they are presented in the framework for TIMSS 2019 (Mullis & Martin, 2017).

Mathematics and science have only shared the same domains in the cognitive dimension since 2007 (Mullis & Martin, 2012, 2017).[3] This change is an example of the way the assessment frameworks have been updated over time.

2 In this curriculum analysis, the content of a country's curriculum guides and textbooks was analysed and mapped to the IEA's curriculum frameworks. In addition, questionnaires about the educational system, decision-making processes for learning goals, and general contexts for mathematics and science teaching and learning were administered, and the data from these questionnaires were analysed (Martin, 1996).

3 Prior to that there were some differences between the cognitive dimensions for mathematics and science, but the same principle applied i.e. there are domains in the cognitive dimension that aim to cover cognitive skills of different levels, from knowing facts to reasoning and analysing in more complex and unfamiliar contexts.

In 2015 the IEA offered a less difficult mathematics assessment at the fourth grade called TIMSS Numeracy. This enabled the extension of the TIMSS mathematics achievement scale, providing better measurement at the lower end of the scale (Mullis & Martin, 2017). From 2019, a 'less difficult mathematics' component is presented in booklets alongside the mainstream items as an option for countries to participate in, with the scale fully merged together.

TIMSS Advanced is also intermittently conducted alongside TIMSS, having been conducted in 1995, 2008 and 2015. TIMSS Advanced is for students in the final year of secondary school enrolled in special advanced mathematics and physics programs or tracks. TIMSS Advanced has its own Assessment Framework, and includes the following content domains: algebra, calculus, and geometry in mathematics; mechanics and thermodynamics, electricity and magnetism, and wave phenomena and atomic/nuclear physics in physics (IEA, n.d.-a).

Updating the TIMSS frameworks is a collaborative process involving research coordinators from participating countries/benchmarking entities, the TIMSS & PIRLS International Study Centre, chief subject consultants and international expert committees. It is an exercise in balancing two competing interests: on the one hand, the frameworks must maintain continuity to enable trend measurement; on the other hand, they must be adjusted if they are to be relevant in changing educational contexts (Mullis & Martin, 2012).

The assessment frameworks guide item development and item selection, and these two activities also involve collaboration, this time between research coordinators from participating countries/benchmarking entities, expert test development teams and item review committees.

With respect to item development and item selection, the competing interests of continuity to enable trend measurement and change to maintain relevance are balanced through a design in which old items are progressively replaced with new items so that no items feature for more than three assessment cycles (Mullis & Martin, 2012).[4]

TIMSS assessments include both multiple-choice items and constructed-response items. In the early cycles about two-thirds of the total score points in the assessment came from multiple choice questions (Martin, Gregory, & Stemler, 2000; Martin & Kelly, 1999); in the last few cycles, this number has been close to half (Martin, Mullis, & Chrostowski, 2004; Mullis, Drucker, Preuschoff, Arora, & Stanco, 2012b; Olson, Martin, & Mullis, 2009).

Contextual information

The TIMSS context questionnaires aim to facilitate a better understanding of the contextual factors that affect student learning in mathematics and science.

Since the first cycle of TIMSS, questionnaires have been administered to assessed students, their teachers and the principals of their schools (Martin et al., 2000; Martin & Kelly, 1999; Martin et al., 2004; Mullis, Drucker, Preuschoff, Arora, & Stanco, 2012a; Olson et al., 2009). The student questionnaire aims to find out about students' basic demographic characteristics and their home and school experiences, as well as their attitudes towards learning mathematics and science. The teacher questionnaire aims to find out about teachers' education, professional development, and teaching experiences. The school questionnaire aims to find out about school resources, types of programs, and overall learning environments (Mullis, Drucker, et al., 2012a).

For the first time in TIMSS 2011, a home questionnaire was also administered to the parents/guardians of students assessed in the fourth grade. The home questionnaire aims to find out about resources for learning in the home, languages spoken in the home, parental expectations of children's education, and opportunities provided at home for early literacy, numeracy and science activities (Mullis & Martin, 2013).

The contextual questionnaires are updated, ensuring that they reflect changes in education practices and contexts. Existing scales are updated and new scales are added. In 2019, students who undertook TIMSS via the digital format also undertook a questionnaire about their experiences with computers and the Internet (Mullis & Martin, 2017).

When taken together, the student, home, teacher and school questionnaires cover the attained curriculum, the implemented curriculum and the intended curriculum. Additional information about the system-level elements of the intended curriculum – including information about policies related to school entry, promotion and pre-primary education; and information about the structure and components, format and emphases of the primary and secondary curricula – is obtained through the administration of a curriculum questionnaire. This questionnaire has been administered since the second cycle of TIMSS, and is completed by the research coordinators of participating countries/benchmarking entities (Martin et al., 2000; Martin et al., 2004; Mullis, Drucker, et al., 2012a; Olson et al., 2009).[5]

4 For example, in TIMSS 2011, the final item pool consisted of 20% items that first featured in TIMSS 2003, 40% items that first featured in TIMSS 2007, and 40% new items; by this time all items from TIMSS 1995 and TIMSS 1999 had been released (Mullis & Martin, 2012).

5 In the first cycle of TIMSS, data about system-level elements of the intended curriculum were collected through the curriculum analysis undertaken by IEA.

Since TIMSS 2007, the study has obtained further information about the intended curriculum by requiring each participating country/benchmarking entity to contribute a chapter to a TIMSS encyclopaedia. These chapters summarise the education system, the policies related to mathematics and science instruction, and the mathematics and science curricula in primary and secondary school. They are usually written by staff from ministries of education or research institutions (Mullis, Martin, et al., 2012).

Target population and sampling method

The target populations for TIMSS are defined with reference to UNESCO's International Standard Classification of Education (ISCED) scheme. The technical report for TIMSS 2015 defines the international target populations as:

> **Fourth–grade**: All students enrolled in the grade that represents four years of schooling counting from the first year of ISCED Level 1, providing the mean age at the time of testing is at least 9.5 years.
>
> **Eighth–grade**: All students enrolled in the grade that represents eight years of schooling counting from the first year of ISCED Level 1, providing the mean age at the time of testing is at least 13.5 years.
>
> (Mullis & Martin, 2017)

Participating countries/benchmarking entities use these international definitions as the basis of their own target population definitions. In some cases, target population definitions are the same as the international definitions. In other cases, factors such as the standard age for starting school and practices of grade promotion and retention mean that the international definitions are not suitable, and the participant target population definitions are different. For example, if the mean ages of children in the fourth and eighth grades are less than the thresholds stated in the international definitions given above, then the target population definitions will specify the fifth and ninth grades instead (Joncas & Foy, 2012; Martin, Mullis, & Hooper, 2016).

Though the above international definitions and the participant population definitions that are derived from them specify all students, participants are permitted to make school-level and student-level exclusions for political, organisational and operational reasons, providing these exclusions do not exceed set limits (Joncas & Foy, 2012).

There are technical standards for the sampling precision of estimates. These standards are usually met with a sample of 150 schools that yields approximately 4000 students for each target population. For countries participating in both TIMSS at the fourth grade and TIMSS Numeracy, the required student sample size is doubled, resulting in around 8000 students (Martin, Mullis, & Hooper, 2016). TIMSS samples intact classes of students, so this sample size works out to sampling one class of approximately 27 students from each of the 150 sampled schools (Joncas & Foy, 2012).

Some participating countries/benchmarking entities choose to sample more than one class per sampled school, either to obtain a larger student sample or to enable the better estimation of school-level effects. A participating country/benchmarking entity may also be required to sample more than 150 schools if the standard class size is particularly small, if streaming students by ability is a common practice, if high levels of non-response are expected, or if the TIMSS standards for sampling precision have not been met in previous cycles (Joncas & Foy, 2012; Martin, Mullis, & Hooper, 2016).

The National Research Coordinator from the participating country/benchmarking entity are responsible for developing and implementing the national sampling plan, with support from IEA. The research coordinator constructs a complete and accurate sampling frame.[6] This sampling frame should consist of all schools from all different parts of the education system that have full-time students within the agreed target population definitions (Joncas & Foy, 2012).

In the first stage of sampling, schools are sampled from using systematic sampling with probability proportional to size. Field trial and main survey stage 1 sampling is usually undertaken at the same time. Two replacement schools are sampled along with each main sampled school. School sampling is conducted by IEA/TIMSS staff (Joncas & Foy, 2012; Martin, Mullis, & Hooper, 2016).

In the second stage of sampling, one or more intact classes are sampled with equal probability of selection using systematic random sampling.[7] Class sampling is undertaken by the research coordinator using software developed by IEA for use in its surveys. TIMSS samples intact classes rather than directly sampling students for two reasons. Firstly, the study examines students' curricular and instructional experiences, and these are often organised at the classroom level. Secondly, it minimises disruption at the schools if the assessment includes all students in some classes rather than some

6 If explicit stratification is used then one sampling frame must be constructed for each explicit stratum.

7 Since small classes increase the risk of obtaining unreliable estimates, if a sampled school is identified as having small classes, these classes are grouped together into pseudo-classes that have adequate numbers of student before the second stage of sampling.

students from all classes (Joncas & Foy, 2012; Martin, Mullis, & Hooper, 2016).

Assessment administration

In 2019, TIMSS began the transition to a digital format; in 2023 the digital platform will be the default. However, countries unable to transition to the digital form will be offered a paper-based option comprising trend items only (IEA, 2019).

Within a participating country/benchmarking entity, after schools have been sampled, the national research centre is responsible for identifying and training school coordinators. The school coordinators are tasked with providing the national research centre with information for within-school sampling of classes; identifying and training test administrators; updating tracking forms; organising the time and place for test administration; distributing questionnaires; maintaining the security of test booklets; and managing the receipt and return of all assessment materials.

Managing the activities on the day of test administration is the responsibility of the test administrators. At the start of test sessions the test administrators must read instructions that are standardised across all participating countries/benchmarking entities. For both fourth grade and eighth grade a test session is divided into two parts. The duration of each part and the duration of the break between the two parts are also standardised across all participating countries/benchmarking entities.[8]

Both the school coordinators and the test administrators are supported in their work by manuals that are developed by the TIMSS & PIRLS International Study Centre and translated and adapted by national research centre staff as required.

Since TIMSS is a comparative international survey, assessment booklets must be standardised across countries. The first step in this standardisation is a structured and documented process in which skilled translators from participating countries/benchmarking entities translate and adapt test items from the source language to the target language(s) under the direction of the TIMSS & PIRLS International Study Centre.[9] Independent expert translators engaged by the IEA verify the translations. Second, the translated booklet layout is verified by the International Study Centre. Feedback is provided, which the participating countries/benchmarking entities are expected to review and act on where necessary. After translation and verification, participating countries/benchmarking entities are expected to follow further standard, internationally agreed-upon procedures to complete the preparation of their materials (Yu & Ebbs, 2012).

In addition to the preparation of assessment booklets, other assessment activities, including test administration, scoring, and data entry and processing, must also be standardised as much as possible. To achieve this, TIMSS has developed and documented procedures, protocols, software and training, and also initiated an independent quality assurance program (Johansone, 2012).

In 2023, TIMSS will implement a group adaptive assessment approach. This is where items graded on difficulty will be administered to populations at different rates, depending on the country's overall achievement (IEA, 2019).

Reporting and dissemination

After each assessment cycle, TIMSS results are reported in international reports prepared by the TIMSS & PIRLS International Study Centre. Mathematics and science results are mostly presented separately, with both fourth grade and eighth grade results are presented in the same report. Each report begins with some introductory information about the history and context of TIMSS, the nature of the current assessment, and the range of participating countries/benchmarking entities. Student achievement results are presented next, followed by the background questionnaire data.[10]

Cognitive results are given separately for fourth grade and eighth grade. Results are reported for each participating country in terms of means and distributions of student achievement. Trends in achievement over multiple cycles, cohort comparisons (i.e. comparisons involving the fourth grade results of the previous cycle and the eighth grade results of the current cycle), achievement differences by gender and trends in achievement differences by gender are also reported.

Student achievement results are reported with reference to a handful of unchanging points on the TIMSS scale: the midpoint of the scale at 500, and the points at 625, 550, 475 and 400. The points at 625, 550, 475 and 400 are the TIMSS International Benchmarks indicating respectively advanced achievement,

8 For fourth grade each of the two parts of a test session is 36 minutes; for eighth grade each of the two parts is 45 minutes. For both fourth grade and eighth grade there must be a break between the two parts but it cannot exceed 30 minutes (Martin, Mullis, & Hooper, 2016).

9 English has been a source language for TIMSS since the first cycle. Since TIMSS 2007, Arabic has also been a source language.

10 All TIMSS reports can be downloaded from http://timssandpirls.bc.edu/isc/publications.html.

high achievement, intermediate achievement and low achievement.[11]

The TIMSS International Benchmarks are given not only as numerical proficiency scores but also as detailed proficiency descriptions. These descriptions of what benchmark scores mean in terms of knowledge and skills are developed by the TIMSS & PIRLS International Study Center and the item review experts through data analysis and conceptual analysis of the assessment items. Examples of anchor items (i.e. items that function best for students with achievement at or near a benchmark) are also provided.

Student achievement is also reported using the scales for each of the cognitive and content domains for mathematics and science. Average achievement on the cognitive/content scale is compared to average achievement on the relevant overall scale (i.e. for mathematics or science). Trends in average achievement on the cognitive/content scale, and average achievement disaggregated by gender on these scales are also reported.

TIMSS reports present background data organised by themes. In the TIMSS 2015 mathematics report, chapter headings included: student engagement and attitudes, class room instruction, teachers' and principals' preparation, school safety, school climate, school composition and resources, and home environment and support (Martin, Mullis, Foy, & Hooper, 2016). A variety of background data for students, teachers and schools is reported and linked with average achievement scores. In 2011 TIMSS began presenting a number of policy-relevant questionnaire scales using Item Response Theory. These scales were developed by applying Item Response Theory techniques to responses to multiple questionnaire items that had been designed to measure the same underlying construct (Mullis, Drucker, et al., 2012a). The scales cover areas including resources available at home for learning and education, resources available at school, teacher working conditions, school climate and students' attitudes towards learning.

While the results reports present the data from the student, teacher and school questionnaires, the data from the curriculum questionnaire are presented in the TIMSS encyclopaedias. These data are not analysed, but simply presented in a way that enables easy comparison.

In addition to the results reports and encyclopaedias that are produced each cycle, TIMSS also produces technical reports (also called 'Methods and Procedures') that describe in detail all technical aspects of the assessment.

TIMSS results reports, encyclopaedias, technical reports, assessment frameworks and other documentation for all cycles can be downloaded from the website of the TIMSS and PIRLS International Study Centre.[12] The international databases for all cycles, and accompanying user guides, can be downloaded from the TIMSS and PIRLS website.[13] IEA's Data Processing Centre has developed the IEA IDB Analyser and IEA Data Visualiser software applications to facilitate the analysis and visualisation of data from IEA studies. These applications can be downloaded from IEA's website.[14]

Influence

As the longest running, and one of the most widely applied, mathematics and science assessments, TIMSS has had considerable global influence on curricula development. The impact of TIMSS commenced from its first cycle, with some countries initiating policy changes in response to low TIMSS results. For example, at least partly in response to TIMSS results, the New Zealand Government introduce a comprehensive numeracy policy and strategy in 1998 (Mullis, Martin, & Loveless, 2016).

In general, countries that perform relatively poorly, or whose performance has declined, have tended to initiate policy changes (Cresswell, Schwantner, & Waters, 2015). Governments have also used TIMSS to identify and target disparities within countries (Cresswell et al., 2015).

In the report '20 years of TIMSS' it is observed that more than half of TIMSS countries have revised their mathematics and science curriculum since their participation commenced. Moreover, this revision of curriculum uses TIMSS data to monitor and assess outcomes (Mullis, Martin, & Loveless, 2016).

One example of a country that was influenced by TIMSS and used TIMSS data to make policy reform was Chinese Taipei. Authorities were concerned by two results from TIMSS 2003. First, they had more 8th grade students in the "Below Intermediate" group than their four East Asian neighbours; second, the number

11 In the earliest cycles of TIMSS the reference points were not fixed on the TIMSS scale but instead derived from the achievement results for that cycle – the international achievement average was used instead of the midpoint of the scale, and achievement scores that corresponded to the 95th, 75th, 50th and 25th percentiles of student achievement were used instead of the values of 625, 550, 475 and 400. This practice was discarded because cross-cycle comparisons of achievement, where achievement is referenced to these values that vary from one cycle to another, could lead to erroneous conclusions about improvement or decline in performance. The fixed values for the four benchmarks were set at a research coordinators' meeting in 2004. These values were chosen because they were close to the percentile-derived benchmarks used in the previous cycle.

12 http://timssandpirls.bc.edu/isc/publications.html.

13 http://timss.bc.edu/

14 http://www.iea.nl/data.html

of fourth and eighth graders who did not enjoy and had low confidence in mathematics was considerably higher than the international average. In fact, it was more than twice as bad as the international average. (Lin, 2018)

The first response to these results was to implement a remedial 'After Class Support' program. This involved out-of-school supplementary mathematics courses for struggling students. However, subsequent cycles of TIMSS showed that this program failed to improve students' enjoyment and confidence in mathematics (Lin, 2018).

Consequently, a second project was developed, called 'Just Do Math' (JDM). In regular classes, JDM helped struggling students establish fundamental prerequisite concepts before learning a mathematics topic, through the development of mathematics grounding activity (MGA) modules. Additionally, JDM incorporated gamified learning, mathematics camps and professional development activities for teachers (Lin, 2018).

By mid-2017, more than 45,000 students had participated in the mathematics camps, more than 10,000 teachers had participated in JDM professional development, about 150 MGA modules had been developed, and about 60 MGA designers had been trained. The success of this program is being measured using TIMSS data (Lin, 2018).

By the second decade of the twenty-first century, TIMSS appeared to be influencing a degree of policy convergence amongst participating countries, in regards to topics covered in the curriculum and skills to focus on developing (Kadijevich, 2019). For example, TIMSS cognitive domains (i.e. knowing applying and reasoning) have become widely adopted in mathematics curricula (Mullis, Martin, Goh, & Cotter, 2016).

The ACER Global Education Monitoring Centre supports the monitoring of educational outcomes worldwide, holding the view that the systematic and strategic collection of data on educational outcomes, and factors related to those outcomes, can inform policy aimed at improving educational progress for all learners.

References

Cresswell, J., Schwantner, U., & Waters, C. (2015). *A Review of International Large-Scale Assessmentsin Education: Assessing Component Skills and Collecting Contextual Data*. Melbourne: Australian Council for Educational Research (ACER).

Department of Education USA. (2018). TIMSS USA: *Frequently Asked Questions*. Retrieved from

IEA. (2014a). IEA: Brief history. Retrieved from http://www.iea.nl/brief_history.html

IEA. (2014b). IEA: Completed studies. Retrieved from http://www.iea.nl/completed_studies.html

IEA. (2014c). IEA: Studies. Retrieved from http://www.iea.nl/studies.html

IEA. (2014d). TIMSS 1995 Participating Countries. Retrieved from http://timss.bc.edu/timss1995i/t95_countries.html

IEA. (2019). TIMSS 2023 Brochure. Retrieved from https://timssandpirls.bc.edu/timss2023/downloads/T2023_TIMSS_Brochure.pdf

IEA. (n.d.-a). About TIMSS Advanced 2015. Retrieved from http://timssandpirls.bc.edu/timss2015/international-results/advanced/timss-advanced-2015/about-timss-advanced/

IEA. (n.d.-b). About Us [Pamphlet]. Retrieved from http://www.iea.nl/fileadmin/user_upload/IEA_Documents/IEA_Brochure.pdf

Johansone, I. (2012). Operations and Quality Assurance. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Retrieved from http://timssandpirls.bc.edu/methods/pdf/TP_Operations_Quality_Assurance.pdf

Joncas, M., & Foy, P. (2012). Sampling Implementation. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Retrieved from http://timss.bc.edu/methods/pdf/TP_Sampling_Design.pdf

Kadijevich, D. M. (2019). Influence of TIMSS Research on the Mathematics Curriculum in Serbia: Educational Standards in Primary Education. *The teaching of mathematics*, XXII(1), 33-41.

Lin, F. L. (2018). How Chinese Taipei used TIMSS data to reform mathematics education. In *IEA Compass: Briefs in Education No.2*. Amsterdam.

Martin, M. O. (1996). Third International Mathematics and Science Study: An Overview. In M. O. Martin & D. L. Kelly (Eds.), *Third International Mathematics and Science Study (TIMSS) Technical Report Volume I: Design and Development* Chestnut Hill MA: Boston College.

Martin, M. O., Gregory, K. D., & Stemler, S. E., (Eds.). (2000). *TIMSS 1999 Technical Report*. Retrieved from Chestnut Hill, MA: http://sling11.bc.edu/timss1999i/pdf/T99_TR.book.pdf

Martin, M. O., & Kelly, D. L., (Eds.). (1999). *Third International Mathematics and Science Study Technical Report Volume I: Design and Development*. Retrieved from Chestnut Hill, MA: http://timssandpirls.bc.edu/timss1995i/TIMSSPDF/TRall.pdf

Martin, M. O., Mullis, I. V. S., & Chrostowski, S. J., (Eds.). (2004). *TIMSS 2003 Technical Report*. Retrieved from Chestnut Hill, MA: http://timss.bc.edu/PDF/t03_download/T03TECHRPT.pdf

Martin, M. O., Mullis, I. V. S., Foy, P., & Hooper, M. (2016). *TIMMS 2015 International Results in Mathematics*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

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.

Martin, M. O., Mullis, I. V. S., & Hooper, M. (2016). *Methods and Procedures in TIMSS 2015*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Boston College.

Mullis, I. V. S., Drucker, K. T., Preuschoff, C., Arora, A., & Stanco, G. M. (2012a). Assessment Framework and Instrument Development. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Retrieved from http://timssandpirls.bc.edu/methods/pdf/TP_Instrument_Devel.pdf

Mullis, I. V. S., Drucker, K. T., Preuschoff, C., Arora, A., & Stanco, G. M. (2012b). Instrument Development. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Retrieved from http://timss.bc.edu/methods/pdf/MP_Assessment_T11.pdf

Mullis, I. V. S., & Martin, M. O. (Producer). (2012, 19 February 2014). Developing Assessment Frameworks for Measuring trends in a Changing Context [PowerPoint slides] Retrieved from http://www.iea.nl/fileadmin/user_upload/General_Assembly/53rd_GA/GA53_Developing_Assessment_Frameworks.pdf

Mullis, I. V. S., & Martin, M. O. (Eds.). (2013). *TIMSS 2015 Assessment Frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).

Mullis, I. V. S., & Martin, M. O. (Eds.). (2017). *TIMSS 2019 Assessment Frameworks*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College and International Association for the Evaluation of Educational Achievement (IEA).

Mullis, I. V. S., Martin, M. O., Goh, S., & Cotter, K. (2016). *TIMSS 2015 Encyclopedia: Education Policy and Curriculum in Mathematics and Science*. Retrieved from http://timssandpirls.bc.edu/timss2015/encyclopedia/countries/armenia/use-and-impact-of-timss-2/

Mullis, I. V. S., Martin, M. O., & Loveless, T. (2016). *20 Years of TIMSS: International Trends in Mathematicsand Science Achievement, Curriculum,and Instruction*. Chestnut Hill, MA: TIMSS & PIRLS International Study Center, Lynch School of Education, Boston CollegeandInternational Association for the Evaluation of Educational Achievement (IEA).

Mullis, I. V. S., Martin, M. O., Minnich, C. A., Stanco, G. M., Alka, A., Centurino, V. A. S., . . . (Eds.). (2012). *TIMSS 2011 Encyclopedia: Education Policy and Curriculum in Mathematics and Science*. Retrieved from Chestnut Hill, MA: http://timssandpirls.bc.edu/timss2011/downloads/TIMSS2011_Enc-v1.pdf

Olson, J. F., Martin, M. O., & Mullis, I. V. S., (Eds.). (2009). *TIMSS 2007 Technical Report*. Retrieved from Chestnut Hill, MA: http://timss.bc.edu/timss2007/PDF/TIMSS2007_TechnicalReport.pdf

Yu, A., & Ebbs, D. (2012). Translation and Translation Verification. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. Retrieved from http://timssandpirls.bc.edu/methods/pdf/TP_Translation_Verif.pdf

The ACER Centre for Global Education Monitoring supports the monitoring of educational outcomes worldwide, holding the view that the systematic and strategic collection of data on educational outcomes, and factors related to those outcomes, can inform policy aimed at improving educational progress for all learners.

https://www.acer.org/au/gem