

IEA's Rosetta Stone: Measuring global progress toward the UN Sustainable Development Goal for quality education by linking regional assessment results to TIMSS and PIRLS International Benchmarks of Achievement

This IEA proposal is to address the need to measure progress toward the UN Sustainable Development Goal 4: Ensure inclusive and quality education for all and promote lifelong learning. In particular, the proposal describes a strategy for developing Indicator 4.1.1: Proportion of children and young people: (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex. As set forth in “Unpacking Sustainable Development Goal 4: Education 2030: Guide”, the principles, strategies, actions for this target go beyond the simple dichotomy of “literate” versus “illiterate” and are underpinned by the contemporary understanding of literacy as a continuum of proficiency levels. More specifically, the guide states that “action for this target aims at ensuring that by 2030 all young people and adults across the world should have achieved **relevant and recognized proficiency levels in functional literacy and numeracy skills**”.

IEA's TIMSS and PIRLS international assessments provide widely recognized proficiency levels in numeracy and literacy, respectively, for students at the end of primary schooling. TIMSS has been measuring trends in mathematics and science at four-year intervals since 1995. PIRLS has measured trends in reading literacy at five-year intervals since 2001. With 50-70 countries participating in each assessment cycle, the TIMSS and PIRLS achievement scales and their International Benchmarks are well established and used by countries all around the world. Especially pertinent to measuring progress to the SDG goals, both TIMSS and PIRLS have devoted considerable resources to extending their achievement scales to provide high quality measurement for countries where most children still are developing basic numeracy and literacy skills. For example, the PIRLS assessment has been doubled in scope with the same amount of coverage allocated to a less difficult version of PIRLS that assesses literacy with shorter and simpler texts. It also has reading passages in common with PIRLS such that students can participate primarily with literacy passages and items and still be reported on the PIRLS achievement scale. Similarly, TIMSS mathematics now includes a less difficult assessment providing a comprehensive measurement of basic numeracy skills.

1.1. Objective

This proposal presents a strategy for providing information about the proportions of primary school students that have achieved established proficiency levels in literacy and numeracy. The aim is to establish a link between the results on regional assessments conducted at the primary level and the TIMSS and PIRLS International Benchmarks for numeracy and literacy.

There are five regional assessments planning reading and mathematics assessments at the end of primary schooling in 2018 or 2019:

- SACMEQ – Southern and Eastern Consortium for Monitoring Educational Quality
- PASEC – Programme for the Analysis of Educational Systems
- LLECE – Latin American Laboratory for the Assessment for the Quality of Education
- SEA-PLM – Southeast Asia Primary Learning Metrics
- PILNA – Pacific Island Literacy and Numeracy Assessment

The reading and mathematics assessments planned for 2018/19 provide a perfect opportunity to link these regional assessment results to IEA's TIMSS and PIRLS achievement scales. These regional assessments measure achievement at the sixth grade, except SEA-PLM which is at the fifth grade. The content of the regional mathematics assessments align well with the TIMSS fourth grade assessments of numeracy and mathematics. Similarly, the content of the regional reading assessments align well with the PIRLS fourth grade assessment of literacy and reading comprehension.

The overarching idea is to construct a concordance table that translates between the scores on each of the regional assessments in mathematics and reading and scores on TIMSS and PIRLS, respectively. The concordance table is the "Rosetta Stone" that provides a translation from the countries' regional assessment results to the TIMSS and PIRLS achievement scales. Similar to the original Rosetta Stone, which provided a link between Greek and Egyptian hieroglyphics, the concordance table provides a link between regional assessments and the TIMSS and PIRLS achievement scales. The countries participating in the regional assessments can use the translations to determine what percent of their students could be expected to reach the TIMSS and PIRLS International Benchmarks.

1.2. Implementation

IEA will work with the study centers for each of the five regional assessments. The proposal is to have a subset of countries (3-5) from each regional assessment administer selected booklets of TIMSS and PIRLS achievement items at the same time as their upcoming regional assessments. Depending on the level of mathematics and reading achievement in a region, the booklets can be tailored to contain primarily items assessing TIMSS Numeracy and PIRLS Literacy.

The same students should take the regional mathematics and reading assessments and then also the TIMSS and PIRLS booklets, preferably on the following day. The combined data across the 3-5 countries will provide scores on both the regional assessment and TIMSS and PIRLS for approximately 15,000 students from the region that can be used to construct the "Rosetta Stone" concordance tables for numeracy and literacy achievement.

For each regional assessment, because the concordance tables provide a projected TIMSS or PIRLS score for all possible regional assessment scores, it will be possible to determine the regional assessment scores equivalent to each of the TIMSS and PIRLS International

Benchmarks. TIMSS and PIRLS each have four International Benchmarks – Low (400), Intermediate (475), High (550), and Advanced (625). For each country participating in a regional assessment, progress toward an International Benchmark can be estimated by the percentage of students reaching the regional assessment score equivalent to the International Benchmark. For example, a country may want to determine the percentage of students reaching the Low International Benchmark. Hypothetically, if the concordance table showed that a regional assessment score of 562 in reading was equivalent to 400 on the PIRLS reading scale then all students in the country reaching 562 could be considered to have reached the Low International Benchmark.

Although based on data from the 3-5 countries that participate in the linking study, the concordance table and the Benchmark equivalent scores can be applied in all the countries in the regional assessment (whether they participated in the linking study or not).

1.3. **Schedule**

The Rosetta Stone Linking Project for regional assessments will take four years: 2018-2021.

- 2018 – Meet with regional study centers to plan operations; prepare TIMSS and PIRLS assessment booklets and data collection manuals
- 2019 – Conduct linking data collection in accordance with regional assessment schedules; conduct training in constructed response item scoring
- 2020 – Prepare for and conduct psychometric scaling of regional assessment and TIMSS and PIRLS data and construct concordance tables
- 2021 – Produce the reports to regional assessment study centers, including technical documentation about the match between the assessment frameworks and assessment items for the regional assessments and TIMSS and PIRLS and the methodology employed