



United Nations
Educational, Scientific and
Cultural Organization



UNESCO
INSTITUTE
FOR
STATISTICS



GLOBAL
ALLIANCE
TO MONITOR
LEARNING



Rosetta Stone

Measuring global progress towards
SDG4 by linking assessments results
to TIMSS and PIRLS International
Benchmarks of Achievement

GAML6/REF/4



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

IEA, December 19, 2018



Contents

1.	Summary	3
2.	Introduction	3
3.	Implementation.....	4
4.	Study management and coordination	5
5.	Potential timeline.....	6
6.	Survey operations	7
7.	Scaling methodology.....	7
8.	Related documents	8

1. Summary

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 of regional assessments conducted at the primary level, and the TIMSS and PIRLS achievement scales for numeracy and literacy.

There are five regional assessments currently planning reading and mathematics assessments targeted at the end of primary schooling in 2019. This proposal suggests to start working with the regional assessment ERCE – the Fourth Regional Comparative and Explanatory Study, conducted in 19 Latin American and Caribbean countries.

ERCE, the reading and mathematics assessment planned for 2019, provides a perfect opportunity to link the results from a regional assessment to IEA's TIMSS and PIRLS achievement scales. ERCE measures achievement at the sixth grade.¹ The content of ERCE is expected to align well with the TIMSS fourth grade assessments in numeracy and mathematics; likewise, the reading component of ERCE is expected to align well with the PIRLS fourth grade assessment in literacy and reading comprehension.

The overarching goal is to construct a concordance table that “translates” the scores resulting from ERCE in mathematics and reading to scores on TIMSS and PIRLS, respectively. The concordance table would therefore represent the “Rosetta Stone”, analogous to the original Rosetta Stone which provided a link between Greek and Egyptian hieroglyphics, providing a translation between countries' ERCE results and the TIMSS and PIRLS achievement scales. Countries participating in ERCE can then use the translations to determine percentages of their students that could be expected to reach the TIMSS and PIRLS international benchmarks or any other benchmarks that can be measured on the TIMSS and PIRLS scales². Only three to five countries would need to administer the linkage material in order for all participating countries to be able to use the concordance table – four countries are proposed for the main data collection cycle in 2019.

2. Introduction

The IEA provides countries with important trend results about their students' reading literacy and mathematics and science achievement in a global context. The IEA's TIMSS (Trends in International Mathematics and Science Study) and PIRLS (Progress in International Reading Literacy Study) are international large-scale assessments aiming at helping countries make informed decisions about how to improve teaching and learning in their national context. More than 70 countries are participating in the IEA's TIMSS 2019, which has regularly assessed mathematics and science achievement at grades 4 and 8 since 1995. More than 60 countries participated in the IEA's PIRLS 2016, which has regularly assessed reading achievement at grade 4 since 2001. To encourage participation by a broad range of countries, the IEA's TIMSS and PIRLS use mathematics and science assessments of different difficulty designed to match the educational achievement of students from countries at different stages of development in terms of their education

¹ ERCE also measures achievement in lower grades. However the assessed grade level is different (3rd grade for ERCE). Linking TIMSS and PIRLS to the 6th grade for those regions which are expected to have mainly significantly lower achievement levels compared to the TIMSS/PIRLS averages is deemed as most appropriate.

² Please have a look at the following document: <http://unesdoc.unesco.org/images/0026/002606/260607E.pdf>

systems. The assessment results from the different countries, however, can be projected on the same global achievement metric for TIMSS, and PIRLS, respectively.

TIMSS and PIRLS can be seen as a tool to help monitor progress towards Sustainable Development Goal (SDG) 4, which aims to “ensure inclusive and equitable quality education and promote lifelong learning opportunities for all.”³ According to Target 4.1 of the global indicator framework to monitor SDG4 targets, all girls and boys should complete free, equitable, and quality primary and secondary education leading to relevant and effective learning outcomes. In this regard, TIMSS and PIRLS can contribute by providing the necessary data for indicator 4.1.1b demands, i.e., the proportion of children and young people at the end of primary education achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex.

Despite active steps taken since 2016 by the UNESCO Institute for Statistics, which has convened the Global Alliance to Monitor Learning, there is no global standard for proficiency yet. As a result, countries and dependent territories report data according to the benchmarks defined by the respective cross-national assessments in which they take part. In TIMSS and PIRLS, there are four different benchmarks (low, intermediate, high, and advanced). In the case of PIRLS, these benchmarks describe the types of reading strategies children undertake and the cognitive processes of which they are capable. For example, students at the low international benchmark can retrieve explicitly stated information and make straightforward inferences⁴.

Rather than requiring all countries to participate in TIMSS and PIRLS for SDG 4 reporting purposes, the Rosetta Stone initiative seeks to link the results of regional assessments with the TIMSS and PIRLS achievement scales, thus making it possible to project the location of each country participating in one of the linked regional assessments onto the TIMSS and PIRLS scales.

While this proposal encompasses aims to link ERCE scores to the TIMSS and PIRLS achievement scales as a starting point, there are four other regional assessments which also measure reading and mathematics skills at the end of primary schooling. These are:

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

As a later step or as an additional option, the Rosetta Stone initiative could be broadened so as to also link these studies with TIMSS and PIRLS.

3. Implementation

The IEA will work with the study center for ERCE. The proposal envisions four countries, which participate in ERCE, administering selected booklets composed of TIMSS and PIRLS achievement items in addition to ERCE instruments during the upcoming ERCE assessment cycle. Depending on the specific levels of mathematics and reading achievement in a region, the booklets can be tailored to contain primarily items assessing TIMSS Numeracy and PIRLS Literacy (see Chapter below on “Methodology” for details on TIMSS Numeracy and PIRLS Literacy).

³ <http://unesdoc.unesco.org/images/0024/002463/246300E.pdf>

⁴ <http://unesdoc.unesco.org/images/0026/002606/260607E.pdf>

According to the proposal, the same students would take the ERCE mathematics and reading assessment, and then also complete the TIMSS and PIRLS booklets – ideally on the following day. The combined data gathered from the four countries taking part in administering both assessments, approximately 15,000 students, will consist of scores on *both* ERCE and TIMSS and PIRLS. This data can then be used to construct the “Rosetta Stone” concordance tables for numeracy and literacy achievement.

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

TIMSS and PIRLS have four International Benchmarks each: Low (400), Intermediate (475), High (550), and Advanced (625). For each country participating in ERCE, 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 an ERCE 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. The estimations would work in the same way for any other benchmark classifications that can be measured on the TIMSS and PIRLS scales. More detailed information on the Scaling Methodology can be found in chapter 7.

Although the concordance tables will be created based only on data from the four countries that participate in the linking study, both the concordance table and the International Benchmark equivalent scores can be applied to *all* the countries participating in ERCE (whether or not they participated in the linking study).

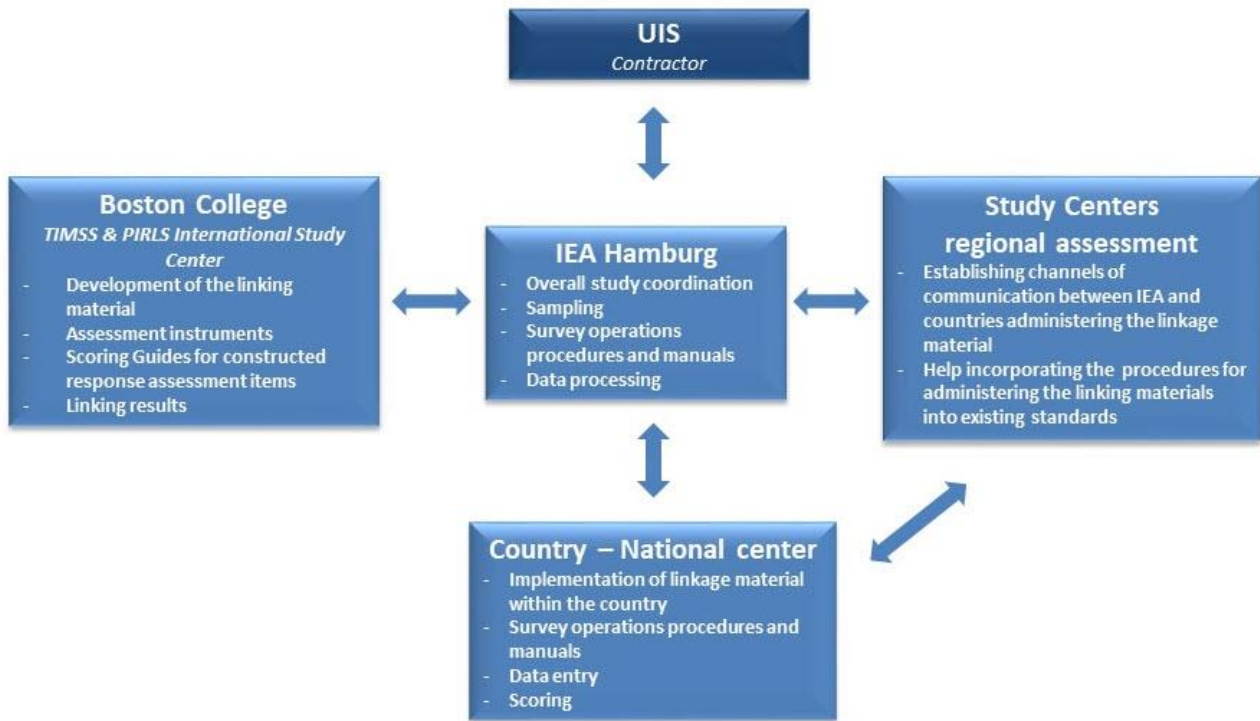
4. Study management and coordination

IEA’s Rosetta Stone would be managed by an IEA-led International Coordination Team, in cooperation with the TIMSS & PIRLS International Study Center at Boston College, and the study center for ERCE.

The IEA Hamburg will be responsible for the overall project management, including the supervision and coordination of all survey operations procedures. Communication with the ERCE study center will be one of the key management activities in order to ensure that the project is progressing in a way that is consistent with the specified standards, the agreed study timelines, and the general goals of the initiative. In addition, the IEA will be responsible for the survey operations, translation, and adaptation of all linking materials in collaboration with the countries. The administration and processing of the linking materials will follow standards used in TIMSS and PIRLS. In order to avoid opening up completely separate branches of procedures, and to ensure comparability of data, the IEA will cooperate closely with the ERCE study center – ensuring in turn that the linkage material will be administered according to the well-established TIMSS and PIRLS standards.

The TIMSS & PIRLS International Study Center at Boston College will be responsible for the design of the linking material, including the development of the international source versions of the instruments in English, as well as the scaling, linking to TIMSS and PIRLS scales, analyses of the data, and reporting of the results.

Figure 1: Rosetta Stone management structure



5. Potential timeline

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

- 2018 – Meet with ERCE study centers to plan operations; prepare TIMSS and PIRLS assessment booklets and data collection manuals and conduct a pilot with the Rosetta material in combination with the ERCE field test
- 2019 – Conduct data collection of linking material in accordance with ERCE assessment schedule; conduct training in constructed-response item scoring
- 2020 – Prepare for and conduct psychometric scaling of regional assessment and TIMSS and PIRLS data;
- 2021 – Construct concordance tables and produce short technical documentation about the match between the assessment frameworks and assessment items for the regional assessments and TIMSS and PIRLS, and the methodology employed.

The first year is dedicated to the planning of the survey operations procedures activities, and establishing contact with the ERCE study center and the 4 countries participating in ERCE which are willing to administer the linkage material. This work is well underway, with one country already having administered the linkage material for the ERCE field test - giving the IEA the possibility to further evaluate the linking material.

In 2019, ERCE will conduct their main data collection phase. The work in this year will focus on evaluating the Field Test linking results, preparing the linkage material for the main data collection phase, preparing the linkage material for each country participating in the linkage study, and trainings in test administration, scoring, and data entry activities.

Once the data from the main data collection phase in 2019 has been collected, the IEA will process, clean, and weight all linkage material data received from the respective regional study centers. As soon as all data has been processed, the TIMSS & PIRLS International Study Center will conduct psychometric scaling of processed regional assessment and TIMSS and PIRLS data, and construct the concordance tables.

The final reports (including the concordance tables) will be published in 2021, along with comprehensive technical documentation.

6. Survey operations

The IEA will provide standards, manuals, and guidelines for the study, to be used as the basis for the work of the national test administrators. These documents define the rules national centers and representatives are asked to follow when preparing and administering linkage material within their country in order to obtain high-quality survey data. In accordance with these documents and standards, special attention will be given to the training of the national centers and test administrators in order to enable them to fulfil all required tasks and activities according to the IEA's high quality control standards. Moreover, the IEA will provide standards relating to quality control in terms of data entry and verification. Constant support of and contact with the ERCE study centers as well as countries participating in the linking exercise is envisioned to assure smooth operation procedures.

7. Scaling methodology

Information regarding how achievement on regional assessments would “translate” to the TIMSS and PIRLS metrics represents great benefit to all countries participating in the regional assessments – not only those countries or populations who administered both assessments. After linking analyses are performed, measurement of progress towards the UN SDG goals, using TIMSS and PIRLS international achievement scales, would be possible for countries participating in those regional assessments included in the proposal. Moreover, access to the rich array of documentation and policy interventions which have been produced using TIMSS and PIRLS data over the past decades represents invaluable help for countries looking to measure progress and effect change in their education systems.

The linkage of regional assessments to the TIMSS and PIRLS grade 4 mathematics trend scale and reading trend scales, respectively, will be based on well-formulated and robust statistical methodology. Starting in 2015, the IEA introduced the TIMSS Numeracy assessment, which was composed of newly-developed items specifically developed to better address mathematics competencies in countries where the TIMSS assessment was judged to be too difficult. In order to project scores obtained on the TIMSS Numeracy Assessment onto the TIMSS grade 4 mathematics trend scales, all countries participating in TIMSS 2015 at the fourth grade level were divided into three groups: those which administered the TIMSS 2015 fourth grade

assessment, those which administered the TIMSS Numeracy 2015, and those which administered both assessments.

After the data collection period, psychometric data obtained from those countries who participated in both assessments was used to place all item parameters (i.e. for items from both the TIMSS 2015 fourth grade assessment and the TIMSS Numeracy 2015 assessment) on a common scale. As the item parameters for TIMSS trend items were already set, having been previously administered and with known item characteristics, unified item calibration was used to estimate item parameters for TIMSS Numeracy items and items new to the TIMSS 2015 assessment. Conditioning was then used to estimate student achievement on the TIMSS Numeracy assessment as plausible values, enabling IEA scaling experts to set linear transformations in order to place TIMSS Numeracy items and scores onto the TIMSS fourth grade mathematics trend scale.

Likewise, the PIRLS 2016 Literacy assessment, composed of reading passages and questions more targeted to countries in which the PIRLS assessment was deemed to be too advanced, was administered to a subset of countries participating in the PIRLS 2016 assessment, and a sub-population of students was administered booklets containing reading passages and questions from both PIRLS 2016 and PIRLS Literacy 2016. A similar methodology was used to link PIRLS Literacy passage questions to the PIRLS to the PIRLS reading trend scale.

This logic would serve as a template for the methodology used in linking regional assessments (i.e., ERCE). In this context, a subset of students participating in ERCE would be administered the Rosetta linking booklets, in a rotated design, in addition to their regional testing days. IEA scaling experts would then use unified item calibration to estimate item parameters for the ERCE items. Conditioning would enable estimations of student achievement on the regional assessments to be placed on the TIMSS and PIRLS trend scales.

According to the above-described procedures, the IEA will produce concordance tables and technical documentation for participating countries and included assessments. These materials resulting from the analysis phases could then be used by countries and the international study centers of the regional assessments to their own ends – including progress made towards the UN SDG goals, policy interventions modeled by other countries participating in TIMSS and PIRLS, etc.

8. Related documents

- Activities Chart Rosetta
- Rosetta international cost estimate
- IEA Annual Report 2016 & 2017
- IEA Statutes
- Commercial business register extract

The IEA considers this entire proposal (including related documents) to be commercially sensitive and does not wish the content to be disclosed at any time, including any request for information relating to unsuccessful tenders. The entire information included here and in the cost estimate is hence flagged as proprietary and confidential.