

ACTRC

Large scale assessment review

Esther Care

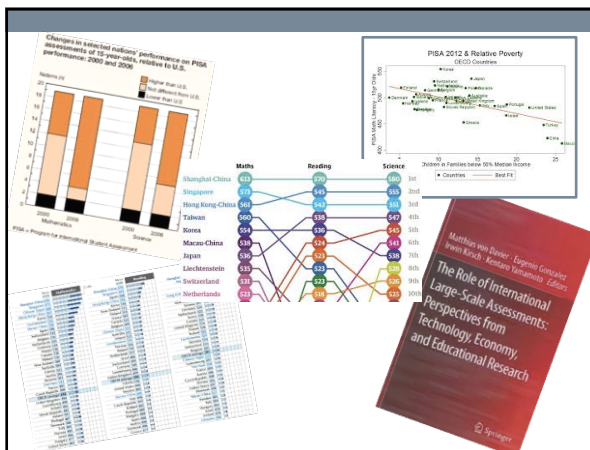
PATEF-UPDATE

Manila, 23 October 2015

The Assessment, Curriculum and Technology Research Centre is a partnership between the University of Melbourne and the University of the Philippines supported by the Australian Government.



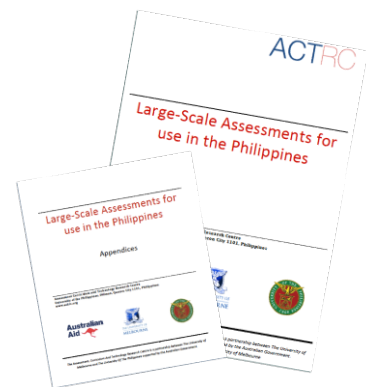
ACTRC has been established as a joint research centre between the Assessment Research Centre at the University of Melbourne and the University of the Philippines. The Centre is supported by the Australian Government.



Overview

This review provides:

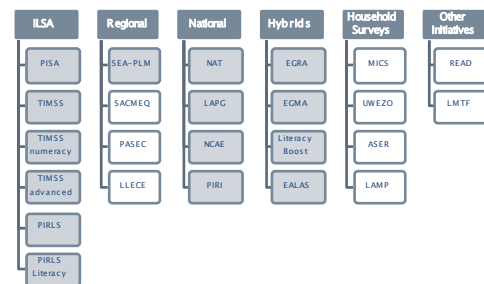
- Context for use of LSA
- Descriptions of selected LSA programs
- Comparisons of selected LSA programs
- Overview of use of LSA in Philippines previously
- Considerations and cautions in electing to participate in LSA
- An Appendix with some example survey questionnaires which collect contextual data



Large scale assessment defined

- 'large-scale assessment'
 - measurement of student learning designed to describe the achievement of students in particular areas of learning across an education system
- 'international large-scale assessment'
 - assessment across different countries not defined by a region
- 'regional large-scale assessment'
 - assessment across different countries within a particular geographic region (subset of international large-scale assessment programs)

Figure 1.1 Types of Assessment Programs and Initiatives



Usefulness

Usefulness of LSA depends on:

- who gets tested
- what gets tested
- when tests occur
- how a test takes place
- why a test takes place

Who

Who is assessed?
Population and sample approaches

Population approach	assessment of all students within the target range in a participating country
Sample approach	assessment of selected students within the target range - an assessment can be administered to a (selected) sample of students and findings can be extrapolated statistically to describe the population

What

What gets tested

Each LSA varies, and a country needs to decide what competencies are of interest, for example:

- Reading
- Maths
- Science
- Global citizenship

When?

- use of assessment data as a baseline measure before a country implements system-wide change
- degree to which the Philippines can rely on its current and previous national assessments to evaluate progress is minimal, due to:
 - implementation of the K to 12 reform is well underway
 - comparability of previous assessment data with current is tenuous given that both the curriculum and the assessments that are aligned with these differ

Why

Why a test takes place

- What are the questions of primary importance that stimulate the country to participate?

PISA, TIMSS and PIRLS

The major international large scale assessment programs

Curriculum, application, and understanding

TIMSS

- TIMSS is designed to provide information about mathematics and science education across time and across grades.. to improve teaching and learning of mathematics and science by providing information about student achievement in relation to different types of curricula, instructional practices, and schools

PISA

“PISA looks at students’ ability to apply knowledge and skills in key subject areas and to analyse, reason and communicate effectively as they examine, interpret and solve problems” (PISA, 2015a)

Details

TIMSS

- Targets: Gr 4, 8
- Countries: ~ 49 in 2015
- Areas: Maths, Science
- Min # students: 4,000
- Sampling: School sampling followed by whole class
- Cycle: 4 years

PISA

Target: 15 year olds
Countries: ~60 countries in 2015
Areas: Reading, Maths, Science, Problem solving
Min # students: 1,500 (pilot), 6,000 (main study)
Sampling: School sampling followed by student sampling
Cycle: 3 years

New versions for Grade 4+

TIMSS Numeracy

- assesses fundamental mathematical knowledge, procedures, and problem solving strategies
- test items are similar to TIMSS Grade 4 items, but numbers are simpler and procedures are more straightforward
- can be administered to students in Grades 4-6

PIRLS Literacy

- reflects the same conception of reading as PIRLS but is less difficult and is designed to assess basic reading skills that are a prerequisite for PIRLS
- reading passages are shorter, with easier vocabulary and syntax

Background questionnaires

System and school

System and school

Mathematics	
1. Number and operations	1-10
2. Algebra	11-20
3. Geometry and measurement	21-30
4. Data analysis and probability	31-40
5. Number and operations	41-50
6. Algebra	51-60
7. Geometry and measurement	61-70
8. Data analysis and probability	71-80
9. Number and operations	81-90
10. Algebra	91-100
11. Geometry and measurement	101-110
12. Data analysis and probability	111-120
13. Number and operations	121-130
14. Algebra	131-140
15. Geometry and measurement	141-150
16. Data analysis and probability	151-160
17. Number and operations	161-170
18. Algebra	171-180
19. Geometry and measurement	181-190
20. Data analysis and probability	191-200

Mathematics	
1. Number and operations	1-10
2. Algebra	11-20
3. Geometry and measurement	21-30
4. Data analysis and probability	31-40
5. Number and operations	41-50
6. Algebra	51-60
7. Geometry and measurement	61-70
8. Data analysis and probability	71-80
9. Number and operations	81-90
10. Algebra	91-100
11. Geometry and measurement	101-110
12. Data analysis and probability	111-120
13. Number and operations	121-130
14. Algebra	131-140
15. Geometry and measurement	141-150
16. Data analysis and probability	151-160
17. Number and operations	161-170
18. Algebra	171-180
19. Geometry and measurement	181-190
20. Data analysis and probability	191-200

Student and teacher

Student and teacher

4 How would you characterize each of the following within your school?

	Not like	High	Medium	Low	Not like
a) Teacher/Class interaction	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) Teacher understanding after class/extra-curricular work	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) Teacher's degree of stress in implementing the school's curriculum	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) Teacher's expectations for student achievement	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e) Parental support for teacher achievement	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
f) Parental involvement in school activities	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
g) Teacher's negative and personality	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
h) Teacher's role in school	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Homework

33 How often does your teacher give you homework in each of the following subject?

	Never	1 or 2 times a week	3 or 4 times a week	5 or 6 times a week	Sometimes
a) Mathematics	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) Biology	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) Earth science	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) Chemistry	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e) Physics	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

7 Thinking about your current school, indicate the extent to which you agree or disagree with each of the following statements.

	Strongly agree	Disagree	Strongly disagree
a) This school is located in a well-organized area	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) I feel safe at this school	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) This school's security policies and practices are sufficient	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) The students behave in an orderly manner in class	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e) The students are respectful of the teachers	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

8 In your current school, how severe is each problem?

	Not a problem	Minor problem	Relatively minor problem	Relatively serious problem	Serious problem
a) The school building needs significant repair	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
b) Teachers are overworked	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
c) Teachers have too many teaching hours	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
d) Teachers have too many administrative tasks that detract from the effective teaching of the students	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
e) Teachers do not have adequate instructional materials and supplies	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

South East Asian
– Primary
Learning Metrics

UNICEF
SEAEMQ-INNOTECH

Regional ILSA Example

SEA-PLM

South East Asian
– Primary
Learning Metrics

UNICEF
SEAEMQ-INNOTECH

- learning outcomes of primary school children (starting with 10 year olds)
- reading, writing, mathematics and global citizenship/civics education
- four countries currently piloting

Preparation, choices, and purposes

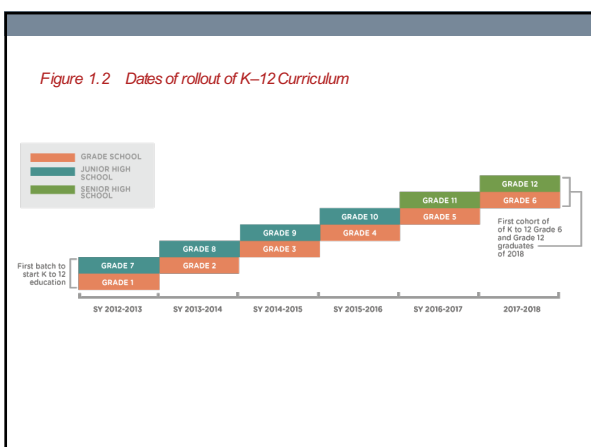
Preparation

Participation in LSA needs to consider:

- What the purpose is: what are the related questions?
- How would either high results or low results be explained in the context of the initial questions?

Research-informed hypotheses need to be made explicit, so that relevant data can be gathered prior to, or simultaneous with the LSA event. The suggestions in the table below are by no means exhaustive, but could act as prompts for discussion and formulation of hypotheses.

Purpose	Possible claims if ranking is higher than designated benchmark countries	Possible claims if ranking is lower than designated benchmark countries
Affirmation of need for reform	Reform is not justified since the cohorts being assessed have not experienced the full K-12 and country results are already fine	Reform is justified
Affirmation of effectiveness of reform	Reform is justified – without reference to the fact that cohorts assessed have not experienced the full K-12	Reform is not yet fully implemented, so this is an expected result
Use for benchmarking within country	Greater pressure for higher results next time to demonstrate effectiveness of the reform	Demonstrates starting point in the reform effort, with assumption of improvement next time around



Summary global ILSA

	PISA	TIMSS			PIRLS		
		TIMSS	TIMSS Advanced	TIMSS Numeracy	PIRLS	PIRLS Literacy	ePIRLS
Occur	Every 3 years	Every 4 years	Every 4 years	Every 4 years	Every 5 years	Start in 2016	Start in 2016
Dates of next studies	2015, 2018, 2021	2015, 2019	2015, 2019	2015, 2019	2016, 2021	2016, 2021	2016, 2021
Who is tested?	15-year-old	4th and 8th grade	12th grade (final year secondary students)	4th, 5th or 6th grade	4th grade (can be 5th and 6th grade)	4th grade	4th grade
What is tested?	Ability to apply skills and knowledge in real life contexts	Mathematics and science	Advance mathematics and physics	Numeracy: fundamental math procedures & problem-solving strategies	Reading	Reading: less difficult version of PIRLS	Computer-based reading assessments; online reading skills

Purposes	ILSA capacity
a) establishing baselines given the implementation of a new curriculum	✓✓✓
b) measuring effectiveness of curriculum implementation (internal national assessments) through international benchmarking and comparability (international large scale assessments)	✓✓
c) evidence-based program management at the system level (planning/budgeting, monitoring and evaluation) with assessment data as critical inputs	✓✓
d) establishing variability within the system both across region, across curricular area, and across year levels to be used as critical input to the over all K to 12 monitoring and evaluation (M&E) system	✓✓

Issues and sustainability

Issues

- Language in Grade 4
- Impact of language on science and maths outcomes
- Need for demonstration of understanding rather than rote learning
- How to engage the public and media prior to the events

Sustainability

- A primary purpose for use of LSA is to ensure capacity building at all levels of a system's educational enterprise
- Sustainability can be promoted through
 - appointment of a dedicated assessment and measurement team that will work not only on the ILSA, but also on national assessment
 - scheduled and maintained professional development for regional directors, principals and teachers
 - repeat participation in the LSA; the value of participation in more than one round provides much greater opportunities than afforded by benchmarking with one round only.

ACTRC

e.care@actrc.org

www.actrc.org

 www.facebook.com/ACTRC.org