
A “Buyer’s Guide” for Student Learning Assessments

Kevin Macdonald

UIS Consultant

GAML February 25th, 2025

Why do we need a Buyer's Guide?

Motivation:

- **Countries participating in assessments that are too difficult:** good for “shock value” but may cause embarrassment and data may not be useable
 - **Lack of sustainable capacity building:** e.g.: assessments financed by different donors repeatedly in the same country but no national assessment unit to sustain skills, expertise or experience
 - **Assessment data not used by countries to improve learning:** countries may not have the resources or knowledge about how to use the data
 - **Power imbalance:** donors and the Global North have significant influence over choice of assessment and design
- ➔ **More broadly:** lack of preparedness to participate and navigate the market for student assessments

Policy makers tend to understand differences between learning assessments but need support on choosing

- UIS receives requests from government counterparts about how assessments differ, and the Buyers Guide offers a response

Objectives of the Buyer's Guide

Objective: advocate for policy makers to better prepare their countries for international assessments

How?

- 1. Reinforce that countries have a choice in which assessments to participate in** (e.g.: that they are buyers in a market)
- 2. Highlight differences in assessment programs that policy makers may not pay much attention to but are important**, e.g.: that assessments differ in...
 - how actionable assessment results are or in how they can be used
 - how difficult they would be and when does difficulty matter
 - the type of capacity building and knowledge transfer
 - the ownership or influence a country has in assessment programs
 - cost (beyond fees but also in “return” on that investment)
- 3. Finally, reinforce that there should be a strategy + institutional set up that all this feeds into**

*Purpose of the Buyer's Guide is not measuring SDG 4.1.1 but on help countries achieve their (broad range of) goals

Five questions to ask yourself before buying

From the perspective of the buyer (a country) for student assessments:

- 1. Why am I participating in a student assessment?** Assessments vary by how well suited they are to different objectives and goals can conflict
- 2. How am I going to use the assessment results?** Different ways to use results (e.g: inform PD, raise awareness, policy) but need to know how you will use the data before choosing
- 3. Do I have a strategy that addresses participation in assessments?** Strategy would specify goals from assessments programs, how to improve capacity for assessment (including enabling environment)
- 4. Am I prepared to work with donors on financing an international assessment?** Governments need to be ready to leverage donor interest towards their own goals---having a strategy or policy is critical here.
- 5. Do I have an institutional home for assessment that can accumulate expertise?** International assessments are often implemented by a team that lacks an institutional home—the result is loss of institutional memory and expertise.

1. How actionable are the assessment results?

How easily can the results of an assessment inform how to improve learning outcomes? E.g.: what can I do specifically to improve results for future assessment rounds?

Secondary school assessments: PISA, TIMSS 8 th Grade, national assessments	Primary school assessments: TIMSS 4 th Grade, PIRLS, regional assessments, national assessments	EGRA
<ul style="list-style-type: none"> Late in students' education cycle, so results may be less malleable (particularly if foundational skills are not well developed) 	<ul style="list-style-type: none"> Early in students' education cycle and learning outcomes may be more malleable Item analysis to identify specific skills and competencies needing improvement evidence on impact of these approaches is limited. 	<ul style="list-style-type: none"> Interventions to improve EGRA results are well evaluated (RCTs), i.e.: offering a clear indication of what needs to be done to improve EGRA scores But implementing these interventions at scale not well developed

Note: that we are differentiated from use in terms of comparing policy and research, evaluating programs, etc.

2. How do assessments differ in difficulty?

Why does difficulty matter? Data may not be usable for identifying the skills and competencies that students need most support on

Trade-off: being able to compare to a large number of countries (e.g.: in TIMSS / PISA) but results may not be as usable for improving learning--though new tools (AMPL) offer alternatives

International assessment	Regional assessments	Others
<ul style="list-style-type: none"> Tend to be most difficult (designed for high-income country contexts) 	<ul style="list-style-type: none"> Tend to be closer in difficulty to participating country students by design GAML proficiency level linking and Rosetta Stone offers insights here 	<ul style="list-style-type: none"> National assessments should (in theory) be well aligned to students' ability EGRA / EGMA measure foundational skills (but still see many "zero" scores) AMPL can be integrated into assessments to create comparisons

3. How do assessments differ in capacity building?

- Generally, assessment programs provide training on implementation (adaptation, quality standards, sampling)
- Now increased capacity building related to the broader enabling environment.
- Regional assessment capacity building often tends to depend on donor financing.

International assessment: PISA / TIMSS / PIRLS

Included: training on implementation (e.g.: sampling, quality standards, adaptation, etc.)

Extra: implementation (additional support), analysis, reporting, system capacity (e.g.: enabling environment, PISA CBIS with ACER)

Missing: hands-on contributions to test item design?

Regional assessments: ERCE, PASEC, PILNA, SACMEQ, SEAPLM

Included: implementation, item design use of assessment, others (*to be updated*)

- Varies by assessment program and often dependent on donor financing;
- Less clear on how hands-on country teams can be

Others: AMPL, EGRA / EGMA

AMPL: provides training on adaptation, implementation

EGRA / EGMA: requires hiring expert in early reading and that person training country teams

4. How do assessments differ in ownership?

- Ownership reflects ability of countries to ensure assessments stay relevant to their contexts, but also for accountability and pride.
- Trade-off between ownership and the number of countries that participate
- Global North donors and experts may have substantial *de facto* influence in regional assessments

International assessments	Regional assessments	Others
IEA (PIRLS / TIMSS): through IEA membership	PASEC: steering committee membership	AMPL (through UNESCO membership)
OECD (PISA): OECD members (with some exceptions, otherwise observer status)	SEAPLM: through SEAMEO? Others? (<i>still collecting information here</i>)	EGRA / EGMA: technically available to anyone to adapt and use, though requires expertise
		National assessment: country owned

5. How do assessments differ in costs?

- **Fees paid to assessment providers only one part of cost**, countries also need to finance implementation
- **Consider benefits too**: e.g.: national assessment may be more expensive but represents an investment in capacity
- **How to best compare costs?**
- Per-student cost offers a good measure, independent of sample size, but data limited
- UIS has been collecting cost data for different assessment programs and compiling studies
- More detailed work being undertaken in Africa by World Bank (which will be added) Others?

Assessment provider fees (contributions) as a percent of total cost of implementing assessment

PIRLS	TIMSS	PISA	PASEC	ERCE	SACMEQ	SEAPLM	PILNA
29%	29%	28%	41%	44%	30%	0%?	27%

Source: UIS data compilation

Buyer beware: pitfalls and how to avoid them

- 1. How to avoid participating in an overly difficult assessment:** (a) understand difficulty relative to ability level of students and (b) be clear on goals / trade-offs
- 2. How to avoid not using the assessment results:** plan for how the data will be used before deciding on which assessment / whether to participate and in budgeting
- 3. How to avoid not retaining capacity from assessments:** having an assessment policy and an institutional home for assessment implementation

*Lots of resources already about these questions, but they need more attention!

A more efficient market for student assessments

- 1. Donor should encourage countries to have a strategic approach to assessment participation prior to financing assessment**
- 2. AMPL offers countries the ability to make international comparisons of minimum proficiency using assessments that are well aligned with student ability**
- 3. The need for the Virtual Fund for International Assessment**

THANK YOU

Learn more:

uis.unesco.org

databrowser.uis.unesco.org

[@UNESCOstat](https://twitter.com/UNESCOstat)

#25YearsOfDataInsights