



unesco

Institute for Statistics

Learning Assessment Buyer's Guide

Concept Note

Draft

2 August 2024

DRAFT

Motivation: The state of the “market” for Learning Assessments..... 1

Purpose of the Buyer’s Guide.....2

Main content of the Buyer’s Guide.....3

How do countries choose assessments in practice?3

Ten questions to ask yourself before “buying” a learning assessment.....3

Learning Assessment Comparison Chart.....4

Comparing (and Documenting) Cost5

Comparing Capacity Building Activities6

Buyer beware: what are some lessons? What are some pitfalls?6

Consultation and research activities7

DRAFT

Motivation: The state of the “market” for Learning Assessments

That there is a market for student assessments may not be immediately apparent, but countries are, in effect, buyers choosing among different assessments (i.e.: products) offered by a number of sellers. Country participation in international student assessments has been aptly characterized as a market: there are buyers (countries), sellers (assessment organizations) and products (the assessments themselves)¹. There are a range of assessments available that differ in skills measured (e.g.: TIMSS versus PISA), age of student being assessed (8th grade TIMSS versus 4th grade TIMSS versus EGRA), the breadth of countries that results can be compared to (e.g.: international versus regional versus national assessments), and so on. Participating in any of these assessments is of course not free: assessment organizations (e.g.: the IEA for TIMSS, the OECD for PISA) charge fees to participate effectively making them sellers and the countries (typically governments but also civil society) buyers. Once purchased, there are additional costs and resource requirements to countries for participating (e.g.: government staff time, school staff time, national workshops, etc.) that vary by the assessment purchased. Even if participation in an international assessment is financed by an external donor, the country is still in effect buying the assessment.

However, the market for learning assessments expresses many of the hallmarks of an inefficient and inequitable market. In economic theory, markets are efficient when there is competition, consumers have all the information about the products needed to make an informed choice including fully understanding the costs and benefits, that production technologies are known to all and can be copied, there are no barriers for new sellers to enter, and prices are transparent and uniform. Markets can be inequitable for many of the same reasons, for instance when wealthier or larger consumers can influence prices or have better information about products, but even if efficient, market inequities can arise from differences in purchasing power. From this perspective, the market for learning assessment appears to fall well short of the conditions needed for an efficient and equitable market². There is limited competition for a specific type of assessment as assessments differ in various ways. It is difficult for new providers of assessments to enter the market because of cost in developing a new assessment but also because countries want to participate in assessments with a large number of participants in order to compare their performance globally. There is also a lack of transparency and discrimination on prices with negotiation on fees as well as negotiation between third parties (e.g.: development partners) and providers.

¹ See Montoya & Crouch (2022) *The learning assessment market: pointers for countries – part 1* <https://world-education-blog.org/2019/04/26/the-learning-assessment-market-pointers-for-countries-part-1/>

² See Montoya & Crouch (2022) *The learning assessment market: pointers for countries – part 2* <https://world-education-blog.org/2019/05/20/the-learning-assessment-market-pointers-for-countries-part-2/>

A critical obstacle is a lack of information about learning assessments especially around how to choose which learning assessment is best suited to a country's context and how to fully benefit from participation. The available learning assessments (including national assessments) differ in a variety of ways. Being able to make a free and informed choice is needed in order for a country to participate in a learning assessment that is highly relevant to its needs and to fully benefit from learning assessments. One obstacle is that many countries lack clear information on not only the various learning assessments compare but also how to make this choice. This information is especially needed for lower income countries relying on external financing in order to help ensure that support from development partners for learning assessments are best allocated to meet countries' needs.

Purpose of the Buyer's Guide

The purpose of the Buyer's Guide is to promote countries' free and informed choice of learning assessments. The goal is to inform senior government policy makers of the following:

1. That there is **choice in student assessment**, each with pros and cons
2. That there are a few **technical differences that high level decision makers need to understand**, particularly related to policy relevance, sustainability and reputational risks (e.g.: embarrassment if the assessment is too difficult)
3. That **capacity building is a significant benefit** to participating in assessments but opportunities to build capacity sustainably differ by assessment provider
4. That assessments differ in **costs (beyond the fees charged by assessment providers)** and what these costs are
5. That the **choice should be made within the context of relevance to a country's educational needs and goals**, that is within the context of an assessment strategy or policy, even (perhaps especially) if relying on external donors

The scope of the guide are all assessment activities in a country. These include participation in the major international student assessment programs and regional assessment programs, but also these include assessments that are specific to projects, that are run by civil society, national assessments and to some extent national examinations that while technically are not assessments do in fact function as a (highly visible) measure of learning to the public and motivator for policy change.

Main content of the Buyer's Guide

The following highlights the main content expected to be included in the Buyer's Guide. The following present the main points that Buyer's Guide would address (this is not an outline for the guide).

How do countries choose assessments in practice?

Better understanding countries' (especially governments') decision-making process for choosing assessments is needed to improve the market for learning assessments.

What we know about why a country chooses to participate in one assessment but not other stems from a variety of sources including official (aspirational) statements, technical documents, policies and wide range of rumor. For example, what is the importance of political risk in choosing an assessment versus policy need? How much influence did development partners have in the choice of the assessment? Better understanding the decision-making process will be key to ensuring that not only the information in the buyer's guide is relevant to policy makers but also to assessment providers, international organizations and development partners.

Ten questions to ask yourself before "buying" a learning assessment

Ensuring that countries are well prepared before, during and after purchasing a learning assessment is essential to ensuring that the benefits of resources invested in learning assessments are fully maximized. The users guide would include a list of ten (or so) questions to help guide countries become better prepared for choosing and benefiting from learning assessments. These questions would revolve around the purpose or goal of the assessment (see **Box 1**), the link between the assessment and the educational context, and finally the strength of the institutional (strategic, policy, legal) environment in which the assessment would be conducted.

Box 1. What are the goals and purposes of participating in an assessment?

The goals and purposes of participating in an assessment (whether an international assessment or national assessment) are central to defining a buyers' guide. The following offers the purposes and goals underlying the proposed work:

- **Raise public awareness:** help the public through the media and civil society to better understand education quality, including international comparisons (e.g.: PISA shock Germany)
- **Research:** use by in-country and international researchers to identify weaknesses in the education system, make recommendations, evaluate programs etc.
- **Diagnosis and identifying interventions and reforms:** by research institutes including government, academic and non-government. For example, Jordan's NCHRD was doing item analysis with TIMSS data and to provide teachers with a guide

on what specific skills Jordanian students lag behind in compared to other countries. EGRA has been used in a number of countries used to design early grade reading interventions because the assessment has clear implications for teaching.

- **Setting national targets and planning:** International and national assessment allow governments to set targets for improving learning outcomes and identify needs for allocating resources.
- **Reporting on international commitments:** Assessments are used by governments who have committed to the SDGs to report on SDG 4.1.1 that measures children achieving a minimum proficiency level as well as other related indicators.
- **Monitoring and evaluating specific investments:** Assessments are used to evaluate specific projects or programs. In many cases these are tied to a specific (often donor-funded program) project, and a challenge is ensuring that these efforts are coordinated and promote capacity building nationally. Assessments can also be used opportunistically to evaluate programs (e.g.: Tunisia's planning unit used TIMSS to assess how well their school support program targets the lowest performing students).
- **Examination (high stakes) versus assessment:** Examination is not a purpose of assessment but in many countries, examination is highly visible, particularly for parents and students, and in countries where pass rates are low, a chief indicator of a problem.
- **Capacity building:** Assessment programs can develop local capacity for assessment design and implementation as well as policy analysis. This varies by assessment providers with some having a clear approach to this while others do not, some involving country experts only in sampling while others involving experts in item development and policy analysis. The challenge here is that this knowledge transfer is not well institutionalized and generally ad hoc, particularly in low- and middle-income countries.
- **Ownership and pride:** This is not often thought about as a purpose for an assessment, but this is one area where countries derive value from assessments, for example, to ensure that assessments are culturally relevant and reflective of a region's values (e.g.: PILNA).

Learning Assessment Comparison Chart

Given the various goals and purposes of participating in a learning assessment, the heart of the buyer's guide would be a comparison chart that would enable a country to identify the assessments that best fit their specific purposes. There are a number of key technical features that differentiate assessments that are relevant to the various goals and purposes of an assessment. Relevant features include the most visible qualities including types of skills assessed, the grade level or age of children assessed, and the number of countries results can be compared with. There are also a number of less visible

but important features. For example, assessments vary in their difficult level, and ensuring that the difficulty of the assessment aligns with the ability level of students in a country is needed in order to accurately assess their skill level and the characteristics of their skills. How actionable the results of an assessment are, also varies. For example, countries struggle with question of how to make gains in major international student assessments including PISA and TIMSS; however, how to improve poor results in the EGRA are well understood and researched. Cost and capacity building opportunities are also critical, and these are discussed subsequently. A comparison chart covering these features (as well as cost and capacity building opportunity) for a variety of assessments (see Box 2) would form the core of the Buyer's Guide content.

Box 2. Learning assessments included in the comparison chart

The Buyer's Guide would compare a number of different learning assessments. The following is the tentative list of assessments that proposed:

Global international assessments: PISA, TIMSS and PIRLS

Regional international assessments: ERCE, PILNA, SACMEQ, PASEC, SEA-PLM

Small scale (typically) assessments: EGRA, EGMA

Assessments targeting developing countries: PISA-D, LANA

Module calibrated to the Minimum Proficiency Level: AMPL

School readiness assessments done at preschools or in communities: EDI, EHCI

Population-based assessments: MICS

Citizen-lead assessment: UWEZO, PAL network

A generalized national assessment would also be included for comparison as well.

Comparing (and Documenting) Cost

Fees paid to assessment providers as well as costs (and resources more broadly) vary by assessment but are not well documented. As discussed previously, fees paid to assessment providers by countries are not uniform and can vary with scope for negotiations and subsidies. External donors also negotiate fees and cover costs of assessments outside of government budgets. Finally, governments incur costs that are clearly monetary (organizing workshops, paying for travel) but also costs that are not recorded as extra costs in government budgets including government staff time allocated to assessment activities. There is also resources consumed at the school level including the time of teachers and students. Because of these different sources of cost, there is very little information currently available for comparing the cost of assessment programs. The Buyer's Guide aims to help estimate and present comparable cost data for assessments.

Comparing Capacity Building Activities

Assessment programs provide capacity building opportunities in varying extents, but the preparedness of countries to benefit from sustainable capacity building also varies. First, assessment programs vary by the types of expertise they develop. For example, most international assessment programs provide training and involve local experts in sampling (e.g.: PISA, TIMSS) while others also involve local experts in item development (e.g.: AMPL). Training on analysis of assessment results also varies with some providing support depending on the availability of financing (e.g.: PASEC historically) and other charging additional fees for producing a national report in collaboration with local experts (e.g.: PISA). They also vary by how intensive the training is, ranging from training seminars to more in-depth, hands-on training (e.g.: Laboratorio). However, to benefit from training and capacity building in a manner that is sustainable, countries need to be well prepared. For example, countries that have an established assessment unit that is responsible for various assessment activities in the country (whether a government unit or official research institute) would generally benefit more from capacity building activities because the expertise is kept within government among individuals who are involved in assessment work routinely. Training would be less beneficial in countries where the team conducting the assessment is an *ad hoc* team from, for example, the planning unit staff lacking the needed qualifications prior to the training and not involved in using the training in the future.

Buyer beware: what are some lessons? What are some pitfalls?

Finally, the Buyer's Guide would offer generalized lessons of best practice and pitfalls drawing on the experience of countries, assessment providers and other partners. One lesson, for example, would be the need for countries to have a clear assessment strategy or policy to help coordinate external donors' assessment activities. There are examples of how external donors have financed student assessments but in implementation have bypassed regular government agencies in favor of temporary government entities for project implementation (i.e.: project management units) resulting in governments not gaining the expertise or institutional structure needed to continue assessment activities. There are also examples of external donors financing assessments whose level of difficulty far exceeds the ability of students resulting in an inability to use the data for meaningful analysis. More broadly, there is a trend in which the major international assessments have been initiated and developed by high income countries and academic circles and subsequently extended to low- and middle-income countries. This creates a *de facto* imbalance in the ownership and accountability of international assessments leading to inequality in the market. There are also positive examples as well; for example, the "ownership" Laboratorio effectively transitioning to Latin American countries and the improvement in the technical quality of the assessment.

Consultation and research activities

In preparing the Buyers' Guide, the following consultation and research work is envisioned:

Consultation of assessment providers: At the heart of the Buyer's Guide is a comparison of assessment programs including specifications and the pros and cons of each for specific purposes. Assessment providers have the opportunity to respond to the claims made in the Buyer's Guide is needed to ensure factual accuracy, for fairness, and for the credibility of the guide itself. One approach may be to publish comments and responses from the UIS to ensure transparency.

Collection and documentation of cost data: There are currently some sources of cost data that are available, and these will be compiled for the User's Guide. Requests for information may also be made to assessment providers and countries.

Understanding assessment decision making in practice: Discussions with senior policy makers and local experts in selected countries are proposed in order to provide a better picture of how assessment choices are made by countries.

Documenting good practice and missed opportunities: Finally, the User's Guide will document good practices in how assessment programs are initiated or evolve across time as well as document missed opportunities. For example, the progress made by Laboratorio both technically and in terms of ownership may be included. The role of high-income countries and academia in the initiation and expansion of assessment programs to low- and middle-income countries may be discussed as well as some country examples.