

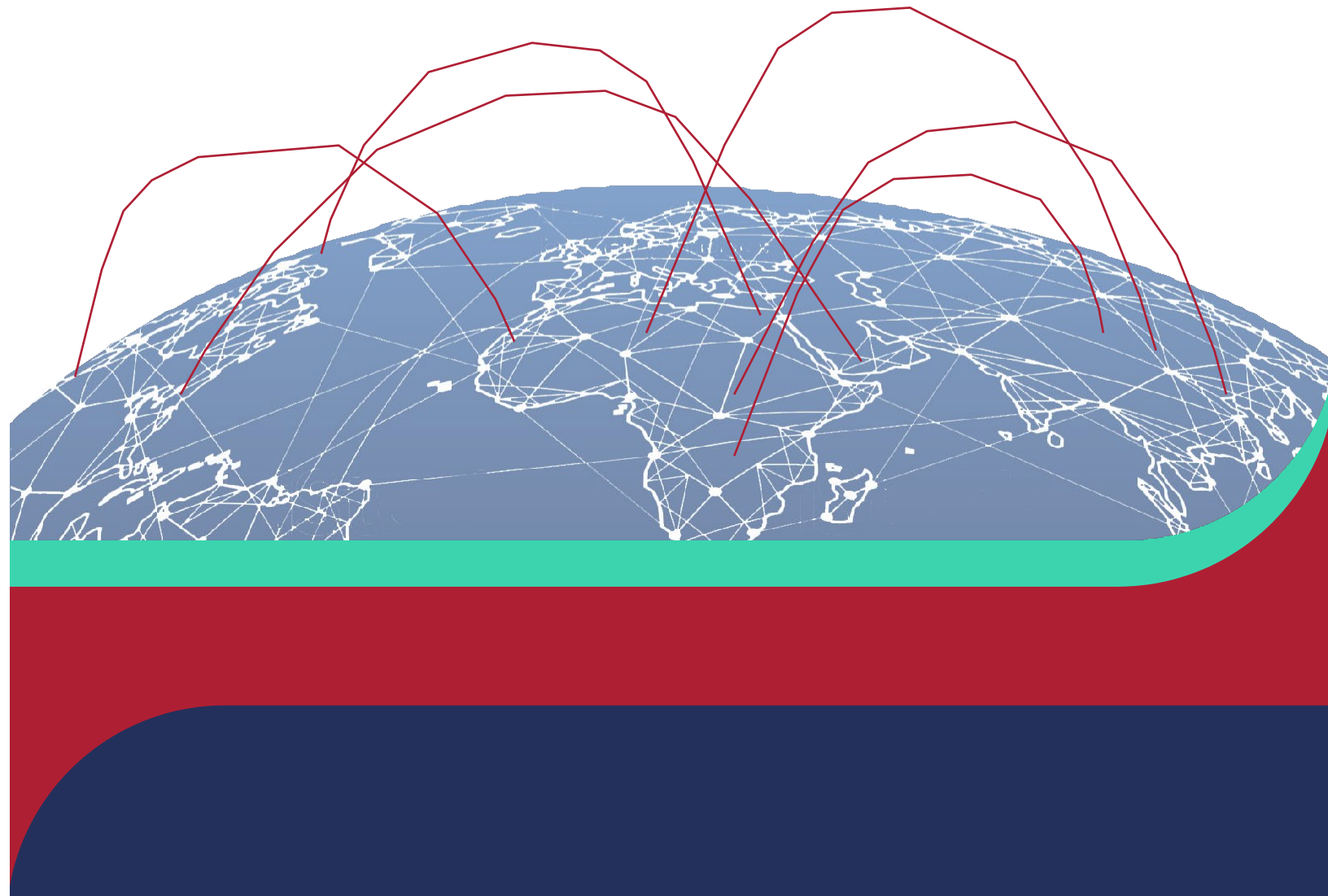
Assessments for Minimum Proficiency Levels (AMPL): an efficient and effective tool for global reporting and strengthening capacity to conduct national assessments

Dr. Silvia Montoya

Director

UNESCO Institute for Statistics

August 2023



Contents

Contents	2
1. Background	3
2. Producing high-quality learning data: challenges and solutions	4
2.1 Instruments to measure learning and report on SDG 4.1.1	4
2.2 Challenges for collection of comparable learning data	4
2.3 Setting standards and expanding options to improve reporting	5
3. Assessments for Minimum Proficiency Level (AMPL)	6
3.1 Development of AMPL modules	6
3.2 Available AMPL modules and language	7
3.2.1 AMPL administration: modalities and languages	8
3.2.2 Combining AMPLs with other assessments tools	9
4. How to integrate AMPL into a National Learning Assessment program	10
4.1 Parallel integration as whole booklet form	11
4.2 Full integration	11
5. References	13

1. Background

The lack of reliable data to measure learning outcomes and progress in learning over time poses a real challenge to policymakers, especially in countries that have not participated in any international or regional learning assessments. In addition, national assessments are not comparable across countries due to different curriculum objectives, coverage of constructs and sub-constructs, assessment frameworks, and items used.

A lack of viable learning assessment systems in countries makes it difficult to monitor learning. Globally, more than 60% of countries do not have data on learning, and this is especially pronounced in Africa. The lack of such data not only hinders countries to monitor their progress in the quality of learning, but also is obstructive in developing appropriate learning strategies for improving quality of education at the international level. According to the UNESCO Institute of Statistics (UIS), 135 countries and territories, accounting for nearly 1.4 billion children aged 14 and under, lack latest quality education data on learning outcomes in reading and numeracy as of November 2021.

‘If we are to measure progress across and within countries across years, we need reliable data that measures what matters and can be compared over time’ (Montoya et al., 2022). Sensing the urgency to establish concrete steps to obtain high quality, globally comparable learning data that can be used to improve national education systems, the UNESCO Institute for Statistics (UIS) started exploring possibilities and limitations of developing a global assessment strategy for SDG Indicator 4.1.1 by comparing different international, regional, and foundational skills assessments of literacy and numeracy, providing the criteria to make comparison across assessments ([Órdenes and Treviño, 2017](#)).

Collaborating with several partners, the UIS developed a [set of tools that would make this possible](#) including: 1) standards would comprise the definition of Minimum Proficiency Levels (MPL) that actors could have as a reference point for reporting on the set of competencies considered as minimum for every child at a certain schooling level, without having to have a single test as the solution to comparability; 2) the [Global Proficiency Framework](#) and its related MPL thus gives guidance as to the skills that students should acquire on the pathway to mastery of reading and mathematics; 3) a set of rigorous methods to align assessments to this common framework to complete the suite (AMPL; Rosetta-stone; Pairwise comparison; policy linking).

In that 2017 study, one of the strategies presented consisted of the creation of a Worldwide Proficiency Assessment on Numeracy and Literacy to serve for global reporting ([Órdenes and Treviño, 2017](#), page 25) that would allow report for Global indicators 4.1.1 without alternating the integrity of the different assessment programs. The ‘Assessments for Minimum Proficiency Level’ - or AMPL – that the UIS has been developing since 2021 responds to this strategy.

After the introduction, this document is organized as follows. Section 2 summarizes the progress regarding the production of high-quality learning data. Sections 3 and 4 are the core of the paper and describe in detail the Assessment for Minimum Proficiency level, its implementation and possibilities of integration within a national learning assessment program.

2. Producing high-quality learning data: challenges and solutions

High quality learning data is necessary to understand the achievements and gaps of learning and to guide policy making. It plays an instrumental role in identifying targeted students who need more help than other students. It also helps in determining the type of support that students need.

High quality learning data is also essential in identifying challenging areas of the curriculum with difficulties in implementation. It highlights the classroom and school factors which hinder and foster learning outcomes.

Ministries and development stakeholders need comparable learning data over time to know whether students are learning, to identify which students are being left behind and to understand where to adjust education policies and practices. However, efforts to collect learning data are often fragmented and irregular, especially in Low- and Middle-Income Countries. Many data investments are made at the country level without consistently adhering to technical guidelines, potentially leading to challenges in achieving comparability.

Worldwide, there is a wide variety of assessments at regional and international levels that assess education performance for children and youth in literacy and numeracy. Each of these instruments fulfills different purposes and gives relevant evidence for informing decision-making in different educational contexts. For global leaders to be able to agree on a common strategy of assessing indicator 4.1.1 and monitoring progress towards it, the information deriving from these assessments needs to be studied to see how comparable it is for defining global indicators.

The process requires the definition of a minimal level of competency in order to comply with the objective of SDG 4 that is that is to have assessment programs that are effective as they place students reliably on a scale, identify key drivers to guide policies, guide instructional responses to improve learning outcomes and track outcomes over time to detect progress.

2.1 Instruments to measure learning and report on SDG 4.1.1

The core of indicator 4.1.1 is measuring the proportion of students at a minimal level of competency. Based on the agreement in 2018, various cross-national programs are used to report the indicator. Those are programmes - international and regional - to measure learning outcomes. International programmes include the Programme for International Student Assessment (PISA), Progress in International Reading Literacy Study (PIRLS), Trends in International Mathematics and Science Study (TIMSS), and Progress in International Reading Literacy Study (PIRLS). Regional learning assessments include the Pacific Islands Literacy and Numeracy Assessment (PILNA), the Southeast Asia Primary Learning Metrics (SEA-PLM), Programme d'Analyse des Systèmes Educatifs de la CONFEMEN (PASEC), Southern and Eastern Africa Consortium for Monitoring Educational Quality (SACMEQ) and the Estudio regional para la calidad educativa (ERCE/PERCE/SERCE/TERCE/ERCE).

Along with these cross-national assessments, there are other ways to measure learning outcomes including national examinations and sampled-based and census-based national learning assessments.

2.2 Challenges for collection of comparable learning data

Collecting comparable learning data over time and between countries is very challenging for several reasons. First, most assessments do not measure what is relevant as they often focus on

measuring the content knowledge without measuring the specific sub-skills that lead to reading with comprehension: measuring these sub-skills is important as it will allow education actors to identify and target specific gaps among students who need help. Second, many assessments are not designed to be psychometrically comparable over time. In addition, comparability is prevented when subject and grade assessed change. Third, it is difficult to compare assessments between countries because different assessments test different skills at different grades and at varying levels of difficulty. Fourth, it is true that international assessments may produce comparable data; however, they have low coverage in low-income and lower-middle income countries, particularly for the early grades of primary. In addition, primary grade international assessments take place every five to six years which is too long to provide meaningful information and inform decisions. Finally, learning assessments within donor projects are not sustainable as they are often limited to the beneficiaries and timeline of the projects (Montoya et al., 2022).

2.3 Setting standards and expanding options to improve reporting

The UNESCO Institute for Statistics (UIS) played a leading role and collaborated extensively with partner organizations to find solutions to enable countries to improve their learning data, building on their existing assessments.

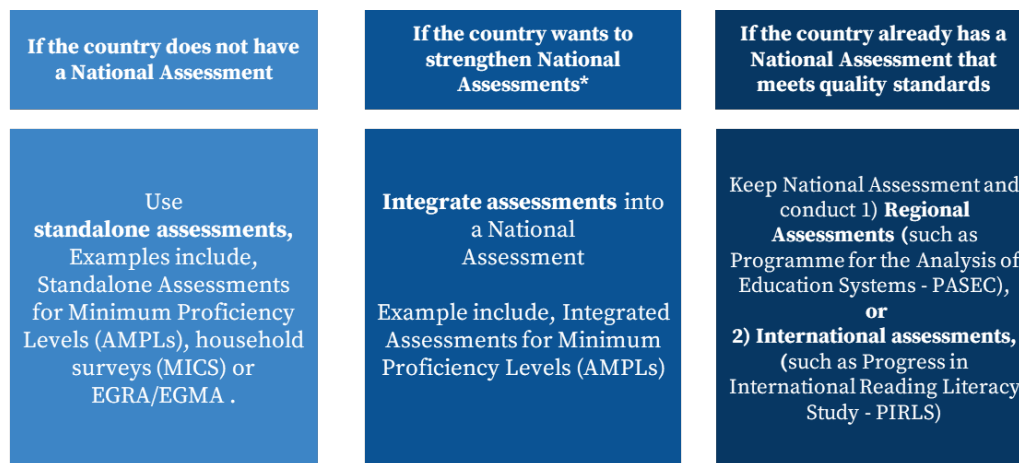
An important amount of work was invested and a time-consuming but necessary consensus process was done to define the Minimum proficiency levels (MPL) which constitute the benchmark of basic knowledge in a domain (mathematics, reading) at a given age/grade measured through learning assessments. This implies the agreement on a set of competencies that students should master and allows assessment programs to report as long as it is possible to identify the Proficiency Level (PL) that is aligned to the agreed definition of the MPL eliminating the unique test requirement.

The set of tools includes a common framework, the Global Proficiency Framework (GPF). The GPF provides internationally accepted definitions of reading and mathematics constructs and defines up to four Proficiency Levels for each grade and domain to help identify a learning transition and guide teaching and learning. One of the proficiency levels is the globally agreed MPL.

To complete the standards, rigorous methods have been developed to strengthen existing learning assessments (national, international and household based) and link them to this common framework both statistically (Assessment of Minimum Proficiency level – AMPL, pairwise comparison and policy linking). These methods expand the options of reporting.

Recommendations have been made according to country's initial situation and assessment history presenting a menu of options that analyzes all different aspects such as the definition of and alignment to the global minimum proficiency level and presents the costs and benefits of the different linking strategies and options (Figure 1).

Figure 1 – Assessment Roadmap.



3. Assessments for Minimum Proficiency Level (AMPL)

Assessments for Minimum Proficiency Levels (AMPL) are ground-breaking and robust tools to measure the attainment of a single proficiency level for reading and mathematics at a given level of education. AMPLs allow to identify the proportion of children and young learners in a level of education achieving at least the MPL.

AMPLs were developed to meet both the need for quality and internationally comparable learning data, and the need for a flexible administration mechanism based on country needs and capacity (UIS, 2022). AMPLs are based on the published technical standards, adoption of modern measurement practice and scientific sampling methodology. The implementation process of AMPLs is technically rigorous and participatory, with a standard setting exercise and hands-on capacity building enabling countries to use the tools and methods provided independently. A country-specific report will be ready within four months after submitting data to the technical partner and/or the UIS. The SDG4.1.1 indicator(s) will be produced by sex, which is a reporting requirement. It is also possible to produce sub-national level tables based on interest and subject to sampling design.

3.1 Development of AMPL modules

The UIS has developed AMPLs in technical collaboration with the Australian Council for Educational Research (ACER) and the AMPL modules were piloted in six African countries in the Monitoring Impacts on Learning Outcomes (MILO) project funded by the Global Partnership for Education (GPE).

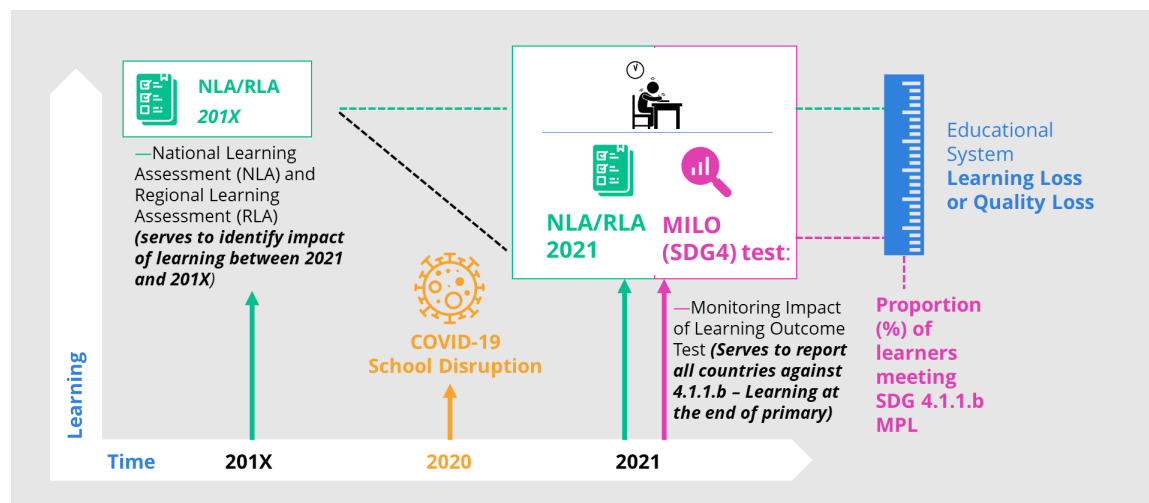
The four overarching goals of the MILO project were to:

- Evaluate the impact of COVID-19 on learning outcomes by reporting against SDG Indicator 4.1.1b
- Identify the impact of different distance learning mechanisms put in place to remediate the learning disruption generated by COVID-19
- Expand the UNESCO Institute for Statistics (UIS) bank of items for primary education
- Generate a toolkit to scale assessment results to international benchmarks, reporting against SDG 4.1.1b

The project consisted in repeating an assessment previously administered to identify the losses and to benchmark against the global definition of MPL for end of primary by adding a calibrated module for the MPL at the end of primary for both reading and mathematics (AMPL). This AMPL module developed to that purpose would allow to understand the percentage of students achieving the MPL as show in the figure.

The development of the AMPL is an important step forward and has the potential to align national and cross-national assessment programs to a single set of global standards in mathematics and reading as articulated in SDG 4.1.1. AMPL modules have technical standards for data processing and sampling and are calibrated to the globally established Minimum Proficiency Levels (MPL) for reading and mathematics and were developed to allow flexibility to measure each of the three levels of SDG 4.1.1 separately, or in combination, depending on the countries' capacity and fit to their agenda (Figure 2).

Figure 2 - Monitoring the impact of COVID in Learning (MILO)



3.2 Available AMPL modules and language

AMPL modules could exist for the different MPLs defined by indicator 4.1.1. **AMPLb** was the first module developed in 2021 and it was administered in English and French in 6 African countries within the MILO project. AMPL b is targeted at the MPL of the end of primary allowing the reporting of SDG indicator 4.1.1b.

In 2023, the UIS engaged into the development of **AMPLa** module that is targeted at the MPL of indicator 4.1.1a, that is lower primary. Given the MPL measures various subskills including listening comprehension, it has a segment of 20 minutes that is administered through a recording produced by the country in the relevant language by native speakers.

This new module targets at MPL 4.1.1a could be administered at the end of primary and would allow policy makers to understand not only how many students are below the MPL for indicator 4.1.1b but also to understand the percentage of students that are at the end of primary who are already mastering the level of MPL of the end of lower primary. This is critical information for any policy maker, to be able to understand what are the set of skills that students not meeting the MPL for the level are not able to master; key information to design appropriate policy interventions. This assembling of modules that allows to identify students in 2 proficiency levels is denominated **AMPLab** as reflected in Table 1 below.

Summarizing, Table 1 shows that AMPLa, AMPLb and AMPLc produce internationally comparable data to report on SDG4.1.1a, SDG4.1.1b and SDG4.1.1c indicators, for early grades, end of primary and end of lower secondary respectively. AMPLc is still under development.

Table 1. AMPLs modules, target population and related indicator

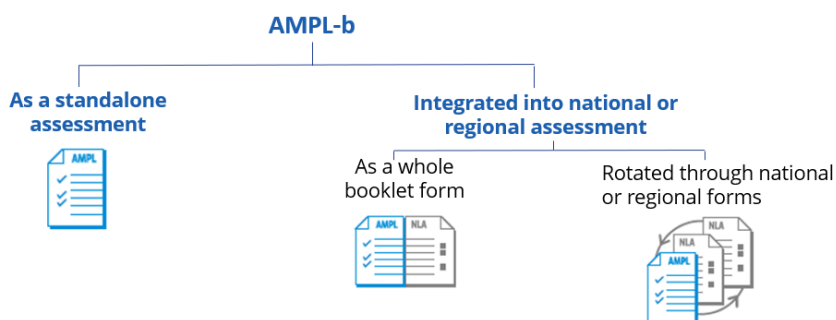
Administration	Objective	Target population	Measures MPL	Serves to global reporting of indicator
AMPL-a	To determine proportion of population above and below MPL-a/measures MPLs at early grades	End of lower primary (early grades)	MPL 4.1.1a	SDG 4.1.1a
AMPL-b	To determine proportion of population above and below MPL-b/measures MPLs at end of primary	End of primary	MPL 4.1.1b	SDG 4.1.1b
AMPL-a+b	To determine proportion of population above and below MPL-b (end of primary) also serves to measure the proportion of population at the end of primary that achieves MPL-a (early grades).	End of primary in low-performing contexts. End of lower-primary in high-performing contexts	MPL 4.1.1a and MPL 4.1.1b	SDG 4.1.1b
AMPL-c (under development) * ¹	To determine proportion of population above and below MPL-c	End of lower secondary	MPL 4.1.1c	SDG 4.1.1c

3.2.1 AMPL administration: modalities and languages

The administration of AMPLs allows countries to respect the integrity of the national assessment and or to build a national assessment using this module calibrated only for one proficiency level. They can be administered as a standalone assessment or integrated into existing national and regional assessments to produce internationally comparable SDG4.1.1 indicators (Figure 3).

¹ A PISA module has been developed by OECD and is targeted at PISA level 2.

Figure 3: Alternative administration of AMPLs



AMPLa, b and ab have been administered in various languages and countries. AMPLb was integrated into PASEC, the regional assessment, in Burkina Faso, Burundi, Côte d’Ivoire and Senegal; it was integrated into the national assessment of Kenya and Zambia (Table 2) as part of the MILO project. AMPLa is finalizing its administration this quarter and close to reporting.

The AMPL modules are further being developed by adding several languages and improving other technical aspects. In fact, they have been translated and administered in Urdu, Hindi, Arabic and Spanish. The AMPLs are continuously replenishing item improvement, cross-linguistic improvements and the equality of item and examining parameters from different populations or measurement conditions (parameter invariance) and adaptation following some [technical parameters](#) (UIS and ACER, 2023).

Table 2: AMPL implementation by country, type, language and modality of administration.

Country	AMPLa	AMPLb	AMPLab	MILO	Language	Administered modality
	Early grades (SDG 4.1.1a)	End of primary (SDG 4.1.1b)	End of Primary (SDG 4.1.1b+MPLa)			
Gambia	X				English	Standalone
India*	X	X			English+Hindi	Standalone
Jordan		X			Arabic	Standalone
Kenya			X		English	Standalone
Lesotho			X		English	Standalone
Pakistan		X			Urdu	Standalone
Sierra Leone		X			English	Standalone
Zambia	X		X		English	Standalone
Burkina Faso				X	French	Integrated AMPL/Regional assessment (PASEC)
Burundi				X	French	Integrated AMPL/Regional assessment (PASEC)
Ivory Coast				X	French	Integrated AMPL/Regional assessment (PASEC)
Kenya				X	English	Integrated AMPL/National assessment
Senegal				X	French	Integrated AMPL/Regional assessment (PASEC)
Zambia				X	English	Integrated AMPL/National assessment

* Piloting for possibility to integrate in next NAS cycle for Grade 3 and 5

3.2.2 Combining AMPLs with other assessments tools

The AMPL assessment is a short assessment in mathematics and reading which can be used to determine a populations attainment of Minimum Proficiency levels in those constructs and that can be used to report against SDG indicator 4.1.1. The MPL proficiency level module AMPL

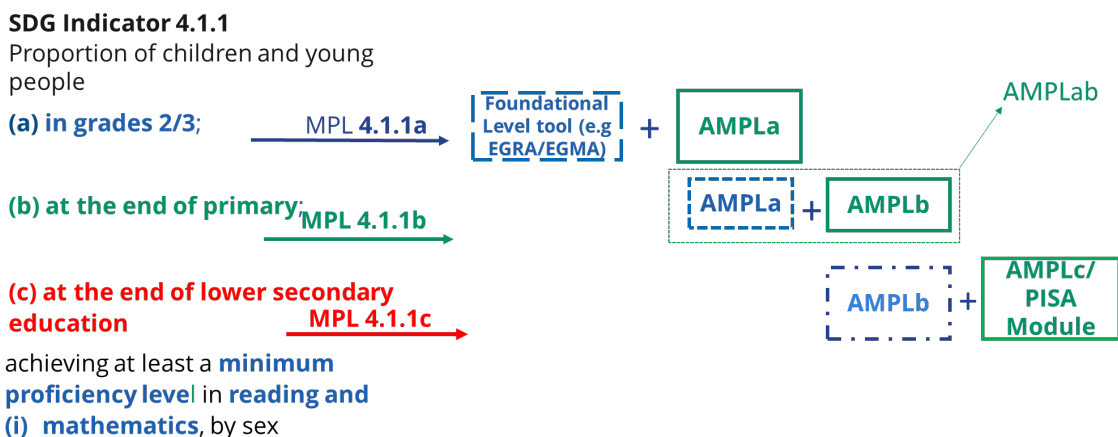
might not be able to offer enough granularity of information to countries where most of the children are below that MPL. This is a very likely situation in Low and Low middle income countries. This is the reason why other tools calibrated to level of proficiency below the MPL could be a good complement to obtain more precise information to structure policies and remedial measures.

The rationale described above was the main reason why AMPLa is administered as well at the end of primary. This facilitates for a country to identify the students above the MPL with module AMPLb while module AMPLa offers the intended details about specific skills and contents mastered for those below the MPL of end of primary (SDG 4.1.1b).

Figure 4 below shows that a possible scenario for countries to take full advantage of available tools is to combine them in a way that allows not only global reporting but also maximizes the quantity of information under common parameters. The type of implementation currently underway for Zambia and Lesotho is also possible to expand to other countries for AMPLab could also be extended to 4.1.1a.

In the case of 4.1.1, the MPL proficiency level achievement is eventually too high for many students and countries could find necessary to obtain granularity in information about each of the steps of early reading such as oral fluency, reading comprehension and listening comprehension. Tools such as EGRA/EGMA or national initiatives of the type that look at early stages of the “learning to read” process could provide valuable insights into students' foundational learning and therefore, combined with AMPL modules, would enhance the country's capacity to monitor learning outcomes, tailor interventions, and make informed policy decisions to improve educational quality and equity (Figure 4)

Figure 4: Integrating AMPLs with other assessments- Potential scenarios



4. How to integrate AMPL into a National Learning Assessment program

National assessments generally assess learning outcomes based on curriculum expectations and may include assessment material in reading and mathematics as well as other constructs (writing or science). In many LMIC, national assessments are applied in single forms, which means that all students receive the same assessment, and results are reported in percent correct.

There are several types of national assessments that are used in LMIC. These include sample-based monitoring assessments, examinations, and census-based diagnostics.

The main purposes of integrating AMPL into a national assessment program are:

- To assist the development of the national assessment by establishing national reporting scales based on objecting measurement methods (IRT).
- To assist the development of the national assessment by ensuring the assessments are strongly aligned to the constructs of reading and mathematics as described in the GPF and are targeted at the globally agreed MPLs.
- To locate the global MPL benchmarks on the established national assessment scales.
- To determine the proportion of target populations meeting the global MPLs.
- To build technical capacity in national assessment centers to proceed with conducting high quality national assessments that can be used to determine the proportion of target populations meeting the global MPLs, independently from incorporating an AMPL.

There are two main types of integrating AMPLs into national assessments.

4.1 Parallel integration as whole booklet form

Refers to administering AMPL with a pre-determined test design to a set of students who are administered a national assessment that follows a second independent test design. AMPL may be administered in parallel to any type of national assessment in a synchronous or asynchronous way, i.e. the AMPL can be administered before, alongside or after the national assessment is administered.

In a parallel design, the AMPL and the national assessments must be linked through students. In other words, the set of, or a sub-set of, students completing the national assessments must also complete the AMPL assessments. It is very important to ensure appropriate data linkage between the national and AMPL assessments through a unique student identifier. Appropriate data management is essential. Because of the high stakes nature of examinations all students must take the same assessment under the same conditions. Therefore, AMPL should be administered in parallel rather than fully integrated to avoid impacting on students' educational careers.

4.2 Full integration

Refers to administering the AMPL within a single test design that incorporates the national assessment. Full integration creates some dependency between the national assessment and AMPL, but parallel integration could be preferable in some circumstances – such as when the research goals are more focused on positioning effects within national assessment. In general, full integration would only be advisable once the AMPL has been validated in the language of administration and preferably after an initial parallel integration within the country.

The integration process involves several steps including:

- review of the national assessment construct alignment to GPF and MPL
- review of quality of items
- translation and linguistic quality assurance
- test design
- sample design
- data management

- assessment implementation
- psychometric analysis
- reporting

ACER (the UIS technical partner) has developed guidelines that assume that a national assessment team will be assisted by a UIS technical support agent (ACER, 2023b, draft in progress). The process has a series of steps that include a pre-screening that attempts to determine whether a process has delivered a satisfactory outcome for integration to proceed; the process needs to be revisited for improvement; integration should be abandoned and consist of reviewing a) national assessment construct alignment to GPF & MPL; b) item quality and c) sample design and implementation.

These pre-screening conditions are not different than the ones to define in an assessment is suitable for Policy Linking (See [Policy Linking toolkit](#)) . Construct alignment of the national assessments to the GPF is an essential pre-requisite to AMPL integration. If national assessments assess other constructs (such as writing or science) beyond the reading and mathematics constructs will not be included in the psychometric alignment of the national and AMPL assessments. The psychometric goal of integrating AMPL and National Assessments is to equate the national assessment scale with the AMPL scale so that the benchmarks from the latter can be applied to the former.

5. References

Australian Council for Educational Research (ACER), (2023). *Guidelines and considerations for integrating AMPL modules into national assessment programs* (Draft in progress).

Montoya et al. (2022). *Measure what matters: Making progress on a common framework to measure learning*. <https://www.globalpartnership.org/blog/measure-what-matters-making-progress-common-framework-measure-learning>

Montoya et al. (2023). *Helping countries get comparable learning data*. <https://www.globalpartnership.org/blog/helping-countries-get-comparable-learning-data>

Montoya, S. (2023). *Reporting learning outcomes in basic education: Country's options for indicator 4.1.1*. UNESCO Institute for Statistics. Montreal, Canada. https://gaml.uis.unesco.org/wp-content/uploads/sites/2/2023/05/Countrys-reporting-option_Zambia_2023.05.15_FINAL.pdf

Treviño, E. and Órdenes, M. (2017). *Exploring Commonalities and Differences in Regional and International Assessments*. Information Paper No. 48. UNESCO Institute for Statistics. Montreal, Canada. <http://uis.unesco.org/sites/default/files/documents/ip48-exploring-commonalities-differences-regional-international-assessments-2017-en.pdf>

UNESCO Institute for Statistics (2022). *Assessments for Minimum Proficiency Levels (AMPLs) Ground-breaking tools to produce internationally comparable data on SDG 4.1 indicators*. Brochure. <https://milo.uis.unesco.org/wp-content/uploads/sites/17/2022/10/ampl.pdf>

UNESCO Institute for Statistics (2023). *Terms of Use for Assessment of Minimum Proficiency Level (AMPL) products*. https://milo.uis.unesco.org/wp-content/uploads/sites/17/2023/08/AMPL-Terms-of-Use_2023.08.11-1.pdf

UNESCO Institute for Statistics (UIS) and Australian Council for Educational Research (ACER) (2023). *Adapting AMPL instruments: Assessment of Minimum Proficiency Levels (AMPL)* (Draft in progress). https://milo.uis.unesco.org/wp-content/uploads/sites/17/2023/07/Instrument-Adaptation_AMPL.pdf

UNESCO Institute for Statistics (2023). *Policy Linking for Measuring Global Learning Outcomes Toolkit : Linking assessments to the Global Proficiency Framework*. https://gaml.uis.unesco.org/wp-content/uploads/sites/2/2023/05/Policy-Linking-Toolkit-version-3.0_FINAL-2023.04.18.pdf