



UNESCO  
INSTITUTE  
FOR  
STATISTICS



GLOBAL  
ALLIANCE  
TO MONITOR  
LEARNING



SUSTAINABLE  
DEVELOPMENT  
GOALS

## Reporting Indicator 4.6.1



## Background

SDG 4 calls for an increased focus on learning outcomes, with five of the ten education targets highlighting learning skills and outcomes for children and adults. The UIS established GAML in 2016 as a platform to convene technical experts, donors and international organizations to provide technical solutions to the learning-related indicators. The GAML work programme includes the development of standards, guidelines and measurement tools to collect data to inform SDG 4 indicators.

Target 4.6 calls on countries to “ensure that all youth and a substantial proportion of adults, both men and women, achieve literacy and numeracy” by 2030. More specifically, Indicator 4.6.1 refers to: “Proportion of population in a given age group achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills, by sex”.

### Framework for reporting indicator 4.6.1

The fitness for use of any data system can only be evaluated against the overall purpose of the data. As documented in **Table 1**, comparative data on the level and distribution of adult literacy and numeracy skills are needed to serve five distinct purposes, which have implications for the data collection strategy.

Comparative data on literacy and numeracy are needed by multilateral and bilateral donors to guide their policies and programmes and to monitor progress towards international and national targets, including SDG Target 4.6. It is also imperative for countries to use the data to better understand their national situation.

Measures of literacy and numeracy need to be compared over time to determine relative needs and to track progress.

#### 6.1.1 How Indicator 4.6 is informed to date

Currently, there are only two internationally administered assessments, OECD’s PIAAC and the World Bank’s STEP, which makes use of a version of PIAAC’s literacy assessment. The UIS Literacy Assessment and Monitoring Programme (LAMP) has a methodological framework and tools that are relevant to low- and middle-income countries, though it is not currently being administered. This assessment will be very valuable, since the tools and methodologies used to assess literacy and numeracy in high-income countries, like PIAAC, are considered inappropriate for lower-income countries.<sup>1</sup> Conventionally, these assessments include: m Administration of an extensive background questionnaire that identifies key population subgroups, documents the determinants of skills differences and allows exploration of the impact that skill differences have on individual outcomes;

- Administration of a direct test of adult literacy and numeracy that covers the full range of skills in the population; and
- Administration of a direct test of the reading skills that support the emergence of fluid and automatic reading that characterise performance at the lower levels.

---

<sup>1</sup> An international report by the UIS (2017i) explores the differences between LAMP and the earlier versions of OECD’s literacy assessments, International Adult Literacy Survey (IALS) and Adult Literacy and Life Skills Survey (ALL), and found that OECD literacy assessments that were conducted in OECD countries and exclusively in European languages do not address the challenges of testing in other contexts. LAMP was implemented between 2007 and 2008 in five countries as a pilot.

**Table 1. Uses for data on literacy**

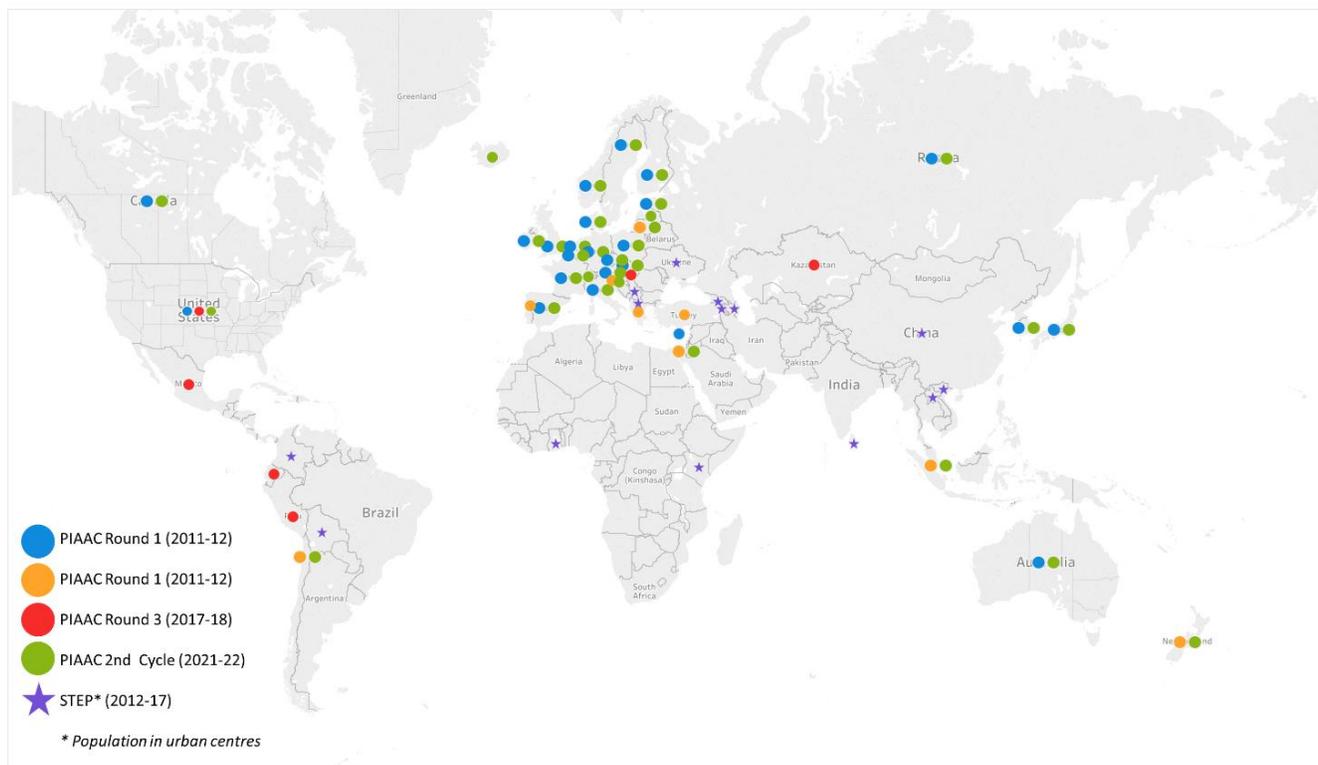
Application type	General purpose	Related policy questions	Implication for data collection strategy
<b>Knowledge generation</b>	Identification of the causal mechanisms that link skills to outcomes	How do individuals acquire skills? How do they lose skills? How are skills linked to outcomes?	Needs longitudinal or repeated cross-sectional data with comparable measures of skills
<b>Policy and programme planning</b>	Planning government response to identified needs to meet social and economic goals	Which groups need to upgrade skills? How many people are in need? Where is need concentrated?	Needs profile of skills for key sub-groups
	Determination of funding levels	How much budget is needed to raise skills at the rate needed to achieve social and economic goals?	Need numbers of adults with different learning needs
<b>Monitoring</b>	Adjustment of policies, programmes and funding levels	Are skill levels rising at the expected rate?	Need repeated cross-sectional skills measures
		Are skills-based inequalities in outcomes shrinking?	Need repeated cross-sectional skill measures for key sub-groups
<b>Evaluation</b>	Formal process to determine if programmes are performing as expected and meeting their objectives	Are government programmes effective and efficient?	Need data on skills gain/loss and costs for programme participants
<b>Administration</b>	Making decisions about specific units: individuals, regions, programmes	What criteria are applied to determine programme eligibility?	Need results that are reliable enough to keep Type I and Type II classification errors to acceptable levels

Source: UNESCO Institute for Statistics (UIS).

There are challenges and constraints associated with each of the two assessments. PIAAC tools may be relevant to the OECD or high-income countries, but they are not relevant and might not be valid for low- and lower-middle-income countries. STEP tools were developed to target low- and middle-income countries. However, STEP focuses on work-relevant skills and does not measure numeracy. Its premise is that numeracy ability is highly correlated to literacy ability. However, using proxies as outcome variables can have a deleterious impact on measurement and behaviour (Gal, 2018). In low- and middle-income countries

especially, it is possible to have respondents who are illiterate but have numeracy skills, so that correlation cannot be taken for granted. These assessment programmes are technically rigorous and respected, with many countries participating.

**Figure 1. Coverage of skills surveys**



Source: UNESCO Institute for Statistics (UIS).

Of the international studies, the RAAMA study by UIL stands apart from the International Adult Literacy Survey (IALS), ALL, PIAAC, STEP and LAMP as the items used to assess literacy skills were not selected in a way to provide systematic coverage of the characteristics that underlie the relative difficulty of tasks, nor were results summarised using methods that confirm the stability, reliability and comparability of measurement. The RAAMA approach to measurement does not provide the needed cross-national comparisons of skills over time. However, the content framework developed for a low-literate population in literacy programmes may contribute to the development of the conceptual framework.

To date, only a handful of countries conduct national adult literacy assessments. Even though many have used the UNESCO definition of literacy as the basis for building their national adult literacy assessment, these assessments vary considerably in terms of content domain definitions and coverage. A number of national assessments were reviewed that measured literacy and numeracy skills indirectly. These assessments rely on self-reports of skills or on performance on very limited numbers of test items.

Research shows that these measures are unreliable, i.e. they are unable to support comparisons within or between countries (Niece and Murray, 1997). The fundamental problem with self-reports is that adult perceptions of their skill levels are conditioned by their use of their skills rather than their actual skill level and at times the social perception of having the skills. To make matters worse, the relationship of self-perceived skills to actual skills varies significantly among subpopulations within countries and across countries and over time. This renders these assessments of limited use to policymakers.

### 6.1.2 What are the challenges to report?

For SDG 4 monitoring and reporting, there is a need for a common definition and a common reference in reporting. In developing a strategy to monitor progress towards Target 4.6, the primary conceptual issue is agreement on the definitions and dimensions of the constructs of (adult) literacy and numeracy to be measured by Indicator 4.6.1. There are several main issues.

The indicator for Target 4.6 implies a need for measures:

- i) of literacy and numeracy; ii) that are statistically-representative of the adult population;
- iii) that capture a range of definitions of functionality across countries;
- iv) that can be compared under some criteria; and v) that provide a set of cost-efficient options for countries.

The indicator specification also includes several subjective elements that require definition, including:

- i) the definition of “functional” relative to literacy or numeracy;
- ii) a menu of options for countries to measure and report; and iii) a linking strategy to compare different options.

#### *Definition of literacy and numeracy*

The definition of literacy from the UN’s Principles and Recommendations for Population and Housing Censuses, Revision 3 states:

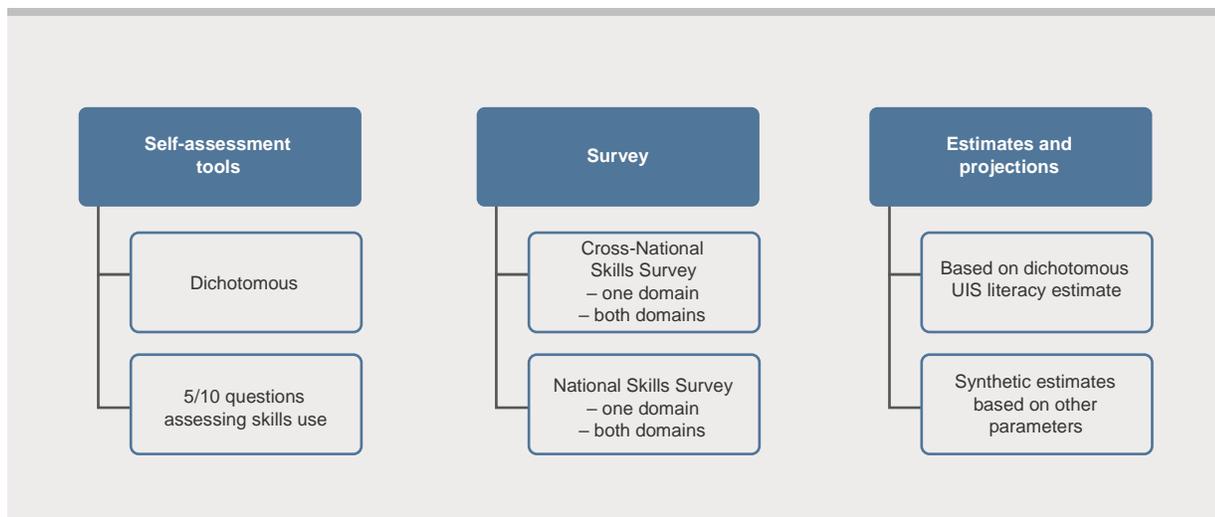
*Literacy has historically been defined as the ability both to read and to write, distinguishing between “literate” and “illiterate” people. A literate person is one who can both read and write, with understanding, a short, simple statement on his or her everyday life. An illiterate person is one who cannot, with understanding, both read and write such a statement. Hence, a person capable of reading and writing only figures and his or her own name should be considered illiterate, as should a person who can read but not write as well as one who can read and write only a ritual phrase that has been memorized. However, a more modern understanding referring to literacy as a continuum of skills, levels, domains of application and functionality is now widely accepted. (UN, 2015).*

In the current generations of comparative assessments, functionality is defined as the level of literacy needed for an individual to cope with the demands that they confront in their daily lives and will differ by country and situation. For this reason, assessment has focused on the use of skills. Therefore, each country must establish its own definition of what level constitutes the functional level(s) that reflects its definition of literacy skill-based inequality in individual outcomes, its targets for the performance of key social institutions, including firms and educational institutions, and its social and macroeconomic demands.

### *No equivalent definition of numeracy exists*

In terms of the conceptualisation of literacy and numeracy as a continuum, the situation in the field of adult assessments differs considerably from that of assessments of school-age children. The framework of the PIAAC assessment draws on a theoretical tradition that has underpinned the conceptualisation of literacy and, subsequently, numeracy in IALS, ALL and LAMP. Therefore, these assessments have a common conceptual framework.

**Figure 2. Summary of reporting options**



Source: UNESCO Institute for Statistics (UIS).

The UNESCO definition of literacy was adopted by GAML with the conceptualisation of literacy and numeracy utilised by PIAAC with adaptations to extend the framework to include foundational skills. This was noted, even though the PIAAC conceptual framework is relatively comprehensive.

#### 6.1.3 Exploring reporting options for Target 4.6.1

Given that the existing assessment tools (PIAAC and STEP) and data collection might be lengthy and costly, there are some alternatives that could be considered to report for Indicator 4.6.1 that include three broad categories: observed data based on self-assessment, administration of skills surveys and synthetic estimates (see **Figure 2**).

#### *Indirect and simplified measures*

The most simplified version is the current dichotomous measure for literacy; it faces the most relevant challenge. For many years, the international definition of a literate person was someone “who can, with understanding, both read and write a short simple statement on his or her everyday life”. This definition has long underpinned the UIS’ regular Survey on Literacy, which produces estimates of the literacy rates in most

developing countries. These estimates, in practice, only distinguish between those who cannot read or write at all and the rest of the population. However, those judged literate relative to this definition can have vastly different levels of skills. Someone who can at best read and understand a simple statement about everyday life is arguably not sufficiently well equipped to cope with the demands of modern-day living. Policy interventions are not only needed for those who are illiterate but also for those with weak literacy skills.

In order to address the needs of people with low literacy skills it is necessary to adopt a more nuanced definition of literacy, which identifies a range of literacy skills and levels of competence. Being able to identify the characteristics not just of the illiterate population but also of those with weak skills will make it possible to better target resources to address their respective needs and increase literacy skills in general.

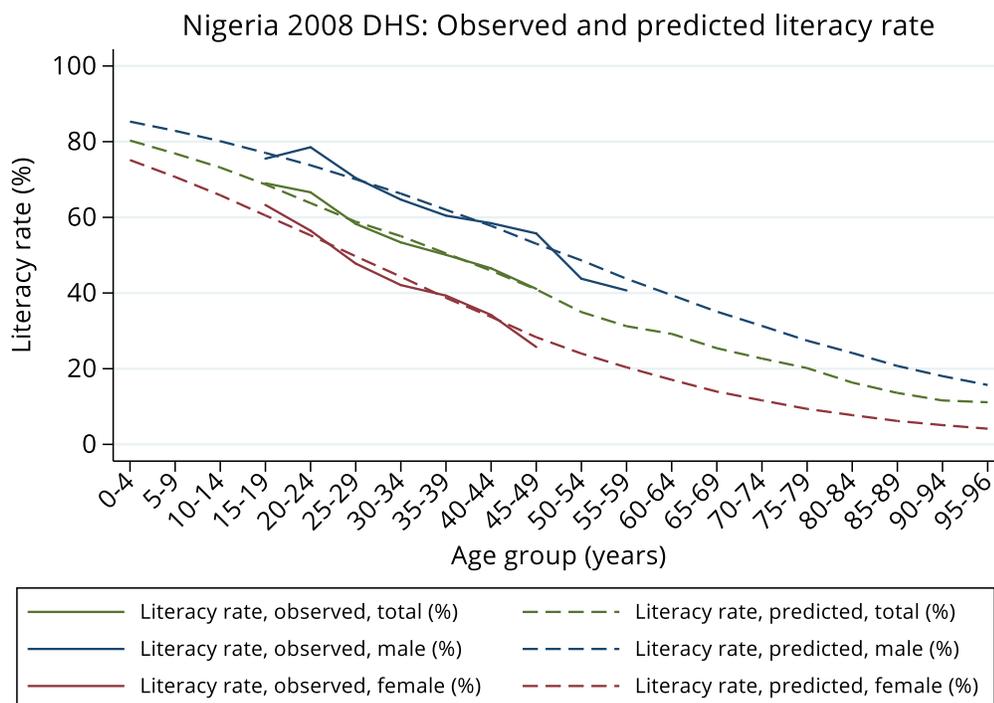
Self-assessment could be a simplified version of the type of DHS and MICS surveys that try to address the dearth of literacy assessments in developing countries by adding a simple test of reading skills to their survey modules. In DHS and MICS surveys, a sample of adult respondents, typically women and men between 15 and 49 years, are asked to read a card with a short, simple sentence in their language. The result is recorded as one of three options: i) cannot read at all; ii) able to read only parts of the sentence; or iii) able to read the whole sentence. The results of these tests are available for nearly all DHS and MICS surveys carried out in the last decade, including a large number of surveys in less-developed countries. The test results are more reliable than self-reported data on literacy and give at least some sense of the level of reading skills. On the other hand, these simple reading tests do not allow the measurement of literacy on a continuum, unlike the assessments mentioned earlier and are therefore only a partial improvement on traditional dichotomous literacy indicators.

### *Skills surveys*

Observed data could be based on either self-assessment or the administration of a skills survey that could have various alternatives and face alternative methodological decisions of the type described below:

- The number of skills domains;
- Testing the whole range of skills or limited to certain parts of the skills distribution;
- Whether the assessment will be conducted as an independent study or added to an existing study;
- Whether the assessment design will provide direct point estimates of skill distributions or support the generation of indirect, synthetic estimates; and
- Whether the assessment tool will be paper-and-pencil or computer-based.

**Figure 3. UIS literacy survey estimates**



Source: Nigeria 2008 DHS

### Synthetic estimates

An alternative is to do synthetic estimates based on observed available data (see **Box 1**). The estimates could find various alternatives but consist, in a simplified version, of combining information regarding

#### Box 1. Synthetic estimates to report for Indicator 4.6.1

A very simplified scheme would need:

- Data for either proportion of totals of persons at each education level by gender and age bands;
- Skills survey database to be prepped for modelling with as many countries as possible;
- The definition of a number of country-level factors according to available data that could be merged with the skills database described in the paragraph above; and
- Parameters to be estimated from the skills database to predict the values of the indicators.

Source: UNESCO Institute for Statistics (UIS).

the distribution of skills from countries that had administered skills surveys according to a defined set of characteristics of the population. This distribution of skills for those categories could be used to predict the levels (and, the indicator, once the minimum fixed levels of numeracy and literacy are defined) to establish the estimate for all countries.

These data would provide national and international users with a regular and current source of evidence, a prerequisite to maintaining policy focus and to adjusting policies and programmes. A large-scale rebasing of the model would be undertaken in 2031 when the next PIAAC collection cycle is undertaken.

Although some country-context parameters can be used, preliminary estimates show that some observable factors such as age, education and participation in some skills activities capture most of the variance. With this bridge, a country that has not administered a skills surveys but still has information about these parameters could have an estimate of the indicator. The degree of precision would vary according to the breadth of the individual information that could serve in the modelling phase.

An example of this type of modelling is the UIS literacy rate. Literacy rates for persons outside the age range with observed literacy rates are estimated using a logistic regression of literacy on age. As an example, see the 2008 DHS data for Nigeria in **Figure 3**. The survey collected information on literacy for women aged 15 to 49 years and men aged 15 to 59 years. The observed literacy rates are indicated by the solid lines and the results of the logistic regression are indicated by the dashed lines.

#### 6.1.4 Reporting on the same scale

Once the definitions for the conceptual framework and levels of proficiency are sufficiently clear to allow options for countries to locate themselves in a continuum, the next step will be to develop an appropriate methodology for creating an internationally-comparable database to report on Indicator 4.6.1, given the use of different tools across countries. This means defining some criteria for linking that could be a combination of strategies or following a stepping-stone approach (as for Indicator 4.1.1).

The indicator requires the following inputs:

- Agreement on a proficiency framework that allows alternative levels of skills or functionality;
- Definition of the reference minimum global level;
- Definition of harmonisation (and linking) strategy that allows location of all efforts into a comparable metric; and
- A modelling strategy to produce an annual comprehensive set of literacy and numeracy estimates.

In order to produce estimates for reporting, there is a need for elaborating a guide and standards for countries that want to measure literacy, such as literacy modules to be added to household surveys and guidelines on the steps and standards for outputs. This type of global public good would facilitate not only a country's measurement but also commonalities in measurement. For this reason it is relevant to reach:

An agreement of the questions in the self-assessment module;

- A definition of the individual background question
- framing that would later serve as common parameters for synthetic estimates; and
- A short literacy numeracy module such as mini-LAMP (see **Box 2**) for those who want to do a shorter skills survey.

### 6.1.5 Laying out a strategy for measuring and reporting

Regarding the definition of literacy, GAML has recommended that the literacy and numeracy indicators be based on the framework of literacy and numeracy used in the OECD's PIAAC adult skills assessment programme. These definitions are precise enough to be measured and broad enough, with added elaboration at the foundational skills, to capture the entire range of skills encountered globally. Although the PIAAC assessment was only administered to 16- to 65-year-olds, the indicator covers 15-year-olds so information from PISA could also be used to report.

#### Box 2. Enhanced and shortened version of LAMP or mini-LAMP

UNESCO's LAMP assessment was developed to better respond to the needs of less-developed countries, while maintaining established proficiency scales. LAMP can be seen as a methodological endeavour to provide sound information, especially concerning the least-skilled in a population. It also shows the complexities of a diverse group of countries facing very different challenges in implementation.

Through LAMP, the UIS has gained a unique perspective on the diversity of human literacy experience. Finally, it has also shown that the methodology, with the necessary adaptations, can be used across different cultures, languages and scripts.

Past experiences suggest the need for alternatives to a full LAMP assessment that would reduce the operational, technical and financial burden of fielding the assessment without compromising the ability to compare results across countries and over time. In this context, the UIS is taking a two-step approach to produce; i) a paper-and-pencil version; and ii) a device- or computer-based version of an enhanced and shortened version of LAMP, referred to as mini-LAMP.

Currently, the paper-and-pencil version of mini-LAMP has been produced and includes:

- Short literacy-relevant background questionnaire;
- Short cognitive modules;
- Administration guide;
- Translation and adaptation guide;
- Sampling guide;
- Scoring guide;
- Data capture and process guide; and
- Software and a data analytical guide.

To help countries with planning, the UIS will produce a national planning report template and memorandum of understanding to initiate discussions with interested countries

The device- or computer-based version of mini-LAMP is still under development.

*Source: UNESCO Institute for Statistics (UIS).*

We propose a strategy for monitoring progress that offers countries a range of options according to their needs and possibilities. Countries on their way to achieving universal secondary education are encouraged to participate in the next round of the PIAAC data collection scheduled for 2021. The PIAAC design and processes are based upon 35 years of development and yield results that are valid, reliable, comparable and interpretable.

**Table 2. Cost of alternative options**

Option	Estimated costs (US\$)	Universe	Needs from countries
PIAAC	2.5 million to 4 million <sup>2</sup> (paper-and-pencil and web-based)	Country with experience in largescale assessment and household survey m Strong technical capacity	Countries near achieving universal secondary education and have strong technical capacity
STEP	500,000 <sup>3</sup> (paper-and-pencil)	Country with experience in largescale assessment and household survey m Good technical capacity	Countries interested in literacy skills in working age population and have technical capacity
Short Literacy Survey <sup>4</sup> (SLS)	200,000–400,000 (web-based)	Country with experience in largescale assessment and household survey	Developed countries that want more skills information beyond self-reporting and self-assessment but do not need a full range of skills estimates
Mini-LAMP <sup>5</sup>	250,000–600,000 <sup>6</sup> (paper-and-pencil) 160,000–300,000 (web-based)	Country with experience in largescale assessment and household surveys	Developing countries that want more skills information beyond self-reporting and self-assessment but do not need a full range of skills estimates
Literacy module (SLS or miniLAMP) attached to DHS/MICS/LFS	150,000–200,000 <sup>7</sup>	Country with experience in largescale assessment and household surveys	Countries that do not want to conduct a separate household survey for adult literacy but regularly conduct household surveys and want a snap-shot of targeted skills distribution
Synthetic estimation	Free-based on UIS methodology paper and set of guidelines on how to produce estimates	Country with technical capacity	Countries that do not want to conduct another assessment but want to project skills using census data and existing assessment data to generate estimators to project future skills by sex and age group

Source: UNESCO Institute for Statistics (UIS).

**Figure 4. Country options from simple to complex**



Source: UNESCO Institute for Statistics

<sup>2</sup> Based on communication with Scott Murray, the cost depends on sample size and cost of implementation within a given country.

<sup>3</sup> Based on communication with the World Bank Group.

<sup>4</sup> Since this is still under development, the estimate is based purely on speculation.

<sup>5</sup> Estimate based on 1,500 cases with varied implementation costs from the UIS 4.6.1 option paper.

<sup>6</sup> The cost is for a paper-and-pencil version and will be substantially smaller if it is attached to existing household surveys.

<sup>7</sup> Estimate based on attaching the literacy module to a sub-sample of an existing household survey and cost of in-country training. No separate sampling cost as main sampling cost is borne by the surveyor.

For countries below this level of educational development, the current PIAAC design offers a limited information return on their investment. Moreover, the technical, operational and financial burdens imposed by PIAAC may be too great for some countries to bear, something that translates into a considerable risk of failure.

Each option comes with advantages and disadvantages, depending on what is valued more: skills coverage, reliability in generated estimates, accuracy in skills estimates and/or consistency in implementation. A combination of selected options could be chosen by countries and used to report on Indicator 4.6.1.

Hence, it is more realistic to consider a strategy that allows:

- Menu of options for countries that includes simpler to more complex alternatives to measure and report and allows countries to find their own model; and
- Use of estimates and projections to serve as a preliminary global picture of adult skills distribution.

In summary, each country has a choice on what works best for them. There are several options depending on socioeconomic development, as well as the technical and financial capacity of the country.

- A developed country that wants full skills distribution of its population could consider PIAAC, which is technically-complex and expensive to implement.
- A developing country interested in understanding the literacy skills distribution of its productive population could consider STEP as it has comprehensive work-related background questions that provide precise skills distribution of the productive population.
- A country interested in only a targeted skills segment could consider SLS or mini-LAMP. Both of these short survey assessments consist of easier items which will provide better skills estimates for the country with a substantially-low literacy population.
- A country that wants a snapshot of its population's skills distribution could consider attaching a literacy module to an existing household survey. This will also reduce operating costs as the sampling cost has already been covered by the household survey.
- A country that has conducted literacy assessment in the past and does not want to conduct another round of adult literacy assessments could consider synthetic estimation. The UIS has developed a methodology paper on the way to produce a synthetic estimation based on basic characteristic variables such as sex, age group, years of schooling, etc. from a census. The relevant assessment data produced from past assessments, relevant characteristic variables and literacy-related questions can generate estimators to project skills distribution.

All options have their own advantages and shortcomings. Each country will need to identify what it considers most important to make the right choice.

## References

Gal, I. (2018). "Developing a monitoring scheme for adult numeracy as part of SDG indicator 4.6.1: Issues and options for discussion". Discussion Paper for the UNESCO Expert Meeting on Adult Literacy and Numeracy Assessment Frameworks, 17-18 May 2018, Paris.

Niece, D. and T.S. Murray (1997). "Literacy Skills and the Readiness of Adults for Lifelong Learning and Further Education and Training: Trends based on the International Adult Literacy Survey".