Concept Note

Measuring Target 4.6: Technical issues and next steps

15th September 2016

Introduction

The international education community has pledged to "ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes", as part of Target 4.6 of the new Sustainable Development Goal for Education (SDG 4).
The focus on quality of education has led to an emphasis on the measurement of learning outcomes at all levels of education. However, measuring learning and skills is not as straight-forward.

There are many similarities in how adults develop skills across cultures, but bigger environment affects the manifestation of skills. Overall, even though cultural and contextual influences matter, adults everywhere show similarities in how they learn to read, write and count. There are strong points of similarity that can form the basis for global measurement, with national standards.
As a UN statistical agency focusing on comparable data and evidence-based decision making, the UIS is providing evidence for developing targeted approaches towards better data collection, and working with the global community in defining robust indicators to report progress in skills development towards the 2030 goals.

The priority is to generate data that are comparable across time and disaggregated by age, sex, disability, socioeconomic status, geographical location (urban/rural areas) and other relevant factors.[1] However, there are technical challenges in generating measurements that are comparable across countries with diverse cultural and economic status.

Globally-comparable measurement of adult skills would be able to track adult learning over time, so that the skills development can be indexed. This measurement should include contextualized information: educational attainment, family background, home and work environment, and opportunity to learn. The global measure should include a range of relevant skills and place emphasis on the importance of basic literacy and numeracy skills.

[1] See: http://www.un.org/ga/search/view_doc.asp?symbol=A/C.3/69/L.9/Rev.1

How have skills been measured to-date?

For many years countries have been sampling and testing adults on skills at the international level. Large-scale, international adult assessment programmes, such as the International Adult Literacy Survey (IALS), the Adult Literacy and Lifeskills (ALL) survey, the Programme for International Assessment of Adult Competency (PIAAC), and the Skills Toward Employment and Productivity (STEP) study, allow countries to compare the skills of their adult population and gain insights into what needs to be improved to have a skilled and productive population. These assessment programmes are technically rigorous and respected. However, the assessments assess higher skill levels and do not measure population in the lower skills continuum well. Some of the assessment tasks are highly demanding and focus on problem solving in a computer environment, which might not reflect the reality of large segments of many least developed countries' low literate populations. This therefore poses a problem for countries with a substantial proportion of the population at a basic literacy level. If these countries are administered the current international assessments, the assessments would not be able to identify whether the low-skilled population possess any literacy skills since they are unable to response to any of the tasks, and hence would provide relatively little guidance as to what skills to improve, and how.

A revised adult literacy assessment is therefore necessary to give the international community a complete picture of adult literacy skills around the world.

The UIS worked with 11 developing countries[2] from 2005 to 2012 through the Literacy Assessment and Monitoring Programme (LAMP) and learnt that conducting household assessments is not as inexpensive or as easy to implement in developing countries as originally thought. The UIS has gained tremendous experience and knowledge in working with these countries, adjusting the reality of operationalizing such a complex undertaking, especially in countries where assessment culture and implementation infrastructure are underdeveloped and that have no experience in collecting assessment data in households. Furthermore, the experience showed that assessing in languages with orthographies or scripts that are not widely available or studied, and/or which are not very familiar to the population that speaks those languages, presents special problems in terms of the cost-effectiveness and relative utility of the assessments.

Some countries also conduct their own adult literacy survey, such as the Kenya National Adult Literacy Survey (KNALS). However, due to differences in construct and framework not all national adult literacy assessments are comparable and it is hard to have a global picture of the skills of adults around the world with non-comparable results.

There are technical challenges in measuring adult skills which include:

[2] The 11 countries are Afghanistan, El Salvador, Jordan, Kenya, Lao PDR, Mongolia, Morocco, Niger, Palestine, Paraguay, and Viet Nam.

- Establishing the applicability of cognitive items to measure skills across all contexts;

- Establishing the relevance of contextual questions to understand the adult skills development;

- Adapting existing or develop new tools that can be administered to countries with different skills distribution at a reasonable cost while still adequately capturing adult skills; and

- Ensuring adequate alignment between global tools and country specific policy contexts, such as the connection between a national literacy framework and items on literacy assessments. As well, data on skills can be more easily interpreted when accompanied by contextual information on personal and individual's environment, which requires additional data collection.

Beyond these issues, the idea of creating a skills scale for all countries necessitates a new level of technical development. Because development is strongly affected by context, a large range of items will be required for accurate measurement of adults in all settings – adult skills may differ greatly, meaning that floor and ceiling effects will be problematic when scales are used across contexts. By creating skills scales with narrow ranges, data will be more likely to demonstrate that almost all adults in some countries are doing well or poorly, which will undercut the overall goal of promoting equity within and between countries. New tools or approaches with wider ranges may be required to capture variation in skills development across contexts.

Challenges in creating a 4.6 measurement strategy

The central challenge of global measurement of Target 4.6 is deciding what is "achieving at least a fixed level of proficiency in functional (a) literacy and (b) numeracy skills" in different contexts, and generating globally comparable tools to describe it. Broadly speaking, "achieving a fixed level of proficiency" is perhaps best understood as the extent to which adults are able to master skills and competencies as defined by their local communities. At present, there is no "absolute" definition of "achieving a fixed level of proficiency" that applies across contexts.

Questions on what it means to be "achieving a fixed level of proficiency" include the following:

- The differences in skill levels may be vastly different from one country to the next. Adults in high-income countries may develop a different set of skills than those in low-income countries. Because the SDGs are intended to apply to all countries, how can such scale be created?

- The overall intent behind Target 4.6 is to index the percentage of adults "globally" who have achieved a fixed level of proficiency. How critical is it to establish predictive validity of influencing factors and skills acquisition, and collect the relevant contextual information?

- How should global measurement appropriately account for cultural influences?

Finally, the political context of measurement means that countries will choose different paths forward for measurement. It may be more acceptable to begin generating technical approaches that allow countries to participate in global monitoring using a tool that is adaptable to the different development levels of all countries. Some countries may choose to participate in global assessments, while others may develop their own national assessments and embed the relevant tool in their assessments to bring their assessment onto the global scale. This means countries can choose to do their own assessment and use one module developed for the global scale. Therefore for global comparison only the results from the global tool are used. If the overarching goal of SDG measurement is to encourage collection of data on adult basic literacy and numeracy skills that can lead to better policies and practices, there is a high priority on generating politically feasible solutions to global monitoring. Therefore, it may be appropriate to generate methods of integrating data across a range of sources (national and regional; household survey and household-based assessment survey) that meet international standards.

Next steps for target 4.6

In sum, the following key points on adult skills must be taken into account when devising strategies for global skills measurement:

- Adult acquisition of skills proceeds in patterns that are in many ways similar across all populations even though the skill levels may be vastly difference across countries. There may be efficiencies in measurement that are gained through reliance on a common set of items or constructs, based on the degree of similarity in existing cross-national measures.

- Aim for construct equivalence, which could provide some degree of population-based tracking using a similar set of constructs and develop a common skills scale.

- Data on adult skills should be accompanied by contextual information, personal, literacy environment and opportunity to learn, to provide a more complete picture of skills development.

Reliance on one tool is most efficient for global monitoring, but will also require additional development to ensure that it is useful across all countries, and in particular, engaging stakeholders and organizations – especially the regional assessment bodies – in the dialogue is necessary. Integrating data from multiple sources, household-based assessments and government administrative data, and develop methodology of alignment, may provide a greater degree of country relevance.

The UIS Learning Outcomes monitoring strategy [3] has taken the following approaches by suggesting:

- a global common content framework for reference (GCCFR) for each of the learning domains, for Target 4.6 - Literacy and Numeracy

- a data quality assurance framework (DQAF) applied to learning outcome data and its implementation tool, the assessment of data process (ADP)

- a reporting package consisting of:
    - a Global Reporting Metric (GRM) and related benchmark to put data onto the same platform for reporting
    - Guidelines for data analysis and data use

The International Code of Practice for Learning Assessment (ICP-LA) and the Data Quality Assurance Framework (DQAF) could serve as a starting point for national assessments for countries to put their data onto the global metric. More work is needed to complete technical development. The work of the UIS, Global Alliance to Monitor Learning (GAML), also highlights the need to build a community around measurement and provide opportunities for sharing ideas, tools, and ultimately data that will lead to more comprehensive and innovative approaches to measurement. This will require strong commitment of stakeholders and data users on how learning outcomes data apply to development policy and literacy practices to improve adult learning conditions or create literacy environment for all population.

Finally, the idea of creating a global learning scale and reporting metric is ambitious. To create scales that would work across populations, we need to track the development of skills in a range of countries, and used that point of comparison to generate a global learning curve.

It is clear that the dialogue should continue among stakeholders and experts on approaches to measurement for 4.6. Next steps could include the following:

1) Create technical and institutional homes for global technical discussion of adult learning and assessment, assisting countries in implementation and use of data to improve policy and practice.

2) Convene stakeholders and experts to define “level of proficiency”; and evaluate the alignment of existing tools with the desired definition and scope of skills.

[3] Implementation strategy for measuring SDG 4 learning outcomes targets concept note in the following link: http://www.uis.unesco.org/Education/Documents/UIS%20LO%20Monitoring%20Strategy%20concept%20note.pdf

3) Define a common set of technical standards for tools used for population-based measurement to inform global monitoring.

4) Explore the process to create learning scales - Literacy and Numeracy.

5) Develop technical approaches to integrate contextual information.

6) Develop approach, methodology and psychometric methods for integrating national and regional data to generate global estimates.

Overall, a challenge for assessing the “level of proficiency” is the difficulty to agree on a standard global definition of “proficiency”. The dialogue on measurement thus has implications for the implementation of strategies to reach Target 4.6, and therefore is an important part of the overall SDG agenda for adult skills.