Will a global learning metric improve equity in education?

Aaron Benavot

Dept. of Educational Policy and Leadership
School of Education
University at Albany-SUNY

International Education Development Forum
Nagoya University, November 3, 2017

Outline of Presentation

1. Historical perspective 1990-2015: global policy interest in learning outcomes grows but is contested
2. Conditions for a global learning metric (GLM) proliferation of, and country participation in, learning assessments
3. The centrality of a learning dimension in SDG4
4. Technical issues in the construction of global learning metrics
5. Can global measures of learning contribute to equity in education and better quality education?

While Universal Primary Education (UPE) has been the main focus of int'l education policy since 1950s, an emphasis on learning emerged at the EFA conference in Jomtien

Jomtien Thailand: Inputs to March 1990 conference, **World Conference on EFA**, especially by the World Bank (see Lockheed and Verspoor 1989) placed considerable emphasis on teaching and learning and not just universal access to basic education. Final document captures key debates around learning issues:

*The focus of basic education must be on **actual acquisition and outcomes**, rather than exclusively upon enrolment, continued participation in organized programmes and completion of certification requirements.*

***Active and participatory approaches** are particularly valuable in assuring **learning acquisition** and allowing learners to reach their fullest potential. It is, therefore, necessary to define acceptable **levels of learning acquisition** for educational programmes and to improve and apply systems of **assessing learning achievement**. (Article IV)*

Quality Education and Learning the World EFA Conference (1)

Final document (1990) captures key debates around learning, including learning for what purposes:

*"Emphasis ought to be on the access to and achievement of learning that is **relevant** to the needs of the individual and his or her community."*

*"One key issue…is the **minimum common level of learning** that must be achieved by . . . learners. A difficult aspect...will be **testing the learning level** or performance of individual learners . . . The ultimate goal...is that children are educated: that they **are literate and numerate, and that they have life skills**.*

Quality Education and Learning the World EFA Conference (2)

"While the emphasis on learning acquisition was welcomed, some interventions ***cautioned against too utilitarian an approach to defining 'an acceptable level of learning'"***
…. "The point about learning is that ***it is a process of growth, and not a product to be acquired: learning is a journey, not a destination.***"

The development of the creative potential of the individual, of imagination, of spiritual and aesthetic values, of community spirit, are justifiable in their own right, and not easily measurable in the short term."

World Education Forum in Dakar Senegal (May 2000): Establishes six EFA goals. One goal devoted entirely to quality education (EFA Goal 6):"*Improving all aspects of the quality of education and ensuring excellence of all so that* ***recognized and measurable learning outcomes are achieved by all****, especially in literacy, numeracy and essential life skills."*

Dakar Framework for Action: Education for All (May 2000)
"Quality is at the heart of education, and what takes place in classrooms and ot her learning environments is fundamentally important to the future well-being of children, young people and adults. A quality education is one that ***satisfies basic learning needs****, and* ***enriches the lives of learners*** *and their overall experience of living."* (para 42)

In the **EFA Global Monitoring Report** (GMR), first published in 2002, learning is viewed as one part of **a multi-dimensional model of quality education**.

MDG Summit (Sept 2000): precursor of SDGs: **quality and learning are not addressed**; emphasis is on universal completion of primary education and gender parity.

Since Dakar (2001), international interest in, and discussions of, education quality are on the rise…

- Several high-level (ministerial) meetings discussed quality issues (Int'l. Conference on Education, 2004; PRELAC, Buenos Aires, 2007).
- Independent Evaluation Group of the World Bank (2006) recommends that countries and development partners emphasize learning outcomes as well as school access.
- The EFA Fast Track Initiative plans (later to become GPE) incorporate the monitoring of learning outcomes as an additional criteria in the endorsement of EFA-FTI country funding.
- UNESCO initiatives: 'But Can They Read', 'From Access to Success'; 'Learning Counts'; 'International Working Group on Assessing and Improving Quality Learning' and 'Teacher Education in Sub-Saharan Africa.'
- Global Action Week (2006): NGOs mobilize around quality issues such as teacher supply and pre- and in-service teacher training.
- Post-Dakar surge of interest in quality education and learning among donor agencies (USAID, DFID, GIZ, etc.)

In early 21st century learning and its measurement become focus of international education policy

- In 2000s World Bank emphasizes **measurement of learning** in all its projects. In 2011 WB publishes a new strategy document for 2020: ***Learning for All***
- Impact of IEA's TIMSS and OECD's PISA in many countries around the world. Mass media attention to the **global competition for skilled labor** and **emerging knowledge societies and role of learning**.
- Hanushek and Woesmann report on the **impact of quality education** (measured by average test scores on international learning assessments) **on economic growth**. Recent OECD book on the need for **universal basic skills.**
- **Learning Metrics Task Force** advances **broad model of lear**ning domains. Millions Learning Project to scale up learning interventions
- **Massive expansion of** international, regional and national **learning assessments since 2000** (see next section)

Tentative conclusion

- Universal access AND learning have been key issues in international education policy discussions and debates since Jomtien. Some emergent consensus.
- Debates focus on how much to emphasize learning outcomes; which types of learning to emphasize; and which learning outcomes can and should be measured.
- These debates were framed by exclusive focus on countries in the Global South in EFA goals, though drawing on evidence from the Global North

2. The conditions for a global metric in learning: The proliferation of and country participation in learning assessments

Types of Learning Assessments

- **International**: IEA-sponsored assessments ie, TIMSS, PIRLS, CIVED (others in the past) and OECD-sponsored assessments ie, PISA
- **Regional**::
  1. SACMEQ: Southern African Consortium for Monitoring Education Quality (sponsored by UNESCO)
  2. PASEC: Programme d'analyse des systèmes éducatifs de la CONFEMEN conducted in 10 SSA and SA countries in 2014
  3. LLECE: Latin American Laboratory for the Assessment of the Quality of Education—SERCE & TERCE (UNESCO sponsored)
  4. PILNA (Pacific Islands Literacy and Numeracy Assessment): administered in 14 Pacific Island countries in 2012
- **National**: in over 150 countries since 2000. e.g., in USA, *National Assessment of Educational Progress* (NAEP); in Japan, *National Assessment of Academic Ability* carried out by National Institute for Educational Policy and Research, Min of Ed, in Japanese and Mathematics (2007-9, 2010, 2012-13) & Sciences (2012)
- **Citizen-led assessments**: Pratham-ASER India (2005-), Pakistan (2008-), Kenya, Uganda and the UR Tanzania (2009-), Mali (2011-), Senegal (2012-), and soon Nigeria and Mexico. Together, Citizen led assessments have reached more than a million children in 2012
- **Early learning assessments:** National Early Grade Reading Assessments (EGRA) or Early Grade Math Assessments (EGMA) conducted in grades 2 or 3 in 13 mostly developing countries. EGRA and EGMA have been carried out in other countries; most are not based on nationally representative samples.

Assessments are one of several approaches to monitoring student learning

- Monitoring actual curricular implementation by subject or topic, through textbook adoption policies, inspectorate, teacher observation, or through indicators of the 'opportunity to learn'
- School-based (formative) assessments that provide detailed (qualitative) information about student progress for teachers, principles, parents; also use of project based assessments
- External 'high stakes' summative examinations at major transition points in the system (known as 'exit' or 'school leaving' exams)
- National learning assessments of knowledge and skills acquired by students from the intended curriculum
- Citizen-led assessments, sampling households, to obtain info on what school and out of school children know, and to hold governments to account for poor learning outcomes
- International and/or regional assessments of student achievement that compare learning levels or cross-cutting skills across education systems

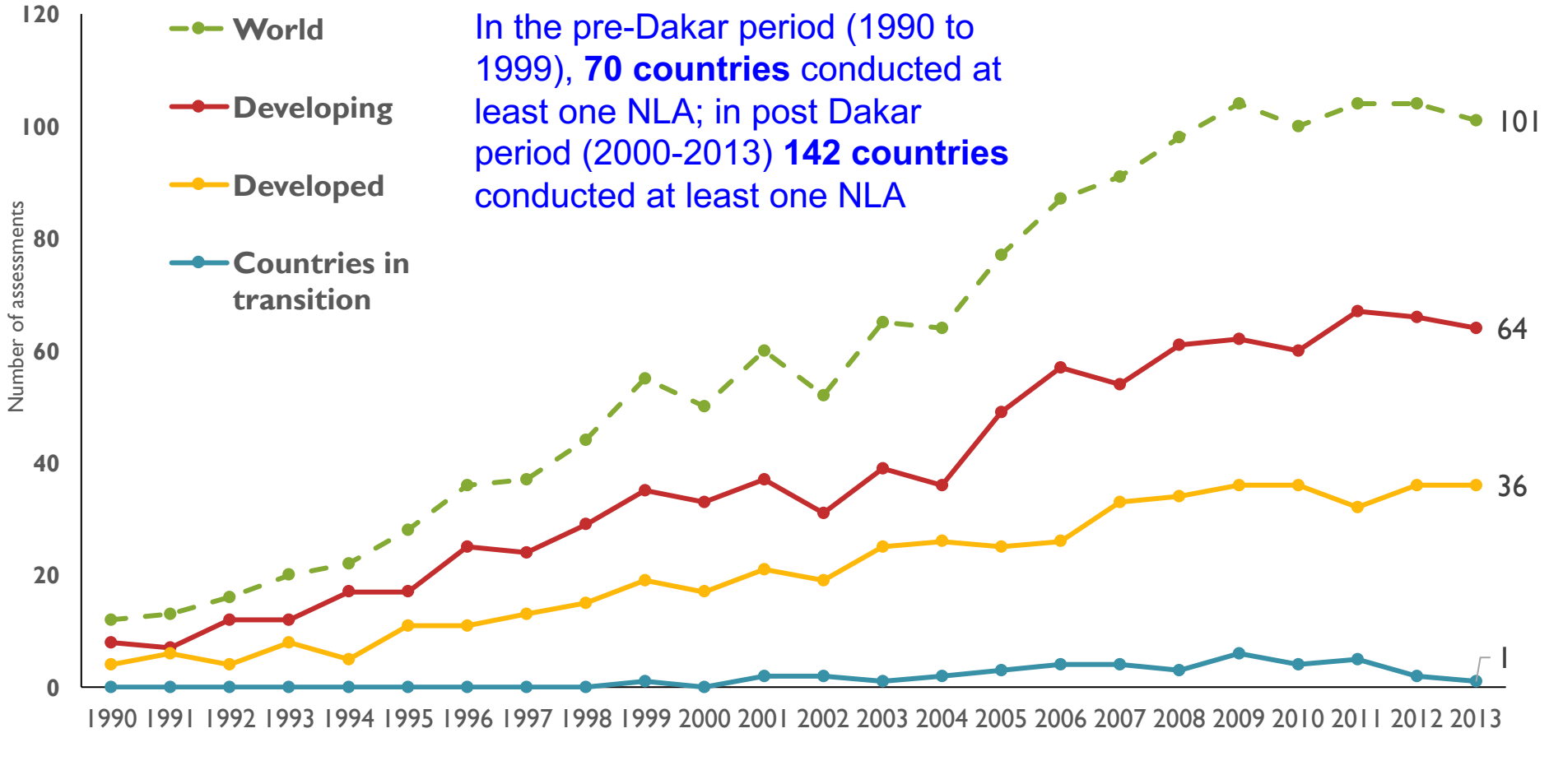
What 'learning' is being assessed in international learning assessments?

- TIMSS, PIRLS and PISA are all **summative assessments**, mainly for stakeholders **external** to the school—e.g., policy analysts, ministry officials and inspectors, curriculum developers, and educational researchers
- TIMSS, PIRLS and PISA place greater emphasis on **cognitive knowledge, skills and competences**, and on the lower levels or stages of learning, mainly in three specific curricular areas: **language, mathematics and sciences**
- Increasing interest in soft skills: problem solving ability; critical thinking; ICT skills
- TIMSS, PIRLS and PISA are all **low-stakes assessments**, and have little direct impact on a child's progress in the educational system or in transition to work.

What is a National Learning Assessment (NLA?

- NLAs typically seek to measure learning on a system-wide basis, for a specified age-based group of students (e.g. 4th graders)
- They provide stakeholders with systematic information about the quantity and quality of learning that has occurred. Usually designed to compare learning levels against a given curricular standard or achievement/ competence level.
- Are predominantly curriculum-based and subject-specific. Rarely examine cross-curricular competencies (like PISA)
- Typically evaluate cognitive outcomes: how much students have learned (knowledge and skills) from what they should have been taught in a specific subject area?
- Mainly created, administered and analysed by national ministries of education or by national evaluation/research institutes. Funding can come from multiple sources.

Big increase in the number of **National Learning Assessments** (NLAs), 1990-2013

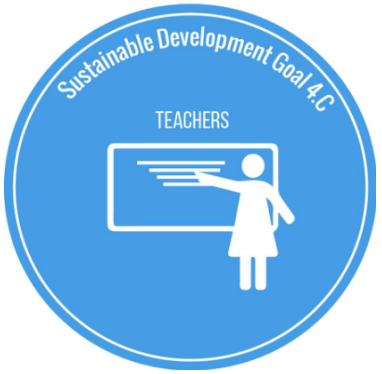


Sources: EFA Global Monitoring Report 2015; Benavot and Köseleci 2015

3. The centrality of a learning dimension in the SDGs, especially SDG4

17 Sustainable Development Goals and 169 Targets adopted by UN Member States (September 25, 2015)

10 SDG 4 targets



What is distinctive about SDG4 and its 10 targets, relative to earlier EFA and MDG policy agendas?

- The SDG4 agenda is **universal**, applicable to all countries not just those in the Global South
- More **ambitious**: seeks to expand and transform education systems (eg, universal primary and secondary education) in ways that are without historical precedent, especially in the time frame allotted
- More **comprehensive** set of targets: includes formal and non-formal education; Early childhood development, ECCE, 12 years of basic schooling, TVET, higher education, lifelong learning opportunities
- More **output** oriented: relevant and effective learning outcomes at primary and secondary education; literacy and numeracy among adults, non-cognitive skills for employment & decent work
- Focused on the **contents** of schooling and not just universal access to, or completion of, primary education

Strong learning dimension in many SDG4 targets

4.1 By 2030, ensure that all girls and boys **complete free, equitable and quality primary and secondary education** leading to **relevant and effective learning outcomes**.

4.2 By 2030, ensure that all girls and boys have **access to quality early childhood development**, care and pre-primary education so that they **are ready for primary education**.

4.3 By 2030, ensure equal access for all women and men to **affordable quality technical, vocational and tertiary education**, including university. [**participation in adult education**]

4.4 By 2030, substantially increase the number of youth and adults who have **relevant skills**, including technical and vocational skills, **for employment, decent jobs and entrepreneurship**

4.5 By 2030, **eliminate gender disparities** in education and ensure equal **access to all levels of education** and vocational training for the vulnerable, including persons with **disabilities, indigenous peoples, and children in vulnerable situations**.

4.6 By 2030, ensure that **all youth and a substantial proportion of adults**, both men and women, **achieve literacy and numeracy**

Learning dimension in SDG4 targets

4.7 By 2030, ensure **all learners acquire knowledge and skills needed to promote sustainable development**, including among others through **education for sustainable development** and sustainable lifestyles, **human rights**, **gender equality**, promotion of **a culture of peace and non-violence**, **global citizenship**, and appreciation of **cultural diversity** and of culture's contribution to sustainable development.

Means of implementation

4.a By 2030, build and upgrade education **facilities** that are child, disability and gender sensitive and provide safe, non-violent, **inclusive and effective learning environments** for all

4.b By 2020, substantially **expand** globally the number of **scholarships** available to developing countries, in particular least developed countries, SIDS and African countries, for enrolment in higher education, incl. vocational training and information and communications technology, technical, engineering and scientific programs, in developed countries and other developing countries

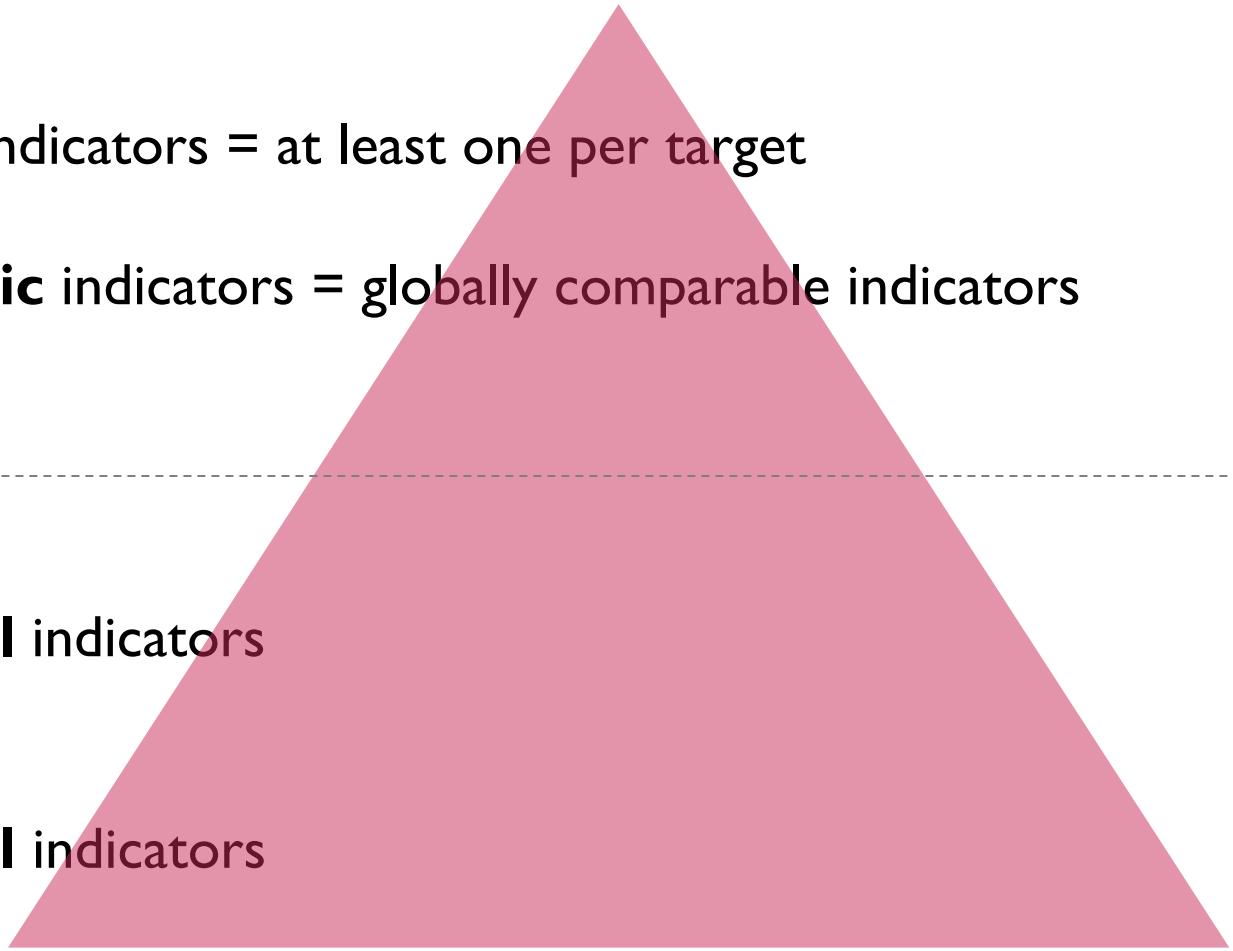
4.c By 2030, substantially **increase the supply of qualified teachers**, including through international cooperation for teacher training in developing countries, especially least developed countries and small island developing States.

SDG4 targets have expanded the learning agenda & reinvigorated earlier debates on measuring learning

- We now have a broader learning agenda in terms of subject domains, relevant skills and competencies for work and life, and relevant values, attitudes and behaviour.
- We now have a learning agenda that is both lifelong and life-wide
- We now have a learning agenda that is relevant in both the Global South and the Global North

Four types of indicators to review and monitor progress in SDGs

- **Global** indicators = at least one per target
- **Thematic** indicators = globally comparable indicators



- **Regional** indicators
- **National** indicators

Source: *Synthesis Report of UN Secretary General* (December 2014)

Global Indicators of Learning

Global Framework

4.1.1 % of children and young people achieving a minimum proficiency in reading and mathematics

4.2.1 % of children under 5 who are developmentally on track

4.6.1 % of youth/adults achieving a fixed proficiency in functional literacy and numeracy

4.7.1 and 12.8.1 Extent to which global citizenship education and education for sustainable development are mainstreamed

Worth noting key aspects of global indicator framework for the SDGs:

1) Was adopted as a "voluntary and country-led instrument"
2) Official statistics constitute the basis…for the global indicator framework, while recommending that national statistical systems explore ways to integrate new data sources…"
3) International organizations shall consult…countries to produce and validate…estimates before publication…"

Thematic Indicators of Learning

Thematic Framework

4.4.2. Percentage of youth/adults who have achieved at least a minimum level of proficiency in digital literacy skills

4.7.4. Percentage of students of a given age group (or education level) showing adequate understanding of issues relating to global citizenship and sustainability

4.7.5. Percentage of 15-year-old students showing proficiency in knowledge of environmental science and geoscience

4. Technical issues in the construction of a global metric of learning

Constructing a Global Metric of Learning (GLM)

Despite the growing number of countries participating in or conducting learning assessments, there is currently no framework to report findings from different assessments in a globally comparable manner.

To construct a GLM, which would create a international reporting framework, there needs to be agreement on:

- Conceptual learning domains in each learning area
- Procedural standards for compiling data
- Reporting scales and benchmarks

Three major options in constructing a GLM

1. Create a new high quality global assessment platform to be used in all countries

2. Expand geographical coverage of existing (high quality) international and regional assessment platforms (e.g., PISA for development; IEA), taking care to embed common learning domains and similar items across platforms.

3. Find ways to use all current assessment platforms (international, regional, national, citizen-led) if they meet quality standards, and produce valid data, and find ways to place assessment results on a universal reporting scale

Options 1 and 2 are expensive, politically infeasible or both. Most current efforts are focusing on Option 3.

UNESCO INSTITUTE FOR STATISTICS

The challenges of reporting on SDG indicators

Using data from existing learning assessments creates challenges since they:

- Vary in construct and content learning domain coverage
- Vary in test design, item format
- Vary in operational and implementation that might affect quality
- Differ in scaling methodology that affects reporting metric
- Vary in quality assurance undertaken in country
- Differ in data analysis undertaken that might affect outcomes

UNESCO INSTITUTE FOR STATISTICS

The challenges of reporting on SDG indicators

Defining **a cut off point** that remains **valid in different platforms and national contexts is difficult**. Without this, it is hard to describe the level of proficiency.

- In Indicator 4.1.1 what counts as meeting a **'minimum proficiency level'** in reading and mathematics for three education levels: lower primary education, upper primary education and lower secondary education.
- In Indicator 4.2.1 what counts as **developmentally on track** with attention to three areas including learning under the age of 5.
- In indicator 4.4.2 how to define youth and adults achieving **minimum level of digital literacy and numeracy**
- In indicators 4.6.1 how to define youth and adults achieving **functional literacy and numeracy**
- In indicators 4.7.4. & 4.7.5 how to measure whether students show adequate understanding of issues relating to global citizenship and sustainability, and knowledge of environmental science and geoscience.

UNESCO INSTITUTE FOR STATISTICS

The 2030 Agenda and reporting

No matter what approach is used, certain assumptions need to be met…

- Learning domains assessed and target populations need to be similar to have valid outcomes.
- There is a need to ensure procedural consistency

Good design and systematic implementation are needed to ensure data collected met minimum quality before analysis and linking.

- Compliance with a minimum standard of quality in the data generating process

UNESCO INSTITUTE FOR STATISTICS

Indicator 4.1.1: Mapping specific benchmarks and a minimum threshold.

Harmonized proficiency levels	Original name of levels by assessment														
	PISA		TIMSS		PIRLS	PASEC				TERCE				SACMEQ	
	Grade and subject														
	8		4	8	4	2	2	6	6	3	3	6	6	6	6
	Reading	Math	Math	Math	Reading	Reading	Math	Reading	Math	Reading	Math	Reading	Math	Reading	Math
Level 8 (L8)	Level 6													Level 8	Level 8
Level 7 (L7)	Level 5	Level 6												Level 7	Level 7
Level 6 (L6)	Level 4	Level 5												Level 6	Level 6
Level 5 (L5)	Level 3	Level 4	Advanced International Benchmark	Advanced International Benchmark	Advanced International Benchmark	Level 4		Level 4						Level 5	Level 5
Level 4 (L4)	Level 2	Level 3	High International Benchmark	High International Benchmark	High International Benchmark	Level 3	Level 3	Level 3	Level 3	Level IV	Level IV	Level IV	Level IV	Level 4	Level 4
Level 3 (L3)	Level 1a	Level 2	Intermediate International Benchmark	Intermediate International Benchmark	Intermediate International Benchmark	Level 2	Level 2	Level 2	Level 2	Level III	Level III	Level III	Level III	Level 3	Level 3
Level 2 (L2)	Level 1b	Level 1	Low International Benchmark	Low International Benchmark	Low International Benchmark	Level 1	Level 1	Level 1	Level 1	Level II	Level II	Level II	Level II	Level 2	Level 2
Level 1 (L1)	Below Level 1b	Below Level 1b	Below Low International Benchmark	Below Low International Benchmark	Below Low International Benchmark	Below Level 1	Below Level 1	Below Level 1	Below Level 1	Level I	Level I	Level I	Level I	Level 1	Level 1

A Long-term solution: The UIS reporting scale

Level 10 and above

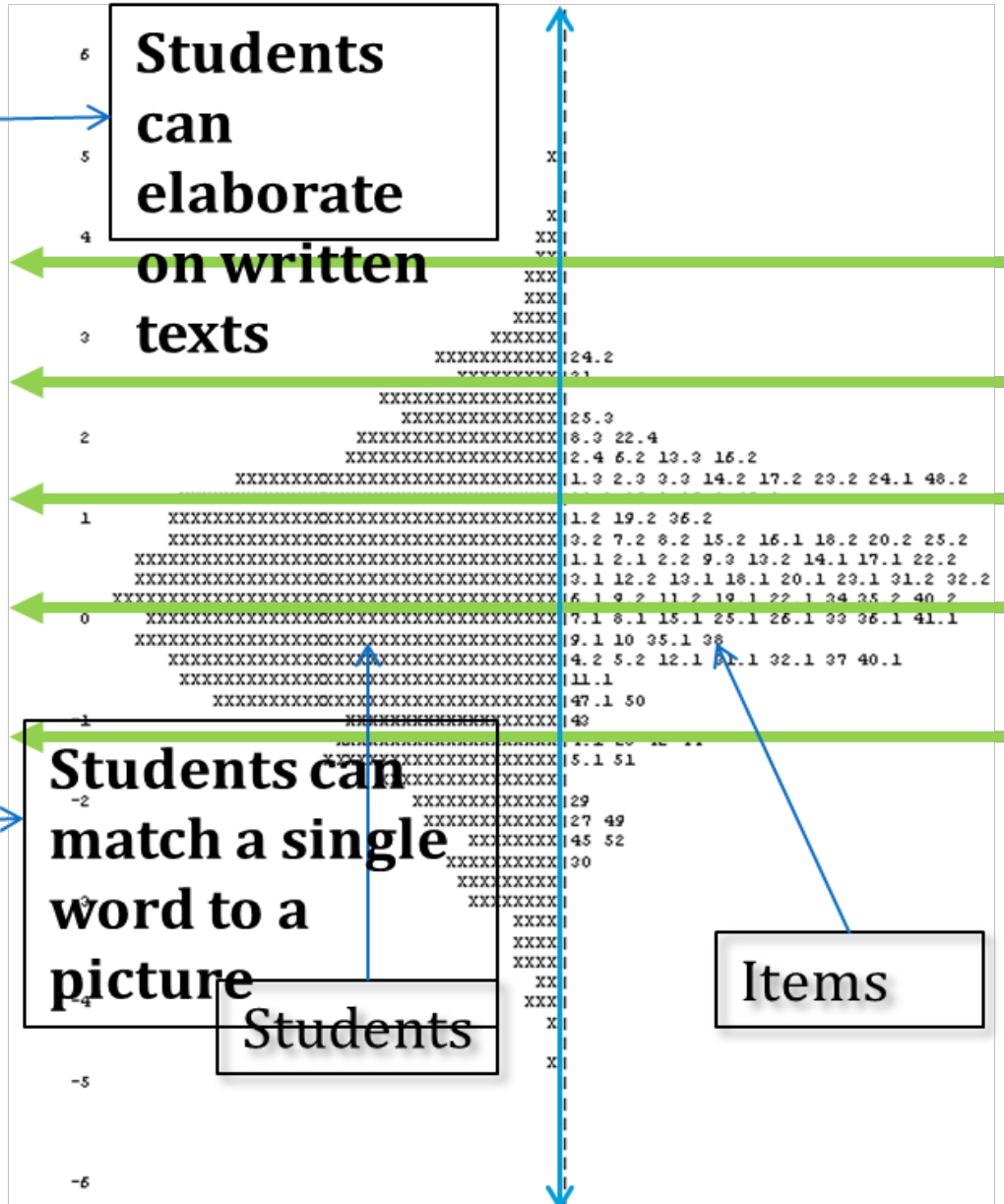
Level 9

Level 8

Level 7

Level 6

Level 5 and below



Recap of Phase I

Phase I Inputs

Use existing cross national assessment datasets

- Determine relative difficulties of items within each assessment program
- Determine how particular items behave in particular linguistic, cultural and curriculum contexts.

Collect views of expert test developers

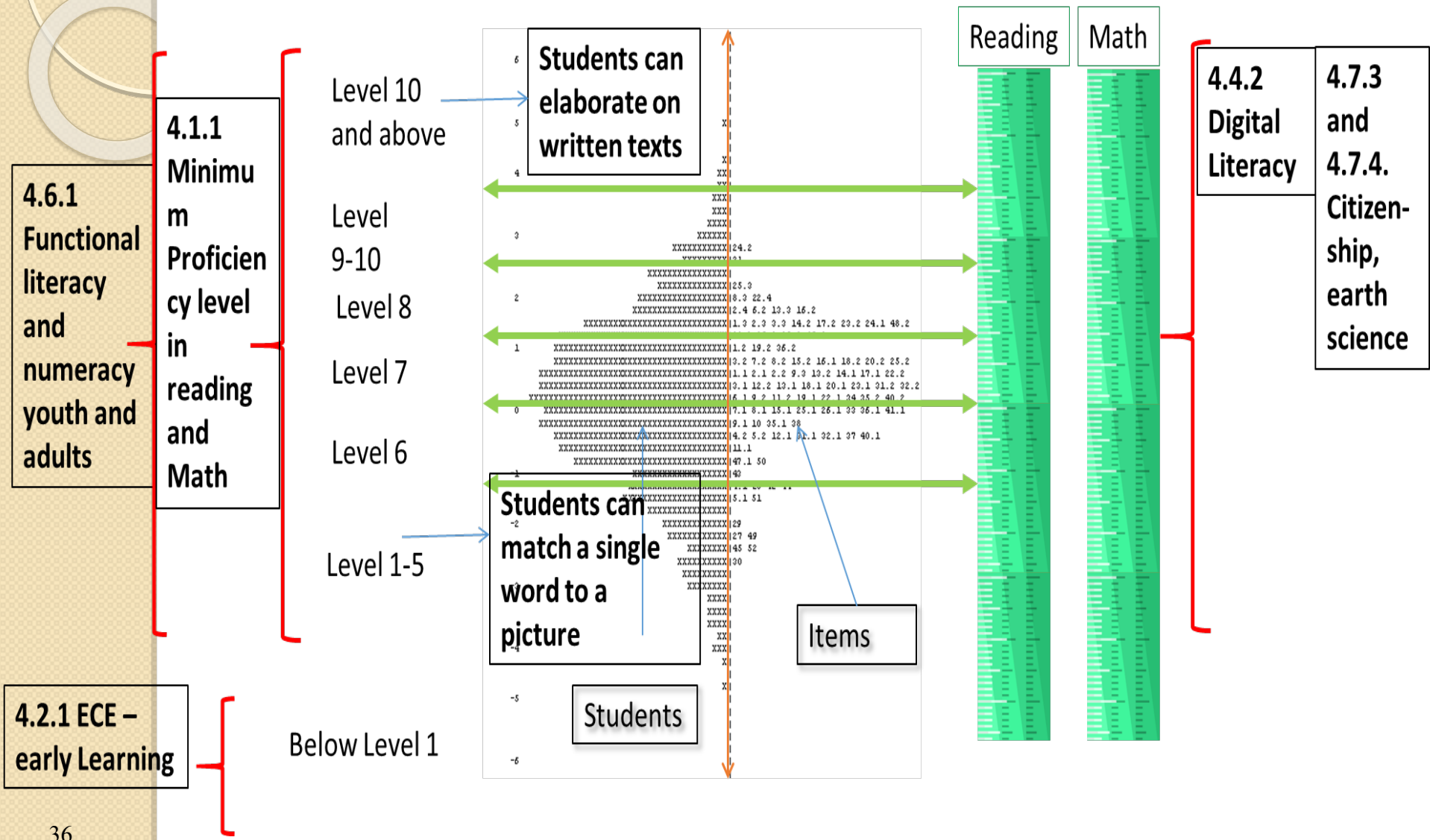
- Judgments on relative difficulties of items across assessment programs
- Views on how particular items behave in different linguistic, cultural and curriculum contexts

UNESCO INSTITUTE FOR STATISTICS

Phase II : Validation and refinement of the scale

- Phase II: Develop a more robust reporting scale by testing, checking and further exploring earlier instruments and judgments
- Collect new data to obtain empirical evidence from student responses
- Two approaches:
  - **Test-based:** administer assessments in their intact forms, either to the same group of children or equivalent groups of children
  - **Item-based:** draw on items from all assessments to create new test forms with items in common; administer new test forms to non-equivalent groups

Phase III: Extending a Lifelong Learning Approach



5. Concluding section

As we know, equity is deeply embedded in the SDG4 agenda

The final Incheon Declaration from the World Education Forum (May 2015) calls inclusion and equity "the cornerstone of a transformative education agenda" and commits governments to develop and implement education policies that address "all forms of exclusion and marginalization, disparities and inequalities in access, participation and learning outcomes."
It makes emphatic this point by stating: "no education target should be considered met unless met by all", and focuses on the disadvantaged and marginalized, "to ensure that no one is left behind."

en.unesco.org/world-education-forum-2015/incheon-declaration

Key Questions

Can efforts and resources being employed to construct a GLM contribute to greater equity in learning -- namely, a reduction of inequalities in learning outcomes between and within countries?

Can a global learning metric that ranks countries based on the percentage of students achieving minimum proficiency in reading or mathematics at the end of primary or lower secondary contribute to improved learning outcomes among the more marginalized and disadvantaged students?

Arguments

A wide array of learning is invoked in the 10 SDG4 targets. Such learning varies by age group – from pre-primary and primary school age children to secondary school age adolescents in and out of school, and adults in different life stages and contexts. The demand for new forms of lifelong learning (job re-training, professional development, ICT training, self-directed learning, cultural knowledge, etc.) is especially pronounced given changing technologies and growing competition for skilled labor.

However, the proposed GLM only examines a limited range of SDG4 learning outcomes – reading and math levels in basic education. GLMs in other subject areas are unlikely in the foreseeable future. Improvements in a broad array of learning outcomes are not only important to ensure progress in SGD4, but also for progress in other SDGs.

Arguments

Creating a GLM that exclusively focuses on cognitive proficiencies in reading and mathematics, while certainly important for learners from disadvantaged backgrounds, diminishes and devalues their knowledge and skills in other domains. This is true in all countries, and even more so in low-income countries.

Thus, to achieve SDG4 targets, and to ensure that no one is left behind, it is critical that countries build capacity to measure learning outcomes across a wide array of domains, and not just in literacy and numeracy. Countries need policies, strategies and programs that address the learning challenges of the marginalized and, concurrently, create valid and reliable sources of evidence to determine which public policies are more or less effective and efficient and why.

On-going concerns

- A global scale for measuring learning levels across countries is increasingly likely.
- Interim data on global and regional deficits in learning, based on standardizing data from many assessment platforms, have already been published (UIS, 2017)
- Powerful actors (eg, World Bank, GPE, European donors, Education Commission, OECD) are supporting (financially and otherwise) the construction of GLMs.
- Some organizations/companies have a monetary interest in creating GLMs and expanding their geographical coverage.
- Current assessments tend to undervalue the knowledge and skills of learners from poor or marginalized groups and ethnic minorities.
- Assessments often convey a deficit model of certain learners, or an impoverished view of learning acquired in and outside of school, partly due to weak models of non-formal and informal learning.
- GLMs are likely to result in unintended consequences: for example, unethical school practices to increase test scores, and undermine the autonomy and professionalism of teachers.

Key questions

- Given the above arguments and concerns, to what extent are GLMs fit for purpose - specifically do they serve the interests of governments, and other educational stakeholders, seeking to make progress on the SDGs and SDG4?
- How will countries react to seeing their students placed on a GLM?
- How, if at all, will countries and other stakeholders use the information from a GLM?
- Will data from GLMs promote improvements in the provision of quality education (eg, teacher training, curriculum development, classroom practices and pedagogy, ICT use)? If so, how?

Conclusions

- If SDG4 is going to be a driver of change for the 2030 Agenda, we need to explore learning synergies across sectors, and develop a more holistic view of relevant knowledge, attitudes and skills beyond foundational skills in reading and mathematics.
- Given that current assessments are less reliable and valid measures of learning among marginalized groups, then the creation of GLMs may inadvertently strengthen deficit models of such learners, impoverish our understanding of they ways they acquire knowledge and skills and exacerbate inequalities.
- Our real challenge is determining: What types of learning, among which groups of learners, at which points in the life courses, are likely to contribute to SDG4 targets and the other SDGs? Currently we have only very limited answers to this critical question.

Thank you!
ありがとうございました

For more information:
Contact Aaron BENAVOT, PhD

abenavot@albany.edu