



United Nations
Educational, Scientific and
Cultural Organization



UNESCO
INSTITUTE
FOR
STATISTICS



GLOBAL
ALLIANCE
TO MONITOR
LEARNING

Global Alliance to Monitor Learning (GAML) Task Force 4.2:

GAML4

November 29 2017 – Madrid, Spain



Scope - GAML SDG 4.2.1 Task Force



- **Target 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

- SDG 4.2.1 - indicator

Proportion of children **under 5** years of age who are developmentally on track in **health, learning** and **psychosocial well-being, by sex**

- a Tier III Indicator

- Three key challenges for **GAML Task Force 4.2**
 1. **Global comparability**
 2. **Definition of “Minimum proficiency level”**
 3. **Periodicity**



United Nations
Educational, Scientific and
Cultural Organization



UNESCO
INSTITUTE
FOR
STATISTICS

UPDATE AND REVIEW

About Task Force SDG 4.2



- UIS is responsible for defining and measuring globally-comparable indicators of Goal 4
- UNICEF is the custodian agency for the provision of data and associated methodological developments for 4.2
- GAML Task Force SDG 4.2 will focus on:
 - ‘learning’ tools and methodologies
 - (ensuring) close links with other GAML initiatives...for early primary grades through target 4.1.1 (a and b in particular)
 - advise and support UNICEF as the custodian agency of SDG 4.2.

Members (30) by Categories & Institution



Categories

Country-Member States

Institution

Estonia (IAEG); Slovak Republic; Qatar ; Uganda(TCG), Phillipines (TCG)

Civil Society

ITA, PAL Network ; Education International (EI)

Academia & Experts

New York University-EQUAL Global Network; McMasters Univ.-Offord Centre for Child ; Univ. of Hongkong; Univ. of Nebraska-Lincoln , Brookings; ACER; FHI 360 ; RTI

Multi-Laterals

UNICEF; UNESCO; GEM Report

Bi-Laterals

USAID, DFID

Regional

OECD

INGOs /Foundations:

Save the Children, OSF; CICED

IAEG

Tiina - ANNUS - Estonia

TCG

Philippines; Uganda,

Chair

Baela R. Jamil - ITA Pakistan; Ed Commission- PAL Network

Co-Chair

Tiina ANNUS - Ministry of Education- Estonia

Assisted by

UIS- Programme Specialist- Omneya .

Update on Progress (June -November 2017)



- Revised TORs finalised in August 2017 aligning GAML's role in Advisory capacity to UNICEF as Custodian of SDG 4.2.
- Encouraged by progress by GAML TF SDG 4.1.1 on an interim measurement strategy, we began conversations with members on the 'interim reporting strategy' for SDG 4.2.1
- Consultation with members for urgency of in-person meeting in October 2017
- Brookings hosted in-person meeting for GAML SDG 4.2.1 on October 27, 2017
- Outcomes from in-person meeting led to expert consultations (early Nov.) on a doable GAML TF action plan for 2018

GAML SDG 4.2 TF Outputs Agreed in August 2017



- (a) Providing inputs to the UIS, GAML and technical partners on the development of the UIS Reporting (Learning) Scale for pre-school ages-
 - Status - completed
- (b) Providing inputs to the UIS and GAML on the revision of modules I and II of the Catalogue of Learning Assessment (CLA, 2.0 version) to ensure that ECE is properly integrated
 - Status - ongoing ..
- (c) Proposing an interim measurement/reporting strategy for 4.2.1. until the ECDI revision is complete and being implemented. This could include, for example, **exploration of coverage and comparability of existing measures.**
 - Status - completed/proposed set of action

Technical Papers Commissioned



1. Key Measurement Questions for SDG 4.2.1 - Discussion Paper for GAML Task Force 4.2 - by Kate Anderson & Abbie Raikes (circulated for discussion) finalized Sept. 2017
2. Options for Development of Indicator 4.2.1 Discussion Paper for GAML Task Force 4.2 - by Hirokazu Yoshikawa Abbie Raikes & Alice Wuermli
3. SDG 4.2.1: Connecting Early Learning to the UIS Reporting Scales
by Dan Cloney - ACER
4. Key questions on the domains of measurement for SDG 4.2.1 Recommendations from GAML Task Force 4.2 by Abbie Raikes

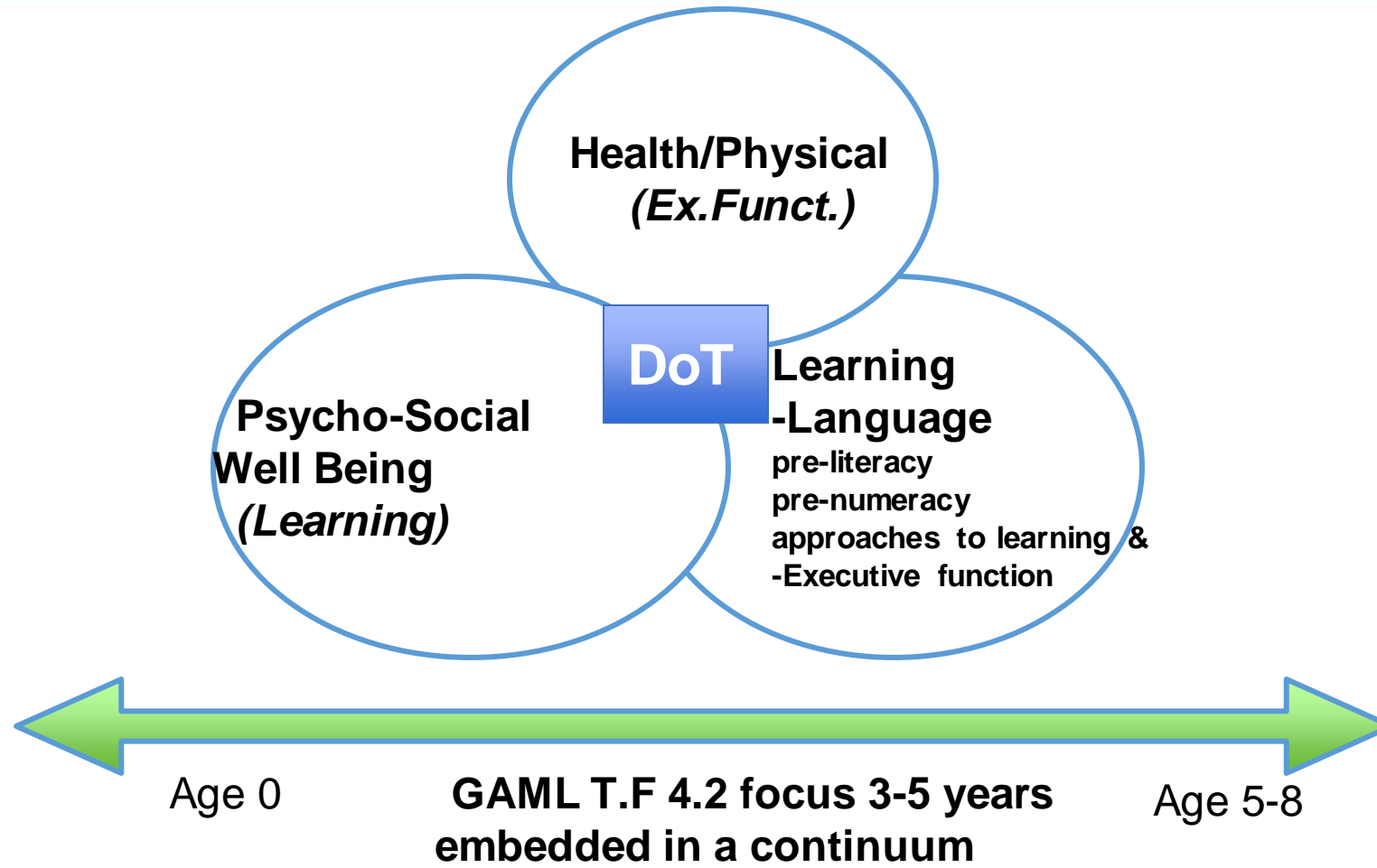


Four Key Questions for our Task Force

1. What to measure? Developing a strong conceptual framework that includes the content and the population covered by the assessments.
2. How to measure? Defining a methodological framework to support rigorous data collection.
3. How to analyse? Determining which approaches to data analysis to use.
4. How to report? Developing a reporting framework that allows results to be compared internationally.

What to Measure?

All 3 Domains -Holistic Def. of Developmentally on Track(DoT)



How to Measure?

All 3 Domains -Developmentally on Track(DoT) at what level

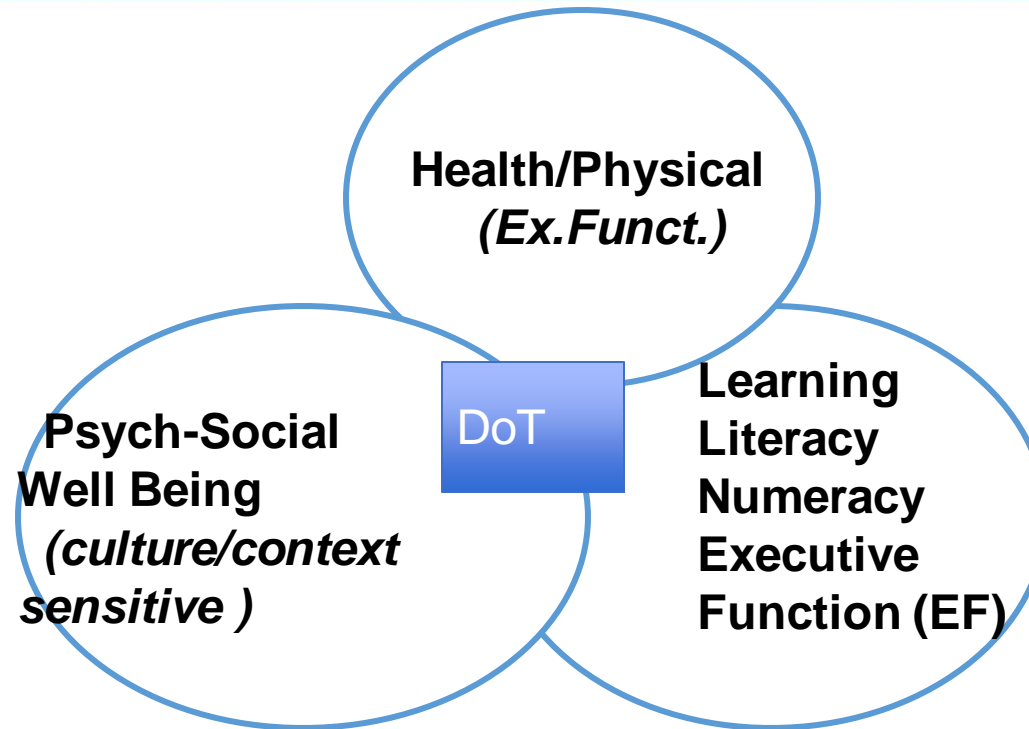


Options

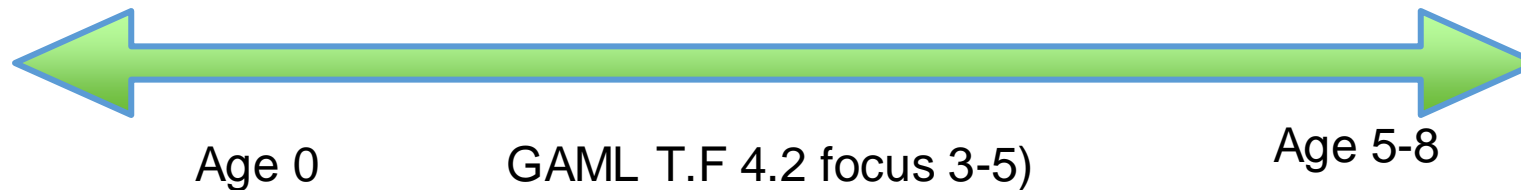
- National Standards
- Global Scale
- Undefined 'evolving'

Hybrid Approach

National Standards reviewed to develop global definition of DoT & a possible Global Scale



ECDI (MICS) Unicef
IDELA (Save the Children)
ELDS(Mc.Master)
MELQO(Brookings,WB, UNESCO UNICEF)
EAP-CDS (Unicef)
PRIDI (Inter-American Dev.Bank)
UNICEF WCARO Early Learning Assessment
Early Human Capability Index
Early Development Instrument



Medium to Long Term Explore Adding to UIS Reporting Scale (4.1.1 (a) with links to 4.2.1)

Decision : Hybrid; National Standards for a Global Definition of DoT & Global in scale



Table 1. Options for defining developmentally on track

Method of Comparison	National Standards	Creation of Global Scale	Leave undefined
Absolute	% children reaching agreed-upon set of skills/competencies, using national standards as starting point	Set of skills defined by experts, but no "absolute" threshold because would be structured as a relative scale	Up to countries to define standard set of skills to measure against, could look across countries over time to identify points in common
Relative	Using national, regional or global population, define "not on track" as 1 or 2 SDs below the mean. With this option, the definition of "on track" would be empirically defined according to a set of items agreed upon using national standards.	% of children reaching development on par with most advantaged children in the world, with set of skills defined by scientists/experts	Create integrated scale that would combine data from range of country-level measures; could then use to create global profile

UNICEF Study on ELDS in 35 countries 2017

Anderson & Raikes
Sept. 2017



United Nations
Educational, Scientific and
Cultural Organization



UNESCO
INSTITUTE
FOR
STATISTICS

REPORTING STRATEGY FOR SDG INDICATOR 4.2.1



Global Alliance to Monitor Learning (GAML) Task Force 4.2: Action Plan 2018

GAML4

November 2017 – Madrid, Spain





Introduction

- Task Force 4.2 Expert Meeting in Washington, D.C., October 27th, 2017
- Action plan for interim reporting
- Using existing data to define developmental milestones across countries → defining “developmentally on track”
- Supporting work of UNICEF - the MICS ECDI



- Hosted generously by Brookings in D.C.

17 Members in Attendance

12 in person & 5 virtual

(ACER, Brookings, ITA, GEM Report; McMaster; Nebraska, PALNetwork, NYU Steinhart, RTI, Save the Children, UIS, UNICEF, WB)

- Preceded by Virtual Meeting with All Members of the Task Force

- After the experts meeting, 3 additional virtual meetings to consolidate the 2018 work plan & costing



- To propose the technical framework required for practical recommendations with a consensus for the ‘interim measurement/reporting strategy’ for SDG 4.2.1-
- To agree on a conceptual framework and process for appropriate measures building of a minimum criteria for the ‘interim measurement strategy’ at the national, regional, and global levels
- To determine the methodological framework based on what is considered as ‘valid assessment practice(s)’ currently for SDG 4.2.1
- To agree on a Reporting Framework for the ‘interim strategy’ that is sensitive to variations in contexts of what ‘developmentally on track’ means
- To explore possibilities of alignment of proposed GAML SDG TF 4.2.1 interim measures for reporting with SDG 4.1.1 a that is underway
- To finalize a work plan for 2018



Two key points of agreement

1. It was noted that given the definition of “interim” as starting at the present moment, interim reporting will use only existing data. Acknowledging that existing data may be flawed in important ways, it is not possible to wait for interim reporting until basic improvements in measurement and data collection are made.
2. It was agreed that for a child to be “developmentally on track” the child would have to be on track in all three domains not just in one or two of the three domains. However, what it means to be “on track” still needs further work

Strategy for Interim Reporting on SDG 4.2.1



SHORT TERM -

- a)** Describe the learning domain and its ties to other domains—general areas of early language/literacy, early numeracy, social/emotional, physical. Describe what developmentally on track means between 0-5
- b)** Identify suggested criteria for reporting through UNICEF to UN for data and measures (and document them in writing and with examples); including necessary and desirable. In the interim, non-ideal measures would be accepted
- c)** Suggested criteria for review by IAEG-ECD, TCG
- d)** Provide suggested criteria and guidelines for member states to report, both for interim and long-term reporting
- e)** Report data with annotations
- f)** Explore what existing data can tell us about benchmarking

Long Term for 4.21. Reporting



Long term strategy for 4.2.1 reporting, to improve comparability:

- i. ECDI is under review, being improved - UNICEF - IAEG-ECD & Expert Group
- ii. Develop empirical approaches that could use data to start identifying commonly-measured “benchmarks” for learning and use the items to build an empirical model over time
- iii. Explore adding to UIS Reporting Scales
- iv. Keep in mind need to have data on the same child for health, learning, psychosocial well-being.

Proposed Process for Interim Reporting for 4.2.1



- Identify ideal criteria for data and measures (and document them in writing and with examples);
- Evaluate existing data sources against those criteria and integrate criteria into CLA and other mechanisms
- Outline a reporting system with two possible approaches:
 1. To describe available data using a format similar to the equating of instruments in 4.1.1; and
 2. To explore possible empirical approaches that could use data to start identifying commonly-measured “benchmarks” for learning and to explore using the items to build an empirical model over time

Identifying Ideal Criteria for Data and Measures



- The interim strategy would be to take all tools, regardless of whether they meet these criteria, and report them using annotations for those that do not meet all the criteria, similar to what is proposed for 4.1.1.
- The long-term strategy would be that all tools would need to meet these criteria to be included in global reporting, including MICS ECDI.
- Encourage/induce some convergence of tools, with an eye towards achieving more comparability.
- Three areas of attention: 1) does the measure cover the necessary domains? 2) what are the properties of the tools? 3) what are the properties of the data?



UNESCO
INSTITUTE
FOR
STATISTICS



GLOBAL
ALLIANCE
TO MONITOR
LEARNING

Global Alliance to Monitor Learning (GAML) Task Force 4.2: Action Plan

1

Steps for completing analytical work



Define characteristics of data sets that will be used in analyses, and define a similar approach to be used across data sets

- Agree on technical characteristics of data sets to be included in analyses
- Outline an analyses plan that investigators working with each data can follow

Define developmental milestones using individual country data files from each assessment

- Use existing data to explore timing of developmental achievements
- Item-level analyses
- Identify similarities and variation in timing of developmental milestones across countries

Examine cross-country functioning of items from each assessment, looking across countries

- Identify items demonstrating cross-country relevance
- Defining “developmentally on track”
- Informing the UIS Reporting Scale

Examine cross-country functioning of items, using data from multiple assessments

- Identify common items across data sets
- Create an approach to examine timing of developmental achievements
- Develop methodologies for “linking” items across data sets



UNESCO
INSTITUTE
FOR
STATISTICS



GLOBAL
ALLIANCE
TO MONITOR
LEARNING

Global Alliance to Monitor Learning (GAML) Task Force 4.2: Action Plan

2

Proposed process and timeline



Convene researchers/holders of large-scale data sets to map out methodology and approach

**January -
March 2018**

Identify psychometricians from each team to work in partnership with expert group to complete the analyses and make recommendations using the findings to inform the interim reporting strategy

**March -
August 2018**

Finalize deliverables and generate recommendations for the interim reporting strategy and defining "developmentally on track"

**Sept- Oct
2018**

- Paper outlining methodology and/or guidance from psychometricians
- Interim report on results and preliminary findings
- Final report and implications for reporting scale and other ECD instruments



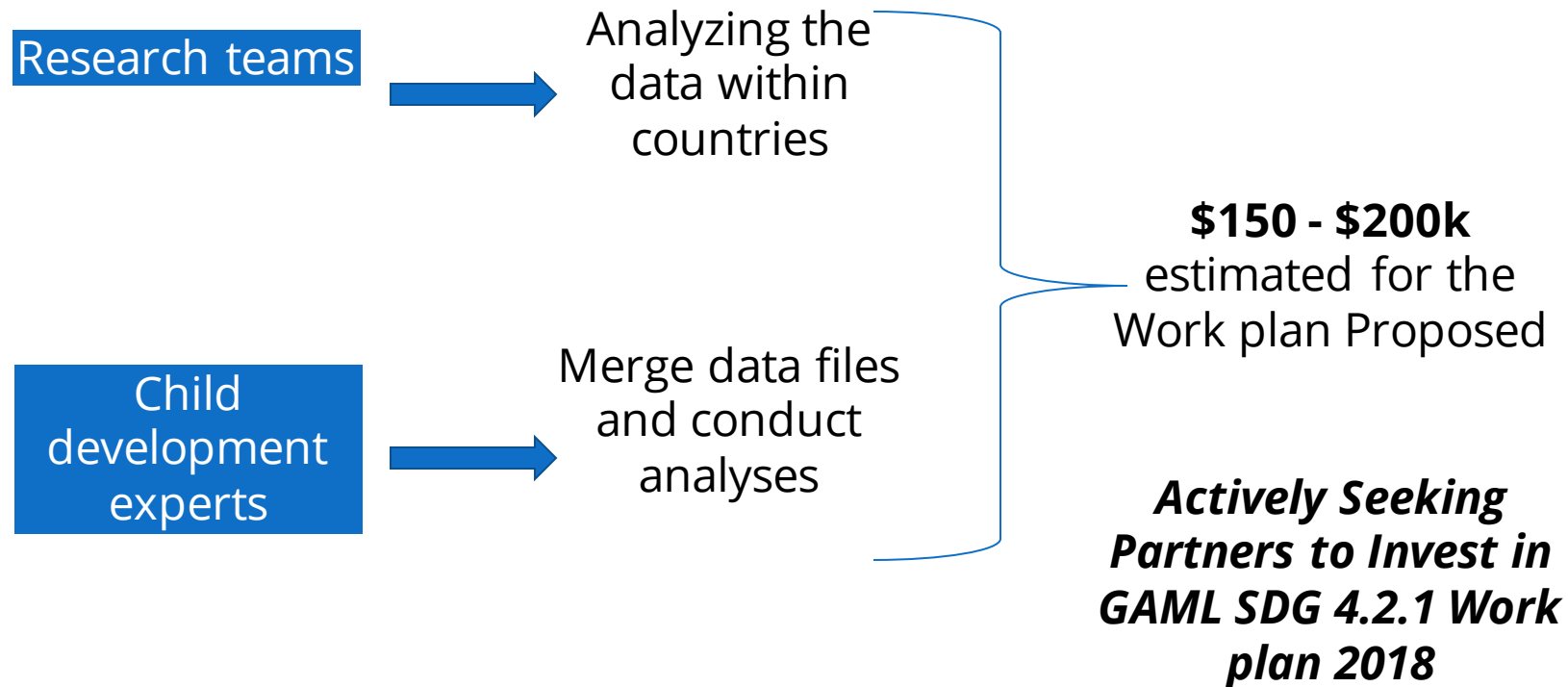
UNESCO
INSTITUTE
FOR
STATISTICS



GLOBAL
ALLIANCE
TO MONITOR
LEARNING

3

Structure and funding





United Nations
Educational, Scientific and
Cultural Organization




UNESCO
INSTITUTE
FOR
STATISTICS



GLOBAL
ALLIANCE
TO MONITOR
LEARNING

Thank you!

Learn more: <http://uis.unesco.org/>

 @UNESCOstat

The following list of technical optimality criteria, proposed during the 27 October 2017 meeting, with one or two which could be added, are to be used for:

- a. Annotating in the interim period
 - b. Deciding on non-reporting in the interim period if a survey or instrument does not meet some crucial criteria. (There may be sub-criteria.)
 - c. Guiding technical developments during the interim period in order to create improved measurement, data collection, and reporting post-interim.
1. Include definition of “developmentally on-track”.
 - Criterion-referenced or, if not criterion-referenced, at least using clear, empirically well-based, and agreed norms
 - May start with a definition for each domain, but note that “on track” means that the child has to “on track” on each domain
 2. Measure learning in a holistic way – that is, measure encompasses all three domains in the SDG.
 - Health, psychosocial well-being, learning
 3. Population-based; that is, representative of the whole population in question, or, if not available for the whole population (e.g., not all age groups), then for representative sub-segments. It would be noted in particular if the measures are representative only of self-selected members or clearly non-representative parts of the population.
 4. Conducted on a representative sample basis.



Proposed Criteria cont..

5. Useful to countries given national standards (or at least not be inconsistent with what countries are working towards for their own purposes).
6. Be globally comparable, or have items and definitions that allow one to determine its comparability with a determinable degree of accuracy.
7. Include background work that allows one to determine reliability and validity.
8. Administered at a variety of ages so that growth curves can be seen for the measures where it is relevant.
9. Have a well-defined reporting framework.
10. Follow the standards in the UIS Good Practices in Learning Assessment (GP-LA) and other "standard" codes of good measurement practice, incorporated by reference.



UNESCO
INSTITUTE
FOR
STATISTICS



GLOBAL
ALLIANCE
TO MONITOR
LEARNING

Participants of Oct 27 In-Person Meeting

Participants in attendance on 27 October 2017

Name	Organization
Abbie Raikes*	University of Nebraska
Alvin Vista	Brookings Institution
Amanda Devercelli*	World Bank Group
Amber Gove	RTI
Amy Jo Dowd	Save the Children
Baela Raza Jamil*	ITA/PAL Network
Claudia Cappa	UNICEF
Dan Cloney*	ACER
Esther Care	Brookings Institution
Hiro Yoshikawa*	NYU Steinhardt
Kate Anderson	Brookings Institution
Luis Crouch	RTI
Magdalena Bendini	World Bank Group
Magdalena Janus	McMaster University
Manos Antoninis	Global Education Monitoring Report (GEMR)/UNESCO
Manuel Cardoso	UNICEF
Silvia Montoya	UNESCO Institute for Statistics

Tools in use to measure 4.2.1

- **Tool; Region Purpose; Method of administration**

- **Early Development Instrument (McMaster University)**- Canada has been adapted and used in representative samples in other countries: Population-level measurement of children's development for 4- to 6-year-olds
Teacher report

- **East Asia Pacific Child Development Scales (UNICEF)**: East Asia region; used in representative samples in 9 countries to date: National level and regionally-comparable data on the development of children aged 3 to 5 years; Direct assessment; short form of scale now developed and ready for use

- **IDELA (Save the Children)**: Global tool; used in at least 30 countries; Programme and national- level data on children's development between 3 and 6 years: Direct assessment

- **MICS Early Child Development Index (UNICEF)**: Global tool; used in representative samples in at least 50 countries: Globally-comparable and national-level data on the development of children aged 36 to 59 month - Parent report through household survey

- **Measuring Early Learning and Quality Outcomes (MELQO)** (Brookings Institution, UNESCO, UNICEF, World Bank): Designed for use as a global "core" to integrate into existing tools and national-level assessments; Globally-comparable and national-level data on children's development between 4 and 6 years;
Direct assessment, teacher or parent survey

- **PRIDI (Inter-American Development Bank)**: Latin America region; used in 4 countries Regional and national- level data on early childhood development and household contexts: Direct assessment; parent survey

- **West and Central Africa Regional Office Regional Prototype (UNICEF)**; West Africa; used in representative samples in 8 countries

- National-level and regionally-comparable data on children's development in the first year of school (6-year- olds);
- Direct assessment of children through groups and individual assessment in schools
- Source: UNESCO Institute for Statistics, concept note by Raikes, 2016



Elements to Consider

Many elements of child development that follow a neurobiologically-driven, universal pathway- General domains like cognitive development/learning, social/emotional development and health all have elements that are universally relevant.

Deciding upon standards for /domains of measurement for SDG 4.2.1 for international comparability in early childhood data, to assess existing data sources against these standards.

Potential tensions between feasibility and precision.

For population-wide measurement of children who have not yet started formal schooling, household surveys are the most viable form of data collection for capturing all children, if the overall goal is international comparability.

Direct, oral assessment of children will yield the most accurate information on specific aspects of children's skills and knowledge, but requires trained observers.

Cost is another consideration. Household surveys are typically more expensive than center or school-based assessments

Direct assessments of learning and development typically involve a longer process of reliability training and can take more time to administer than parent or caregiver surveys.

Some assessments used internationally have licensing fees or require countries to pay for training by the test developers, which can make some internationally- developed assessments more expensive than locally-developed ones; now increasing number of freely available assessments available for early years, but require ongoing investments in technical development and staffing for quality of assessment- costs must be covered.

What is feasible to measure in an internationally-comparable manner across all domains is likely quite general, and may not be specific enough to drive national-level policymaking, but is useful for status of children's development.

Internationally-comparable data has the advantage of spurring global action in ways that country or regional data is not able to do- but generality is achieved a very rough overview; the value of those data for policymaking should be evaluated.

Country and regional data, on the other hand, may be able to provide a more nuanced look at child development, by allowing the inclusion of nationally and regionally relevant constructs and items.

To achieve a global picture of equity in child development, it will be necessary to create a method for integrating the measures to generate estimates of child development across countries – and high-income countries may end up on a different scale than low-income countries.