



United Nations  
Educational, Scientific and  
Cultural Organization



UNESCO  
INSTITUTE  
FOR  
STATISTICS

# Measurement/Reporting strategy for SDG Indicator 4.2.1

## Proposal by GAML Task Force 4.2

Global Alliance for Monitoring Learning  
Fourth meeting  
28-29 November 2017  
Madrid, Spain

GAML4/12



**SDG 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

## Introduction

On 27 October 2018 an expert meeting hosted by GAML at the Brookings Institution, in Washington, D.C., took place. (See list of attendees present). Presentations were made and a rich discussion ensued. The 4.2.1. Task Force made recommendations after the workshop. The following are the conclusions and steps forward as seen from the Task Force, based on the discussion on 27<sup>th</sup> October and various report-backs and discussions post meeting.

This document, therefore, comprises recommendations for a strategy for interim *reporting* on 4.2.1 and at the same time outlines a strategy for medium-term improvement of both measurement *and* reporting. We define “interim” as the period starting essentially with the present moment, and extending until a sufficient improvement in *measurement* and data *collection* has been made to warrant a change in the *reporting* process. It is estimated that this could take two or three years. This does not mean that improvements in measurement and collection of data would await until the end of the interim period. Various partners will proceed to analyze, develop better measurement and data collection procedures, etc., starting at the present moment and going on until such a time as sufficient changes and evidence have accumulated so as to warrant a change in reporting policy.

There were two key points of agreement or “sense of the meeting” that are worth highlighting before moving on to the recommendations.

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, because interim reporting begins now, 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 as described below.

The following steps are provided as logically emanating from the meeting and the discussions.



## Steps

| Time period   | Suggestions  |
|---|--|
| <b>Interim – the present moment</b>   | <p>Countries report according to the instrument or reporting method of their choice, subject to two caveats:</p> <ol style="list-style-type: none"> <li>a. Reporting is annotated according to a set of technical characteristics that measurement and data collection specific to early childhood should possess. Using this, the global community can judge how to use the reported information. These criteria (referred to as the “optimality” criteria) will be elaborated over the next few months, but a sketch for them was provided in the presentations discussed on 27 October. The sketched criteria are provided as an annex to this document for easy reference.</li> <li>b. In some cases, if there are serious technical issues with the underlying instruments or data, the data, or components thereof, might not be reported. An annotation to that effect will be made, to distinguish these cases from cases where there was no reporting.</li> </ol> <p>It would have to be accepted that in some cases there will be no reporting during the interim period.</p> <p>Given UNICEF’s mandated role in the SDG 4.2.1 reporting process it is expected that a common default reporting would be based on data from the ECDI. However, it is expected that as the ECDI improves, it too should fit the optimality criteria for all reporting. Furthermore, it is expected that some countries might not choose to participate in the ECDI or might choose to report using a different approach, even if they do participate.</p> |
| <b>Interim – starting at present and with urgency, but not necessary in the next few months</b> | <p>The following steps are listed in approximate order of urgency and sequentially.</p> <ol style="list-style-type: none"> <li>1. Analysis of existing databases including ECDI, but also others, plus an expert meeting, to define benchmarks (based on the empirical work plus expert opinion) for “developmentally on track.” This could be, but need not be, initially be used as part of the optimality criteria to annotate countries’ reporting (that is, do the country benchmarks resemble the GAML-recommended ones?) but, in the less-urgent interim period, to guide creation of country benchmarks and encourage a convergence to benchmarks that are as common as possible. A consultant or firm could be contracted to carry out the analysis and present the results to a specialized expert panel. The Secretariat or the Task Force would then comment, amend, and recommend the benchmarks. This step would “receive” and recognize analyses under way by various actors in the Task Force or outside it, but it is recommended that a person or team be designated to ensure that no information is “orphaned” and who can collate and “receive” these research inputs. The process would specifically help inform the ECDI and for that would need to be provided to the ECDI process by mid-year 2018.</li> </ol>  |



| Time period  | Suggestions   |
|--|---|
|  | <p>2. Clearer specification to be developed of the three domains (and sub-domains) covered, including the ECDI, but also providing domain definition useful for countries reporting using measures other than the ECDI. This would require a consultant to examine the main instruments that already exist and their background literature, collating and working out commonalities, in a manner similar to that done by UIS and IBE via a consultancy, for mathematics for 4.1.1.</p> <p>3. Finally, as the above work proceeds, or with those results in hand, engage a consultant or firm to develop equating procedures across various instruments, as well as guidelines for making country-based assessments “equitable” via the inclusion of suitable items. This procedure can include item-based equating or “social moderation” or “conceptual equating.” It would be similar to work described for 4.1.1 by ACER in the context of its work with UIS on a universal reporting scale, and by consultants appointed by UIS for that purpose who have provided some views on “social moderation.”</p> <p>It is recommended that at least steps 1 and 2 above be carried out in sufficient time to affect UNICEF’s process of ECDI improvement. But note that since it is assumed that many countries might not initially participate in the ECDI, or may participate but may not report based on the ECDI, the steps outlined here (and in the optimality criteria in the previous row) would feed a process that would:</p> <ol style="list-style-type: none"> <li>a. Continue to improve annotation</li> <li>b. Guide processes whereby countries and assessment agencies that work with countries would improve their measurement tools, data collection, and reporting</li> </ol> |
| <p><b>Longer term, within-interim and post-interim</b></p> | <p>Official and unofficial agencies and countries utilize the measurement and reporting guidelines to gradually improve the quality of measurement and its comparability.</p> <p>Measurement and reporting based on improved measurement. The optimality criteria for interim <i>reporting</i> also serve as a guide for ongoing <i>improved measurement</i> (i.e., measurement with known reliability and validity analyses, comparability analysis or comparable items, etc. as per the optimality criteria in the annex.)</p> <p>At fixed points in the future, such as the start of a new round of ECDI, reporting comes to include improved practices.</p> <p>At that point (and at various points in the process) the GAML 4.2.1 Task Force, as well as the Secretariat, take stock of how the various steps are proceeding.</p> <p>Sometime during 2020 a more intense process of scrutiny would take place to see how well the various measures are converging and how the quality of measurement is improving, and corrective action would take place.</p>   |



## Annex A.

### *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                  |

*\*denotes participation via videoconference*

## Annex B.

### *Technical optimality criteria*

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.
  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.