SDG 4
Exploring Options for Reporting

Silvia Montoya
Director, UNESCO Institute for Statistics
Reporting and measuring progress in SDG4

• What are now accepted as standard features of assessment systems that produce comparable results?

• What are some of the implementation problems detected even in relatively well-designed testing programmes?

• How to build comparable results?

• Whole debate around measuring learning outcomes.
  • cross-country comparability has been over-emphasised relative to comparability over time within countries.
  • It is the latter the one of greatest importance for national policymakers
The reporting format aims to communicate two pieces of information:

- the percentage of students meeting minimum proficiency standards for the relevant domains (mathematics and reading) for each point of measurement (grades 2/3; end of primary and end of lower secondary); and

- when different programs can be considered comparable and the conditions under which the percentage can be considered comparable to the percentage reported from another country.
Content Alignment

Activities

Outputs to Inform Reporting

1. Global Content Framework (GCF)
2. Content Alignment Tool
3. Online Platform
Procedural Alignment

Conceptual Development

Quick Guides for implementation

Procedural Alignment Tool

Activities

Outputs to Inform Reporting

1. Manual of Good Practices
2. Quick Guides for Implementation
3. Procedural Alignment Tool and Online Platform
Conclusions from the expert meeting on Indicator 4.1.1. Representatives from regional and international learning assessments met to agree on minimum proficiency levels (MPLs) in Paris, France.

SDG Indicator 4.1.1: Proportion of children and young people in (a) Grade 2 or 3; (b) at the end of primary education; and (c) at the end of lower secondary education, achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex.
Conclusions from the expert meeting on Indicator 4.1.1

**Definition of Minimum Proficiency Level**

- Grade 2/3
- End of primary
- End of lower secondary

**Facilitate alignment between assessment programmes;**
**Enable countries to pursue different options for assessments; and**
**Allow for some harmonization of results**

**Reading & Mathematics**

- Mapping proficiency levels
- Mapping proficiency level descriptors
- Identify common recommended benchmarks for MPLs → Harmonizing national assessments
Agreement on comparability for reporting indicator 4.1.1
- General Policy Level descriptor (what the child/youth) is able to do
- Multi-fold reporting

<table>
<thead>
<tr>
<th>Assessment/ Level</th>
<th>End of Lower Secondary</th>
<th>End of Primary</th>
<th>Grades 2/3</th>
</tr>
</thead>
<tbody>
<tr>
<td>End of Lower Secondary</td>
<td>• PISA level 2</td>
<td>• SACMEQ level 3 / level 4</td>
<td>• PASEC grade level 2</td>
</tr>
<tr>
<td></td>
<td>• TIMSS low international</td>
<td>• PASEC grade 6 level 1</td>
<td>• TERCE grade 3 level 2</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• PILNA level 6</td>
<td>• Further work to align MICS/ASER/Uwezo</td>
</tr>
<tr>
<td></td>
<td></td>
<td>• TERCE grade 6 level 1</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• TIMSS 4 intermediate international benchmark</td>
<td></td>
</tr>
</tbody>
</table>
Linking: Several ways forward...

• Approaches that have been put forward differ most obviously in terms of their technical complexity, financial cost, and implied comparability of national statistics.

• Less obvious differences relate to
  • their sustainability over time,
  • their impact on the politics, planning and operations of national education authorities,
  • their ability to contribute to capacity building within these authorities, and
  • their persuasive power in the media and policy debates.
The 2030 Agenda and reporting (II)

• A dialogue about definitional issues
  • What is the construct (for instance, reading?)
  • What are the contents?
  • What is the minimum proficiency?
  • How to express everybody in same scale?
• No matter what methodology that is used there are assumptions need to be met...
  o Learning domains and target population needs similarities to have valid outcomes.
  o Ensure procedural consistency
• Respect to national ownership, meet national needs and sensitivity to cultural values
SDG4 reporting – Innovative Solutions to comparability
3 options to link assessments to a global scale

Non - Statistical Methods

Social Moderation
Experts as moderators

- Alignment of constructs
- Link through proficiency levels policy descriptors

Statistical Methods

Test-based linking
Common individuals, different tests
- Psychometric calibration based on 3 countries that sit for two different tests
- Concordance expanded table that includes all countries

Item-based linking
Common items, different individuals
- Psychometric calibration based on common items in different test for different individuals
- Alignment of standards and levels of difficulties
Rosetta Stone Proposal

International Assessment
- Argentina
- United States
- South Africa
- Ghana
- El Salvador
- Botswana
- Tunisia
- Pakistan

Regional Assessment
- Bolivia
- Argentina
- Venezuela
- Cuba
- El Salvador
- Costa Rica
- Ecuador
- Guatemala

Countries in Regional Assessment not Participating in IEA

Expanded Table
- Burkina Faso
- Mozambique
- Namibia
- Burundi
- Swaziland
- South Africa
- Senegal
- Botswana

* Assessment rankings are hypothetical for descriptive purposes only
Item-based linking: The Learning Progression Explorer

Students can elaborate on written text

Level 10 and above

Level 9

Level 8

Level 7

Level 6

Level 5 and below

Students can match a single word to a picture

Items
Thank you!

Silvia Montoya
s.montoya@unesco.org

Learn more:
http://gaml.uis.unesco.org/index.html/
@UNESCOStat