

TASK FORCE 4.1 UPDATE

Fourth Meeting of the Global Alliance to Monitor Learning

November 28-29, 2017

*General Secretariat of the Organization of Ibero-
American States, Madrid, Spain*

OVERVIEW

1. Task Force 4.1 composition and mandate
2. Progress to date
3. Inputs to 4.1.1 measurement and reporting strategy
4. Conclusions and next steps

1. TASK FORCE 4.1 COMPOSITION & MANDATE

61 volunteers representing various stakeholders, countries, and organizations

Identify technical approaches to measurement of learning under Target 4.1, particularly Indicator 4.1.1.

Target 4.1: By 2030, ensure all complete quality primary and secondary education leading to relevant learning outcomes

Indicator 4.1.1: Proportion: (a) in grades 2/3; (b) at end of primary; and (c) end of lower secondary achieving at least minimum proficiency in (i) reading, (ii) math, by sex

2. PROGRESS TO DATE

- ❖ 5 meetings – March 8; April 3, 10; August 17, October 10
- ❖ 3 GAML technical products/outputs reviewed
- ❖ 2 Learning Progression Explorer Webinars
- ❖ 2 cross-national assessments expert meetings
- ❖ 1 Subgroup on 4.1.1a

Report: *“Task Force 4.1 Inputs to the Measurement and Reporting Strategy for Indicator 4.1.1”*

3. INPUTS TO 4.1.1 MEASUREMENT & REPORTING STRATEGY

Overall recommendations for next steps

- ❖ GAML Secretariat/UIS to convene diverse group of content experts, developmental psychologists, assessment experts, and others who can bring latest research, evidence, and data to bear on drafting of longer-term strategy, particularly for 4.1.1 a.
- ❖ Countries to be more actively brought into discussions on 4.1.1 to ensure that proposed measurement and reporting approaches are sufficiently adaptive and responsive to their contexts.

3. INPUTS TO 4.1.1 MEASUREMENT & REPORTING STRATEGY ctd

Recommendations for 3 phases

- I. Conceptual framework: Who and what to assess?*
- II. Methodological framework: How to assess?*
- III. Reporting framework: How to report?*

I. Conceptual framework: Who and what to assess?

Status

- Most assessment programs provide grade-based data relevant to 4.1.1b and 4.1.1c

Key Issues

- 4.1.1 a assessments should focus on precursor and early skills; emphasize accuracy, comprehension, automaticity/speed. Very few cross-national assessments measure these precursor and early skills.

Recommendations

- Short-term: (i) Continue with mapping frameworks, but more focus on grades 2/3; (ii) Consider drawing on EGRA, EGMA, household/citizen tools for 4.1.1 a
- Longer-term: Develop set of purpose-built tools for countries for 4.1.1 a

Task Force feedback on content reference frameworks for mapping

- ✓ Extend/test against other languages and cultures
- ✓ Incorporate other disciplines/perspectives
- ✓ Use more explicitly research-based approach and more specialist input
- ✓ Incorporate more concrete examples
- ✓ Explain how framework might be adapted over time

II. Methodological framework: How to assess?

Status

- Most assessments emphasize sample-based and group-administered approaches and primarily focus on children and youth in school.

Key Issues

- Most early-years assessments designed for one-on-one administration
- How (i) include OOSC, (ii) determine acceptable minimum data quality requirements, (iii) decide which assessment to use?

Recommendations

- Short-term: Be flexible and focus more on encouraging countries to get in habit of submitting data on learning
- Longer-term: Move towards more school-based and group-administered approaches and more rigorous standards and criteria

III. Reporting framework: How to report?

Status

- Most cross-national assessments convert raw to scaled scores using IRT and report scaled scores and % of students reaching specific proficiency levels

Key Issues

- Many national assessments still report mean raw scores or % correct.
- Comparability of results across systems and languages is issue for all assessments.

Recommendations

- Short-term: (i) Prioritize comparisons within languages. (ii) Use hybrid approach of translation and adaptation to balance relative difficulty of instruments across languages and enhance comparability
- Longer-term: Enhance comparability of results by linking assessments

Task force feedback on UIS reporting scale

- ✓ Clarify relationship to Content Reference Frameworks
- ✓ Be clearer about objective and target audience
- ✓ Pay more attention to representation, inclusion, efficiency in empirical validation
- ✓ Consider using more traditional reporting scale
- ✓ Consider developing three scales, one for each measurement point
- ✓ Consider giving more attention to development of existing cross-national assessments

Task Force feedback on setting benchmarks on reporting scale

1. Global or national “minimum proficiency” benchmarks?

➤ **Even split**

➤ 2. One or three “minimum proficiency” benchmarks per domain?

➤ **Majority in favor of 3 benchmarks**

3. Existing “minimum proficiency” benchmarks or new benchmarks?

➤ **Slightly more in favor of existing**

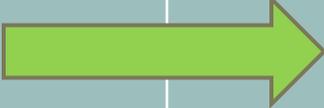
➤ 4. Global or national expectations for % of students to reach “minimum proficiency”?

➤ **Even split**

➤ 5. Status- or progress-based expectations?

➤ **Slightly more in favor of status-based**

4. CONCLUSIONS & NEXT STEPS

Short-term		Longer-term
Continue mapping assessment frameworks, but with more focus on grades 2/3		
Consider drawing on EGRA, EGMA, household/citizen tools for 4.1.1 a		Develop set of purpose-built tools for 4.1.1 a
Be flexible, focus on encouraging countries to get in habit of submitting data		Move towards more school-based and group-administered approaches and more rigorous technical standards and criteria
Prioritize comparisons within languages. Use hybrid approach of translation and adaptation to balance relative difficulty of instruments.		Enhance comparability of results through linking assessments

4. CONCLUSIONS & NEXT STEPS

• Content reference frameworks

- ✓ Extend/test against other languages and cultures
- ✓ Use more explicitly research-based approach, more specialist input, other disciplines
- ✓ Incorporate more concrete examples
- ✓ Explain how framework might be adapted over time

• UIS Reporting Scale and Benchmarks

- ✓ Clarify relationship to Content Frameworks
- ✓ Be clear about objective and audience
- ✓ Pay attention to representation, inclusion, efficiency in empirical validation
- ✓ Consider using more traditional reporting scale
- ✓ Develop 3 scales, one per measurement point
- ✓ Use 3 benchmarks per domain, one for each measurement point
- ✓ Give more attention to development of existing cross-national assessments