



United Nations
Educational, Scientific and
Cultural Organization



UNESCO
INSTITUTE
FOR
STATISTICS

SDG 4 Reporting: Defining benchmarks and proficiency levels in the UIS Reporting Scale

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

GAML4/3



The UNESCO Institute for Statistics (UIS), through its Global Alliance to Monitor Learning (GAML), is working on an approach to monitoring learning outcomes for Indicator 4.1.1 of the UN's Sustainable Development Goal 4 (SDG 4): Quality Education:

Proportion of children and young people: (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics. (Inter-Agency and Expert Group on SDG Indicators, 2016)

Monitoring progress against Indicator 4.1.1 will require the reporting of student outcomes at several different stages of learning in a broadly consistent way across education systems, to enable meaningful international dialogue about learning progress and how it may be supported.

Learning and what is measured about learning varies widely across local contexts. Agreeing on common contents and competencies is a first step. Education systems make independent interpretations and decisions about what learning means, how it is described in curriculum, and how it is assessed and reported.

This short summary describes the process to develop a reporting scale and recommended benchmarks.

Reporting scale and recommended benchmarks

A process is needed to define what students are required to learn in each level, domain and point of measurement, in this case, more specifically, what students are expected to learn in reading and mathematics at the three levels of education defined in Indicator 4.1.1 – grades 2/3, end of primary, and end of lower secondary.

The following process is proposed to construct the global reference for content and the UIS proficiency scales.

1. Decide number of performance levels and labels on the UIS Proficiency Scales

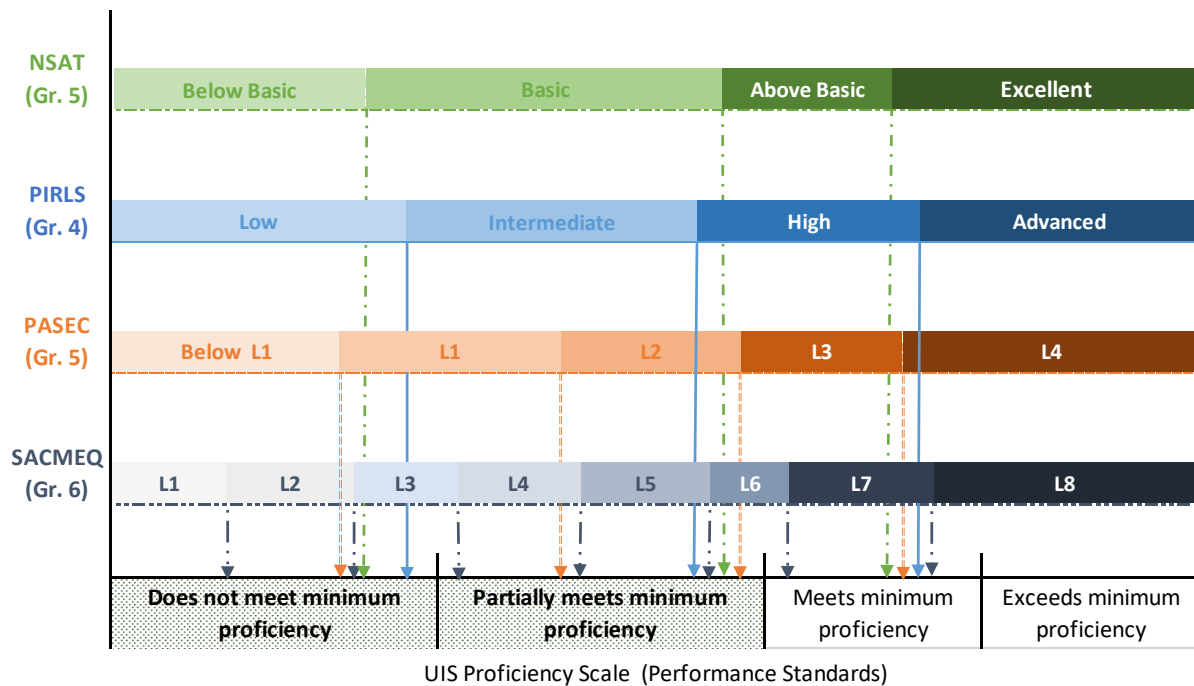
In this step, the number of levels to be used and their names on the scales are determined. This could be done by the GAML. Typically, no more than four performance levels are needed (Perie, 2008). Beyond four levels, it becomes difficult to describe meaningful differences across the levels. Three is probably advisable for the UIS proficiency scales. After determining the number of levels, the next task is to name the levels. There are no clear-cut guidelines on how to develop names for the levels; however, it is recommended that they be thoughtfully chosen to relate to the purpose of reporting and supportable inferences arising from the classifications (Cizek & Bunch, 2007).

Below are some examples of labels, based on those used in various assessment programmes, which the UIS could consider adopting for the UIS proficiency scales:

- Does not meet minimum proficiency/partially meets minimum proficiency/**meets minimum proficiency**/exceeds minimum proficiency.¹
- Does not meet standards/partially meets standards/**meets standards**/exceeds standards.
- Below basic/basic/**proficient**/advanced.
- Beginning step/nearing proficient/**proficient**/advanced.
- Level 1/level 2/**level 3**/level 4.
- Novice/apprentice/**proficient**/distinguished.

The levels in bold are considered the “desired” level of student performance that policymakers expect all students to achieve.

Figure 1. Mapping proficiency levels for national and cross-national assessments: an example



2. Write policy definitions for the performance levels of the UIS Proficiency Scales

The next step is to develop a generic policy definition for each performance level. These definitions are not linked to content but are more general statements that assert policymakers’ position on the desired level of performance. They are particularly useful in the context of reporting multiple assessments. First, they facilitate the articulation of performance levels across grades by ensuring the same level of rigor at each level across each grade. Second, they allow a reader to interpret

¹ Although labels for four performance levels have been proposed, it is recommended that the UIS consider using *three* levels, given the complexity of defining proficiency levels for a diverse set of countries to report against and in linking NAs and CNAs to the UIS scales. Four levels may suggest a level of precision that is not supported.

proficiency in a similar manner regardless of the subject assessed. The policy definitions need to be written for each level, not including the lowest level of performance. Table 1 presents some examples from assessment programmes in the United States.

Table 1: Illustrative policy definitions for performance levels

Assessment	Performance levels		
National Assessment of Educational Progress (NAEP)	<p>Basic: This level denotes partial mastery of prerequisite knowledge and skills that are fundamental for proficient work at each grade.</p>	<p>Proficient: This level denotes solid academic performance for each grade assessed. Students reaching this level have demonstrated competency over challenging subject matter, including subject-matter knowledge, application of such knowledge to real-world situations, and analytical skills appropriate to the subject matter.</p>	<p>Advanced: This level signifies superior performance beyond proficient.</p>

3. Develop full descriptions for the performance levels of the UIS Proficiency Scales

After the policy definitions have been adopted, content descriptions should be added to develop full descriptions of the performance levels. The full descriptions express the knowledge and skills required to achieve each performance level. They can be used to provide stakeholders with more information on what students at each performance level know and are able to do, and what they need to know and be able to do to reach the next performance level.

To develop the full descriptions, a performance-level descriptor (PLD) writing workshop should be conducted with subject matter experts (SMEs) for each domain. Five to eight SMEs per subject and grade span will suffice (Perie, 2008). To write the PLDs, those with an understanding of the policy context should work alongside those with an understanding of teaching and learning the subject-matter content. The SMEs will start with the policy definitions (see Step 2b) and expand those definitions in terms of specific knowledge, skills, and abilities at each level for each domain and for each grade. The PLDs should be very detailed and reflect the content standards defined in Step 1.

Since the PLDs of the UIS proficiency scales will be the basis for linking with national assessments and cross-national assessments, it is essential that they are fully elaborated and include details related to each content standard identified in Step 1.

4. Proposed minimum levels

Proficiency levels are not defined by the imposition of arbitrary cut points on the scale but are defined by points where the underlying cognitive processes shift using the proficiency level and performance-level descriptors available. The experts should propose to the GAML Plenary and to the Technical Cooperation Group on the Indicators for SDG 4 – Education 2030 (TCG) a set of proficiency levels for the regional and global levels from the ones agreed in Step 2.