

# NOT FOR CIRCULATION



UNESCO  
INSTITUTE  
FOR  
STATISTICS



## Methodology for Ordering Performance Level Descriptors

10-11 September 2018  
UNESCO Headquarters  
7, place de Fontenoy,  
75352 PARIS 07 SP, France



The information in this report may not be reproduced, in part or in whole, and by any means, without permission from the Organization. All requests should be submitted to: [uis.LO@unesco.org](mailto:uis.LO@unesco.org)

This paper is presented to describe the methodology utilized in analysing, comparing, simplifying, and ordering the performance level descriptors for various national and multinational assessments against the UIS Proficiency Scale in mathematics

## Background

### Indicator 4.1.1

The UNESCO Institute for Statistics' (UIS) goal as a custodian agency for reporting against the Sustainable Development Goals in Education (SDG4) is to develop standards, methodology and guidelines to enable countries in the production of data for the reporting of indicators Indicator 4.1.1. requires member countries to report on the "proportion of children and young people....to achieve at least a minimum proficiency level in reading and mathematics". In order to define the minimum proficiency for report indicator 4.1.1, the UIS has developed the Global Framework for Mathematics, organized and compiled cross-national assessment performance level descriptors, with the goal of building consensus on the number of performance levels, definition of policy and performance descriptors, as well as of the minimum proficiency levels for each education level.

### List of Assessments

The assessments for which PLD's were analyzed for this project are shown in Table 1, The assessments were grouped into three grade-level bands, or measurement points: 2-3; 4-6; and 8-9.

**Table 1. Assessment Information.**

Assessment Name	Assessment Type	LSA	Year Administered
ASER	National Citizen-Led	Grades 2-3	2017
EGMA	National	Grades 2-3	Not provided
PASEC	Regional	Grades 2-3	2014
TERCE	Regional	Grades 2-3	2014
UNICEF MICS6	Household Survey	Grades 2-3	Not provided
Uwezo	National Citizen-Led	Grades 2-3	Not provided
PASEC	Regional	Grades 4-6	2014
PILNA	Regional	Grades 4-6	2015
SACMEQ	Regional	Grades 4-6	2007
TERCE	Regional	Grades 4-6	2014
TIMSS	Cross-national	Grades 4-6	2015
PISA	Cross-national	Grades 8-9	2015
PISA-D	Cross-national	Grades 8-9	Not provided
TIMSS	Cross-national	Grades 8-9	2015

## Performance Level Descriptors

### Definition

Each assessment in Table 1 has a number of performance level descriptors (PLD's) associated with it. These PLD's delineate one or more mathematical skills and/or processes that are associated with test takers who achieve that performance level. The number of PLD's varies by assessment, as does the format in which the PLD's are written. Examples of mathematical skills include counting, adding fractions, solving equations; examples of mathematical processes include employing basic formulas, interpreting problem situations, and communicating reasoning.

### **Analysis, comparison, and ordering**

The primary, if not sole, criterion for analysing PLD's is the *cognitive demand* required by the mathematical skills and/or processes contained in each PLD. This is complicated by the fact that most, if not all, PLD's contain multiple skills and processes. Thus, comparing PLD's becomes a matter of determining and comparing the *overall* cognitive demand of each PLD. This requires a high level of careful analysis, and is as much art as science. Successively comparing PLD's against each other eventually resulted in a list of the PLD's within each measurement point, arranged from lowest to highest overall cognitive demand. As an additional point of information, each PLD was given a one-sentence summary, which may facilitate easier comparison for future work.

### **Proficiency Scale**

Once the list of PLD's for each measurement point was completed, it was then necessary, and possible, to create the overall Proficiency Scale for mathematics. This was begun by placing *all* the PLD's from *all* three measurement points into a single list, from the lowest of grades 2-3 to the highest of grades 8-9. However, it could not be assumed that the highest-level PLD of one measurement point had a lower cognitive demand than the lowest-level PLD of the next-highest measurement point. The next step was therefore to compare the high-level PLD's of grades 2-3 against the low-level PLD's of grades 4-6, utilising the same process of comparing the overall cognitive demand of the PLD's, and re-arranging PLD's as appropriate. This was then repeated with the PLD's at the border of grades 4-6 and grades 8-9. This resulted in a list of *all* PLD's across all three measurement points.

### **Ordering within measurement points**

The final step after creating the Proficiency Scale was to identify which PLD's contained grade-level appropriate (GLA) skills and processes for each measurement point. For this step, cognitive demand was *not* a criterion, as each measurement point contains a range of skills from low to high cognitive demand. The Proficiency Scale includes a number of PLD's that did not contain GLA skills or processes even at the lowest measurement point. It also included many PLD's that were GLA at more than one measurement point.

Once the Proficiency Scale was complete, it was then possible to set the performance levels at each measurement point, using the list of GLA PLD's. Each measurement point used the same four performance levels—Below Basic; Basic; Proficient; and Advanced. As with the first step in the process, determining where to set each performance level required a good deal of careful analysis, especially since the skills and processes taught at each grade can vary, in some cases widely, from nation to nation. Finally, at each measurement point, the lowest PLD in the Proficient performance level was marked as the dividing line between proficient less than proficient test takers. See Figure 1 for an excerpt of the Proficiency Scale.

Figure 1. UIS Proficiency Scale (excerpt).

Assessment Name	Assessment Level	Domain	Performance Level	Descriptor	One line descriptor	Order	GR 2-3	GR 4-6	GR 8-9
SACMEQ 2007 (grade 6)	Regional	Mathematics	Level 1	<u>Pre Numeracy</u> Applies single step addition or subtraction operations. Recognizes simple shapes. Matches numbers and pictures. Count in whole numbers.	Students at this level can perform only the simplest of computations and can recognize simple shapes.	42	X		
PILNA 2015 (grades 4/6)	Regional	Mathematics	Level 2 425-449	<ul style="list-style-type: none"> <li>Write a three-digit number not involving zero in words and in numerals, and write a three-digit number involving zero in words only.</li> <li>Compare prices of items and calculate the total cost of two items.</li> <li>Subtract a two-digit number from a two- or three-digit number without regrouping and solve simple word problems involving addition.</li> <li>Identify hands of a clock and know the relation of days and weeks.</li> </ul>	Students at this level can write, compare and compute with whole numbers.	43	X		
TERCE 2014 (grade 3)	Regional	Mathematics	Level 4 843+	Students can: <ul style="list-style-type: none"> <li>Solve more complex problems in the area of natural numbers.</li> <li>Solve problems involving comparison and conversion of measures</li> <li>Solve more complex problems involving elements of geometric figures or flat representation of geometric shapes.</li> </ul>	Students at this level are able to utilize reasoning skills to solve complex mathematical and real-world problems.	44	X	X	

### Mapping

Once the PLD's were placed in order, the final task was to create a graphical display, or mapping, of each assessment's PLD's against the performance levels at each measurement point, as well as an overall mapping of all assessments. This overall mapping is not compared against performance levels, but is mapped against the grade-level progression, in order to show where the individual PLD's for each assessment lie in comparison to each other. It should be noted that those PLD's that were considered to be below the minimum level for the grades 2-3 measurement point were *not* included in the mapping for that measurement point, or for the overall mapping.

As is typical of assessments, each performance level represents a range of abilities on the part of test takers. This range is usually represented by scale scores, which are provided for most of the assessments in this project. However, each assessment uses a different scale, so a direct comparison between scale scores is not possible. Because the performance levels were set without the benefit of scale scores, a decision was made to map the space for the performance levels proportionally to the ordered placement of the PLD's at each measurement point. For example, at grades 2-3, there are 3 spaces separating TERCE Level 1 and Level 2. Thus, the TERCE Level 2 bar takes up 3 columns in the mapping.

The final step in creating the overall mapping was to "fill in the blanks" that existed between performance levels within an assessment when the mappings for all three measurement points were placed onto the overall mapping. For instance, for PASEC grade 6, the bar for "Below Level 1" goes part way across grades 2-3, while "Level 1" begins in grades 4-6. In order to "fill in the gap" on the overall mapping, the "Level 1" bar was extended *backwards* until it "met" the "Below Level 1" bar. This was done as a way of indicating that test takers can, and most likely will, achieve different levels of achievement across the grade-level continuum.

### Policy Level Descriptors

Previously, policy level descriptors in the area of mathematics were developed to characterize (in general terms) the difference in ability between mathematically proficient test takers and those who achieve at a level below proficiency. These policy level descriptors reflect the dividing line between proficient and non-proficient test takers, even though they

do not delineate between the two sub-categories at each level: Below Basic vs Basic, and Proficient vs Advanced. The policy level descriptors are an exceedingly useful and important tool that can be used to validate that the content described at each measurement point is an accurate reflection of the mathematical skills and processes for which students around the world should be expected to demonstrate a certain degree of mastery.

**Figure 2. Policy Level Descriptors for Mathematics.**

<b>Performance Level Policy Descriptors</b>	
Proficient/ Above Proficiency	Students at this level possess a basic, or better, level of mathematical knowledge. They also demonstrate a basic, or better, level of competency with mathematical skills and abilities. These includes the recall of mathematical facts, formulas, and algorithms, the ability to solve application problems, and varying levels of aptitude in using problem-solving strategies and communicating mathematically.
Below Proficiency	Students at this level possess a limited level of mathematical knowledge and demonstrate a lack of competency with most mathematical skills and abilities. They tend to struggle with all but the most routine and straightforward aspects of mathematics.

NOT FOR CIRCULATION