



United Nations
Educational, Scientific and
Cultural Organization



UNESCO
INSTITUTE
FOR
STATISTICS



GLOBAL
ALLIANCE
TO MONITOR
LEARNING



Benchmarking for learning

Group discussion

GAML6/REF/13



Indicator 4.1.1: Minimum Proficiency Levels in Reading and Math

The objective of this session is to develop a consensus on some pending issues related to indicator 4.1.1, which reports on the percentage of students achieving the minimum proficiency levels of reading and math in a country.

1. Indicator Definition and Pending Issues

4.1	Target 4.1: By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes
4.1.1	Proportion of children and young people (a) in 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

Indicator 4.1.1 is one of the most important indicators measured and reported by SDG 4, since it addresses the main objective of any education system: Student learning. In terms of measurement and reporting there is a need to reach a consensus.

There are several issues still pending about this indicator, but these issues have to be resolved at the country level.¹ These issues include the following:

- **Coverage:** around 80 countries do not have comparable data on learning outcomes for any breakdown of indicator 4.1.1. Specifically, at the end of primary school, data are not available for around 100 countries, while at the grade 2 or 3, only a few countries have comparable data.
- **Comparability:** various assessments produce data which allows comparison between different countries taking the same assessment. However, data from countries using different assessments cannot easily be compared. There are seven major regional assessments and, while many of them test broadly the similar learning items, no robust framework exists at the moment to compare the data they produce. Similarly, some countries use national assessments to monitor progress, however, in most cases, the data from these assessments cannot currently be compared.
- **Frequency:** the time gaps between two consecutive administrations of different assessments can be up to six years or more. For instance, six years passed between the LLECE assessments in 2013 and 2019, and a larger gap is expected for SACMEQ.
- **Development and maintenance of a country's capacity to undertake, analyse and report proficiency results.** In many low income countries there are problems with existing capacities for undertaking proficiency assessments on a regular basis. In some cases, regular assessment of proficiency at the end of ISCED 0 (Preschool), ISCED 1 (6 years of Primary), and ISCED 2 (lower secondary) is crowded out by the measurement of large international assessments, such as PISA. Hence, countries would have to examine their financial and human resources in order to determine their measurement and reporting capacity.

¹ Gustafsson, Martin, 2019. Costs and Benefits of Different Approaches to Measuring the Learning Proficiency of Students (SDG Indicator 4.1.1). Montreal: UNESCO Institute for Statistics.

1.1 What the data show

Table 1 shows the latest results for those countries reporting indicator 4.1.1.

Table 1. SDG 4, Target 4.1 – By 2030, ensure that all girls and boys complete free, equitable and quality primary and secondary education leading to relevant and effective learning outcomes

Indicator 4.1.1 Percent of students Achieving minimum proficiency in Reading and Math (baseline)						
Region	Early Grades		End of Primary		End of Lower Secondary	
	Reading	Math	Reading	Math	Reading	Math
Sub-Saharan Africa						
Average (for reporting countries)	38.4	32.3	34.8	18.5	32.4	26.0
Percent countries reporting	49.0	47.1	29.4	33.3	7.8	9.8
Population coverage	74.0	73.5	41.2	46.9	8.8	11.6
Northern Africa and Western Asia						
Average (for reporting countries)	42.4	27.9	56.7	48.7	47.6	32.5
Percent Reporting	42.3	46.2	3.8	11.5	34.6	69.2
Population coverage	42.6	29.0	1.9	18.6	34.5	74.9
Central and Southern Asia						
Average (for reporting countries)	30.9	27.3	49.8	52.2	33.2	27.7
Percent Reporting	50.0	50.0	42.9	42.9	28.6	42.9
Population coverage	95.1	95.1	90.8	90.8	19.8	24.9
Eastern and South-eastern Asia						
Average (for reporting countries)	72.2	62.5	72.7	70.0	68.1	67.4
Percent Reporting	50.0	55.6	16.7	11.1	61.1	61.1
Population coverage	77.1	87.8	6.3	5.6	91.3	91.3
Oceania						
Average (for reporting countries)	69.9	51.2	75.3	56.2	73.7	68.5
Percent Reporting	12.0	8.0	4.0	4.0	8.0	8.0
Population coverage	72.5	71.8	60.2	60.2	71.8	71.8
Latin America and the Caribbean						
Average (for reporting countries)	68.3	58.5	56.1	49.6	52.2	35.1
Percent Reporting	46.9	44.9	34.7	34.7	28.6	28.6
Population coverage	87.7	87.4	54.6	54.6	85.5	85.5
Europe and Northern America						
Average (for reporting countries)	89.6	68.6	82.7	76.4	78.9	73.4
Percent Reporting	33.3	29.8	19.3	24.6	64.9	66.7
Population coverage	60.5	60.1	30.5	32.7	94.0	98.0

Source: UNESCO Institute for Statistic (UIS) database

Reporting achievement of minimum proficiency levels for early grades is more prominent in Latin America and the Caribbean and, to some extent, Sub-Saharan Africa, two regions that have embraced Early Grade Reading Assessments (EGRA) and Early Grade Mathematics Assessment (EGMA) more than other regions of the world. In many high income countries, and in some large countries with decentralized administrative structure like Indonesia, and Nigeria, there is little reporting on early grades assessment. The lack of data from these countries have an impact on regional averages and, by implication, on the regional thresholds for minimum proficiency levels.

Measuring and reporting minimum proficiency levels at the end of primary is prominent in Latin America and the Caribbean, where almost 60% of the countries in the region publish assessment results.

In the case of minimum standards for lower secondary, 80% of countries in Europe and Northern America report this indicator, followed by more than half of the countries in Eastern and South-eastern Asia. Almost 40% of Latin America and the Caribbean countries also report results for lower secondary, while the majority of countries in other regions of the world fail to report this indicator.

The above results show a wide variation in the number of countries reporting across the three levels of assessment. This variation confirms that assessment instruments are yet to be consistently or uniformly applied across countries. For now, some countries emphasize early grade and primary, while others emphasize lower secondary.

2. Practical Issues 1: definition of Minimum Proficiency Levels (MPL)

One of the most important issues in the definition of the scales is the proficiency benchmarks or levels embedded within the numerical scale and their cut points on that numerical scale. These benchmarks are typically associated with Proficiency Level Descriptors, which describe in some detail the skills that are typical of students at any given cut point in the scale. Typically, an overarching policy statement or policy definition gives meaning to the succession of cut scores and the proficiency levels but most importantly for defining what constitutes a minimum. In the case of indicator 4.1.1, an agreement has been reached in 2017 about the definition of the MPL and contents that define the minimum proficiency level in each point. That definition “operationalizes” descriptors that can be used to drive comparisons (e.g., that can be used in policy linking).

Table 2. Minimum Proficiency Levels for Mathematics

Educational Level	Descriptor	Assessment PLD's that align with the descriptor	MPL's in the Assessments
Grades 2-3	Students demonstrate skills in number sense and computation, shape recognition and spatial orientation.	<input type="checkbox"/> PASEC 2014 – Level 2 <input type="checkbox"/> TERCE 2014 – Level 2 <input type="checkbox"/> SERCE 2006 – Level 2	Level 2 Level 2
Grades 4-6	Students demonstrate skills in number sense and computation, basic measurement, reading, interpreting, and constructing graphs, spatial orientation, and number patterns.	<input type="checkbox"/> PASEC 2014 – Level 3 <input type="checkbox"/> SACMEQ 2007 – Level 5 <input type="checkbox"/> PILNA 2015 – Level 6 <input type="checkbox"/> SERCE (6to grade) - Level 3 <input type="checkbox"/> TERCE 2014 (6to grade) – Level 3 <input type="checkbox"/> TIMSS 2015 (4 th grade)– Intermediate International	Level 2 Level 3 Level 5 Level 2 Intermediate International
Grades 8 & 9	Students demonstrate skills in computation, application problems, matching tables and graphs, and making use of algebraic representations.	<input type="checkbox"/> PISA 2012/2015/PISA4D – Level 2 <input type="checkbox"/> TIMSS 2015 (8 th grade) – Intermediate International	Level 2 Intermediate International

Note: alignment for the Math component of some assessment is pending upon analysis of items and test.



Table 3. Minimum Proficiency Levels for Reading

Minimum Proficiency Levels for Reading

Educational Level	Descriptor	Assessment PLDs that align with the descriptor	MPL in the assessment, if available
Grade 2	They read and comprehend most of written words, particularly familiar ones, and extract explicit information from sentences.	<input type="checkbox"/> PASEC (Gr. 2) – Level 3	<input type="checkbox"/> Level 3
Grade 3	Students read aloud written words accurately and fluently. They understand the overall meaning of sentences and short texts. Students identify the texts' topic.	<input type="checkbox"/> Uwezo – Std. 2 (Story with meaning)	<input type="checkbox"/> Std. 2 (Story with meaning)
		<input type="checkbox"/> PASEC 2014 (Gr. 2) – Level 4	<input type="checkbox"/> Level 3
		<input type="checkbox"/> SERCE (Gr.3) – Level 2	<input type="checkbox"/> Level 1 (appears that way from Technical reports)
		<input type="checkbox"/> TERCE (Gr. 3) – Level 2	<input type="checkbox"/> Level 2
		<input type="checkbox"/> UNICEF MICS 6 – -Proficient Level	"demonstrated foundational reading skills"
		<input type="checkbox"/> EGRA – Level 9	<input type="checkbox"/> Not specified
Grades 4 & 6	Students interpret and give some explanations about the main and secondary ideas in different types of texts. They establish connections between main ideas on a text and their personal experiences as well as general knowledge	<input type="checkbox"/> ASER – Std. 2 (story)	<input type="checkbox"/> Std. 2 (story)
		<input type="checkbox"/> SACMEQ 2007 – Level 5	<input type="checkbox"/> Level 3
		<input type="checkbox"/> PASEC 2014 (Gr. 6) – Level 4	<input type="checkbox"/> Level 3
		<input type="checkbox"/> PIRLS 2011 and 2016 – Low	<input type="checkbox"/> Low
		<input type="checkbox"/> PILNA 2015 (Gr. 4 & 6) Level 5	<input type="checkbox"/> • Level 4 (grade 4) and Level 5 (grade 5)
		<input type="checkbox"/> SERCE 2006 (Gr. 6) – Level 3 <input type="checkbox"/> TERCE 2014 (Gr. 6) – Level 3	<input type="checkbox"/> Level 2
Grades 8 & 9	Students establish connections between main ideas on different text types and the author's intentions. They reflect and draw conclusions based on the text.	<input type="checkbox"/> PISA 2011/2015/PISA4D – Level 2	<input type="checkbox"/> Level 2

3. Practical Issues 2: aligning to a global scale

There are numerous ways and different contexts in which reading and mathematics are measured at the national level. There is a basic distinction between assessments that are informal, formative, short, or designed by teachers, inspectors and district authorities, versus formal, typically summative, longer assessments. These distinctions are important for educators because implementing short, formative assessments to monitor progress can lead to the development of more complete summative assessments.

Table 4. Comparability of Learning Assessments

		<i>Comparability</i>	<i>Needs of action</i>
School-based			
	Global	Yes, for participating countries	None
	Regional	Yes, for the countries	Translate into a global scale
	National	No but could be scaled	Translate into a global scale
Population Based		Yes but alignment of verbal definition should be completed	Translate into a global scale
National Examination		No	-----

There are two main types of linking: statistical and non-statistical. Statistical linking is more accurate, but it has greater requirements, i.e., common students either taking different assessments or having common items across assessments. Non-psychometric linking based on policy descriptors is less accurate, but acceptable when requirements of statistical linking are not met due to issues such as design, logistics, or cost with some potential benefits.

Table 5. Summary table of alternatives for aligning to a global scale

<i>Level</i>	<i>Non psychometric calibration</i>	<i>Psychometric Calibration</i>	
		<i>Test Based Calibration</i>	<i>Items based linking</i>
Early Grades	Yes	Unlikely	Unlikely
End of Primary	Yes	Yes	Feasible
End of lower secondary	Yes	No	Feasible

In all of this, the UIS vision, at least initially, is to try carry out these tasks using a portfolio approach that adjusts over time. For instance: non-psychometric calibration might be less accurate, but might yield results faster and at much lower cost. Plus, much of what is learned, and the raw materials used for non-psychometric calibration (e.g. sample items) can be re-used for item-based linking. And no one has firmly established how much rigor can be gained by using a test-based calibration for example, and at what cost. It could be that for certain combinations of grade level and subject, non-psychometric costs 1/10th as much as some other methods, but is 80% as accurate. These factors are still unknown. Thus the rationality of a portfolio approach that can vary over time but wastes no resources in that the resources for one method can be re-purposed or simply re-used for other methods.

4. Practical Issues 3: the Out-of-School Children and the adjustment to indicator 4.1.1

In 2016, 263 million children, adolescents and youth were out of school, representing nearly one-fifth of the global population of this age group. 63 million, or 24% of the total, are children of primary school age (typically 6 to 11 years old); 61 million, or 23% of the total, are adolescents of lower secondary school age (typically 12 to 14 years old); and 139 million, or 53% of the total, are youth of upper secondary school age (about 15 to 17 years old). Not all these kids will be permanently outside school, some will re-join the educational system and, eventually, complete late, while some of them will enter late. The quantity varies per country and region and demands some adjustment in the estimate of indicator 4.1.1.

Table 6. Rates of out-of-school by SDG regions and levels of education, both sexes, 2017 or latest year available

Region	Rate of out-of-school (%)		
	Primary	Lower secondary	Upper secondary
Sub-Saharan Africa	20.51	35.51	56.96
Northern Africa and Western Asia	10.58	14.22	32.17
Northern Africa	10.65	10.14	33.33
Western Asia	10.30	17.24	31.23
Central and Southern Asia	6.21	16.62	47.25
Central Asia	2.51	5.30	18.07
Southern Asia	6.32	17.22	47.83
Eastern and South-eastern Asia	3.78	8.55	19.27
Eastern Asia	3.04	6.55	15.91
South-eastern Asia	5.04	11.57	25.75
Oceania	9.21	3.66	22.81
Latin America and the Caribbean	4.72	7.15	22.79
Caribbean
Central America
South America
Europe and Northern America	2.94	1.78	6.54
Europe	2.37	2.25	7.43
Northern America	3.76	0.74	5.03
World total	8.88	15.65	35.82

Source: UNESCO Institute for Statistic (UIS) database.

Note: '...': data not available

To the light of these numbers, we proposed some adjustments in the current expression of the indicator to reflect all the population and the target that involves access, completion and learning. In line with UIS (2017)², it was assumed that children not in school would not have reached the minimum level of proficiency. The

² UIS (2017a). *More than one-half of children and adolescents are not learning worldwide*. Montreal and UIS (2017b). *Counting the number of children not learning: Methodology for a global composite indicator for education*. Montreal.



following equation could be applied to arrive at a percentage of the lower primary-aged population not being proficient:

$$\text{indicator 4.1.1 lower primary} = S \times (1 - \text{Out of school rate}) \quad (1)$$

The percentage proficient among enrolled students is multiplied by one minus the percentage of the population not in school.

However, not all kids currently out-of-school are going to be permanently out-of-school and in the same way, not all the kids that are currently in school would necessary complete the levels which is relevant for the end of primary and the end of lower secondary. In terms of estimation of indicator 4.1.1 and to better understand school exposure and its implications on measuring indicator 4.1.1, a basic classification would divide children and youth into two main groups, those who complete the level and those who do not complete it, where each group is further composed by sub groups:

1. Completers
 - Those who are in school and who complete their respective level of education;
 - Those who will start school late (now out-of-school) and are expected to complete the last grade;
2. Non Completers
 - Those who are in school but would drop out before reaching the last grade;
 - Those who will start school late and would drop out in the future;
 - Those who were in school but dropped out; and
 - Those who were never in school and would never enter.

Therefore, the soundest option is to use the rate of completion by level and make the adjustment on indicator 4.1.1b and 4.1.1c as follows using again the proficiency among enrolled students and adjust by the non-completion rate using the following formula

$$\text{indicator 4.1.1 at the end of level} = S \times (\text{completion rate of level } i) \quad (2)$$

producing the following adjusted Proficiency levels for Population:

Table 7. Indicator 4.1.1 adjusted

<i>Indicator 4.1.1</i>	<i>Indicator 4.1.1 adjusted</i>
Early Grades	$4.1.1a = S \times (1 - \text{rate of out - of - school})$
End of Primary	$4.1.1b = S \times \text{rate of completion}$
End of Lower Secondary	$4.1.1c = S \times \text{rate of completion}$

5. Benchmarks: proposed options

Below are different options for proposed minimum **targets for the percentage of students meeting the minimum proficiency levels in reading and math**. The three options proposed in this document are in line with the strategic objectives of the SDG 4 initiative for education. They are based on reported data reported by countries and collected by UIS.

5.1 Option 1: Simplified Minimum Regional Targets

This option eliminates the distinction between reading and math, which facilitates monitoring, as shown in Table 8.

Table 8. Indicator 4.1.1 Proposed Targets for the Percent of Students Achieving minimum Proficiency in Reading and Math, by SDG Regions

Region	Early Grades	End of Primary	End of Lower Secondary
Sub-Saharan Africa	50	60	75
Northern Africa and Western Asia			
Central and Southern Asia	40	50	70
East and South-eastern Asia	90	70	75
Oceania	95	95	95
Latin America and the Caribbean	80	80	80
Europe and Northern America	99	99	99

5.2 Option 2: Regional Minimum Targets by Economic Vulnerability

Two key empirical findings in education are relevant for this proposal. The first one is the close relationship between economic development and education performance (Hanushek and Woessman 2008)³, which in turns relates student learning to poverty, and the relationship between vulnerable ethnic groups and poverty, (Hall and Patrinos 2010)⁴ which in turns leads to lower educational performance. Table 9 proposes two targets, one for the national average, and one for students living in extreme poverty. For practical purposes, extreme poverty is defined as per capita income in the bottom 20% of the income distribution. Poverty is used as a proxy for human vulnerability, which can come through ethnicity, religious affiliation, gender, and other societal markers where vulnerable groups tend to be marginalized and, as a result, left to suffer discrimination in education.

³ Hanushek, Eric A., and Luther Woessmann, 2008. The Role of Cognitive Skills in Economic Development. Journal of Economic Literature, Vo. 46, No. 3, pp. 607-668.

<http://hanushek.stanford.edu/sites/default/files/publications/Hanushek%20Woessmann%202008%20JEL%2046%283%29.pdf>

⁴ Hall, Gillete, and Harry A. Patrinos, 2010. *Indigenous People, Poverty, and Development*. Washington DC: World Bank. http://siteresources.worldbank.org/EXTINDPEOPLE/Resources/407801-1271860301656/full_report.pdf



Table 9. Indicator 4.1.1 Proposed Targets for the Percent of Students Achieving minimum Proficiency in Reading and Math, by SDG Regions

Region	Early Grades	End of Primary	End of Lower Secondary
Sub-Saharan Africa	50	60	75
Northern Africa and Western Asia	50	60	75
Central and Southern Asia	40	50	70
Eastern and South-eastern Asia	90	70	75
Oceania	95	95	95
Latin America and the Caribbean	80	80	80
Europe and Northern America	99	99	99
Target Percent of Students in Extreme Poverty Achieving the Minimum Proficiency Levels for Reading and Math			
Region	Early Grades	End of Primary	End of Lower Secondary
Sub-Saharan Africa	35	42	53
Northern Africa and Western Asia	35	42	53
Central and Southern Asia	28	35	49
East and South-eastern Asia	63	49	53
Oceania	67	67	67
Latin America and the Caribbean	56	56	56
Europe and Northern America	69	69	69

6. Issues for discussion

Given these three options the discussion should answer some key questions:

- I. How can non-reporting countries be enticed to report data on indicator 4.1.1? Is the threat of reduced external aid to education a good option? Who should work with countries at the individual country level to improve capacity and ensure data quality?
- II. How high should be a target before it becomes a fantasy? Can targets be revised for each country or should regional targets be enough to monitor country performance?
- III. Which of the three options better addresses the essence of Indicator 4.1.1 while motivating countries to act, and to report results?
- IV. How should Out-of-school and Completion be factored-in indicator 4.1.1? Should the indicator being published in the in-school and population based versions?