

SDG Goal 4.1, Indicator 4.1.1

Reading and Mathematics Assessment Blueprint

**COVID-19: Monitoring Impacts on Learning Outcomes
(MILO)**

VI.0, 31 March 2021

The Global Education Monitoring (GEM) Centre supports education stakeholders to collect, analyse and use high-quality data to improve learning outcomes. The GEM Centre is a long-term partnership between the Australian Council for Educational Research (ACER) and the Australian Government's Department of Foreign Affairs and Trade (DFAT).



**Transforming
Education**



Acknowledgements

This project, the Assessment and Study of COVID Impact on Learner Progress, is referred to as the COVID-19 MILO (Monitoring Impacts on Learning Outcomes) project. This UNESCO Institute for Statistics (UIS) project is funded by the Global Partnership for Education (GPE).

The Australian Council for Educational Research (ACER) is the technical partner for this project. Support is provided from the Global Education Monitoring (GEM) Centre, an ACER initiative in partnership with the Australian government's Department of Foreign Affairs and Trade. The GEM Centre is also contributing to the UIS Global Item Bank. Technical and implementation support, and contribution to the assessment item pool, is provided by CONFEMEN.

Contents

Introduction	1
Learning areas and target population.....	1
Domain and construct targets	3
Item types	5
Item sources and criteria for selection	5
Test booklet design	7
Appendix.....	8

Table of Tables

Table 1: Target population for assessment	2
Table 2: Classification and targets for reading assessment items	4
Table 3: Classification and targets for mathematics assessment items	5
Table 4: The results of the review process for reading	6
Table 5: The results of the review process for mathematics	7
Table 6: MILO Assessment: test design	7
Table 7: GPF reading domains, constructs and sub-constructs	8
Table 8: GPF mathematics constructs and sub-constructs	9

Introduction

The COVID-19: Monitoring the Impacts on Learning Outcomes (MILO) project aims to measure learning outcomes in six countries in Africa, in order to analyse the long-term impact of COVID-19 on learning and to evaluate the effectiveness of distance learning mechanisms utilised during school closures. In addition, this project will develop the capacity of countries to monitor learning after the crisis.

The four overarching goals of the project are to:

- Evaluate the impact of COVID-19 on learning outcomes and measure the learning loss by reporting against SDG indicator 4.1.1b
- Identify the impact of different distance learning mechanisms put in place to remediate the learning disruption generated by COVID-19
- Expand the UIS bank of items for primary education
- Generate a toolkit so that assessment results can be scaled to international benchmarks, reporting against SDG 4.1.1.b.

This document provides a detailed assessment blueprint for the COVID-19 MILO assessment instruments. A separate framework is provided for the contextual questionnaires. This assessment blueprint describes the learning areas and target population for the MILO project, the domain and constructs assessed, the item types included, the sources of items and a description of how they were selected, and the final MILO test booklet design.

Learning areas and target population

The two learning areas being assessed in the MILO project are Reading and Mathematics. In line with the Global Proficiency Frameworks (GPFs), Reading and Mathematics are referred to as “learning areas”, which are then broken down into domains, constructs, and sub-constructs.

The focus of the MILO project is on students at the end of primary school. However, as outlined in the MILO Study Design, the definition of the end of primary schooling differs across systems and countries. In reporting against SDG indicator 4.1.1(b) the UIS also allows the flexibility for countries to report at grade levels close to the end of primary schooling if the quality and appropriateness of the outcomes data is more suitable for purpose.

The benchmark that will be used to indicate learning outcomes in this study is aligned with Sustainable Development Goal (SDG) indicator 4.1.1(b):

*the proportion of children and young learners ... at the end of primary ... achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex.*³

The Minimum Proficiency Level in reading for end of primary schooling is described in as:

Students independently and fluently read simple, short narrative and expository texts. They retrieve explicitly-stated information. They interpret and give some explanation about the main and secondary ideas in different types of texts, and establish connections between main ideas in a text and their personal experiences. (ACER, 2020, p6)

The Minimum Proficiency Level in mathematics for end of primary schooling is described as:

Students recognise, read, write, order, compare and calculate with whole numbers, simple fractions and decimals. Students can measure length and weight using standard units, calculate the perimeter of simple 2D shapes and area of rectangles. They read, interpret and construct different types of data displays such as tables, column graphs and pictographs and recognise, describe and extend number patterns. They can solve simple application problems. (ACER, 2020, p4)

The assessment will be administered to children in six countries in Africa as shown in Table 1.

Table 1: Target population for assessment

Participating Country	Language of assessment	Target grade
Burkina Faso	French	G6 (CM2)
Burundi	French	G6
Cote d'Ivoire	French	5/6*
Senegal	French	G6 (CM2)
Kenya	English	G7
Zambia	English	G5

*Grade to be confirmed

Items selected for inclusion in the MILO test booklets include a range of items around the end of primary minimum proficiency level, as well as some items below the end of primary minimum proficiency level. This range of difficulty levels is in order to be able to gain some understanding of children's learning outcomes within the target population who have not yet attained the end of primary minimum proficiency.

Domain and construct targets

As defined by the GPF, the Reading learning area comprises three domains: “Comprehension of spoken or signed language”, “Decoding” and “Reading Comprehension” (see Table 7). The first of these is discussed in the GPF only in relation to Grades 1-3, so items focussing on this domain are not included in the assessment. Items focussing on decoding are similarly not the emphasis of this assessment, partly because although this domain is outlined in the GPF for grades 1-9, these skills are most relevant in the early years of school. In addition, these skills are elicited most easily in assessments that are administered one to one rather than to a classroom, as will be the case for the COVID-MILO assessment. The emphasis of this assessment, therefore, is on the third domain, “Reading Comprehension”. The domain of “Reading Comprehension” is further broken down into three constructs: R1- Retrieve information, R2 – Interpret information and R3- Reflect on information.

The targets for each of the three constructs within the domain of “Reading Comprehension” are as follows:

- Retrieve information: 35-45%
- Interpret information: 45-55%
- Reflect on information: 15-25%

These targets were developed with reference to existing large-scale and regional assessments, and the work of the GPF alignment group. In relation to the former, this breakdown is analogous to that used in the large-scale international assessment PIRLS¹, in which the equivalent breakdown is 20%, 60% and 20%. The slightly greater emphasis on items relating to retrieving information is explained by the likely greater familiarity of this style of item for the target population. The Draft Alignment Criteria developed for reviewing assessments’ suitability for reporting against SDG 4.1.1 specify that in order to be considered “strongly aligned” with the GPF for Grades 4-9, at least 5 items for each of the three constructs should be included. The proposed targets for the MILO assessment allow this specification to be met. Table 2 shows the classification of the items chosen for the assessment against the specified targets, revealing that all targets were met by the selection.

¹ The IEA Progress in International Reading Literacy Study

Table 2: Classification and targets for reading assessment items

Construct	Number of items	Percentage of items	Target percentage
Retrieve information	11	35%	35-45%
Interpret information	14	45%	45-55%
Reflect on information	6	19%	15-25%

The mathematics assessment comprises five domains and targets for the mathematics assessment have been set at the domain level. These targets are:

- Number and operations 35-45%
- Measurement 15-20%%
- Geometry 15-20%%
- Statistics and Probability 5-10%
- Algebra 5-10%

The mathematics targets were also developed with reference to existing large-scale and regional assessments, and the work of the GPF alignment group. The content breakdown is analogous to that used in the large-scale international assessment TIMSS2 at Grade 4, in which the equivalent breakdown is 50% Number, 30% Measurement and Geometry and 20% Data. The slightly variations between the TIMSS targets and the MILO targets is explained by the fact that the Algebra domain of the GPF at end of primary is content that would have been incorporated into the Number domain within the TIMSS Framework (in the topic area called Expressions, simple equations, and relationships).

The GPF advisory group on alignment specified that in order to be considered “strongly aligned” with the GPF the following targets need to be met:

- At least 5 items from the Number and operations domain
- At least 5 items from the Measurement and Geometry domains
- At least 5 items from the Statistics and Probability and Algebra domains
- A total of 50% of all the sub-constructs in the mathematics GPF that are relevant to the target grade level. For example, if there are 20 sub-constructs at Grade 5, at least 10 of the sub-constructs should be included in the assessment.

² The IEA Trends in Mathematics and Science Study

The proposed targets for the MILO mathematics assessment allow this specification to be met. Table 3 shows the classification of the items chosen for the assessment against the specified targets, revealing that all targets were met by the selection.

Table 3: Classification and targets for mathematics assessment items

Domain	Number of items	Percentage of items	Target percentage
Number and operations	11	37%	35–45%
Measurement	5	17%	15–20%
Geometry	6	20%	15–20%
Statistics and Probability	4	13%	10–15%
Algebra	4	13%	10–15%

Additionally, out of the 22 relevant sub-constructs in the GPF at end of primary, a total of 16 different sub-constructs have been used in the mathematics assessment, well above the target of 50% set by the GPF advisory group on alignment. A detailed list of all the GPF constructs and sub-constructs is provided in Table 8 in the Appendices. The sub-constructs selected for the mathematics assessment are highlighted in Table 8.

Item types

All items for the MILO project were paper-based rather than digital. Only closed item types – those which require no expert judgement to score – were included. The closed item types included were multiple choice, in which students select the correct answer from a set of options, and complex multiple-choice, in which students select the correct option for each of a series of statements (eg a select true or false for a number of different statements). Restricting the assessment to closed item types allows for efficient administration and data entry, since there is no need to train coders to score the items.

Item sources and criteria for selection

All items selected for inclusion in the COVID MILO test booklets were from the UIS Global Item Bank. In the first stage of item selection, a subset of items from the Global Item Bank were identified for further review. The criteria used to select items for review were that:

- i) the items were suitable for students working at the level of the upper primary MPL or below
- ii) the items were multiple-choice (or another closed item format)

iii) the items did not use a sentence fragment as the item stem (since this format can be difficult to translate)

iv) the items originated in either English or French, and

v) (for reading) the item or stimulus did not rely heavily on language-specific features that would not translate well (eg, a poem based on rhyming).

For reading, this resulted in a set of 206 items in English and 31 items in French proceeding to the next stage. (Note that in the tables below, because the pieces of stimulus to which reading items related were also rated, for English 265 ratings were given, and for French, 40). For mathematics, a set of 228 items in English and 15 items in French met the criteria to proceed to the next stage.

In the third and final stage of item selection, test booklets were constructed which drew on the results of the review process, and met the targets according to construct. Table 4 shows the results of this process for reading. Of the stimulus pieces and items chosen, the average rating out of 5 from stage 2 of the review was 4.2 for both the English material and the French material. For the stimulus pieces and items not chosen, the average rating was 3.4 for the English material, and 3.3 for the French material. The items selected for inclusion were contributed to the Global Item Bank by PASEC, the Gambia, Dominica, Brunei, Ghana, ACER, Jamaica, Canada (Ontario) and Hong Kong.

Table 4: The results of the review process for reading

	Number of ratings given (stimulus and items)	Number of ratings that met the quality criteria (stimulus and items)	Number chosen for inclusion
English	265	173	35
French	40	21	8

The final stage of item selection for mathematics took a similar approach to reading. However, for mathematics at this stage, an additional criterion for selection was that, for ease of administration, any variation in item type or format should not necessitate a large number of practice questions. This meant that 68 of the reviewed items that did not follow a conventional multiple choice or complex multiple choice format were excluded from further consideration (for example, items requiring students to draw lines to match objects or numbers in one set to objects or numbers in another set). Table 5 shows the results of this process for mathematics.

The average rating of the mathematics items chosen for inclusion in the test booklet was 4.2 out of 5 for the English material and 4.3 out of 5 for the French material. For the material not chosen, the average rating was 4.1 out of 5 for the English material (noting that this includes the 68 items with an inadmissible item type) and 2.9 out of 5 for the French material. Items

selected for inclusion were contributed to the Global Item Bank by PASEC, Antigua and Barbuda, St Vincent and Grenadines, the Gambia, Zambia, Jamaica, Canada and Australia.

Table 5: The results of the review process for mathematics

	Number of ratings	Number of ratings that met the quality criteria	Number chosen for inclusion
English	228	142	20
French	15	10	10

Test booklet design

There is one cluster of 31 reading items. There is one cluster of 30 mathematics items. There are two test booklets, to be rotated across learners. Each booklet contains both the reading and mathematics clusters, but the ordering of the clusters is reversed across the two booklets. This allows any position effects, such as the fact that students might be more fatigued when completing the second cluster, to be dealt with. The MILO test booklet design is shown in Table 6.

Table 6: MILO Assessment: Test design

Test booklet	Introduction	Part 1	Part 2
Test Booklet 1	Introduction and practice	Mathematics Cluster	Reading Cluster
Test Booklet 2	Introduction and practice	Reading Cluster	Mathematics Cluster

Appendix

Table 7: GPF reading domains, constructs and sub-constructs (with constructs included in the MILO assessment highlighted)

Domain		Construct		Subconstruct	
C	Comprehension of spoken or signed language	C1	Retrieve information at word level	C1.1	Comprehend spoken and signed language at the word or phrase level
				C1.1	Recognize the meaning of <u>common grade-level words</u> in a short, <u>grade-level continuous text</u> read to or <u>signed for</u> the learner
		C2	Retrieve information at sentence or text level	C2.1	Retrieve <u>explicit information</u> in a short <u>grade-level continuous text</u> read to or signed for the learner
				C3	Interpret information at sentence or text level
D	Decoding	D1	Precision	D1.1	Identify symbol-sound/fingerspelling and/or symbol-morpheme correspondences
				D1.2	Decode isolated words
		D2	Fluency	D2.1	Speak aloud or sign a <u>grade-level continuous text</u> at pace and with accuracy
R	Reading comprehension	R1	Retrieve information	R1.1	Recognize the meaning of <u>common grade-level words</u>
				R1.2	Retrieve <u>explicit information</u> in a <u>grade-level continuous text</u> by <u>direct- or close-word matching</u>
				R1.3	Retrieve <u>explicit information</u> in a <u>grade-level continuous text</u> by synonymous matching
		R2	Interpret information	R2.1	Identify the meaning of <u>unknown words and expressions</u> in a <u>grade-level continuous text</u>
				R2.2	Make <u>simple inferences</u> in a <u>grade-level continuous text</u>
				R2.3	Identify the main and secondary ideas in a <u>grade-level continuous text</u>
		R3	Reflect on information	R3.1	Identify the <u>purpose</u> and audience of a text
				R3.2	Give an overall evaluation of a text, and justify that evaluation
				R3.3	Evaluate the status of claims made in a text
				R3.4	Evaluate the effectiveness of a text

Table 8: GPF mathematics constructs and sub-constructs (with sub-constructs included in the MILO assessment highlighted)

Construct		Subconstruct	
N1	Whole numbers	N1.1	Identify, count in and identify the relative magnitude of whole numbers
		N1.2	Represent whole numbers in equivalent ways
		N1.3	Solve operations using whole numbers
		N1.4	Solve real-world problems involving whole numbers
N2	Fractions	N2.1	Identify and represent fractions using objects, pictures and symbols and identify relative magnitude
		N2.2	Solve operations using fractions
		N2.3	Solve real-world problems involving fractions
N3	Decimals	N3.1	Identify and represent decimals using objects, pictures and symbols and identify relative magnitude
		N3.2	Represent decimals in equivalent ways (including fractions and percentages)
		N3.3	Solve operations using decimals
		N3.4	Solve real-world problems involving decimals
N4	Integers	N4.1	Identify and represent integers using objects, pictures or symbols and identify relative magnitude
		N4.2	Solve operations using integers
		N4.3	Solve real-world problems involving integers
N5	Exponents and roots	N5.1	Identify and represent quantities using exponents and roots and identify the relative magnitude
		N5.2	Solve operations involving exponents and roots
N6	Operations across number	N6.1	Solve operations involving integers, fractions, decimals, percentages, and exponents
M1	Length, weight, capacity, volume, area and perimeter	M1.1	Use non-standard and standard units to measure, compare, and order
		M1.2	Solve problems involving measurement
M2	Time	M2.1	Tell time
		M2.2	Solve problems involving time
M3	Currency	M3.1	Use different currency units to create amounts
G1	Spatial visualizations	G1.1	Compose and decompose shapes and figures
G2	Properties of shapes and figures	G2.1	Recognize and describe shapes and figures
G3	Position and direction	G3.1	Describe the position and direction of objects in space
S1	Data Management	S1.1	Retrieve and interpret data presented in displays
		S1.2	Calculate and interpret central tendency
S2	Chance and probability	S2.1	Describe the likelihood of events in different ways
		S2.2	Identify permutations and combinations
A1	Patterns	A1.1	Recognize, describe, extend and generate patterns
A2	Expressions	A2.1	Evaluate, model and compute with expressions
A3	Relations and functions	A3.1	Solve problems involving variation (ratio, proportion, and percentage)
		A3.2	Demonstrate an understanding of equivalency
		A3.3	Solve equations and inequalities
		A3.4	Interpret and evaluate functions