

Study Design

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

8 April 2021, Version 3.1

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).

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

Acknowledgements	1
Introduction	3
Background and purpose.....	3
Design overview.....	4
Roles and responsibilities	7
Method.....	8
The population of interest	8
Survey and sample design.....	8
Technical standards	9
Assessment of learning outcomes	9
MILO test design.....	11
Contextual information.....	11
Materials preparation.....	12
Data collection	12
Ordering and timing	13
Data management.....	13
Standard setting exercise	13
Psychometric analysis	14
Data analysis.....	14
Reporting.....	14
Toolkit and looking forward to learning recovery	16
Capacity development	16
References.....	18

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 an overview of the study design and an outline of the method that will be used to implement the project which will take place in 2021. Detailed information on procedures will be made available progressively as part of the deliverables for other product areas, including sampling, field operations and data management.

Background and purpose

The COVID-19 pandemic has severely impacted all areas of life. The ambitious education goals of the 2030 Sustainable Development agenda are threatened by nationwide and localised school closures around the world¹. In response to the pandemic, governments around the world, are putting measures in place to mitigate the impact of school closures, and to facilitate the continuity of *education for all* through remote learning.

An increased reliance on parent-led and self-guided learning, staggered school attendance and the lack of adequate home educational resources have caused substantial disruptions to learning for many children and are exacerbating learning inequities².

¹ [UNESCO](#), 2020

² Kuhfeld, et al, 2020

This study will examine reading and mathematics outcomes in six countries in Africa, in order to determine the impact of COVID-19 on learning and to evaluate the effectiveness of distance learning mechanisms utilised during school closures.

The focus of this study is on students at the end of primary school. 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

This project will further provide a toolkit to participating countries that enables them to monitor learning outcomes against SDG indicator 4.1.1(b) after the crisis.

Design overview

The main aim of this study is to determine the **impact** of COVID-19 on learning outcomes at the end of primary schooling. To quantify **current** learning outcomes an assessment of reading and mathematics will be administered to students at the end of primary school in 2021. To quantify the impact on learning outcomes, data on learning outcomes **prior** to the pandemic must be available for comparison. The knowledge about current and prior learning outcomes will also lay the foundation for the participating countries to compare **future** learning outcomes⁴. In the ideal situation, prior, current and future data can be used to determine **learning loss** due to COVID-19 and **learning recovery** following COVID-19. This is illustrated in Figure 1.

The benchmark that will be used to quantify learning outcomes will be the proportion of students in the end of primary population, in each country, that meet the Minimum Proficiency Levels (MPL) referred to in SDG indicator 4.1.1(b) and described in ACER-GEM (2020). The link to this benchmark will be established in the 2021 assessments as part of this study and will be retrospectively applied to prior assessment results.

³ Development Goals and targets of the 2030 Agenda for Sustainable Development:

<https://unstats.un.org/sdgs/metadata?Text=&Goal=4&Target=4.1>

⁴ Measuring learning recovery is beyond the scope of this project. However, the toolkit provided will allow the participating countries to compare future learning outcomes should they choose to do so.

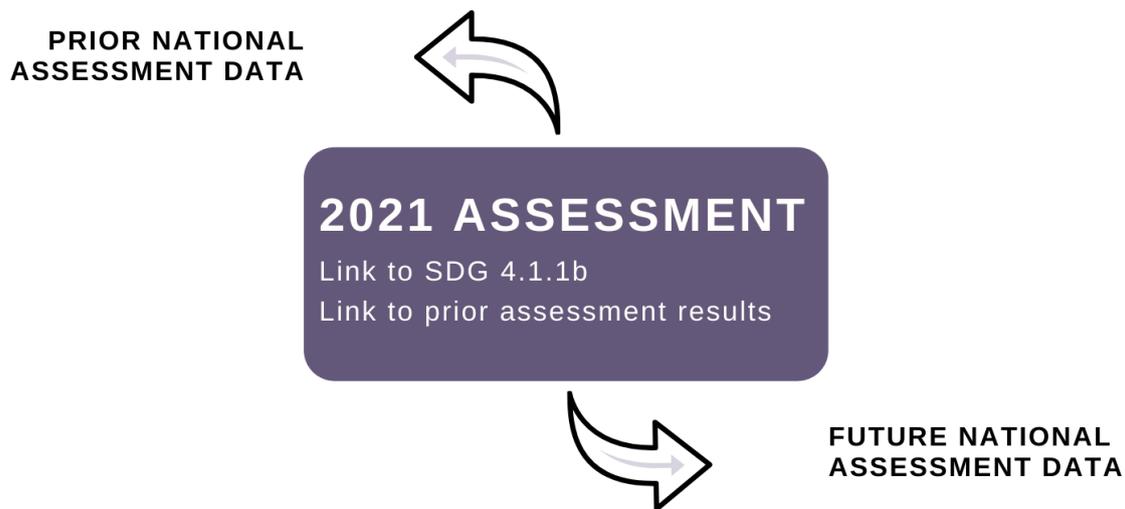


Figure 1: Study design overview

In order to measure learning loss the study must have access to prior baseline assessment data, so that a psychometric link can be established between the MILO and the prior national assessments. Figure 2 illustrates a potential decision path in determining how learning loss will be measured.

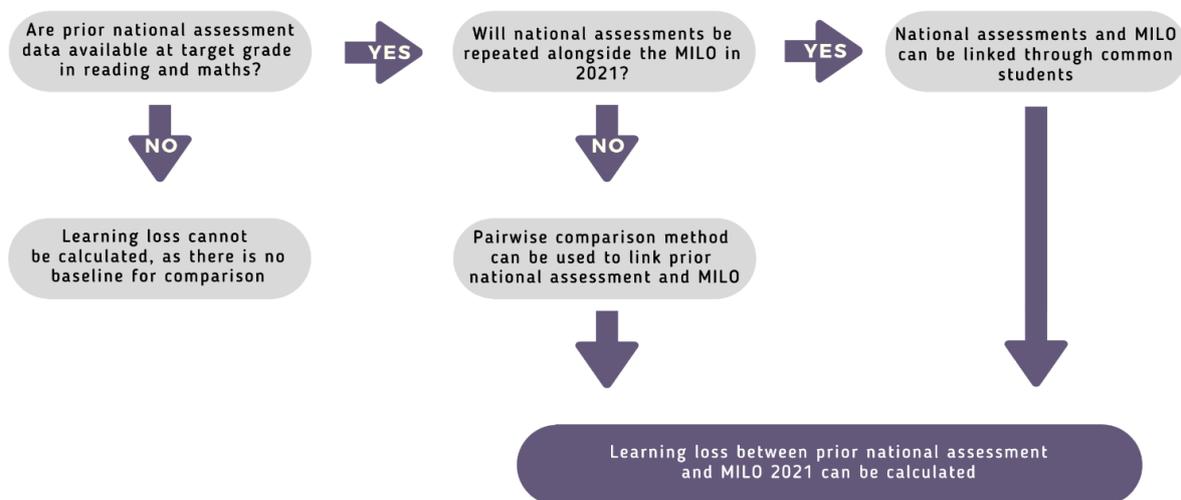


Figure 2: Decision path for measuring learning loss

In addition, to characterise the educational impacts and responses to COVID-19 contextual information will be gathered alongside the assessments at the student, school and system levels. This information will be helpful in exploring the relative efficacy of educational response. It will also help in exploring differential impacts of COVID-19 on learning outcomes for different subpopulations, including by gender, socioeconomic status or geographic remoteness.

Burkina Faso, Burundi, Cote d’Ivoire, and Senegal are Francophone participants. Kenya and Zambia are Anglophone participants. The following study design by country is proposed. These designs reflect that all participating countries have prior learning outcomes data available and are repeating at least part of a prior learning assessment in 2021 alongside the MILO assessments.

Table 1: Study design by country

Participating Country	Target grade	Link to Prior Learning Outcomes	2021 Learning Outcomes	Contextual information
Burkina Faso	6			
Burundi	6	PASEC, 2019 Booklet(s)		
Cote d’Ivoire	6			
Senegal	6		MILO assessment	MILO questionnaires
Kenya	7	NASMLA G7, 2019 Booklet(s)		
Zambia	5	NAS G5, 2016 Booklet(s)		

Roles and responsibilities

The COVID-19 MILO project is organised around four distinct roles, reflecting the entities that are involved and their different responsibilities. These roles are study manager, technical partner, PASEC coordination partner and participating country. An outline of the responsibilities of each entity in the different roles is provided in Table 2.

Table 2: Roles and responsibilities of participating entities

Role	Entities	Responsibilities
Study Manager	UIS	Determine project parameters and funding Liaise with participating countries Liaise with technical partners
Technical Partner	ACER	Design and provide all documentation and materials for procedures and production Liaise with countries to design and implement school samples Provide data management systems and services Undertake analysis Reporting Quality assurance Capacity development
PASEC coordination partner	CONFEMEN	Liaise with Francophone participants Coordinate implementation of Francophone participants Provide psychometric data and instrumentation from PASEC 2019
Participating Countries	Burkina Faso Burundi Côte d'Ivoire Senegal Kenya Zambia	Liaise with UIS Liaise with technical partners, as appropriate Provision of prior assessment booklets, framework and data Provision of school sample frame Recruitment of schools Materials production and distribution Data collection, including organising and training test administrators Data entry and data management Documentation of all technical processes Quality assurance

Method

The population of interest

In line with SDG indicator 4.1.1(b) this study aims to investigate learning outcomes at the end of primary schooling. However, 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.

In determining the population of interest in each participating country, two main factors will be considered:

- 1) the structural definition of the end of primary schooling within the country, aligned to international standards (ISCED); and
- 2) the presence of previous learning outcomes information from primary schooling, and, in particular, the information from grade levels that have been used to report previously against SDG 4.1.1(b).

ACER and the UIS will work with each participating country to determine the most appropriate target population within that country. To assist this, the *COVID-19 MILO Sampling Framework* and *School Sampling Preparation Guide* documents will be provided and decisions recorded using *COVID-19 MILO Sampling Forms*.

Survey and sample design

To support the goals of this program to report the impact of COVID-19 on end of primary populations, a survey design framework will be developed. The survey design framework will outline the overall approach to obtaining reliable estimates of outcomes and recording survey outcomes, drawing on the methods used in established large-scale assessments.

The sample will be designed to yield an effective sample size of 400 students. As the sample design involves a first stage of sampling schools and a second stage of sampling students, it is expected that a sample of at least 250 schools and 5000 students will be drawn. More precise estimates of school and sample size will be made on a country by country basis once the effects of clustering students within schools are explored during the sample design phase.

ACER will work with countries to document and quantify key sample design decisions including the target population, coverage and exclusions, explicit and implicit stratification, and school and within-school level exclusions. Negotiation and formal documentation of key sample design decisions is expected. These discussions will be supported by the *COVID-19 MILO Sampling Framework* document and a *School Sampling*

Preparation Guide. Sample design decisions are recorded in the *COVID-19 MILO Sampling Forms*.

Technical standards

Technical standards will help ensure the precision and validity of the sample, the quality of the operations and resulting data and the quality of the psychometric analysis.

Technical standards will be provided for an in-school assessment, considering the administration of the 2021 assessments and questionnaires alongside an existing regional or national assessment so as to provide a link to prior assessment results.

The standards will include the following key areas:

- Sampling standards relating to the level of precision and validity of the sample, including sample size, response rates and sample coverage of the target population.
- Linguistic quality control in the production of the French and English versions of the 2021 assessments and questionnaires.
- Data standards relating to: assessment and questionnaire construction and adaptation; test administration, quality monitoring and security of assessments, questionnaires and data; data management, data cleaning and data protection and security.
- Psychometric standards relating to the methods used to scale the results and to align the results to the SDG 4.1.1.b Minimum Proficiency Levels to ensure that the data is valid and reliable.

The standards will be expected to be adhered to by participating countries and technical partners. However, recognising the potential ongoing disruption of the COVID-19 pandemic in 2021, the attainment of some technical standards may be problematic. ACER will document technical partners' adherence to the standard. Participating countries will document their own adherence to the standards. ACER will critique documented deviations to the standards and will judge the impact of any deviations on the quality of the resulting data.

Assessment of learning outcomes

Participating countries will be provided with assessment instruments that will measure the attainment of SDG 4.1.1(b) Minimum Proficiency Levels (MPL) in reading and mathematics in students at the end of primary education. The assessments will be paper-based, provided in print-ready format, and available in English and French.

The assessments will follow an *Assessment Blueprint* that defines the coverage of the learning areas, domains and constructs as referenced in documentation of the MPLs. The assessment blueprint will specify the coverage of learning areas and the relative

proportion of domains. The assessment items will be drawn from the UIS Global Item Bank and will include French and English source items.

The Minimum Proficiency Level in mathematics for end of primary schooling is described in ACER (2020) 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.

The Minimum Proficiency Level in reading for end of primary schooling is described in ACER (2020) 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.

An outline of the selected domains and constructs covered by the MILO assessments appears in Table 3.

Table 3: Domains and constructs in the MILO assessments

Learning areas	Reading	Mathematics
Domains	Reading comprehension	Number and operations Measurement Geometry Statistics and probability Algebra
Constructs	Retrieving information Interpreting information Reflecting on information	Whole numbers Fractions Decimals Integers Exponents and roots Operations across number Length, weight, capacity, volume, area and perimeter Time Currency Spatial visualizations Properties of shapes and figures Position and direction Data Management Chance and probability Patterns Expressions Relations and functions

MILO test design

There is one cluster of 30 reading items. There is one cluster of 30 mathematics items. There are two MILO 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.

Table 4: MILO Assessment: test design

Booklet	Introduction	Cluster 1	Cluster 2
MILO Booklet 1	Introduction and practice	Mathematics Cluster	Reading Cluster
MILO Booklet 2	Introduction and practice	Reading Cluster	Mathematics Cluster

Contextual information

A *Contextual framework* will provide a rationale for the contextual data collected as part of this study. Three questionnaires will be constructed, each focusing on a different level: student, school and system. The student and school-level questionnaires will be paper-based and print-ready. The system-level questionnaire will be online. The

student-level questionnaire will be completed by the students undertaking the assessments, the school-level questionnaire will be completed by school principals, and the system-level questionnaire will be completed by governments on their country's COVID-19 education response. The questionnaires will be paper-based, provided in print-ready format, and available in English and French.

All three questionnaires will gather information on teaching and learning during the COVID-19 pandemic. Other factors such as community engagement and support for teachers to deliver effective remote pedagogy will be included in the school and system questionnaires.

The questionnaire will also provide information that will allow the exploration of inequitable impact of COVID-19 on learning. To this end there will be a focus on gender, socioeconomic status, and geographic remoteness.

It is highly likely that school enrolment and attendance has attenuated in some locations as a result of the pandemic. The student and school questionnaires will include items that help characterise the students that have remained engaged in schooling, to help determine the nature of impact of COVID-19 on school attendance. ACER will need corresponding information on these factors from the prior assessment administration to make the comparisons between samples.

Materials preparation

Printing of the booklets, questionnaires and administration manuals will be handled by the institutions within participating countries. ACER will provide print ready MILO assessments and questionnaire, and test construction guidelines. Participating countries will prepare their own national / historical assessments with guidance from ACER.

Some adaptation of the international source version of the questionnaires will be required and will be the responsibility of the participating institution and the technical partner.

Data collection

Institutions within participating countries will be responsible for the data collection. Data on student learning outcomes and contextual information will be collected through administering traditional pencil and paper tests and a questionnaire to children at the end of primary school⁵, in school settings. ACER will provide supporting materials such as field operations manuals. Participating countries will prepare their own instructions for administering the national / historical assessments.

⁵ The definition of end of primary schooling will likely vary across countries. Defining the end of primary school for the purpose of reporting against SDG 4.1.1b will be determined on a case by case basis in consultation with each country and the UIS.

Ordering and timing

The ordering and timing of the assessments is as follows:

Day 1

Session 1. First hour MILO (+introduction/practice)

Session 2. Second hour MILO

Session 3: 30-35 min Questionnaire

Day 2

Session 4: First hour historical (+introduction and practice if needed)

Session 5: Second hour historical

Data management

ACER will provide a data management tool called ACER Maple to each participating country. It is essential that this tool is used to:

- Track school participation
- Draw the within-school sample
- Produce student tracking forms
- Enter the MILO assessment data and the MILO contextual data from the questionnaires
- Record the participation of students as per the tracking form
- Submit the data to ACER.

ACER Maple will also be used to enter the national assessment data. **Participating institutions will be requested to adjust their national / historical assessments to include item codes and ID information so that they are consistent with ACER data management practices.**

Standard setting exercise

In parallel to this project, a subset of the MILO assessment items will be used in a Pairwise Comparison exercise involving items from the UIS Global Item Bank, the Pacific Islands Literacy and Numeracy Assessment, and other items from ACER's Item Bank. This exercise will result in a substantial set of items being placed on a single scale (by domain), which is a precursor to a standard setting exercise.

With the items ordered by difficulty a bookmarking method will be used to establish the Minimum Proficiency Level cut-points on each scale.

Psychometric analysis

The MILO assessment data will be psychometrically scaled, using a two-dimensional model to produce estimates for mathematics and reading learning outcomes. Using the benchmark set in the standards setting exercise, the proportion of students meeting or exceeding the MPL for SDG 4.1.1b will be estimated.

To estimate learning loss will require bespoke psychometric treatment for each participating country, although some efficiencies may be possible for the three countries repeating PASEC.

Data analysis

Analysis of the data will focus on the following research questions:

- What proportion of the target grade are reading SDG4.1.1b Minimum Proficiency Levels in reading and mathematics, by gender?
- What is the impact of COVID-19 on learning outcomes in reading and mathematics, by gender?
- What changes are there to the population attending the target grade in 2021 and the population that attended the target grade previously?
- What is the relative effectiveness of different distance learning mechanisms utilised during school closures?

Reporting

The reporting will clearly articulate the outcomes of the study, in a way that is clear to policy and practitioner stakeholders.

ACER will provide the following study reports:

- A final report on the impact of COVID-19 on learning outcomes, with a focus on gender
- Individual country summaries on the impact of COVID-19 on learning, with a focus on gender.

The final report will include:

- A brief description of the study purpose and design, the target population, country samples and response rates
- A statistical summary of the scale score distribution by country and by gender
- A statistical summary of the proportion of students reaching/exceeding SDG 4.1.1.b Minimum Proficiency Level cut-point by country and by gender
- Reporting on COVID-19 impact through comparisons of the proportions of students reaching/exceeding SDG 4.1.1.b Minimum Proficiency Level cut-point

on current national assessments linked to the MILO tests and previous national assessments outcomes, where possible

- Descriptive statistics of contextual factors at the student, school and system levels related to COVID-19 impact
- Inferential statistics of associations of contextual factors and achievement, aggregated at different levels.
- Technical descriptions of psychometric methods used to link with SDG 4.1.1b
- Technical descriptions of data analyses used to link with past national assessment results.

If comparisons of the proportions of students reaching/exceeding SDG 4.1.1.b Minimum Proficiency Level cut-point are not able to be determined due to the lack of prior national assessments, reporting on COVID-19 impact will not be possible. Descriptive and inference statistics of associations of contextual factors and achievement, aggregated at different levels, will instead be used to describe the context of learning outcomes/proportion of students reaching/exceeding SDG 4.1.1.b Minimum Proficiency Level.

A schema of how the national / historical assessment results and the MILO test results will be used to report any learning gap due to the impact of COVID-19 and expressed as a proportion of the population meeting SDG 4.1.1b is presented in Figure 3.

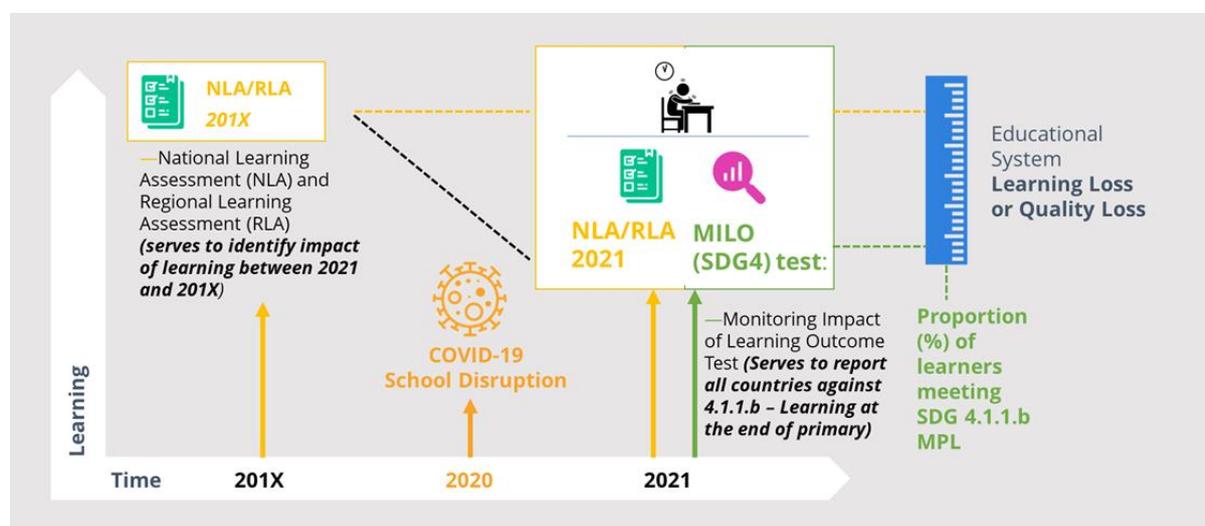


Figure 3: Schema for analysing and reporting learning loss

The individual country summaries will describe and highlight country-specific findings based on the final report. The final report and individual country summaries will be provided by ACER to the UIS for dissemination.

Toolkit and looking forward to learning recovery

ACER will provide a toolkit of materials that together will enable the UIS to scale up the study to more countries to allow them to report empirically against international benchmarks, reporting on SDG 4.1.1b, and providing a method to estimate the effects of COVID-19 on learning outcomes. Importantly, the toolkit will provide a technical description of the method so that the study is reproducible and can be scaled up to include other countries and so that the participating countries can monitor recovery from COVID-19 following this project, should they choose to do so.

The toolkit will be comprised of 13 elements (see Table 5).

Table 5: Toolkit elements

Toolkit element
Study design [this document]
Technical standards
Target population and sampling design
Assessment blueprint
Assessments in French and English, for reading and mathematics
Test construction guidelines
300 items to assess mathematics outcomes from pre-primary to upper secondary levels, in English
300 items to assess reading outcomes from pre-primary to upper secondary level, in English
Contextual framework, focused on emergency learning response
Student, school and system questionnaires
Report detailing the analysis method and results of aligning the scores from the MILO assessment and the national assessment, including: <ul style="list-style-type: none">• Technical descriptions of psychometric methods to be used, to link with SDG 4.1.1b• Technical descriptions of data analyses to be used to link with past national assessment results
Test administration guidelines
Data management guidelines

Capacity development

Capacity development will be an integral part of the project implementation. The overall aim of the capacity development is to build sustainable capacity of national teams in developing, implementing and using data from large-scale learning assessments for education system monitoring. A particular emphasis will be on analysing the impact of COVID-19 on learning outcomes, and preparing for the analysis of post-pandemic learning recovery.

The capacity development activities will be based on the Principles of Good Practice in Learning Assessment⁶ and aligned to the different products described in this proposal.

For the areas of Sampling, Test Administration, and Data Management, capacity building activities have been integrated and costed as part of the product. A focus is thereby on skill development through working directly with, and in consultation with, ACER experts on key tasks. The documentation of technical aspects such as key sample design decisions, quality assurance of standardised administration procedures, and data validation procedures, is important to ensure the technical rigour of the assessment. Enhanced skills in using software tools for sampling and data management that have in-built quality assurance mechanisms, are essential to increase and ensure the accuracy and quality of the data.

In order to design additional capacity building activities that will support the national teams, ACER will develop a capacity development overview. This overview will present 7 options for capacity building modules:

- Developing high quality assessment items
- Psychometric methods – Educational Measurement
- Psychometric methods – Pairwise comparison method
- Psychometric methods – Standard setting in assessment
- Psychometric methods – Common item linking
- Psychometric methods – Common person linking
- Data analysis – Controlling for differences in socioeconomic status

For each module, the capacity development overview will describe the scope, target audience, difficulty level, delivery modes and cost estimates, to allow the UIS and participating countries to prioritise the two modules to be developed within the capacity development provision that has been made in the project budget.

⁶ ACER & UIS, 2017

References

- ACER & UIS. (2017). *Principles of good practice in learning assessment*. <http://uis.unesco.org/sites/default/files/documents/principles-goodpractice-learning-assessments-2017-en.pdf>
- ACER-GEM. (2020). *Refinement of minimum proficiency levels. GAML-7: WG/GAML/6* http://tcg.uis.unesco.org/wp-content/uploads/sites/4/2020/10/WG-GAML-6-Minimum-Proficiency-Levels_revised.pdf
- Kuhfeld, M., Soland, J., Tarasawa, B., Johnson, A., Ruzek, E., & Liu, J. (2020). *Projecting the potential impacts of COVID-19 school closures on academic achievement*. (Annenberg Institute Working Paper No. 20-226.) <https://www.edworkingpapers.com/sites/default/files/ai20-226-v2.pdf>
- UNESCO. (2020). *Education: From disruption to recovery*. <https://en.unesco.org/covid19/educationresponse>