

Quality Assurance Guidelines for the UIS Global Item Bank

**COVID-19: Monitoring the Impacts on Learning
Outcomes (MILO)**
28 January 2021

The ACER Centre for Global Education Monitoring supports the monitoring of educational outcomes worldwide, holding the view that the systematic and strategic collection of data on education outcomes, and factors related to those outcomes, is required to inform high quality policy aimed at improving educational progress for all learners.



Acknowledgments

This project, the Assessment and Study of COVID Impact on Learner Progress, is referred to as the COVID-19 MILO (Monitoring the 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 ACER's Global Education Monitoring Centre (GEM Centre), an ACER initiative in partnership with the Australian government's Department of Foreign Affairs and Trade.

Contents

Acknowledgments	2
Introduction	4
Construct validity	4
Translatability	5
Cultural issues	5
Technical criteria	6

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.

The quality assurance guidelines presented in this document are developed to support the UIS bank of items (the Global Item Bank). The UIS proposed the establishment of a Global Item Bank, to include assessment items that countries may use to monitor their progress against indicator 4.1 of the Sustainable Development Goals. It is envisaged that material in this item bank will come from many different sources. In order to ensure that material included in the global item bank is of high quality, a set of quality assurance procedures should be established. This document (Deliverable 6.4) provides a set of four key quality assurance areas that are essential to consider – construct validity, translatability, cultural issues and technical criteria. These four areas will also be included in the set of guidelines for item review (Deliverable 7.1)¹ to support the item selection for the COVID-19 MILO testlets.

Construct validity

Construct validity refers to the extent to which a test measures that which it claims to be measuring. A necessary step in ensuring construct validity is the definition of what is to be measured. The draft Minimum Proficiency Levels (MPLs) are accompanied by a glossary defining key terms for each of the constructs of reading and mathematics. Taken together, these documents provide the necessary guidance about how the constructs being measured have been defined.

¹ This deliverable is due to be submitted by 31 March 2021.

In the context of review of possible items for the Global Item Bank, it will be necessary to consider whether:

- each assessment task measures an element of the skills, knowledge and understanding contained in the construct, as represented in the MPLs
- each assessment task provides a piece of evidence about the extent to which students have developed the skills, knowledge and understanding defined in the MPLs and their associated documentation.

Translatability

It is envisaged that participants will provide material for consideration in the Global Item Bank in one of a small group of major world languages. Therefore an important part of the quality assurance procedure will involve experts considering whether the material can be translated into the local language. Translatability is not only about whether an item can easily be translated into a given language, but whether the same level of difficulty can be maintained, and the item can be translated in such a way that the same skills are tested in the different language versions.

Common problems for translation include use of a word in the original language of the item that does not exist in a local language; the substitution of another term may change the difficulty of the item because the alternative term gives more, or fewer clues to students. Stimulus material for reading items sometimes includes language-specific qualities that are difficult to capture in translation (e.g. word play, rhythm). Item format and issues of grammar and syntax also need to be considered. For example, sentence completion items can pose difficulties in translation because word order in sentences varies across languages (e.g., Japanese= subject-object-verb, English= subject-verb-object, Tagalog= verb-subject-object). Use of the passive voice can also be an issue, as many languages do not have an equivalent.

Cultural issues

Since it is envisaged that the Global Item Bank will contain items from multiple sources, it will be important for reviewers to consider whether there are any cultural concerns about items or pieces of stimulus. That is, reviewers should reflect on whether the topic is acceptable and appropriate for students in their country. This question is not about student familiarity with a context, as when items are drawn from multiple sources, it is not possible to ensure settings that are familiar to students in all countries. Rather, cultural concerns may arise when the contexts of references would not be acceptable in a given country for reasons based on, for example, social norms or religious beliefs.

Technical criteria

Items being considered for inclusion in the Global Item Bank should be evaluated against a set of technical criteria, including clarity and correctness, centrality and difficulty. These concepts will be expanded on in the set of guidelines for item review (Deliverable 7.1).

For example, all items should be as transparent as possible, so that the challenge for students is in interpreting the stimulus material, and the item itself does not pose extra difficulty. This means the item should be succinct, should have no tricks, ambiguity or difficult language and should avoid negatives. Multiple-choice items should have only one correct answer, and there should be nothing in the way the item is written that would point an experienced test-taker to the correct answer. If the item being submitted has already been administered to groups of students, then evidence of how the item has functioned should also be reviewed.