Capacity Development for AMPLab

Overview
May 2023

The Global Education Monitoring (GEM) Centre drives improvements in learning by supporting the monitoring of educational outcomes worldwide. 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).

Acknowledgments

Table of contents

Background

As part of SDG 4, Indicator 4.1.1 aims to measure the "proportion of children and young people: (a) in grades 2/3; (b) at the end of primary; and (c) at the end of lower secondary achieving at least a minimum proficiency level in (i) reading and (ii) mathematics, by sex." To meet this goal, UIS has coordinated efforts to establish common reading and mathematics scales for all three points of Indicator 4.1.1, building on existing cross-national and national assessments. As a result of these efforts, two important points of consensus have been reached: the definition of the Minimum Proficiency Level (MPL) and the Global Proficiency Framework (GPF). The MPL has been so far identified in each of the major cross-national assessments, based on the proficiency level descriptors in each of them.

The overarching objective of the AMPLab project is to measure and analyze the performance of students at the end of lower and upper primary using an assessment that aligns with GPF Level 2. This will:

- enable the collection of more informative data about where students are performing in terms of the GPF in reading and mathematics,
- produce baseline measures to set targets and compare learning gains/losses
- facilitate reporting on SDG 4.1.1
- aid the tracking of learning progress over time
- complement tools that had been already developed in 2021 in the Monitoring the Impacts on Learning Outcomes (MILO) study.

Capacity Development

Two capacity development courses have been designed to equip participants with an understanding of the theory of Rasch measurement, its application for assessment, and for estimating characteristics of a population. The courses build advanced level skills in educational measurement, including large-scale assessment analysis with plausible values and dealing with complex survey designs. These courses are designed to build expertise to enable high-level analysis and reporting of large-scale educational data and trends.

These courses are delivered online (via ACER's learning management system (Moodle)) and are designed to be self-paced with some accompanying live facilitated sessions over Microsoft teams/Zoom (3 x 90 minutes). Questions and comments can be submitted either via the Moodle discussion forums (preferred method as all participants can see responses) or via email to facilitators. Course 1 is to be undertaken from 3rd to 16th July and Course 2 from 17th to 30th July. An overview of the dates and details of the live sessions is presented in the Table 1 below:

Table 1: Overview of course delivery

Date	Session	Details
July 3	Welcome/ Introduction to Course 1	This session will cover: • Introduction to participants and facilitators • Courses overviews • Moodle orientation • Software requirements • Introduction to key concepts for Course 1
July 14	Review of Course 1/ Introduction to Course 2	This session will cover: • Q & A session • Review of key concepts covered in Course 1 • Introduction to key concepts for Course 2
July 28	Review of Course 2/ Wrap-up	• Q & A session • Review of key concepts covered in Course 1 • Final notes

Each course has a learning assurance task (LAT) to be completed at the end, with marking and feedback provided following submission. Each participant who accesses the Moodle content and engages with the online live sessions will receive a certificate of participation. Those who successfully complete the LATs will receive a certificate of course completion.

An overview of the two courses is presented below.

Course 1: Understanding population characteristics: working with plausible values

In this course participants will develop the skills necessary to work with data generated from large scale assessments, including the AMPLab program, Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS) and Programme for International Student Assessment (PISA). The data used in this course to illustrate concepts and that participants will work with will be taken from publicly available data sets.

The course content includes an introduction to the theory behind different item response models that can be used to reliably estimate population parameters. Participants will apply this knowledge to fit an item response model to undertake secondary analysis to yield unbiased population parameter estimates.

This course is suitable for those knowledgeable about assessment systems and who wish to improve their knowledge of educational measurement. Participants will build skills in analysing educational data and using statistical software.

Duration: 2 days (or equivalent)

Software: ConQuest, R and RStudio are required for completion of this course.

Learning outcomes:

- Understand the difference between item response models that assume some underlying ability distribution (e.g., the normal distribution) in the population from which a sample is drawn and item response models that do not make such assumptions.

- Discriminate between different kinds of ability estimates produced by item response models, including point estimates appropriate to provide scores to individuals including students and imputations used in large scale estimates to describe population parameters including means and variances. Participants will be able to evaluate the pros and cons of different kinds of ability estimates in the context of an analytic or research plan.

- Independently implement and analyse the results of an analysis using ACER ConQuest.

- Apply theoretical knowledge to fit an item response model that is appropriate to use secondary analysis to understand aspects of the populations. This will include conducting descriptive and correlational analysis using plausible values.

Course 2: Complex survey designs in large-scale assessment

In the previous course we introduced plausible values as a method of accurately estimating population parameters. In this course we introduce complex survey designs, used in large scale assessment programs to accurately estimate population parameters that are representative of national or sub-national groups of interest.

In large scale assessments it can be very inefficient to draw simple random samples from a population: it can be costly to visit a distant school to assess one student. Typically, a two-stage sampling approach is taken (e.g., Stage 1 – select a sample of schools, Stage 2 - sample students within these schools). This results in clustering that needs to be taken in to account when estimating uncertainty in parameter estimates.

Participants will apply this knowledge to fit an item response model typically used in large scale assessment programs like AMPLab program, Trends in International Mathematics and Science Study (TIMSS), Progress in International Reading Literacy Study (PIRLS) and Programme for International Student Assessment (PISA). Appropriate secondary analysis methods will be used which account for measurement and sampling error to yield unbiased population parameter estimates.

This course is suitable for those knowledgeable about assessment systems and who wish to improve their knowledge of educational measurement. Participants will build skills in analysing educational data and using statistical software.

Duration: 2 days (or equivalent)

Software: ConQuest, R and RStudio are required for completion of this course.

Learning outcomes:

- Understand the difference between a simple random sample and a complex sampling design and apply this knowledge to understand why two-stage sampling is often undertaken in large scale assessment.

- Discriminate between the sources of uncertainty in population parameter estimates (i.e., sampling error and measurement error) and describe the methods used to account for each in large scale assessment designs.

- Independently implement and analyse the results of an analysis using ConQuest and R (or alternative statistical software).

- Apply theoretical knowledge to interpret results from secondary analysis and to describe the relative magnitude of sampling and measurement error in the total uncertainty of a parameter estimate.