Summary of scale score distribution

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

11 January 2022

Acknowledgments

Introduction

The COVID-19: Monitoring Impacts on Learning Outcomes (MILO) project aimed 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, the project developed the capacity of countries to monitor learning after the crisis.

The four overarching goals of the project were 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 the summary distribution—mean and variance—of the reading and mathematics IRT scale scores in the MILO project.

Summary scale score distribution

All assessments using an Item Response Theory (IRT) based method produce a scale calibrated in logits. Typically, the logit scale will have a mean of approximately 0 and a standard deviation of approximately 1. When the Assessments for Minimum Proficiency Levels (AMPL) data was scaled using the Rasch model, a logit scale was produced. In most large-scale assessments the logit scale is transformed into a reporting scale to aid interpretation. For example, studies like PISA and TIMSS transform the 0;1 logit scale to a 500;100 reporting scale. That is, a linear transformation of the mean from 0 to 500 and the standard deviation from 1 to 100. In such a reporting scale, under a normal distribution, 99.7% of scores will fall between 200 and 800 scale score points. Other studies use different values for the reporting scale transformation. However, in each study, the reason behind a reporting transformation is simply to help aid interpretation: the difference between scale scores of 400 and 550 (reporting scale) is much more intuitive to many than the difference between scale scores of -1.0 and 0.5 (logits) even though these example figures represent the same ability estimates.

In the MILO project, the goal was to report outcomes in terms of the proportion of the population who met or exceeded the cut-points of SDG4.1.1b in reading and mathematics. To achieve this, there was no need to produce a reporting scale. In fact, ACER advised that with so few countries undertaking the MILO project, and the

likelihood of expansion of the AMPL tools to other countries in the future, the development of a reporting scale at this stage would be premature. It is likely that a more stable reporting scale can be developed after the inclusion of more representative data from some higher performing populations.

The summary distribution—mean and variance—of the reading and mathematics IRT scale scores in MILO are provided in Tables 1 to 4. The tables provide estimates for both the AMPL and the historical assessments by country and gender.

Table 1: Mean reading scale score overall and by gender

Country	Mean performance in reading, logits					
	2021 AMPL			Historical assessment		
	All	Boys	Girls	All	Boys	Girls
Burkina Faso	-0.15 (0.05)	-0.14 (0.05)	-0.17 (0.05)	-0.38 (0.06)	-0.41 (0.07)	-0.36 (0.05)
Burundi	-0.92 (0.03)	-0.89 (0.03)	-0.94 (0.03)	-0.92 (0.03)	-0.89 (0.03)	-0.94 (0.03)
Cote d'Ivoire	-0.79 (0.07)	-0.82 (0.07)	-0.75 (0.08)	-0.84 (0.08)	-0.85 (0.08)	-0.82 (0.08)
Kenya	0.90 (0.07)	0.84 (0.07)	0.97 (0.07)			
Senegal	-0.12 (0.05)	-0.18 (0.05)	-0.08 (0.06)	-0.11 (0.07)	-0.16 (0.08)	-0.06 (0.08)
Zambia	-0.90 (0.04)	-0.93 (0.05)	-0.86 (0.04)	-0.92 (0.03)	-0.95 (0.03)	-0.90 (0.03)

Standard errors (SE) are reported in brackets.

Table 2: Variance in reading scale score overall and by gender

Country	Variance in performance in reading, logits					
	2021 AMPL			Historical assessment		
	All	Boys	Girls	All	Boys	Girls
Burkina Faso	0.67 (0.08)	0.66 (0.09)	0.67 (0.07)	0.78 (0.06)	0.82 (0.09)	0.74 (0.04)
Burundi	0.27 (0.02)	0.27 (0.02)	0.27 (0.03)	0.25 (0.02)	0.25 (0.02)	0.26 (0.03)
Cote d'Ivoire	1.95 (0.12)	1.94 (0.13)	1.95 (0.15)	1.59 (0.10)	1.55 (0.11)	1.64 (0.11)
Kenya	1.50 (0.08)	1.57 (0.09)	1.42 (0.08)			
Senegal	0.87 (0.08)	0.82 (0.07)	0.91 (0.09)	0.96 (0.09)	0.97 (0.10)	0.95 (0.09)
Zambia	0.49 (0.06)	0.49 (0.07)	0.49 (0.05)	0.46 (0.03)	0.44 (0.03)	0.48 (0.04)

Standard errors (SE) are reported in brackets.

Table 3: Mean mathematics scale score overall and by gender

Country	Mean performance in mathematics, logits					
	2021 AMPL			Historical assessment		
	All	Boys	Girls	All	Boys	Girls
Burkina Faso	-0.48 (0.04)	-0.43 (0.04)	-0.51 (0.04)	-0.67 (0.04)	-0.67 (0.05)	-0.67 (0.04)
Burundi	-0.75 (0.04)	-0.65 (0.04)	-0.83 (0.04)	-0.66 (0.03)	-0.55 (0.04)	-0.75 (0.03)
Cote d'Ivoire	-1.54 (0.06)	-1.53 (0.06)	-1.54 (0.07)	-1.64 (0.06)	-1.58 (0.06)	-1.72 (0.07)
Kenya	0.47 (0.04)	0.48 (0.05)	0.47 (0.04)	0.59 (0.10)	0.71 (0.12)	0.56 (0.10)
Senegal	-0.39 (0.04)	-0.38 (0.04)	-0.40 (0.05)	-0.35 (0.06)	-0.34 (0.06)	-0.35 (0.06)
Zambia	-1.31 (0.03)	-1.31 (0.03)	-1.30 (0.03)	-1.16 (0.02)	-1.14 (0.02)	-1.17 (0.03)

Standard errors (SE) are reported in brackets.

Table 4: Variance in mathematics scale score overall and by gender

Country	Variance in performance in mathematics, logits					
	2021 AMPL			Historical assessment		
	All	Boys	Girls	All	Boys	Girls
Burkina Faso	0.39 (0.04)	0.39 (0.04)	0.39 (0.04)	0.46 (0.03)	0.49 (0.05)	0.43 (0.03)
Burundi	0.41 (0.03)	0.40 (0.04)	0.40 (0.03)	0.41 (0.03)	0.42 (0.03)	0.38 (0.03)
Cote d'Ivoire	1.30 (0.09)	1.31 (0.09)	1.29 (0.10)	1.07 (0.07)	1.05 (0.07)	1.08 (0.09)
Kenya	0.69 (0.03)	0.74 (0.04)	0.64 (0.03)	0.84 (0.10)	0.66 (0.07)	0.64 (0.09)
Senegal	0.58 (0.04)	0.56 (0.04)	0.60 (0.05)	0.61 (0.04)	0.64 (0.06)	0.59 (0.04)
Zambia	0.26 (0.03)	0.26 (0.03)	0.26 (0.03)	0.27 (0.01)	0.27 (0.02)	0.27 (0.02)

Standard errors (SE) are reported in brackets.

The tables in this document may be used as a reference for analysts to check calculations using the international datasets.