

POLICY LINKING FOR MEASURING GLOBAL LEARNING OUTCOMES UPDATES

The following document provides an update on the development and pilot testing of the Global Proficiency Framework and the Policy Linking for Measuring Global Learning Outcomes Toolkit, which were designed to support country reporting of reading and math outcomes to Sustainable Development Goal 4.1.1.

BACKGROUND

Countries face challenges with comparing, aggregating, and tracking (CAT) learning outcomes over time and reporting outcomes to Sustainable Development Goal (SDG) 4.1.1. The three challenges most commonly faced by countries include:

- 1) Differences in the level of difficulty of assessments due to: variations in programming, interventions, and contexts across a country; the need to have several different languages of assessment to reach all learners; and/or the need to change assessments over time due to test leakage issues. When assessments are not equated to ensure comparability, these challenges make CAT outcomes difficult;
- 2) A lack of benchmarks (or cut scores) defining global minimum proficiency on national assessments;
- 3) Limited resources to buy into regional or international assessments that allow comparability.

Ministries of Education, international education donors, partners, and other stakeholders need a method for accurately determining how learning outcomes compare both between contexts in a country and across countries and for reporting on progress in reading and mathematics. This information is critical for identifying gaps in learning outcomes so that resources can be focused on those areas and populations most at need. Linking can be done either statistically, using common items between assessments or having common learners take more than one assessment, or non-statistically, using expert judgments. Although statistical methods are often associated with higher levels of precision, they are not always practical or financially feasible and involve several methodological prerequisites.

Policy Linking for Measuring Global Learning Outcomes (policy linking for short) offers countries a way to address this challenge by providing a low-cost, quick, easy-to-implement method that allows countries and other stakeholders to link existing national and sub-national assessments to a common scale (the Global Proficiency Framework) and set benchmarks for reporting the percentage of learners in a country meeting global minimum proficiency requirements.

Policy linking was first proposed as a method for linking assessments to SDG 4.1.1 at a Global Alliance to Monitor Learning (GAML) meeting in September 2017. Since then, with support from other donors and partners, the UNESCO Institute of Statistics (UIS) has been working to assess country and partner interest and buy-in to the methodology as well as to determine its viability and validity for comparing learning outcomes across diverse contexts and countries.

UIS and its partners have made a great deal of progress in this regard. Updates on that progress are included in the following sections of this summary document.

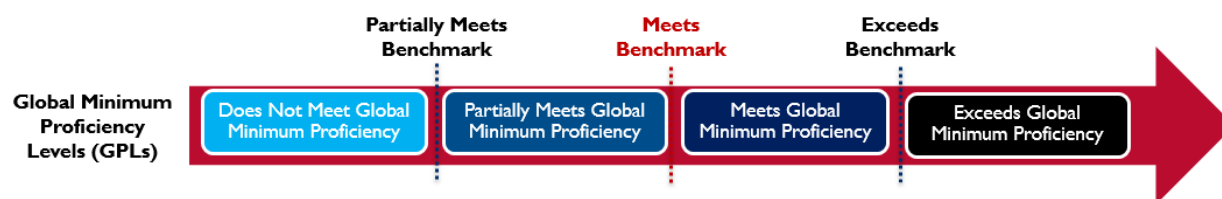
GLOBAL PROFICIENCY FRAMEWORK DEVELOPMENT AND PILOTING

Overview - A group of 50 global reading and math content experts came together in April and June of 2019 to develop a common scale for linking assessments to SDG 4.1.1 through policy linking. Using the UNESCO International Bureau for Education (IBE) Global Content Frameworks of Reference for Reading and Math¹ (the GCFRR and GCFRM, respectively) as a basis, the content experts developed a research-based progression of the minimum knowledge and skills learners in Grades 2 to 6 should be able to demonstrate with respect to the key domains of reading and math. The Global Proficiency Framework includes four Global Proficiency Levels (GPLs) as shown in **Figure 1** below: *Does Not Meet Global Minimum Proficiency, Partially Meets Global*

¹The GCFRR and GCFRM synthesize content and assessment framework information from more than 50 countries around the globe, providing a picture of the common expectations countries have for learners' performance in reading and math.

Minimum Proficiency, *Meets Global Minimum Proficiency*, and *Exceeds Global Minimum Proficiency* and can be used to set three benchmarks on assessments, with the *Meets Global Minimum Proficiency* level being the most critical since it aligns with the requirements for SDG 4.1.1. Stakeholders argued for the inclusion of four levels to allow countries to show progress toward reaching the 4.1.1 indicator.

Figure 1. GPF Global Proficiency Levels



Source: Davis, Ferdous, Kelly, et. al., Policy Linking Toolkit for Measuring Global Learning Outcomes, October 2020

Piloting - Following completion of the GPF in June 2019, UIS and its partners pilot tested it in nine countries either as part of a policy linking exercise or to inform country modifications to its assessment and education systems.² Pilot countries included Bangladesh, Djibouti, the Gambia, Ghana, India, Madagascar, Malawi, Nigeria, and Senegal during the 2019-2020 academic year. The lessons learned from those pilot tests informed the organization, beginning in May of 2020, of a second round of consultations with reading and math educators, curriculum experts, and psychometricians from the global community, many of whom had participated in the first round. During on-line deliberations between May and October 2020, experts revised the initial GPF and added Grades 1 and 7-9 to the Framework. The result is a GPF that covers the entire nine years of basic education and can be used to set benchmarks for each of the three SDG 4.1.1 education levels—the end of lower primary (Grades 2/3), the end of primary, and the end of lower secondary (Evans, et. al., 2020).³

POLICY LINKING METHODOLOGY DEVELOPMENT AND PILOTING

Overview - To establish numeric thresholds for each proficiency level for different assessments, policy linking requires aligning those assessments to the GPF, matching assessment items to the global proficiency descriptors (GPDs), and setting benchmarks. Since the GPF is used as a reference—or common criteria—for policy linking, these benchmarks represent the same standard of performance on those different assessments as defined by the GPDs, regardless of the difficulty or language of the assessments. This means that the benchmarks are set at different places (numbers) on the different assessments (unless the assessments are of equivalent difficulty).

To set the benchmarks, policy linking uses an internationally recognized, standardized, test-centered, Angoff-based benchmarking procedure. The Angoff procedure requires groups of national subject-matter experts, called panelists, to make judgments on the assessments. The panelists include master teachers and curriculum experts from the country who are familiar with the performance of learners for specific grades

² While the GPF was created for use with policy linking and is not intended to be prescriptive in nature, countries have used it as a tool to inform the development or adaptation of national performance standard frameworks for guiding the construction of new or adapted national assessments and to inform country content standards and curriculum frameworks, teacher training, and text and materials development in countries that are looking to modify their education systems.

³ Evans, Watson, Chiappetta, et. al., *Global Proficiency Framework for Reading*, October 2020.

and subjects. They follow the Angoff procedure to 1) examine the country's assessment instrument(s) in relation to the GPDs and 2) estimate how learners in each of the GPL categories would perform on the assessment. Planners and facilitators organize and conduct separate workshops by grade, subject, and language with different groups of panelists to set the equivalent benchmarks for those assessments.

Piloting - The United States Agency for International Development (USAID) contracted with Management Systems International (MSI) to develop the Policy Linking Toolkit in 2018. In close coordination with technical experts, international donors, and stakeholders, MSI developed a first draft of the Toolkit in September 2019. Between October 2019 and September 2020, UIS and USAID, with support and collaboration from MSI, FHI360, Evans and Associates, the Bill and Melinda Gates Foundation (BMGF), the UK Department for International Development (now called the Foreign, Commonwealth and Development Office), the World Bank, the Australian Council for Educational Research (ACER), the People's Action for Learning Network (PAL Network), The Education Partnership (TEP) Centre, and Zizi Afrique, pilot tested the Toolkit in Bangladesh, India, Malawi, Nigeria, and with the PAL Network's International Common Assessment of Numeracy (ICAN) (through remote engagement with panelists from Kenya and Nigeria as a result of the COVID-19 crisis). Those pilots informed changes to the Toolkit, including lessons about facilitating remote policy linking workshops, that UIS, USAID and MSI incorporated between April 2020 and October 2020.

DEVELOPMENT OF THE 4.1.1 REVIEW PANEL AND DRAFT CRITERIA FOR POLICY LINKING VALIDITY

Following feedback from multiple stakeholders, including those in attendance at GAML 6 in 2019, UIS developed a set of Criteria for Policy Linking Validity that require country assessments proposed for policy linking meet a minimum set of criteria, including criteria regarding sufficient alignment of the assessment with the GPF and measures of assessment reliability and validity. UIS has begun recruitment of psychometricians and content experts to sit on the 4.1.1 Review Panel that will review country assessments and related documents to determine feasibility of policy linking for linking those assessments to SDG 4.1.1. The panel may also review the validity of other linking exercises for reporting to SDG 4.1.1.

POLICY LINKING EVALUATION

The BMGF contracted the National Foundation for Educational Research (NFER) in early 2020 to evaluate the policy linking methodology to determine its feasibility and validity as a method for linking assessments to SDG 4.1.1. NFER has shared findings and recommendations for the methodology that have been incorporated into the Policy Linking Toolkit and GPF and will deliver the final results of the evaluation in the Spring of 2021.

NEXT STEPS

UIS, USAID, and the World Bank have plans to pilot the new versions of the Policy Linking Toolkit and GPF in at least eight additional countries between November 2020 and May 2021. Lessons from those pilots as well as from the final evaluation report that NFER will produce will inform final revisions to the Policy Linking Toolkit and GPF in the Spring of 2021. Further, UIS plans to bring on up to fifteen 4.1.1 Review Panelists before the end of 2020. Those panelists may also provide feedback and input into the documents and

especially into the Criteria for Policy Linking Validity.