unesco
Institute for Statistics

25 YEARS

Education Data and Statistics Commission

Strengthening MICS FLS Module (FLM2.0)

February 2025

João Pedro Azevedo

Eva Quintana

Sakshi Mishra

Global Alliance to Monitor Learning

#25YearsOfDataInsights

1. Child Centric Surveys: Multiple Indicator Cluster Survey Program (MICS)

MICS: The Largest Source of Global Child Data

- Over 360 surveys in 120 countries in the last 30 years
- Conducted by National Statistical Offices with UNICEF support
- Surveys repeated every 4-5 years (or more frequently if needed)

Features:

✓ **Representative sample** of households

✓ **Comprehensive**: 5 Questionnaires (household; children under five; women 15-49; men 15-49; children aged 5 to 17) covering a wide range of topics: child survival, nutrition, early childhood, education and skills, protection + more

✓ **Cost-effective**: One survey generates ~200 indicators, including 40 SDG indicators – extremely important consideration in an increasingly resource constrained environment

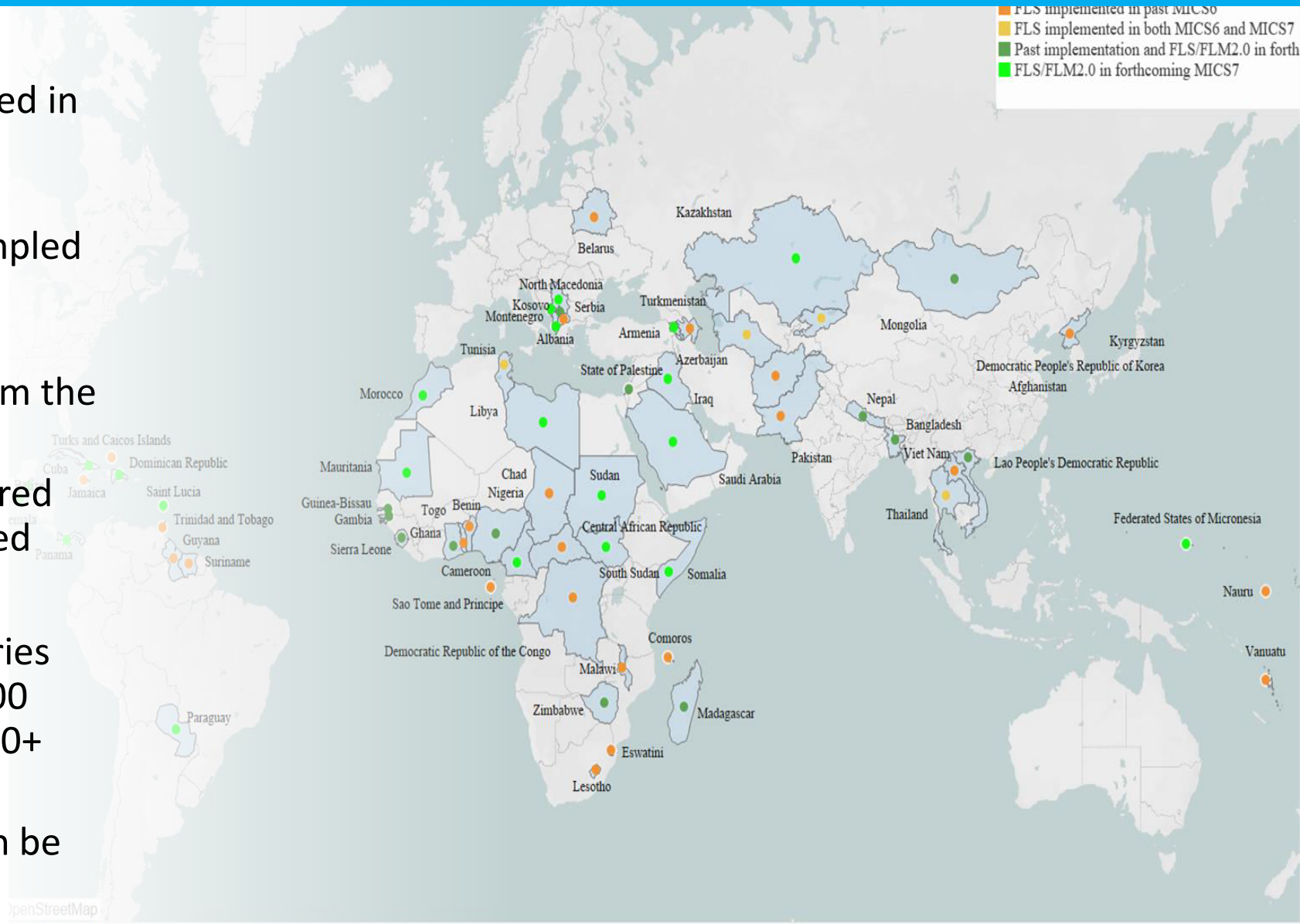
✓ **Expertise in creating statistically valid tools** Examples include: Early childhood development index, Mental health module, Time use module, Water Quality Testing

✓ **Global network of experts**: supporting capacity strengthening on sampling to implementation and data processing

✓ **New innovations** include longitudinal phone-based household surveys (MICS-Plus) and linking household and administrative data (MICS-Link)

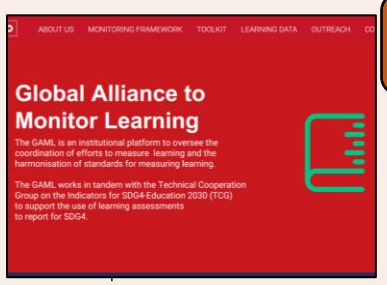
2. Foundational Skills Module (FLS, 2016 - current)

- **MICS FLS module features**:
  - Assessment languages determined in collaboration with MoE
  - Administered to one randomly selected child (age 7-14) per sampled household
  - Requires informed consent from primary caregiver and assent from the child
  - Individually and orally administered by an interviewer specially trained



- Implemented in more than 50 countries and territories to more than ~ 234,000 children between ages of 7 to 14 in 10+ languages, since 2016
- All survey data and questionnaire can be downloaded from the MICS website:
- https://mics.unicef.org/surveys

3. Multi-topic survey allows learning outcomes to be unpacked in different ways



Minimum proficiency levels (MPLs) released

GPF development process begins for reading and mathematics; field test in 9 countries

GPF for reading and mathematics: Grades 1 to 9 finalized and launched Policy Linking Toolkit(PLT) released

GAML agrees on MPL mapping against the GPF and PLT revised in April 2023

UIS SDG 4.1.1 TAG; draft criteria shared for review; final criteria released

2014 2015 2016 2017 2018 2019 2020 2021 2022 2023 2024 2025

UNICEF begins work on foundational skills module and convenes a TAG

Four rounds of field test in Ghana, Belize, Kenya and Costa Rica

Module finalized and released in end-2016

Foundational skills module included in the proposed protocol for SDG4 reporting GAML4/REF/10 and accepted for interim reporting.

Initial discussions on development of instrument for SDG4.1.1a and b reporting through household survey data

UNICEF launches work on revising early grade items and developing a new module for end of primary

Items for end of primary developed using AMPL (b) as blueprint; items shared for review w/ partners

Cognitive testing of instruments in Zambia, April 2023

UNICEF FLM 2.0 TAG meeting

Quantitative testing: Kenya in school and Qualitative testing in households

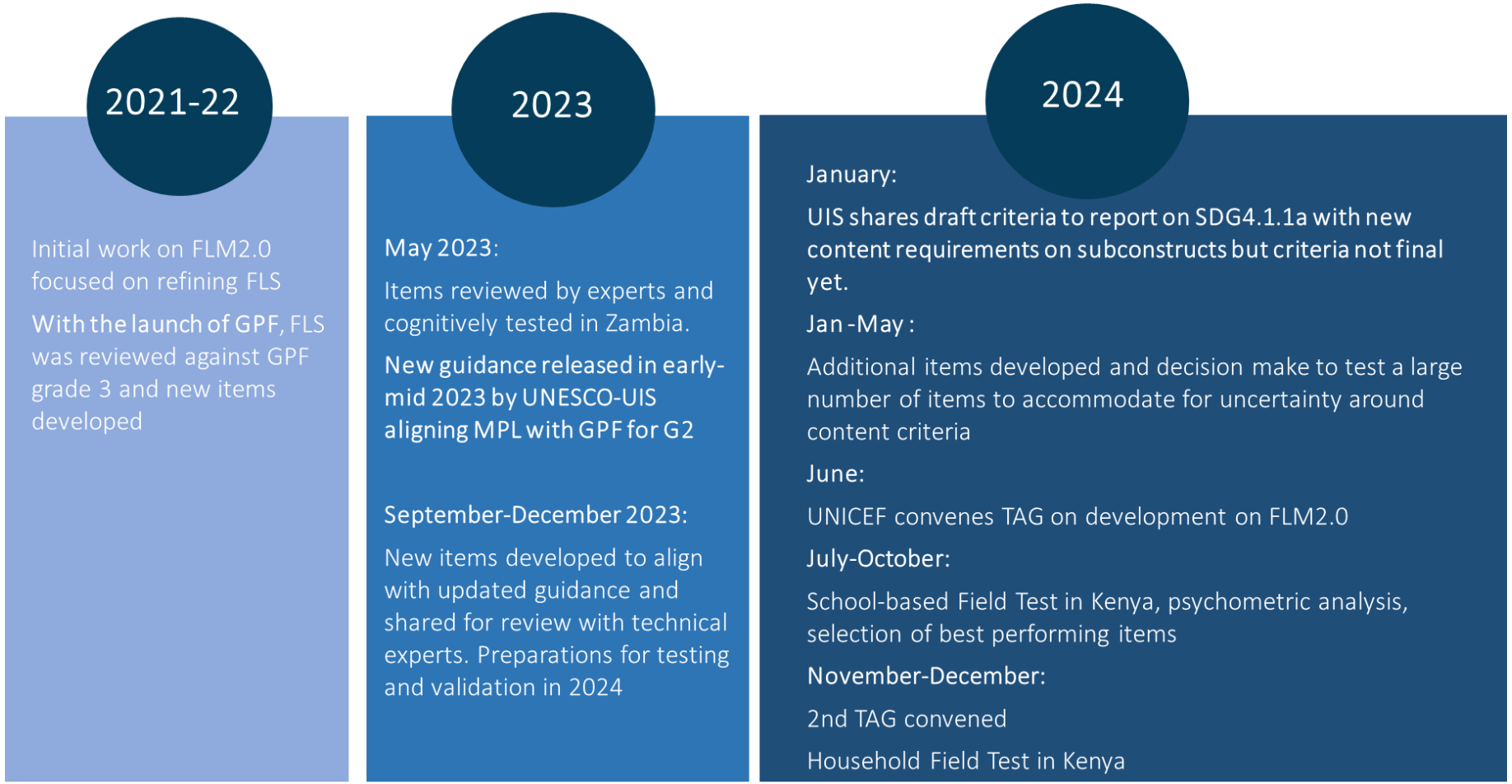
2nd Quantitative validation in Zimbabwe in schools. **Roll out of early primary module**

4 Foundational Skills Module 2.0 (FLM2.0)

Initially in 2021 Aligning with the Global Proficiency Framework (GPF) for grade 3

Adjustment in alignment: Based on guidance on Minimum Proficiency Levels (MPL) aligning with grade 2 GPF in 2023 & UIS criteria for SDG 4.1.1 reporting in 2024

Development: Supervised by a Technical Advisory Group (TAG) of experts, economists, psychometricians, and assessment specialists from ACER, AIR, World Bank, University of Amherst, University of Oregon



5 Foundational Skills Module 2.0 (FLM2.0):Criteria 1 Content

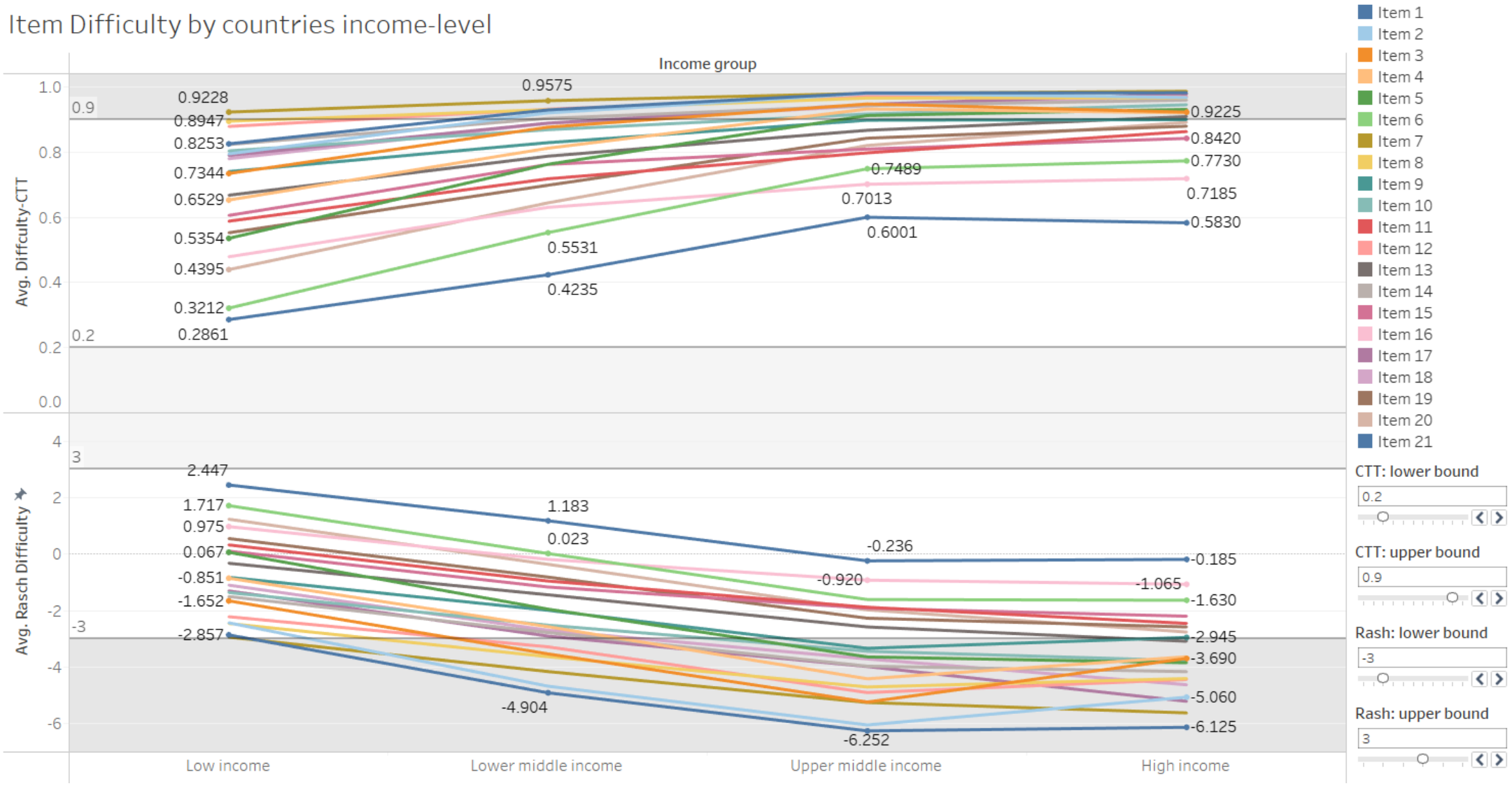
Module	Items by form	Domains	Items by form	Module
		Listening Comprehension	7	FLM2.0
FLS	1 long para	Decoding	2 short + 1 long para	
	5	Reading Comprehension	15	

Module	Items by form	Domains	Items by form	Module
FLS	10	Number and operations: Identify and count in whole numbers, and identify	6	FLM2.0
	5	Number and operations: Solve operations using whole numbers	5	
	5	Algebra: Recognize, describe, extend, and generate patterns	2	
		Number and operations: Represent whole numbers in equivalent ways	1	
		G1.1: Recognize and describe shapes and figures	2	
		G2.1: Compose and decompose shapes and figures	1	
		G3.1: Describe the position and direction of objects in space	2	
		M1.1: Use non-standard and standard units to measure, compare, and order	1	
		M2.2: Solve problems involving time	1	
		M3.1: Use different currency units to create amounts	2	
		N1.4: Solve real-world problems involving whole numbers	2	
		S1.1: Retrieve and interpret data presented in displays	1	

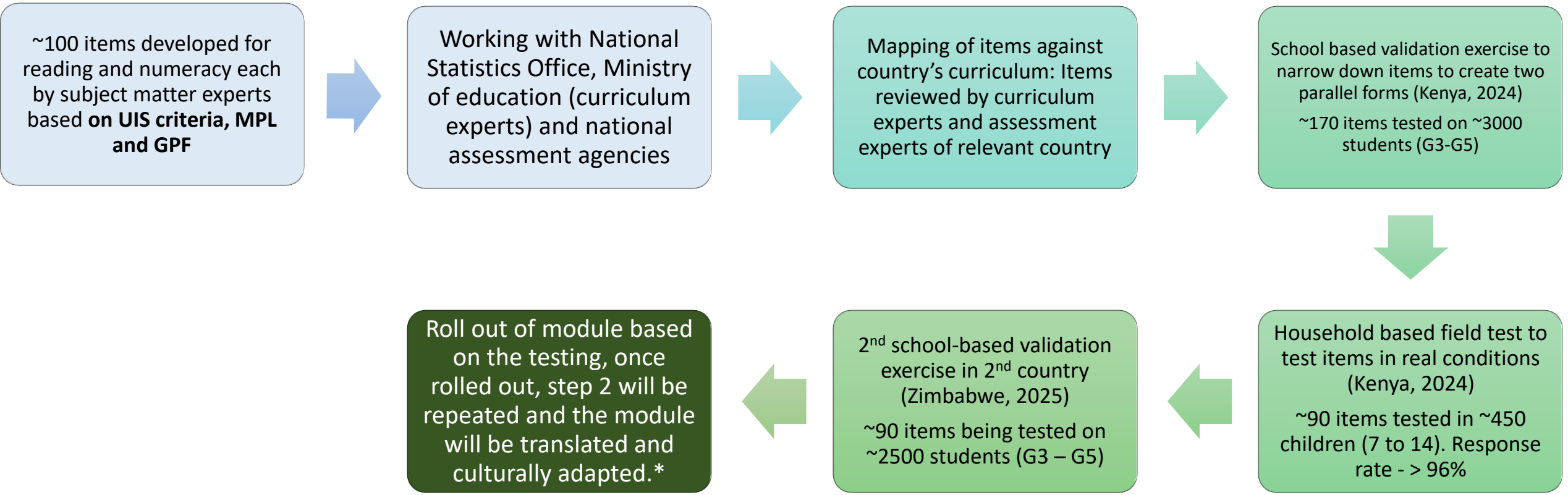
- ✓ **Wider** content coverage than FLS; some tasks like listening comprehension will be new for the MICS
- ✓ **Aiming to** create two parallel forms in reading and numeracy
- ✓ **Tested ~**70 reading and ~100 numeracy items to meet SDG content criteria
- ✓ **Form Validation:** The final two forms: reading contains 20+ items; numeracy has 24-25 items
- ✓ **Development** of items and content in English language

6. Criteria 2: A reflection on item performance based on FLS

Same items display different difficulties by level of effectiveness of the education system



7 FLM2.0: Criteria 2 Qualitative and Quantitative testing



**Items will be reviewed by local experts in implementing countries, translated, and culturally adapted. UNICEF will roll out the module based on testing the reliability in two countries, in each country reliability will be assessed ex-post (after data collection and before dissemination of results).*

7. Multiple Indicator Cluster Surveys (MICS) and FLM2.0: Criteria 3 Sampling

- MICS sample is based on probability sampling methods to generate representative and unbiased results.
- Typically, a two-stage design:
  - Enumeration Areas (EAs) selected from appropriate sampling frame (e.g., national census) using probability proportional to size.
  - Random systematic sampling used for household selection within listed EAs
- Every household and member has a known probability of selection
- Valid inferences to the population or subgroups are possible through weighted data
- Sampling errors can be estimated
- Oversampling of specific population subgroups or age subsets (e.g., under 5 children) for reliable indicators if needed
- FLS sampled one child of age 7 to 14. Final sampling for FLM2.0 is **under discussion (one child/ all children/ age bracket)**

7 FLM2.0: Criterion 3: Sampling efficiency and Fit for Purpose

Assessment Validation versus Reporting

Kenya school-based field test (July, 2024)

KNBS KENYA NATIONAL BUREAU OF STATISTICS *Keeping you informed*

REPUBLIC OF KENYA
MINISTRY OF EDUCATION, SCIENCE AND TECHNOLOGY

THE KENYA NATIONAL EXAMINATIONS COUNCIL

- ✓ Field test led by Kenya Bureau of Statistics (KNBS), Kenya Ministry of Education (MoE), Kenya National Examinations Council (KNEC) and UNICEF
- ✓ **Purpose**: Validate item performance based on sample focused around MPL (Minimum Proficiency Level) to gauge initial item performance against the reference population
- ✓ **Sampling**:
  - ✓ Grade Selection (Grades 3 to Grades 5) – Why? :
    - i. Field test conducted mid-academic year, therefore Grade 2 students would essentially have been in Grades 1 + (0.5) of grade 2 and so on.
    - ii. KNEC's review of alignment of items against Kenya's national curriculum: 45% of math items aligned with Grade 2-3 standards, 55% aligned with Grade 3-5 standards
  - ✓ Outcome based sampling:
    - a. Identifying two districts with enough schools to reach desired sample size to test 4 forms
    - b. Merging MoE's enrolment data with KNEC's examination data to create a list of school **(List of 992 schools)**
    - c. School and child eligibility criteria:
      - i. Minimum enrolment of 30 students per grade (Grades 3, 4, and 5)
      - ii. Exclusion of: All-boys, all-girls, and special needs schools
      - iii. Schools where at least 60% of students passed the Kenya Primary National Examination **(68 schools)**
      - iv. Variation in rural/urban location and in public/private management **(15 schools)**
      - v. Children randomly selected for administration of assessment from difference classes in each grade

10. FLM2.0: Criteria 4 and 5 Forthcoming releases and plans

- **Report from field tests :**
  - Report on item performance and administration of items summarizing the results

- **Scoring guide:**
  - Evaluate how revised FLM2.0 instrument align with a common scale

- **Field test and interviewers manual:**
  - MICS already has an elaborate manual, the FLM2.0 protocols will be added to these

- **Roll-out (Q2, 2025)**:
  - Countries who are in questionnaire design or earlier phases of MICS to start implementing FLM2.0
  - Updated module, guidelines and protocols
  - New template for each country including item and reliability statistics and other information required to meet the criteria on SDG reporting as much as possible

11. Recommendations for GAML

Requests GAML to consider the value of measuring learning outcomes in household surveys. These surveys reveal drivers of learning and the correlation across children, parents, communities, and schools while being cost-effective—one survey informs ~40 SDG indicators. They also enable age-based sampling, helping track learning gradients. Leveraging these data is essential.

Request GAML flexibility in implementing the standards so we create the much-needed market contestability in the learning assessment space and for governments to be able to track and report progress while strengthening national assessments

Request GAML to provide additional clarity on benchmarking criteria (which methodologies to use in different languages)

Request GAML to consider that when assessing item performance (criteria 2 and 5), it is important to consider what should be the valid population of reference. As the nationally representative population might itself be biased against the expected level of difficulty of the items.

unesco
Institute for Statistics

25 YEARS

Education Data and Statistics Commission

THANK YOU

Learn more:

uis.unesco.org

databrowser.uis.unesco.org

@UNESCOstat

#25YearsOfDataInsights

Global Alliance to Monitor Learning