MELQO Planning and Implementation Toolkit

Version-April 2018

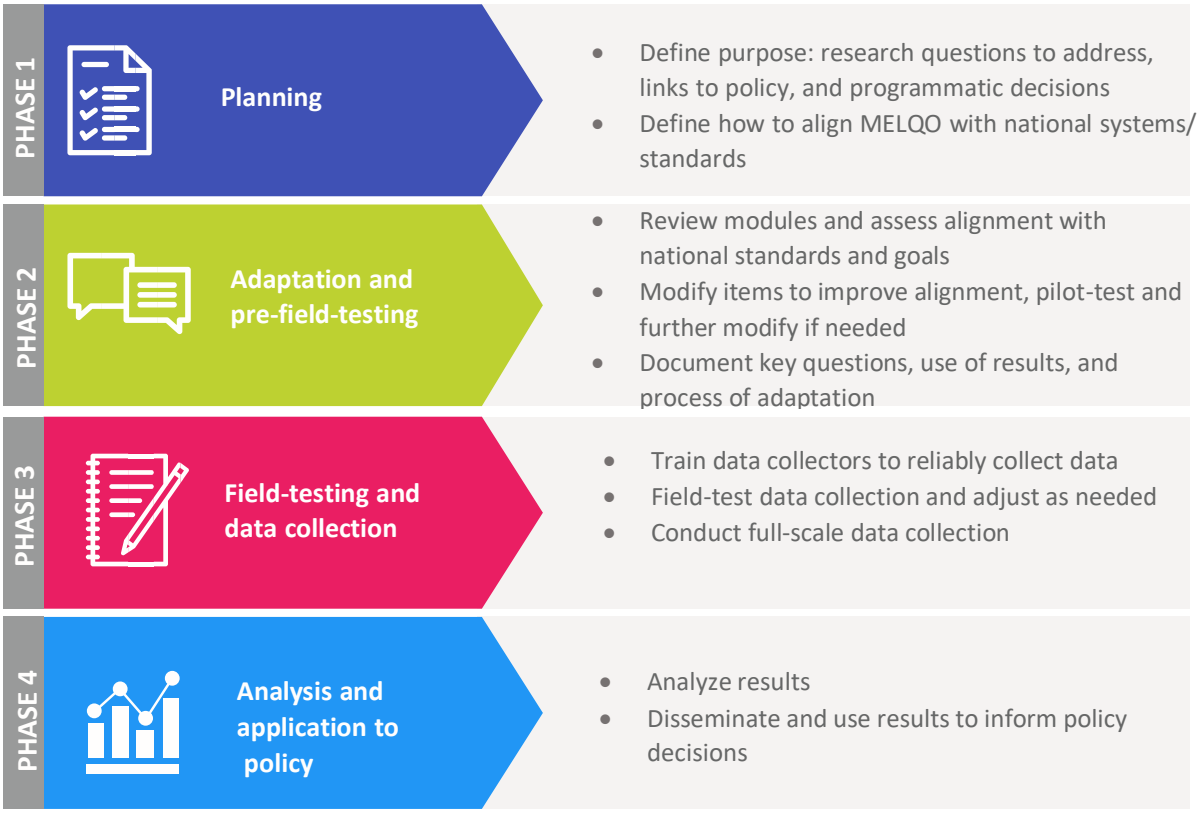
Overview of the process and resources for MELQO measurement studies

The Measuring Early Learning Quality and Outcomes (MELQO) modules are designed to provide country-level evidence on child development and learning, to meet country demands for national data and potentially to help inform global monitoring. MELQO includes a suite of 6 tools in two modules designed to measure child development and learning (MODEL) and quality learning environments (MELE). The MODEL Module of tools assess early learning and development through: 1) direct child assessments, and 2) surveys of parents and teachers. The MELE Module of tools assess early learning environments through 1) classroom observations, and 2) surveys of parents, teachers and school administrators. These multiple sources of information serve to create a more complete picture of child development and learning environment across multiple indicators.

Depending on the needs and goals of individual projects, the tools can be used together or separately as part of population surveys, national monitoring, impact evaluations, or program evaluations.[1]

Based on the experiences of the MELQO consortium so far, the activities involved in piloting the MODEL and MELE modules generally fall into four phases: 1) planning, which includes defining the purpose of the exercise and linking to existing systems; 2) adaptation and pre-field-testing; 3) field-testing and data collection; and 4) analysis and application to policy.

This toolkit provides practical planning and implementation resources for country teams using the MELQO modules. Supplementary resources can be accessed through registering for the MELQO portal at ecdmeasure.org.



[1] MODEL scores should NOT be used as Kindergarten or primary entrance exam or diagnostic or screening tools. MELE scores should NOT be used as evaluations of individual teachers.

PHASE 1

Planning and study design

During the planning phase, it is important to define the purposes for using the modules, including identifying the main research questions and determining how the data will inform policy and programmatic decisions.

The following pages include implementation resources for the planning stage, including: (a) description of core modules; (b) research questions and sampling considerations; (c) recommended capacity needed for a country-level MELQO study; (d) sample study timeline; (e) detailed study task list; (f) sample budgets for MODEL and MELE.

By the end of this phase, your team should have:

1. Key research questions defined
2. Sampling framework developed
3. Budget and team in place, with roles, responsibilities, and timeline defined
4. Plan for logistics in place: recruitment/section of sites for pre-field-testing, field testing, and larger study, arrangements for training, translation
5. Database system and how data will be collected (paper vs. tablet) and personnel needed (programmers; data entry workers)
6. Identified human resources needed (number and capacity of trainers, supervisors, and enumerators needed for study)

Supplemental Resources for Planning Phase

- MELQO overview presentation
- Sample ToRs

(Please visit ecdmeasure.org and register for the MELQO Portal to access these materials)

MELQO PLANNING RESOURCE: **CORE MODULES**			
Module	**Tool**	**Description**	**Materials**
MODEL	**Direct Assessment**	Collects information on children's early learning. *Approximately 25 minutes to administer.*	Core Direct Assessment Enumerator Booklet
			Core Direct Assessment Stimulus booklet
			Core Direct Assessment Manual
	Teacher survey	Asks teacher about children's behavior and learning in school. *Approximately 10-15 minutes to administer*	Core Teacher Survey
			Core Teacher Survey manual
	Parent/Caregiver survey	Asks parent our caregiver about children's behavior and learning at home, and collects information on family backgrounds and home learning environments. *Approximately 15-25 minutes to administer*	Core Parent Interview
			Core Parent interview Manual
MELE	**Classroom observation**	Collects information on classroom quality through direct observation *recommended 2 hour minimum*	Core Classroom Observation Enumerator Booklet
			Core Classroom Observation Manual
	Teacher Interview	Collects information from teacher on background, compensation, motivation. *Approximately 10-15 minutes to administer*	Core Teacher Interview
			Core Teacher Interview Manual
	Head teacher interview	Collects information from head teacher/director on program. *Approximately 5-10 minutes to administer*	Core Head Teacher Interview
			Core Head Teacher Manual

MELQO PLANNING RESOURCE:
RESEARCH QUESTIONS AND SAMPLING CONSIDERATIONS FOR STUDY DESIGN

Research questions

A central goal of the first phase of MELQO is to clarify what research questions will be addressed through use of the MELQO tools. Depending on the mix of tools used, research questions can include

- Are children reaching national standards for learning at the start of school? What skills do children have, and what skills require more work?
- Are there differences in quality of learning environments or in children's overall learning based on region or other characteristics?
- How do children and teachers spend their time in pre-primary classrooms? Are children using materials and if so, how are they used?

Equity considerations

A central reason for collecting data on early childhood development and learning is to determine patterns of equity, or the extent to which children from different demographic groups experience disparities in access to quality learning environments, and ultimately, in learning. Documenting patterns in equity is an important goal for the MELQO Global Project, and it is recommended that analyses are undertaken specifically to address patterns of equity in learning and access to quality early childhood programs.

Sampling

It is important to note that sampling plays a critical role in generating reliable results on equity. As the study is designed, decisions will be made on how to ensure adequate representation of schools and children. Two types of sampling frameworks can be developed:

1. **school-based sampling framework**, which is drawn from a list of schools maintained by the education ministry.
2. **household-based sampling framework**, which requires partnership with the national statistics offices to estimate population-level sampling that will yield representative samples of young children.

In both cases, sampling frameworks are based on available information from government ministries that ideally contain the full population of schools or children. School-based sampling has generally been more feasible than household-based sampling in MELQO projects to date. In many countries, information on all schools in a region or country is available, which allows creation of a sampling frame that will allow comparisons between regions, or if information is available, on public vs. private schools. Once the data have been collected, this approach allows for representative findings from all children enrolled in school, and comparisons between regions and/or private vs. public schools, if schools are evenly selected by region or private vs. public status.

Sampling influences the analyses of the data for equity, because samples that are more representative will lead to greater confidence in the findings on differences based on region, private vs. public schools, and other factors that may be associated with equity (see equity data analysis guidelines in Data Analysis section on page 28.

MELQO PLANNING RESOURCE:
RECOMMENDED TEAM BACKGROUND/EXPERIENCE

Role	Basic Scope of Work	Ideal Qualifications and Experience
Project coordinator *Note: Project coordinator scope of work may also be taken on by team (to cover scope of technical ECD inputs and coordination of logistics)* *Project Coordinator may also take on Master Trainer role, depending on competencies*.	Oversee study implementation and coordinate full MELQO process, from adaptation, to pilot, to full survey. Tasks will include management and hiring enumerators, support training logistics, travel logistics to schools for data collection, operational plan for data collection, oversight of data collection, and compiling database from pilot and full survey. Also responsible for liaising with national stakeholders and (usually MoE) and MELQO team.	◦ Advanced degree, preferably in ECD, education, public health, or related field ◦ Excellent management and organizational skills and strong quantitative skills ◦ Experience living and conducting field research in developing country strongly desired ◦ Ability to present positions and to negotiate with senior officials and manage high-level relationships with partner organizations ◦ Fluency and excellent verbal and written communication skills in English (and local language, if applicable) ◦ Knowledge of Stata, (strongly preferred), SPSS, SAS, R, or other data analysis software. ◦ Well organized, detail-oriented, able to prioritize, and manage multiple tasks simultaneously with minimal supervision. ◦ Project management experience, as well as experience supervising, designing, and implementing data collection and/or field work activities, and managing teams of field workers, is strongly preferred. ◦ Experience in designing survey questionnaires, and analyzing quantitative survey data, is preferred ◦ Familiarity with national context, ministries and other stakeholders ideal/desired
Master trainer/ECD expert	Review tools and lead adaptation process; pilot and revise tools, train enumerators and supervisors, oversee the data collection process; provide advisory support during data analysis	◦ Advanced degree in early childhood development ◦ Experience in applied qualitative/quantitative research techniques, particularly in direct assessment ◦ Experience in training and supervising enumerators in field work ◦ Familiarity with local ECD context preferred ◦ Experience working with children ◦ Proficiency in English ◦ Proficiency in local language strongly desired
Translator	Translates tools and training manuals (will be required before piloting and after revisions)	◦ Experienced translator fluent in English and local language ◦ Experience translating documents for education/child development settings strongly desired
Enumerator	Collects data through direct child assessment, teacher/caregiver interview, teacher and head teacher interview, classroom observation	◦ Experience in data collection surveys ◦ Knowledge of the local environment and education settings ◦ Ability to collect and gather information in an objective way ◦ Ability to communicate verbally and written in local language ◦ Experience/knowledge of early childhood development strongly desired ◦ Experience working with children ◦ Basic understanding of applications on mobile technology (if study requires data collected on tablets)
Statistician/Psychometrician	Oversees data analysis	◦ Psychometric expertise ◦ *Experience with data analyses* relevant to the construction and evaluation of scales (e.g., IRT, CFA, measurement invariance), preferably within the field of child development ◦ Experience with national learning studies in low- and middle-income countries preferred
ECE Stakeholder group	Provides input on adaptation process and application to policy. Essential for contextualization and political buy-in. Some countries establish formal MELQO Technical Committees to lead national process, others convene stakeholders at each phase.	◦ Ministry of Education staff/ Gov't officials, ECD researchers, policy experts, experts in child development and measuring quality, other partners and stakeholders (NGOs, research teams, higher education, etc.)

MELQO PLANNING RESOURCE:
Sample timeline for MELQO study

Steps	Task	Weeks																				
		1-2	3-4	5-6	7-8	9-10	11-12	13-14	15-16	17-18	19-20	21-22	23-24	25-26	27-28	29-30	31-32	33-34	35-36	37-38	39-40	41-42
1. INITIAL PLANNING	**1.1** Background research on ECD measurement	■																				
	1.2 Determine research questions & identify partners, logistics	■	■																			
	1.3 Review core versions of MELQO		■																			
2. STUDY DESIGN	**2.1** Determine sampling framework		■	■																		
	2.2 Decide database systems			■																		
3. ADAPTATION/ PRE-FIELD TESTING	**3.1** Align MELQO with ELDS/national priorities		■	■																		
	3.2 Hold Adaptation Workshop			■																		
	3.3 Complete adaptations				■																	
	3.4 Translations					■																
	3.5 Pre-field testing						■															
4. PILOTING/ FIELD TESTING	**4.1** Prepare and train for piloting						■	■														
	4.2 Complete piloting								■													
	4.3 Analyze pilot data									■												
5. PREPARE FOR FULL DATA COLLECTION	**5.1** Revise tools										■											
	5.2 Finalize tools and support materials											■										
	5.3 Identify and hire data collectors										■	■										
	5.4 Recruit and select programs										■	■										
	5.5 Finalize database management & collection procedures											■										
	5.6 Conduct training workshops												■									
6. NATIONAL STUDY/ FULL DATA COLLECTION	**6.1** Collect data													■	■	■	■					
	6.2 Data entry and cleaning																■	■				
	6.3 Analyze data																		■	■		
7. REPORTING & DISSEMINATION	**7.1** Reporting										■										■	■
	7.2 Disseminate findings and follow-up																					

MELQO PLANNING RESOURCE:
DETAILED TASK LIST FOR NATIONAL MELQO STUDIES

Steps	Task	How to complete it	Why it is needed
1. INITIAL PLANNING	**1.1** Background research on MELQO & ECD measurement in your context	• Use existing research and/or data to describe present levels of child development and quality in learning environments. • Learn about the overall structure and background of the MELQO tools. • Complete mapping of relevant data projects, summarize findings.	• Helps identify areas for measures to cover; • Rough estimates of ages at which children might achieve items if child development data is available
	1.2 Determine research questions and develop general project plan	• Using existing data (and gaps in data), determine what information can be gathered using the MELQO tools and how the data may be used. • Determine what questions need to be answered based on existing problem and need. • Develop a general plan for project, including timeline, key partners, and key roles (field coordinator, master trainer, enumerators, translators). In some cases, it may be helpful to establish a MELQO steering committee to ensure government ownership and buy-in throughout process.	• Helps identify which of the MELQO tools to use and plan for how data will be used
	1.3 Review Core versions of MELQO tools and make draft notes on any anticipated changes	• Convene research team to do an initial scan of the measures to determine if items are appropriate; may include some site visits to see if tools align with what is happening in programs	• Provides basis for next phases of work
2. STUDY DESIGN	**2.1** Determine sampling framework	• Sampling framework should be determined for both pilot and full study. • Statistician/sampling expert should provide input based on study design and research questions.	• Ensure representative sample, increase confidence in findings related to equity
	2.2 Database system decisions	• Determine how data will be collected (paper vs. tablet) as well as where it will be stored; set up data entry databases; establish how data will be checked and entered; determine how identification numbers will be assigned	• Minimize errors in data collection and data entry, limit missing data, increase confidence in integrity of data
3. ADAPTATION & PRE-FIELD TESTING	**3.1** Alignment of curricula, standards or policies with MELQO	• Obtain curricula and/or Early Learning & Development Standards; complete table demonstrating alignment (see examples starting on page 13 below). • If some areas are not currently covered by MELQO tools identify the priority of adding items to address these areas during adaptation. • If there are items on MELQO tools not aligned, mark those for consideration of elimination during the adaptation process.	• Measures should be aligned to curriculum standards; if Ministry or other partners have priorities within the standards, this should be identified
	3.2 Hold Adaptation Workshop	• Arrange meeting with MOE, other partners and research team to review priorities for adaptation. • Emphasis should be placed on defining top priorities, clarifying any mismatches with culture and context; making recommendations for changes to wording and content of tools. May include site visits and some limited testing of items.	• Important for ensuring appropriateness of measures and buy-in from MOE and other partners

	3.3 Complete adaptations	• Review results from workshop and make changes to measures (words, terms adapted to cultural context)	• Create an appropriate set of MELQO tools to be piloted
	3.4 Translations	• Once tools are finalized, translate tools and manuals.	• Tools should be in local language(s) of instruction
	3.5 Pre-field testing	• Adapted modules should be pre-field tested on a small sample • Pre-field testing should be conducted by reliable and trained enumerator (may be project coordinator or trainer) • Recommended 25-50 children and teachers/parents for MODEL and min 10 classrooms for MELE to identify any major problems with the items.	• Ensures workability of measures and process; accurate estimates of time, gather materials required for training
4. FIELD TESTING/ PILOTING	**4.1** Prepare and train for field-testing/piloting	• Determine pilot sampling framework and recruit sites to participate in the pilot study/ field-test (at min 20 classrooms for observation, 20 teachers/parents for surveys, 200 children) • Identify enumerators and train on the MELQO tools; • Pilot enumerators should have experience assessing children and should be trained on tools, manuals, MELQO materials	• Prepare to collect pilot data to inform national study
	4.2 Complete field-testing/piloting	• Take tools and complete in field. • Document each item using pictures and videos for training (see list) • Identify areas for further revision and estimate time required for each tool to be completed. Collect feedback on: (a) how well items are working (which items are confusing for parents, teachers, children, or enumerators; what procedures worked well and what procedures are needed for national sample; what areas/items need further attention during training?); and (b) enumerator reliability: enumerator pairs should double score each of the tools for at least 20% of cases to determine inter-rater reliability and potential issues	• Establish distribution of scores on scales of instrument to allow for psychometric analyses and recommendations for revisions prior to final data collection
	4.3 Analyze pilot data	• Enter data into database and clean (check for errors, merge data); • Determine how well database systems works and changes needed; • Pilot data should be analyzed to inform additional tool adaptations/revisions. • Report to be prepared of pilot results and findings	• Informs revisions of tools
5. PREPARE FOR FULL DATA COLLECTION	**5.1** Revisions of tools	• Use pilot results to modify tools as needed; examples of changes could include modifications to difficult items and/or changes in item wording or structure	• Ensures workability of measures
	5.2 Finalize tools and support materials	• Create final version of tools and translate/back translate into language of assessment and surveys. • Ensure full alignment of training manuals to tools; • Document each item using pictures and videos for training. • In some cases a tool approval/validation workshop may be required for political sign-off.	• Prepare for national study
	5.3 Identify and hire data collectors	• Consider hiring more than needed in case some are not able to achieve reliability;	• Prepare for national study

		• Determine which tool each data collector will learn; • Create teams of data collectors and assign to a supervisor	
	5.4 Recruit and select programs	• Select and notify programs based on sampling framework; • Notify/send letters to programs	• Prepare for national study
	5.5 Finalize database management and data collection procedures	• Prepare all materials for data collection including paper copies or use of technology, assign identification numbers and scheduling	• Prepare for national study
	5.6 Conduct training workshops	• In advance of training, secure location for training workshops and other materials (projector, AV equipment for viewing videos)	• Prepare data collectors
6. NATIONAL STUDY/FULL DATA COLLECTION	**6.1** Collect data	• Collect data following procedures established	• Collect data
	6.2 Data entry and cleaning	• Enter data (if paper data collection) and clean data	• Determine results from study
	6.3 Data analysis	• Analyze data	• Determine results from study
7. REPORTING & DISSEMINATION	**7.1** Reporting	• Produce report on findings. In some cases, a comprehensive technical report could be complemented with shorter policy briefs for national stakeholders	• Summarize findings from study
	7.2 Dissemination and follow-up	• Dissemination and follow-up discussions on application to improved policies/programs is recommended at various levels (high-level officials, teacher training institutions, regional/district levels, parents, etc.)	• Application of findings from study

Pre-field-testing refers to the process of testing the adapted modules on a small sample of 25 to 50 children and teachers/parents to identify any major problems with the items.
Piloting/Field-testing refers to data collection among a sample of about 200 children, to fully map the amount of time it takes to administer the instrument, examine the psychometric properties of the instrument, and make modifications as needed before large-scale data collection occurs.
Full data collection refers to data collection on the full sample of children.

MELQO PLANNING RESOURCE:
SAMPLE BUDGET FOR MODEL

	Task	Estimate of time needed	Country A (project evaluation)	Country B (nationally-representative study)
Preparation	Preliminary meeting	1-4 days, including planning and execution	US$500	US$1,000
Adaptation and pre-field testing	Expert time for adaptation and pre-field testing[a]	10 days (often ~$400-600/day)	US$4,000 – US$6,000	US$5,000 – US$7,000
	Local meeting(s) (including per diems, space, pre-field test, etc.)	5 days	US$1,000 – US$3,000	US$2,000 – US$5,000
	Translation of tools	Depends on language, length of tool	US$2,000 – US$4,000	US$2,000 – US$4,000
Training	International expert time for training (Master trainers)	10 days (often ~$400-600/day)	US$3,000 – US$5,000	US$4,000 – US$7,000
	Travel (experts, TTL, etc.)	1 week mission for each participant	US$4,000 – US$20,000	US$4,000 – US$20,000
	Space rental, materials, etc.	Depends on country	US$500 – US$2,000	US$500 – US$2,000
Data collection, analysis, and dissemination	Data collection (transport, enumerators' salary/per diem,[b] etc.), materials (tablets, questionnaires, etc.)	Depends on country	US$30,000 – US$45,000 (sample size[c] = 80 classrooms)	US$120,000 – US$200,000[d] (sample size = 600 classrooms)
	Data analysis[e] and report writing	~4-6 weeks	US$4,000 – US$10,000	US$5,000 – US$20,000
	Dissemination	1 event ($1,000-3,000)	US$1,000	US$2,000
Other	General TTL travel	1-4 missions in one year	US$5,000 – US$20,000	US$5,000 – US$20,000
TOTAL			US$55,000 – US$126,500	US$150,500 – US$288,000

a Includes time to update tools and prepare manual, if needed.
b Depends on enumerators' level of training required.
c Note this is NOT a representative sample.
d Note this amount can vary widely depending on the country context, transportation costs, etc.
e Can include psychometric analysis, statistical analysis, etc.

Note: Budgets and timelines can vary significantly (from US$60,000 - US$500,000) depending on the country context, tool, sample size, and, especially, the purpose of measurement. For example, measurement for project evaluation would be slightly less expensive than measurement for a nationally-representative study, which involves more and longer decision-making steps with government as well as more adaptation to reflect national curriculum and standards. Table 3 shows the types of budget items needed to prepare for and implement an early childhood assessment, with examples from both a project evaluation and a nationally-representative study.

MELQO PLANNING RESOURCE:
SAMPLE BUDGET FOR MELE

	Task	Estimate of time needed	Country A (project evaluation)	Country B (nationally-representative study)
Preparation	Preliminary meeting	1-4 days, including planning and execution	US$500	US$1,000
Adaptation	Expert time for adaptation and pre-field testing[a]	10 days (often ~$400-600/day)	US$4,000 – US$6,000	US$5,000 – US$7,000
	Local meeting(s) (including per diems, space, pre-piloting, etc.)	5 days	US$1,000 – US$3,000	US$2,000 – US$5,000
	Translation of tools	Depends on language, length of tool	US$2,000 – US$4,000	US$2,000 – US$4,000
Training	International expert time for training (Master trainers)	10 days (often ~$400-600/day)	US$3,000 – US$5,000	US$4,000 – US$7,000
	Travel (experts, TTL, etc.)	1 week mission for each participant	US$4,000 – US$20,000	US$4,000 – US$20,000
	Space rental, materials, etc.	Depends on country	US$500 – US$2,000	US$500 – US$2,000
Data collection, analysis, and dissemination	Data collection (transport, enumerators' salary/per diem,[b] etc.), materials (tablets, questionnaires, etc.)	Depends on country	US$30,000 – US$45,000 (sample size[c] = 80 classrooms)	US$120,000 – US$200,000[d] (sample size = 600 classrooms)
	Data analysis [e] and report writing	~4-6 weeks	US$4,000 – US$10,000	US$5,000 – US$20,000
	Dissemination	1 event ($1,000-3,000)	US$1,000	US$2,000
Other	General TTL travel	1-4 missions in one year	US$5,000 – US$20,000	US$5,000 – US$20,000
TOTAL			US$55,000 – US$126,500	US$150,500 – US$288,000

a Includes time to update tools and prepare manual, if needed.
b Depends on enumerators' level of training required.
c Note this is NOT a representative sample.
d Note this amount can vary widely depending on the country context, transportation costs, etc.
e Can include psychometric analysis, statistical analysis, etc.

PHASE 2

Adaptation and pre-field-testing

The MELQO modules should be aligned with national goals and cultural expectations for young children and early learning environments. Each country should follow an adaptation process through which core items can be adapted to country/cultural/language context and any new country-specific items can be added. Before finalizing tools for a larger study, they should be pre-field-tested to understand how MELE items work in local classrooms and MODEL items work with local children.

The following pages include implementation resources for this stage, including: (a) country adaptation/alignment samples; (b) MODEL adaptation guidance; (c) additional data to gather during pre-field test (d) tips for videotaping classroom observations.

By the end of the adaptation and pre-field testing phase:

1. Tools aligned to national/programmatic standards, may include additional items deemed necessary for the goals of assessment
2. Instruments and manuals accurately translated (and back-translated) into language(s) of assessment, maintaining integrity of items
3. Tools have been pre-field tested to reflect on how tools and data collection process will work in larger study (see MELQO Resource below)
4. Local classroom videos and photographs for training collected (see MELQO Resource below)
5. Tools validated by stakeholders, who agree on entirety of instruments as necessary and sufficient in pursuit of the purposes of the assessment
6. Report findings on pre-field testing, including recommendations on revisions to adapted tools
7. Training manuals revised and videos from pre-field testing process incorporated into training materials

Supplemental Implementation Resources for Adaptation and Pre-Field Testing

- MELQO Adaptation and pre-field testing presentation
- MELQO Brief 2: Literacy and Language
- Permission form for photography and video use

(Please visit ecdmeasure.org and register for the MELQO Portal to access these materials)

MELQO ADAPTATION RESOURCE:
COUNTRY ALIGNMENT EXAMPLES FROM NATIONAL ADAPTATION WORKSHOPS

Below are samples from country adaptation workshops to align MELQO tools with national pre-primary curricula, syllabi, and/or standards.

Example 1: MELQO items mapped to national curriculum competencies

Curriculum Competency	Activity performed by child	Alignment with MELQO (MODEL DA)	Workshop discussion & agreements
Listening	Listening to songs, conversations and short stories	#16 Listening Comprehension Story: Listen to story and answer questions (listen go stories; give intended message)	This domain is well covered.
	Listening to instructions/directives	Throughout the assessment #18 Head, Shoulders, Knees & Toes: Follow instructions	
	Playing an information sharing game	X	
Speaking	Discussing ; Expressing ideas	#12 Expressive Vocabulary: Name food items and animals (express himself/herself)	Group recognizes it is helpful to look at pronunciation, but the group recognizes it is difficult to find an items that measure if children are pronouncing words correctly (there are different regional accents)
	Identifying different riddles	X	
	Singing different songs	X	
	Telling different stories	X	
	Explaining daily activities	X	
	Expressing things he/she likes/dislikes	X (with links to Item 12 above)	
Mastering Pre-reading Skills	Reading pictures	X	- It may be difficult for some children to name things from their environment if they are from poor families they do not have access certain items. - Telling stories would be nice to include, but it is difficult to measure this- but should see if there are any items for this
	Developing understanding about books/publications	X	
	Identifying different names/words in his/her environment	X	
	Identifying first sounds of names of people/things in his/her environment	#13 Initial Sound Matching;	
	Mentioning names of things whose first sounds are similar	#14 Initiation Sound Identification	
	Recognize sounds of vowels and consonants; recognize and identify first sound	#13 Initial Sound Matching; #14 Initiation Sound Identification: Phonemic awareness and phonics	
		#15 Letter Name Knowledge	
Mastering Pre-writing Skills	Doing exercises that strengthen arm and finger muscles	#17 Name Writing: Write own name (in any way the child knows) #21 Copying: Copy shapes and lines (hold writing instruments; draw lines)	- Well aligned
	Doing exercises of using drawing and writing instruments		
	Doing exercise of drawing lines from left to right and from top to bottom		
	Moulding/constructing shapes of vowel	X	
	Tracing vowel shapes; writing of vowels step-by-step	#17 Name Writing: Write own name (in any way the child knows)	
	Moulding/constructing consonant shapes	X	
	Tracing consonant shapes; Writing consonants step-by-step	#17 Name Writing: Write own name (in any way the child knows)	

Example 2: National curriculum competencies mapped to MELE Classroom Observation Items		
MELE Classroom Observation Item	**Selected competency statements**	**Adaptation and alignment notes**
The teacher implements activities to support development of literacy skills (letter identification, phonics, vocabulary).	10, 11, 13, 18	• There are many types of literacy activities that could place, suggest that enumerators still indicate which literacy competency was observed • As the literacy activities are related, it is better to use the consolidated item (there won't be any loss of information by consolidating)
expressive language skills through telling stories, having conversations or describing events or objects.	5, 8, 14, 15, 16	• Good item, keep
listening and speaking skills using an age-appropriate illustrated storybook with text (listening comprehension & oral language)	9, 17	• Good item- storybook reads included in syllabus
Art	6.1, 6.3, 6.6, 6.7,	• Syllabus includes art and creativity- item aligns well • However- scale could be adjusted: In O class syllabus- competencies are very specific- every child should draw something, cut specific shapes- so a "4" scale may not be appropriate for context- "3" may be shifted to a "4" • Should be qualified to be clear what an art activity is, as per the syllabus expectations.
Science and nature	5.1, 5.3, 5.6, 5.8, 5.10, 5.11, 5.12, 5.13	• Good item, aligns well with syllabus "development of environmental knowledge" • Should be qualified to be clear what a science/nature activity is, as per the syllabus expectations.
Learning activities that promote free play or open choice	Indirect. Relevant to many competencies	• Keep item-Teacher guide calls for 1 period for 25 minutes of free play • Better to make this specific • Examples of free choice and open play should be provided
Health and socio-emotional	4.1, 4.3, 4.4, 4.7, 4.8, 4.9, 4.13, 1.17, 1.19	• Aligns well with "Individual, Social, and Emotional Development" • Better to make this specific • Examples of health and socio-emotional activities should be provided
Learning opportunities that allow children to engage in Music/Movement activities	1.14, 6.9, 6.10, 6.11, 6.12, 6.14	• Aligns very well with teacher guide/syllabus. • Better to make this specific • Examples of music and movement should be provided
Gross motor and physical activity	1.3, 1.6, 1.7, 1.8, 1.14, 6.10, 6.14	• Good as is. • Better to make this specific • It may be useful to use examples to clarify gross motor (as writing was used for fine motor skills) • Examples of gross motor and physical activity should be provided

MELQO ADAPATION RESOURCE:
MODEL ITEM ADAPTATION GUIDANCE

Direct Assessment Adaptation guidelines

Literacy items (*also see MELQO language and literacy brief)*

Literacy items will respond to the linguistic structure of the language; therefore the items must be adapted for each language. The following steps are recommended:

1. Preparation

- Determine the type of script (e.g. alphabetic, alphasyllabic, abjad).
- If possible, determine the frequency of letters or syllables (will be needed for letter name task).
- If possible, identify a bank of 50 common words (likely to be known by pre-school children).

2. Construction and adaptation of items

- Literacy interest.
  Check the face images to make sure that their intended meaning is consistent with children's interpretation. This requires field testing with a sample of children. It is recommended that if there is any concern that a child does not understand the task, add one or two sample items in which the assessor asks the child how (s)he feels when they are eating something delicious for happy or ask how the child feels if they lost their favorite toy.
- Expressive language.
  For this task, the literal translation of body parts is one step. For this task, the assessors need to understand the specific instances in which they need to query the child further (e.g., mouth/lips for tooth).
- Expressive vocabulary.
  This is a semantic fluency task, and it is important when adapting or translating this task that the category of items selected is one in which young children from a range of socioeconomic backgrounds will be able to generate an adequate number of responses. Animals and things one can eat should be categories in which many responses are likely to be generated by young children as long as they are not intimidated by the test context or by the adult assessor. This task was attempted as part of a different assessment with primary grade children in a country in West Africa with poor results, because the children were reluctant to generate responses. This is a significantly different task than asking a direct question, and it is important that the assessor establish a good rapport with the child before initiating it.
- Letter identification.
  Use the frequency list to identify the easiest and hardest letters appropriate for beginning readers (in alpha-syllabic languages, children would not study some graphemes until second grade, so the letters should be drawn from those they are expected to learn in first grade). Randomly list the easiest letters in the first column and the most challenging in the second column.
  - For languages that read from right to left, reverse the columns; put the easiest letters in the right column and the most challenging ones in the left column, and present to the child in that order.
  - If a child begins by telling the assessor the sound the letter makes, the assessor should prompt by saying, "Tell me the name of the letter and not the sound it makes."
- Letter sound identification.
  This task should only be administered if the team training the assessors in the new language has evidence that the assessors can reliably hear and reproduce the sound of each letter of the alphabet. In some languages, there is no distinction between the sound and the name of the letter, but when the sound differs from the name, letter sounds may not have been taught.
- Initial sound discrimination.
  This is likely to be very difficult for some young children as many cultures do not teach children the sounds they hear in words in preschool or later in the primary grades. Sufficient practice is very important to ensure that the child understands the task. More than one practice item is recommended.
- Listening comprehension.
  The story for this task should be carefully evaluated during the adaptation phase to ensure that it is appropriate for young children in the context with the newly adapted tool. Assessor training is also very important for this task as the story should be read at a reasonable pace with appropriate prosody in a manner

that is engaging to the child. The translation must not result in a passage that is much more difficult or easier than the original passage, and it must have generally the same meaning across cultures.
The approximate length of the story for the listening comprehension task should be maintained. However, the length of a passage in an agglutinative language may be longer but with the same meaning due to the length of the words, so length can vary to some degree. In English, this story contains 110 words. The number of words may vary by language, but do not shorten or lengthen the content of the story (add or delete story events) just to get the same number of words. Also maintain the transitions ("after a while," "so"), since those help children remember the sequence.
The story can be adapted in the following ways:
    - The two animals can be changed to two animals that are common in the country.
    - The story opening can be changed to what is commonly used in the language.
- Name writing.
  Check to see if the child knows how to hold a pencil or piece of chalk and appears to know how to use it prior to beginning this task. A warm-up would be helpful in which the assessor asks the child to draw a line or mark on the page to observe the child's level of skill in using a writing implement.

Math items

- Verbal Counting
  Child can self-correct (e.g. 'four, five, seven, oh I mean six, seven, eight…'). Stop the child if they state a number in incorrect order or once they reach 30 (you can interrupt gently and say, 'Thank you, now I have another game for you' or something like that).
- Number Identification
  For languages that read from right to left, reverse the columns and present to the child in that order. Stop rule is at 5 incorrect consecutively. The font in which the numerals are printed must reflect the local script/print.
- Producing a Set
  Use small countable objects – preferably objects that are used as counters in the school setting. If testing very young children, ensure that the objects are not a choking hazard. Do not use food. If the child cannot give you 3 of the 20 and cannot give you 6 of the 20, do not proceed to 14 (child must miss both in order for the stop rule to be engaged
- Number Comparison
  Because the terminology changes from *greater* to *smaller*, emphasis should be placed on these words in this task. Note that the translation of the words *greater* and *smaller* should be to those specific terms in the language that are related to quantity (not size). Children are not shown the numerals in this task.
- Mental Addition
  The counters should reflect objects that the child is familiar with (and if possible has used in settings as counters). The word *balls* can be replaced with other more familiar objects if necessary.
- Measurement Vocabulary
  If any of the pictures are objectionable or unfamiliar in the culture, choose a similar object/animal and place sizes in identical locations as original (the correct answer is in alternate positions in the four items).
- Spatial Vocabulary
  Items in this picture can be replaced with items that are more familiar to children, if necessary. In adapting to the local language, both easy (above, under) and more difficult (next to, in front of) terms should be included.
- Mental Transformation
  Finger motions indicating that the shapes are to be imagined as joined (like a pinching motion with thumb and index finger) can be used. However, the procedures should be standardized across all children.

Social-emotional items

- Perspective-taking/empathy
  Child should look like a group member

Teacher/Caregiver Report Adaptation guidelines

Parent Report

1. Determine the appropriateness of the items in the background section (e.g. caregiver education categories).

2. Establish relevance of the assets listed and remove/adapt/add when required.

Parent and Teacher Reports

1. If DA-matching literacy and numeracy items are used, ensure they are the same in both assessments.

2. In the adaptation workshop, organize a group activity in which every participant/expert reads each question and
responds with (1) whether they understood it and (2) whether they would be able to answer about a child 3 to 6 years old. The items that are deemed inappropriate or difficult to answer should be discussed and a consensus needs to be reached. Note that often inaccurate translation is a source of misunderstanding. Also please note that definitions of children's social behaviour and emotional responses should be carefully considered within the group to ensure that the items accurately describe children's behaviour within the given context.

3. Records need to be kept of all comments and suggestions.

4. If feasible, a few (3–5) parents/teachers should be interviewed using all items, including those that were revised.

5. The potential comments from that sample group should be incorporated into the feedback on the TCR.

6. The version agreed upon through this process must be back-translated and compared with the original.

MELQO PRE-FIELD TESTING RESOURCE:

ADDITIONAL DATA TO GATHER IN PRE-FIELD TESTING

In addition to collecting data from the MELQO tools, the pre-field testing phase offers an opportunity to reflect on the tools and data collection process to improve the implementation during the larger study. To benefit fully from piloting, we recommend collecting the following information.

General reflections – for all tools

1. General comments on the logistics, including scheduling with programs, coordinating materials, transportation, etc.? What is needed to improve this for the national study?
2. What concepts should the trainer pay particular attention to during training?
3. How did the data collection and data entry process work? What revisions or clarifications are needed?

For the Direct Assessment

For all direct assessments, collect:

1. Time it took to administer.
2. Items that seemed confusing to the child.
3. Items that the enumerators had questions about (either for administration or scoring).
4. Suggestions for rewording items or clarifying procedures.

For at least 10 administrations, please collect the following for use in training:

1. Videotape the full administration of the direct assessment for 10 children, with permission from parents and when this does not interfere with the administration (for example, if it is too distracting for the child, do not videotape). Scan the corresponding enumerator booklet with scores for all items.
2. Select samples of children's responses for the writing (name writing and copying) items that represent a variety of abilities.
3. Photographs of materials used for the counters showing the variety of materials used.

For the Classroom Observation

For all observations, collect:

1. Time it took to administer. We recommend at least 90-120 minutes – was this enough or too much time to see all items on the tool?
2. Items that the enumerators had questions about.
3. Suggestions for rewording items or levels.
4. Were the levels appropriate and observable?
5. List of anything notable about the classroom context that wasn't captured by the tool

For ~15 observations, please collect the following for use in training:

1. Videotape the full observation, with permission from programs, teachers and parents. Select a variety of levels of quality and different models. Match these with scoring of items so that a range of quality is shown and identify clips to use for training. Along with the video, scan the corresponding observation scoring sheet. Examples of the following are especially important:
    a. Book read or story telling
    b. Free play
    c. Different methods of instruction
    d. Teacher-child interactions

*Note, a minimum of **five** 30-60 minute videos will be needed for training. It is recommended that you collect additional videos so that you can select the best five. See "Training & Reliability Procedures for MELE" on website for more details.

2. Photographs of showing the variety of materials used (for example, learning corners, books, toilets, outdoor play space, children's artwork, wall displays, etc.).

For the Surveys

For all surveys, collect:

1. Time it took to administer.
2. Items that seemed confusing to the respondents.
3. Items that the enumerators had questions about (either for administration or scoring).
4. Suggested revisions to items or wording of items or response choices.

For 3-5 interviews (at least one of each type – parent, teacher), videotape the interview and scan the interview sheet for use in training.

MELQO PRE-FIELD TESTING RESOURCE:

TIPS FOR VIDEOTAPING CLASSROOM OBSERVATIONS

Videotaping is an integral part of the MELQO system. Video clips from classrooms enable enumerators to observe and measure quality in ECCE classrooms, provide opportunities to trainers to give enumerators practice, and help document quality in classrooms. In order to ensure the best quality recordings there are several steps videographers should take. This supplemental guide helps videographers understand the equipment, protocol and necessary steps for successful videotaping in ECCE environments.

A. Equipment. In order to conduct successful videotaping, videographers need to have the correct equipment to allow them to capture all of the information needed to measure quality.

- Media to use: Smartphones, digital cameras and video cameras can be used to collect videos.
- Tripods/stable recording: It may also be necessary to use a tripod in order to take stable videos in the classroom environment. If a tripod is not available then it is important to videographers to use two hands when holding the camera.
- Microphones: Microphones are also an important aspect of videotaping quality interactions and environments. Videographers should test the microphones available on the recording devices to see if they are sufficient for recording interactions conversations and noise levels in the classrooms. If the microphones are not sufficient then external microphones will need to be used.

B. Protocol for Videographers

1. TEST/PRACTICE IN ADVANCE
    a. Videographers should practice with their recording devices (including sound check) before they conduct an official video recording session. They can work out the bugs and make sure that all the equipment is working properly. Test runs also help them learn the best angles to capture classrooms and teacher child interactions.
    b. Have partner (supervisor or study coordinator) review the tapes for accuracy can also be helpful in ensuring that videotapes are of the highest quality.
2. VIDEO POSITIONNING:
    a. The best position for the camera is at the back of the classroom or on the side of the classroom depending on the set up.
    b. The target object is **the teacher.**
    c. Make sure that later on the observer/viewer of the video will be able to observe what is happening in the classroom (especially how the teacher interacts with the students).
    d. Do not worry about zooming into document materials in the classroom.
3. VIDEO SETUP
    a. Recording devices should be held at a horizontal orientation in order to videotape using a full screen (landscape instead of portrait orientation).
    b. It is also imperative that there is not backlighting in the camera. If backlighting occurs it can be difficult for observers to view the interactions.
    c. Be aware of microphone aware and make sure microphone is not covered so that the sound can be recorded well.
    d. If videographers are using smart phones to record video sessions, they should place their phone in airplane mode so that text messages and phone calls do not come through during the recording.
    e. Make sure battery is full and sufficient
    f. Video clips should be stopped every 15 minutes and then immediately restarted. Once a video clip goes over 15 minutes it can be difficult to load.
    g. Completely label the video clip with the ID number date of the video and start and stop times of the video.

4. GUIDANCE ON CLASSROOM VISIT
    a. Get to the classroom early and introduce yourself to the teacher. Videographers should introduce themselves and tell teachers what will happen.
    b. Simple comfortable clothing should be worn by the videographers and they should not attract attention to themselves when videotaping.
    c. Do not interact with children or distract children from the classroom.
    d. Do not call or talk to people during the videotape and do not breath too close to the camera.
    e. It is important to record the information that is needed during the session (see guidance above on "Additional data to gather"). Remember it is better to record too much information than not enough.
    f. At the end of the visit, send videos to PI/Project Coordinator. Remember that videos are to be confidential and not to be shared with other teachers/people.

By remembering to follow the guide for videotaping, videographers ensure the best possible clips for classroom quality are documented for both observations and training purposes, and help to further the MELQO community.

PHASE 3

Field-testing and data collection

It is recommended that before large scale data collection, the MELQO tools are field-tested (piloted) to ensure that items assess what they purport to test. For MODEL, field-testing involves an iterative process of testing the modules with children an in classrooms to refine items so that they obtain the information required. For MELE, field-testing is focused on making sure that assessors can be reliably trained in administering the items, and that the items capture the quality in a range of settings. The sample size for field testing will be dictated by the country's needs for full-scale implementation[2].

A key element of data collection (for field testing or large-scale data collection) is ensuring highly trained data collectors. Good training and standardization of assessors is essential for getting meaningful results when using any of the MELQO tools. Measures that are not applied accurately or consistently will yield poor quality data, making it impossible to draw conclusions about what was measured, and its interpretation. A well thought-out, detailed training plan enables trainers and trainees to feel confident that they are collecting data that can be used to truthfully describe a group of children or the quality of a preprimary program. **Training and reliability procedures should be led by a certified MELQO trainer.**

On the following page you will find a brief overview of reliability guidelines. **A complete training package is available on ecdmeasure.org**

During this phase:

1. *Training materials prepared, including local videos with master codes*
2. *Multiple opportunities for live practice on all tools are provided to enumerators*
3. *Enumerators trained by Certified MELQO trainer, with reliability established*
4. *Establish sufficient reliability and validity of instruments*
5. *Monitoring of data collection activities throughout study to ensure continued reliability*

Implementation Resources for field testing and data collection

- Complete training package:
  - MODEL Training Modules
  - MELE Training Module
  - Survey Training Module
- MELE Training & Reliability Brief
- MELQO Training Report Template

Please visit ecdmeasure.org and register for the MELQO Portal to access these materials

[2] *For national studies, it is recommended that tools are field tested among a sample of about 200 children (MODEL) and 20 classrooms (MELE) to fully map the amount of time it takes to administer the instrument, examine the psychometric properties of the instrument, and make modifications as needed before large-scale data collection occurs*

MELQO RESOURCE:
TOOL RELIABILITY GUIDELINES

Training and reliability on the MELQO modules can be completed by one of two ways:

a. Cascading training of the trainer (TOT) model: certfied MELQO master trainers train country trainers, who can then train observers in-country. This is most likely to occur in a centralized workshop, where multiple trainers would receive master training and be evaluated simultaneously.

b. Direct training: certified MELQO global team/master trainers train country-level observers (including supervisors, etc.) directly in country.

Training should only be delivered by a certified trainer. Certified means the person has completed a train-the-trainer workshop with the MELQO team and has all resource materials for conducting training following MELQO guidelines.

Trainers and enumerators (i.e. anyone delivering training on or administering each tool) should complete the same procedures outlined below.

Reliability on each module should be re-established annually by completing a refresher training, reviewing any updated materials and completing the administration check (video or live).

Optional reliability/Inter Rater Reliability during data collection: ongoing reliability checks and IRR may take place during the data collection period by having 10% of administrations be double-scored (live or by videotape) or repeating the reliability checklist (live or by videotape).

Overview of MODEL Direct Assessment (DA) Reliability

Anyone delivering training on or administering the DA should complete the following reliability procedures:

1. **Completion of DA training by a certified trainer**.
2. **Pass DA written quiz** (part of DA Training Module)**.** Recommended that enumerators pass written (T/F and multiple choice) quiz with at least 85% of items correct. Trainers should pass with 100% correct.
3. **Reliably administer DA to a preschool-aged child**. Enumerators in training will administer the complete Direct Assessment to a child between the ages of 3 and 6 years of age. A trained person will evaluate the administration using a checklist (part of DA Training Module). This process may be done remotely (for example, a videotape of the administration can be sent to a trained person to evaluate).

Overview of MELE Classroom Observation (CO) Reliability

Anyone delivering training on or administering the DA should complete the following reliability procedures:

1. **Completion of CO training by a certified trainer**.
2. **Pass CO written quiz** (part of DA Training Module)**.** Recommended that enumerators pass written (T/F and multiple choice) quiz with at least 85% of items correct. Trainers should pass with 100% correct.
3. **Video Reliability practice and quizzes:** Enumerators should pass video quizzes with at least 80% agreement (See *Reliability procedures for MELE* for details).

Overview of Survey Reliability Procedures

Anyone training on or administering survey tools (TCR, PCR, Teacher Interview and Head Teacher/Director Interview) should complete the following reliability procedures:

1. **Completion of Survey training by a certified trainer**.
2. **Pass Survey Administration written quiz** (part of Survey Training Module)**.** Recommended that enumerators pass written quiz with at least 85% of items correct. Master trainers should pass with 100% correct.
3. **Reliably administer survey:** Enumerator trainees should administer the complete survey to another adult. A trained person will evaluate the administration using a reliability checklist (available in Survey Training Module). This process may be done remotely (for example, a videotape of the administration can be sent to a trained person to evaluate).

Many countries are seeking ways to use ECD measurement within the context of national policy planning. This could include purposes such as informing standard-setting, influencing decisions about resource allocation, and quality improvement. It is important to make the process of data analysis and reporting transparent and include multiple stakeholder groups to prevent the study methodology from being criticized if the results are different from what was expected. Having multiple stakeholders involved can also help expedite the process and keep the reporting agency accountable for sharing results within a reasonable time-frame.

Implementation Resources for analysis and application

- MELQO Brief 1: Results from Four MELQO Pilot Countries
- Sample data entry spreadsheet with items, response categories, and missing data
- Scoring guidelines
- Guidelines for data analyses: standard recommendations and procedures
- Country results template
- Link to other country results briefs

Please visit ecdmeasure.org and register for the MELQO Portal to access these materials

MELQO DATA ANALYSIS RESOURCE:

SCORING OF MODEL: DIRECT ASSESSMENT AND TEACHER/CAREGIVER REPORT

Analyses of MODEL data often addresses three main themes:

1) determining the overall functioning of the scale, especially in situations where the scales may be used again for national monitoring;
2) identifying patterns in children's learning, especially in relation to equity, for example, the gaps in learning and development between high- and low-income children; and
3) documenting the extent to which curricula standards or benchmarks are in line with expectations articulated through national standards.

Other research questions using MODEL scores have included the association between family characteristics, early health and nutrition status, and child development and learning; patterns of social/emotional development; and responsiveness to interventions.

Identifying underlying domains with factor analyses. A first step in analyses is to establish the presence of domains or latent constructs that drive children's observed scores on either the direct assessment or teacher/caregiver ratings. Our recommendation is to begin with an empirical approach using confirmatory factor analyses, assuming that items will load onto domain-specific factors including literacy, mathematics, executive function, and fine motor for the direct assessment; and problem behaviors, social skills, and attention/self-regulation for teacher/caregiver reports on social/emotional development. The MODEL scores have been shown to yield reliable factors in representative samples, most recently in Tanzania and Nicaragua. Factor structures can vary, but most typically demonstrate a delineation of pre-academic skills such as literacy and math from social/emotional skills, with fine motor and executive function showing variability in their factor loadings. Exploratory factor analyses can also be used if the results from CFA are inconclusive and fail to reveal a reliable factor structure. A complete description of factor analyses on the MODEL is available for Tanzania.

In some countries, items will be added to reflect country priorities. In that situation, analysts can decide whether to include those items in the factor analyses – which may not fall onto discernable factors – or to discard them when establishing the factor structure. Including the items clearly provides a more complete picture of how the items work together, but the resulting factor structure will not match MODEL results from other countries, which may be a drawback if intending to use saved factor scores as a method for analyzing MODEL results. Using all items may preferable in situations where the government is interested in using the scales again, because it will provide a more comprehensive picture of how all items work together.

If factor analyses fail to reveal a reliable factor structure using either CFA or EFA, it is possible to assign each item to a conceptual domain, such as literacy, mathematics and executive function. This method is arguably less ideal, but possible and theoretically justifiable given the process of selecting MELQO items and emphasis on identifying items with cross-national relevance.

Testing item functioning using item response theory (IRT). IRT can provide useful information on individual item functioning. IRT reveals which items are contributing most reliably to overall scores; for example, for child development and learning, IRT will reveal which items are associated with overall scores on the scale and which items are least associated with overall scores. This can be interpreted to reveal items that are poorly-functioning. If the MODEL or MELE will be used as part of an ongoing national assessment, IRT may be useful in identifying which items could be dropped to make the scale more efficient. However, results from IRT should be used carefully, because items that do not correlate with the overall scale still may contain important

information. IRT should be undertaken in partnership with a psychometrician who can help interpret results. IRT has been used with MODEL results in several countries to help refine the scale, but it is expected that results may vary somewhat by country, especially if items have been added based on national standards.

After establishing the existence of a factor structure, the next step is to calculate scores that can be used to summarize children's overall performance on the MODEL, and to report on children's skills in specific areas.

MODEL DA Scoring. Once the factor structure has been assessed, the next step is to calculate scores for each task. In many studies, summary scores that can be used as predictor or outcome variables have also been created to provide an outcome variable that summarizes children's learning across a domain.
Calculating scores for each task. Within the MODEL DA, scoring approaches must account for the different number of sub-items that contribute to each task: For example, some items have several sub-items that collectively contribute to a score on the task, while others have just one or two. There are four methods for scoring the MODEL DA that have been used in existing analyses to date:

1) Binary item method. To determine domain summary scores, scores are recoded so that each task/skill carries equal weight in the summary score. This required us to re-score some of the items that were not binarily "correct/incorrect" (name writing, Shape copying, expressive vocab, etc). The scores can then be used as individual outcome variables or can be summed to create an overall score.
2) Factor scores. If the factor analyses reveal reliable and interpretable factors, the scores from these factors can be saved and used in future analyses.
3) Z-scores. Another method is to sum the child's score on a given task to create a percent correct or a total score, and then converting into a z-score that provides a standardized score for each task that can be summed to create a final score that doesn't weight one item over another.
4) Item-level reporting by national standards. For policymakers, an effective approach to reporting items may be an item-level reporting of the results as they pertain to national standards. For example, if there is a national standard stating that children should be able to count to 10 by end of pre-primary, results can be reported as the percentage of children who can count to 10.

To date, the z-score and factor score approaches have tended to yield more reliable outcome variables than using binary scores or item-level reporting, but the choice of method should be based on the goals of the investigators and the anticipated ease of interpretation for policymakers.

MODEL TCR Scoring. Scores on the TCR can be created by saving factor scores, as outlined above, or can also be created through calculating average ratings. Because the TCR results are all on the same scale, vs. the DA, which has different scales for each item, it tends to be more straightforward and requires less manipulation to achieve a workable score.

MELQO DATA ANALYSIS RESOURCE:
SCORING MELE

Administrations of MELE have revealed issues in scoring and analyses that require careful consideration. First, in many countries, there are items on the MELE that do not demonstrate adequate variability for use in additional analyses. There can be value in these items – if, for example, no classrooms demonstrate free play for children, the scores on the items may be marked as 1 with little variability, but the scores themselves may be very valuable to policymakers, by demonstrating that few classrooms are engaging in desired activities. At the same time, these items will cause considerable difficulty in later analyses. One suggestion is to identify these items and evaluate their relevance to policy, and then remove them from additional analyses.

Second, the MELE scores may or may not form reliable factors. While the items were selected to reflect domains of interest, it is not clear that the domains represent latent constructs in the same manner as the domains of child development and learning. For that reason, it may be theoretically justified to assign items to domains or to use factor analyses to discern reliable factor structures. The differences in the content and structure of MELE across countries means that we have fewer examples from which to report. In previous samples, factors that emerge may represent constructs like sanitation/cleanliness vs. pedagogy, which then can help clarify what factors are most critical for child development.

Methods for scoring the MELE items are similar to MODEL:

1) Factor scores. If reliable and meaningful factors are identified, this method can be used to create variables to use in further analyses.
2) Z-scores. In many MELE versions, items are on different scales, which prevents creating summed or averaged scores. Conversion of items into z-scores can correct this problem, which in turn allows scores to be summed or averaged either within domains or across the entire scale.
3) Item-level reporting by national standards. As noted above, in situations where items may have limited range but policy significance, it is possible to report MELE items by national standards; for example, the number of classrooms with adequate water supply or dangerous conditions.
4) Item-level reporting without national standards. Some of the MELE items may also be useful independently, as predictors of children's learning. For example, whether children are able to ask open-ended questions may be important for child development, even in the absence of any other variables from the MELE scale.

In general, the analyses for the MELE items will require item-by-item analyses to determine which items have adequate ranges, and which are associated with predicted teacher or school characteristics. To date, results have indicated that factor scores may not adequately capture the items that are most strongly associated with child outcomes in a given country. When MELE is being used for the first time, it is strongly recommended to look carefully at each item, and to consider modifications to the scale before using across a nationally-representative population.

MELQO DATA ANALYSIS RESOURCE:
ESTABLISHING CONCURRENT VALIDITY

Once summary scores have been created, it is necessary to examine the associations between child and family characteristics and other predicted associations with MODEL and MELE scores. This step serves two purposes: first, to test the concurrent validity of the scales in the context in which the scales are being used; and second, to begin to examine patterns in equity as outlined above.

It is expected that the size of the associations between factors associated with children's learning, such as maternal education and household assets, will vary based on country. If no association is seen, it may be due to a problem in constructing the scale and/or a restricted range of responses from parents on assets and education, for example. This step can help identify how well the scales are working and if more than one approach to calculating summary scores may be required to find the method that best reveals the patterns in the data. *MELQO Brief 1: Results from Four MELQO Pilot Countries*, available on ecdmeasure.org, provides an example of how concurrent validity was tested in pilots of the MELQO tools conducted in four countries in 2016.

Many MELQO studies have also expressed interest in knowing the types of associations between child outcomes and quality of learning environments. We recommend that analyses of this question are undertaken in partnership with a data analyst who has experience nested models approaches and structural equation modeling. Defining this association is complex and often requires sophisticated modeling to fully describe; it is also important to note that the size and nature of this association will likely vary quite substantially by country.

MELQO DATA ANALYSIS RESOURCE:
REPORTING RESULTS AND APPLYING TO POLICY

Most MELQO studies have reported results using more than one scoring technique, meaning that descriptive information is provided based on national standards, and summary scores are used to provide more detailed information on factors associated with MELE or MODEL scores. As well, if representative sampling took place, group differences can also be reported, most commonly using some form of summary scores.

Examples of MELQO national reports are available on our website.

MELQO DATA ANALYSIS RESOURCE:

RECOMMENDATIONS FOR EQUITY ANALYSES FROM FHI360/SAVE THE CHILDREN EQUITY RESEARCH INITIATIVE

- Distributions are examined first at the **student level**, according to shape and then through an analysis of whether certain populations are more likely to be clustered at the lower end of the distribution.
- The analysis is then repeated at the **school level**, looking at the shape of distribution across schools and whether schools at the lower end of the distribution differ substantively in terms of student composition from others in the sample.
- The final group of questions looks at differences in **resource and input allocations** at the school level, considering indicators such as pupil-teacher ratios, teacher characteristics and other resource allocations, and whether these inputs are correlated with differences in outcomes at the school level.

Structured Questions Guiding Equity Analysis in Education

Level of Analysis	Dimension of Analysis	Questions
Student	Individual	What are the equity dimensions of interest in this dataset?
Student	Individual	What is the overall student composition in our sample? If this is an impact evaluation dataset, what is their composition, by treatment status?
Student		Descriptive analysis to examine variability across key equity dimensions
Student	Individual	What is the overall shape of the distribution of outcomes?
Student	Background	Are the lowest performing students (at the bottom decile/ quintile of the distribution) substantively different from the higher performing students by gender, socioeconomic status, ethnicity/race, language, urban/rural, geographic location, disability status?
School	Cluster (school, district, subnational unit)	What is the overall school-level distribution for outcomes? How large is the variability in outcomes between clusters (schools, districts, regions)?
School	Cluster (school, district, subnational unit)	What is the student composition, along gender, socioeconomic, ethnic, and other student subgroups, among low performing schools?
Student		Impact evaluation with lens on equity
Student	Program/ Intervention	What is the effect of the program on outcomes of interest? is it homogeneous or heterogeneous across the different student subgroups? Does the program/ policy have an equity building effect?
Inputs		Resource and input allocation
School	Classroom	What are pupil-teacher ratios for different schools or clusters? How large is the disparity between high performing and low performing schools?
School	Classroom	What are the average teacher characteristics in low performing schools versus high performing schools?
School	Resources	How are public resources/ program resources allocated between schools?
Other		Adding a time dimension
Student	Program/ Intervention	How do growth trajectories vary across the profiles of students? Does the magnitude of growth differ by equity dimension?
School	Cluster (school, district, subnational unit)	Has the school-level distribution of outcomes improved over time? Has the percentage of low-performing students changed since implementation? Have the gaps between groups closed/widened?