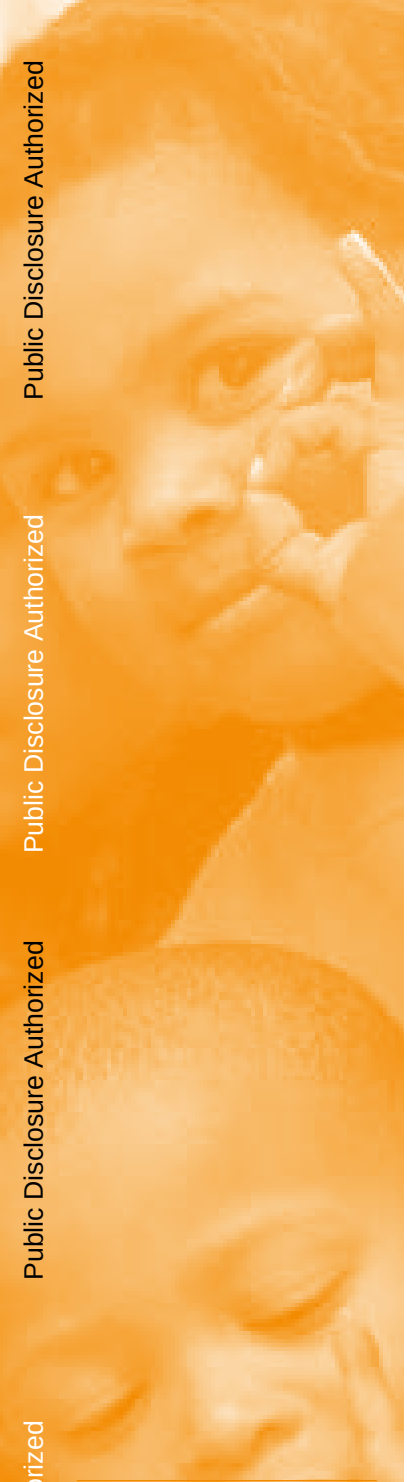

EARLY LEARNING PARTNERSHIP

GUIDANCE NOTE
July 2016

Measuring child development and early learning

This brief is intended to provide basic information on measuring young children's development, specifically focused on the skills and abilities children are likely to need upon entry to primary school.[1]

WHY IS IT IMPORTANT TO MEASURE CHILD DEVELOPMENT?

In response to convincing evidence on the benefits of supporting young children's cognitive and emotional development, investment in early childhood programming is increasing. Parents, policy makers, and funders are increasingly interested in measuring and tracking children's development to understand whether programs are having intended impacts and if children are developing adequately. Achieving reliable and regular measures of young children's development, as well as measures of the quality of their early learning environment and experience, is essential to address the problem of poor learning outcomes in countries around the world.

HOW DO YOU MEASURE YOUNG CHILDREN'S DEVELOPMENT?

There are several domains of young children's development that are considered important to measure. These can be conceptualized in different ways, but generally include the following areas: cognitive, language, motor, socio-emotional, and executive function/self-regulation.[2] These domains contribute to children's long-term well-being and are overlapping and mutually influencing. For example, self-regulation is thought to be relevant across all domains because it is central to what children learn and experience; likewise, language skills influence cognitive and socio-emotional development in addition to the ability to communicate.

This brief was produced by the Early Learning Partnership and reflects significant contributions from the following core members of the Measuring Early Learning Quality and Outcomes (MELQO) Initiative: Amanda Devercelli, Lucy Bassett, Alexandra Solano Rocha, Abbie Raikes, Kate Anderson, Tamar Manuelyan Atinc, Pia Britto, Ana Nieto, and Fabiola Lara.

For more information or to be added to the ELP mailing list, please write to earlylearningpartnership@worldbank.org

Key domains to be considered when measuring young children's development include these[3]:

Cognitive: This includes pre-literacy, problem solving, measurement and comparison, analytical thinking, memory, and early mathematical and number sense.

Language: This includes a child's knowledge and use of words, both in print and in oral form.

Motor development: This includes fine and gross motor skills and measures a child's capacity to control his or her body movements to perform everyday tasks, such as walking, running, or jumping, as well as drawing, writing, holding utensils and picking up objects.

Socio-emotional: This includes a child's awareness of his or her own feelings and those of others. It also measures children's social interactions and how children manage their behaviors.

Executive function/self-regulation: This includes self-control (inhibition and delaying gratification), persistence, and the ability to initiate action and sustain attention.

How do we test children's abilities within these domains?

As elaborated in Table 1, domains are the broad areas to be measured (e.g., self-regulation, social-emotional). Within each domain, there are particular traits or skills to be measured called constructs. These can be things like counting (for early math skills, a subdomain of pre-academic or cognitive skills), or working memory (for self-regulation). Specific survey items are used to measure the constructs, such as a counting test to measure counting skills, or a parent survey question about peer-to-peer interaction.

What are the different methods available to collect data on children's development?

There are a range of methods available to collect data on children's development, each with advantages and disadvantages (see Table 2). Some require more time, some have a higher degree of bias, and others are more costly. By using more than one method, you can triangulate data to increase the strength of the assessment results. However, potential disadvantages of triangulating data are that it can be costly and can sometimes simply yield the same information twice rather than adding considerable value. Some tools use one method only (e.g., just direct assessment), while others use multiple methods.

Different tools are better suited to some purposes; be sure you are clear about the purpose of the measurement effort when you select your tool.

How do you select an instrument?

Defining why you are interested in measuring young children's development will help you determine the type of measurement tool needed. The purpose of the measurement will then have implications for the resources, time, and stakeholders to conduct the assessment. As examples: the purpose could be to perform an impact evaluation of an early childhood development intervention; to gather information on particular domain(s) of interest (e.g. to see how ready children are for primary education); to collect data to inform policy (e.g. the degree of development differences across socio-economic groups) or to improve a system or program variable (e.g. teacher instruction); or to diagnose and assess child progress at a population level. A central question is whether

TABLE 1 Examples of domains, constructs, and items to measure child development outcomes

DOMAIN	CONSTRUCT	ITEM NAME	DESCRIPTION
Cognitive or pre-academic (early math skills)	One-to-one correspondence	Producing a set	• Twenty items/objects are in front of the child. • Ask child to hand you three items. • If child is successful, rearrange the 20 objects randomly. • Ask child to hand you six items. • If child is successful, rearrange the 20 objects randomly again and ask for fourteen items. Stop Rules: If the child cannot give you three items move on to the next item. If the child cannot give you six items move on to the next item.
	Counting	Parent survey question about counting	***Enumerator:*** **How high can (name) count?**
		Counting test to measure counting skills	***Enumerator:*** **Count as high as you can** [stop child at 30]
Language	Phonemic awareness	Initial sound identification	***Enumerator:*** **I want you to tell me the beginning sound of each word.** (The enumerator goes through a practice round first) **What is the first sound in** **"bed"** /b/ **"land"** /l/ **"farm"** /f/ Stop Rule: If the child does not respond after 4 seconds, mark as "No response" and say the next prompt.
Executive function/ self-regulation	Working memory	Forward digit span	***Enumerator:*** **I am going to say a list of numbers. After you hear the numbers, I want you to repeat them after me in the same order. Please listen carefully.** 1...6 5...2...9 8...3...1...4 1...2...4...7...3
Motor development	Fine motor skills	Drawing a person	Give a pencil and paper to the child and say: ***Enumerator:*** **Let's do some drawing. I'd like you to draw a picture of a girl or a boy standing up. Can you try to draw that?**
Socio-emotional development	Awareness of feelings	Understanding feelings	***Enumerator:*** **Now I have some questions about feelings. Please tell me what makes you feel sad or want to cry?** Then ask: **Please tell me what kinds of things make you feel happy?**
	Peer-to-peer interaction		***Enumerator:*** **Does (name) show consideration of other people's feelings?** **Does (name) get along with other children s/he plays with?** **Does (name) offer to help someone who seems to need help?** **Does (name) comfort or assist another child who is hurt, sick or upset?**

TABLE 2 **Methods to collect data to measure child development**

METHOD	DESCRIPTION	ADVANTAGES	DISADVANTAGES
Direct test (or direct assessment	A trained enumerator conducts a one-on-one enumerator-to-child session where a child is required to solve or answer questions in a certain amount of time, sometimes with the use of material provided. This testing, which for preschool children is often framed as a game, can occur in any environment, including a classroom or home visit, but it is important that all children be tested in the same environment to ensure validity of results.	• No problems with recall bias • No problem with bias of respondent (e.g. parent or teacher)	• Can be difficult to test young children (they may be uncomfortable, distracted, tired, hungry) • Requires high level of training of person administering test • Tests that contain unfamiliar elements, such as unfamiliar activities or an unfamiliar language, can compromise quality of assessment
Parent and/or teacher report	Someone who knows the child well (e.g. parent or teacher) can provide the enumerator with important information about the child's abilities. Parent and teacher reports should not be considered substitutions for each other, since they each capture a different part of the child's life and interactions with others.	• Easy to administer and requires less training than direct assessment • There is evidence of correlation with direct assessment • May provide valuable information that cannot be acquired in direct assessment because parents and teachers have frequent, ongoing observations of children in different situations	• Potential bias due to social desirability (telling the enumerator what parent or teacher would like child to be able to do) • There may be systematic differences in interpretation across cultures • Parent may not be able to accurately report child's abilities

BOX 1

RELIABILITY & VALIDITY

Reliability refers to the extent to which a test will consistently provide similar scores when and if administered to a child or group of children over time.

Validity refers to how well a test or assessment measures what it intends to measure.

Concurrent validity refers to the extent to which the results of measurement correspond to those of a previously established measurement for the same construct.

Predictive validity refers to the extent to which results of a test are related to later performance that the test was designed to predict.

the content of the tool is aligned with the goal of measurement, and whether the tool is specific enough to pick up changes in particular skills or designed for use across an entire population.

It is important to keep in mind that the tools to measure child development should have reliable and valid measures (see Box 1). When adapting items for cultural context, you cannot assume that the reliability and validity of the original items will be automatically passed on to the adapted items. Instead, expert guidance will likely be needed to ensure the reliability and validity of the modified tool. For example, when the item "number of friends a child can name" is used in some countries, it effectively represents the construct of socio-emotional development, but in another country this may not be the case; instead, the item could better reflect language ability rather than social understanding. Tools should also have good concurrent validity (measurement results correlate with other established measures) and predictive validity (results are related to later performance).

See Annex 1 for examples of commonly used tools, the domains they measure, countries where they have been used, estimated costs and further contact information. All of the scales featured have been validated, used in more than one developing country, adapted on more than one occasion to be culturally appropriate, and translated into more than one language.

Annex 2 shares information on the Measure of Development of Early Learning (MODEL), which includes modules and support for countries to measure early learning in a globally comparable way.

Key considerations when choosing an instrument to measure children's early learning and development

- Has the instrument been used before for the purpose you've defined?
- What is the purpose of the measurement exercise?
- Was the tool designed for use in the country or region where you are working, or will you need to adapt the tool to country context, culture, language, etc.?
- Has the tool been validated, meaning that there is evidence that what it measures is important for children's development over time?
- How much does the tool cost? Do you need permission to use it? Are there copyright issues to consider? The Bayley Scales, the Denver Development Tests, and the Woodcock-Johnson are some examples of tests that are strictly protected by copyrights. There are also cases where only a licensed psychologist can purchase the tests from the publishing companies.
- Conducting the same assessment repeatedly can give valuable information on trends in child development. Will your project have funds to use the same tool more than once to track progress over time?

How can we integrate child development measurement into a national monitoring system?

Many countries are seeking ways to use measurement of child development within the context of national policy planning. This could include purposes such as informing standard setting, influencing decisions about resource allocation, and/or quality improvement. It is important to align or integrate the assessment efforts into a national monitoring system.

Key questions to consider include these:

- Is there a unit in government that is intimately involved in the design and delivery of the first effort and through this has capacity (and willingness and mandate) to maintain and carry the work forward?
- What information system is currently in use and available in-country? For example, can you integrate indicators of child development into an existing child monitoring system? Can indicators of preschool classroom quality be integrated into school census efforts that are ongoing and channeled into a national education management information system (EMIS)?
- Who is available in-country to collect data regularly over time? Potential sources include local education or health officials who already inspect education facilities or assess children's development (such as district education officers, home visiting professionals, community health workers). Local university graduate students are another source which could be developed through a partnership with a unit in a local university. Training recurrently available workers or forming partnerships with key institutions could facilitate sustainable and ongoing utilization of the data and tool(s).
- Are there other related data collection efforts being carried out with which you could align your effort? Is there an upcoming education survey or assessment that you could leverage? For example, if an Early Grade Reading Assessment (EGRA) is carried out in the country every three years, could you attach to it? Are there household surveys (DHS, MICS, etc.) with which you could align?

How much will an early learning assessment cost and how long will it take?

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.

TABLE 3 Sample budget

	TASK	ESTIMATE OF TIME NEEDED	COUNTRY A (PROJECT EVALUATION)	COUNTRY B (NATIONALLY REPRESENTATIVE STUDY)
PREPARATION	Preliminary meeting	1-4 days, including planning and execution	500	1,000
ADAPTATION	Expert time for adaptation[a]	10 days (often ~$400-600/day)	2,000	6,000
	Local meeting(s) (including per diems, space, pre-piloting, etc.)	5 days	1,000	1,000
	Translation of tools	Depends on language, length of tool	3,000	7,500
TRAINING	International expert time for training (master trainers)	10 days (often ~$400-600/day)	10,000	10,000
	Travel (experts, TTL, etc.)	1 week mission for each participant	20,000	20,000
	Space rental, materials, etc.	Depends on country	44,000 (firm contract)	80,000
DATA COLLECTION, ANALYSIS, AND DISSEMINATION	Data collection (transport, enumerators' salary/per diem,[b] etc.), materials (tablets, questionnaires, etc.)	Depends on country	44,000 (sample size[c] 150 children)	120,000[d] (sample size 1,000 children)
	Data analysis[e] and report writing	~4-6 weeks	8,000	10,000
	Dissemination	1 event ($1,000-3,000)	1,000	3,000
OTHER	General TTL travel	1-4 missions in one year	5,000	15,000
TOTAL			**91,000**	**190,000**

[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 *greatly* depending on the country context, transportation costs, etc.
[e] Can include psychometric analysis, statistical analysis, etc.

TABLE 4 **Approximate timeline**

		MONTH 1	MONTH 2	MONTH 3	MONTH 4	MONTH 5	MONTH 6	MONTH 7	MONTH 8	MONTH 9	MONTH 10	MONTH 11	MONTH 12
LAUNCH AND PLANNING	Hold meeting(s) to identify main research questions of government (local/regional/global experts)	X											
	Appoint task force or focal point from government	X											
	Review existing curriculum and service delivery standards and align with assessment domains		X										
	Propose study design based on government priorities		X	X									
	Map out project timeline and budget		X	X									
	Write and release ToR for data collection				X								
	Get ethical approval as needed				X								
	Gather necessary background data for sampling purposes				X								
	Select contractors and finalize terms of contract				X	X	X						
ADAPTATION AND PRE-FIELD TESTING	Translate and back-translate tools				X								
	Hold in-person meeting(s) with national experts (including curriculum developers, academics conducting ECD research, school inspectorate (or equivalent)) to adapt items/measures and align with curriculum					X							
	Test tools on small sample and further revise and adapt tools as needed						X	X					
PILOTING/ DATA COLLECTION	Train enumerators for data collection (training should include access to children and classrooms so enumerators can practice using the instruments)							X					
	Collect data								X	X			
	Clean data									X			
ANALYSIS AND SYNTHESIS OF RESULTS AND DISSEMINATION	Analyze findings										X		
	Synthesize findings in easily digestible reports for different audiences											X	
	Distill findings for policymakers based on current policy plans											X	
	Disseminate findings through meetings and other events with national policy makers, academics, civil society, donor partners, regional/district education officials, teachers/principals/school managers, and parents/general public												X
	Make findings accessible/understandable for teachers and parents, etc.												X

Note: *Table 4 is an approximate timeline for a national monitoring or research study. Impact evaluations might not need to go through all of these steps unless they are closely tied to curricula. The timeline is meant to be illustrative of the steps needed to undertake the measurement process and how this would be sequenced over a one-year period. In reality, depending on the specifics of the country measurement objective, the steps could be combined or expanded and could happen more or less quickly than noted in Table 4.*

HELPFUL TIPS

1. Get multiple stakeholders involved

It is highly recommended to engage a variety of stakeholders in the measurement process. For adaptation, this may imply local experts such as curriculum developers, academics conducting ECD research, and school inspectors. More broadly, school administrators or service providers where the tool will be implemented, as well as parents, teachers, and government officials, should all be informed of the assessment. In some settings, special permits may be needed in order to conduct assessments, so exploring this and ensuring that all necessary stakeholders are included in the process will reduce the chances of having to take backward steps.

2. Make sure study design is informed by country issues

A robust preparation process that takes into account country priorities, upcoming policy decisions, and political economy issues is essential to ensure that results will be useful for decision-making. Even if this process takes additional time at the outset, it will pay off in the usefulness and relevance of the results.

3. Adapt tools

The tools presented here can be applied in a number of settings, but they should be adapted to be fully responsive to local values and context. Adaptation may involve translation or item modification, for example.

4. Test tools and make any necessary changes

Testing is critical to ensure that a tool functions well and is reliable and valid in your context. Some things to consider during testing include the level of testers' training, test environment, and test procedures (e.g. allowing time for children to get comfortable before starting). It is also useful to document the experience to see if any changes are needed. For instance, the tool may take longer than expected to administer and need to be shortened or an indicator could be poorly received or not culturally appropriate. This is the time to tweak and improve the tool before it is used as a final product.

5. Do not use tests for screening or denying access to learning opportunities

While measurement in early childhood is a potentially helpful tool, it is critical that measurement not be used as a screening process or high-stakes test that could deny some children access to further schooling or services. The implications of a tool that might label a child as "special needs" should be carefully considered, particularly if services are not available to help children with special needs.

6. Ensure enumerators have appropriate skills

Enumerators should be fluent in the language spoken in the communities they will be visiting, should have some familiarity with the education system in the country, and must be comfortable speaking and interacting with young children, school staff and parents, since they will have to conduct interviews and guide focus groups.

7. Get the results out to a broad range of stakeholders

Just as it is important to involve stakeholders at the beginning of the process to define priorities and buy in to the process, it is critical to share results with different types of audiences at different levels. This includes national policy makers and academics, civil society, donor partners, regional/district education officials, teachers/principals/school managers, and parents/general public. For some subgroups (e.g. teachers, parents), it may be necessary convey the findings in a different way to ensure they are understandable.

ANNEX 1 Selected tools to measure children's early learning and development outcomes

Note that all of the scales featured have been validated, used in more than one developing country, adapted on more than one occasion to be culturally appropriate, and translated into more than one language.

TOOL	DESCRIPTION	DOMAINS	LENGTH	COST	COUNTRIES AND LANGUAGES	CONTACT
ASQ (Ages and Stages Questionnaire)	The ASQ is an assessment used by parents, teachers or caregivers for children ages 3 months to 5 years old. It is a parent-teacher self-report, with the assessments being administered in children's natural environments.	Socio-emotional, motor, and cognitive development	30 items and takes 10–20 minutes to administer	US$199 for 19 questionnaires and scoring sheets	Canada, Ecuador, United States, among others. It is available in English, Spanish, French and Korean.	Brookes Publishing, www.brookespublishing.com
Bayley Scales of Infant Development	This is a direct assessment that requires a trained enumerator to be administered.	Motor, language, and socio-emotional development	From 25–35 minutes for children under 15 months and 60 minutes for children 15–42 months	US$1,135 for a set of 25 questionnaires and scoring sheets	Bosnia, Argentina, Bangladesh, Brazil, Chile, China, Costa Rica, Czech Republic, Democratic Republic of Congo, Egypt, Ethiopia, Guatemala, Indonesia, Italy, Jamaica, Japan, Kenya, Lithuania, Malaysia, Mexico, Nicaragua, Nigeria, Philippines, Poland, Seychelles, South Africa, Tanzania, Turkey	Harcourt Assessment, www.psychcorp.com
EDI (Early Development Instrument)	EDI is a checklist of approximately 100 items, completed by educators, applicable for children 4–7 years old, developed to facilitate population-based assessment.	Health and well-being, language and cognitive skills, and socio-emotional development	7 to 20 minutes to administer	US$4.20– US$7.80 per student (negotiable)	Versions of the test can be found in English, Spanish, Bahasa, Amharic. Bengali, Japanese, Kiswahili, Chinese, Italian, Turkish and more.	Magdalena Janus, Associate Professor, McMaster University, edisrl@mcmaster.ca, https://edi.offordcentre.com
IDELA (International Development Learning Assessment)	IDELA is a direct child assessment that measures early learning and development for children ages 3.5 to 6 years with an accompanying questionnaire to assess the home learning environment.	Motor skills, early language and literacy, early numeracy/problem solving, socio-emotional development, and approaches to learning	24 core items (with additional optional items) typically requiring 30 minutes per child	Free with Memorandum of Understanding with Save the Children; Technical Support available with contract to cover labor & travel	Australia, Canada, Chile, Egypt, England, Netherlands, Jamaica, Kenya, Kosovo, Mexico, Moldova, Mozambique, New Zealand, United States	Sara Poehlman, Senior Director, ECCD spoehlman@savechildren.org Amy Jo Dowd, Senior Director, Research, adowd@savechildren.org Lauren Pisani Gorman, Senior Research Specialist, lpisani@savechildren.org
MODEL (Measurement of Development and Early Learning)	The primary purpose of MODEL is to aid in the measurement of groups of children (e.g. at the population level) for child development/learning. It can be used for children ages 3 to 7, and data can be collected using a parent and/or teacher report and a direct assessment instrument.	Socio-emotional skills, pre-academic skills such as language, pre-literacy and pre-numeracy, and executive function	~30–45 minutes per child/ teacher/ caregiver	Free	Bangladesh, Cambodia, Colombia, Kenya, Lao PDR, Madagascar, Mongolia, Nicaragua, Sudan, Tanzania	Lucy Bassett, lbassett@worldbank.org; Abbie Raikes, abbie.raikes@unmc.edu; Kate Anderson, klanderson@brookings.edu Ana Nieto, anieto@unicef.org
PRIDI (Inter-American Development Bank's Regional Project on Child Development Indicators)	PRIDI collects data on children ages 24–59 months old. The survey, administered in households, includes approximately 22 items.	Cognitive, language, motor, and socio-emotional development. The first three domains are measured via direct child observation; the socio-emotional development is measured through a maternal or main caregiver questionnaire.	30–40 minutes to administer per child	Free	Costa Rica, Nicaragua, Paraguay, Peru	Aimee E. Verdisco AIMEEV@iadb.org, http://www.iadb.org/en/topics/education/pridi/home,18292.html
PPVT (Peabody Picture Vocabulary Test)	This is a verbal assessment for children ages 2 years 6 months to adulthood (90+years). It can be administered by anyone with familiarity with scoring and testing materials, and no formal training in psychometrics is needed.	Language (vocabulary and verbal abilities)	174 vocabulary items that take 10–15 minutes	US$379.99 for a complete set of 25	Brazil, Chile, China, France, New Zealand, South Africa, United States, West Indies, among others. Spanish and English versions are available.	Pearson Assessments, http://ags.pearsonassessments.com
Woodcock-Johnston/ Munoz	This tool is a direct assessment test that measures children and adults ages ages 2 years 6 months to adulthood. The tool is administered by examinators/enumerators.	Cognitive development	Depending on the combination of subsets used, it can take from 60-70 minutes to administer with each subset taking approximately 5–10 minutes	The complete battery is US$966.50 plus the additional costs of each individual tested.	Mostly used in the United States but also administered in Costa Rica (in Spanish) and Seychelles.	Riverside Publishing, www.riverpub.com

ANNEX 2 Measure of Development of Early Learning (MODEL)

The Measure of Development of Early Learning (MODEL) was developed by the Measuring Early Learning Quality and Outcomes (MELQO) initiative. This initiative was led by the World Bank, UNICEF, UNESCO, and the Center for Universal Education at the Brookings Institution and included experts and partners from around the world. MODEL was developed through a consultative process, drawing on the best experiences and tools for measuring child development and early learning to date.

The primary purpose of MODEL is to support the measurement of groups of children, for example at the population level. Along with MODEL, another module, Measure of Early Learning Environments (MELE) was developed through MELQO to measure the quality of early learning center-based services. Beyond their use in specific country contexts, the MELQO modules were designed to have sufficient comparability across countries.

MODEL contains items indexing early literacy, mathematics, and executive function, and for the teacher/caregiver report it includes items on socio-emotional development. MODEL includes both a direct assessment module, which requires trained direct observers to assess the children; and a teacher/caregiver report module based on surveys with teachers and/or caregivers. The modules are designed to work together: the direct assessment module provides information on children's early learning, while the teacher/caregiver modules provide insight into children's behavior in school and at home. The caregiver modules provide insight into children's family backgrounds and home learning environments. MODEL measures cognitive, language, and executive function skills, as well as socio-emotional development.

To date, MODEL has been piloted in Bangladesh, Cambodia, Colombia, Kenya, Laos PDR, Madagascar, Mongolia, Sudan, and Tanzania, with more countries (e.g., Nicaragua) in the early stages or joining soon. The content of the assessment is drawn from existing assessments, so it bears similarities with items from many of the tools profiled in Annex 2, including: ASQ, EDI, IDELA, PRIDI, among others. In addition, MODEL is aligned to the Early Grade Reading Assessment (EGRA) and the Early Grade Math Assessment (EGMA).

Ideally, MODEL could be used to present a nationally representative distribution of child development, which could then inform policy planning and budgeting, including resource allocations, curriculum design and teacher training programs, and ongoing monitoring of children's development.

The modules, along with manuals, guides, and other resources, are available upon request.

[i] Note that the MELQO initiative also developed a module, called the Measure of Early Learning Environments (MELE), to assess the quality of early learning environments. The two modules (MODEL and MELE) are complementary and are both meant to be adapted to align with national systems and standards and to be used to inform policy decisions to improve early childhood development.

ANNEX 3 Resources

For more information on instruments or tools to measure children's development and early learning, please consult:

Fernald, Lia C. H., Patricia Kariger, Patrice Engle, and Abbie Raikes. *Examining Early Child Development in Low-income Countries: A Toolkit for the Assessment of Children in the First Five Years of Life.* Washington, DC: World Bank, 2009. http://documents.worldbank.org/curated/en/2009/04/17193896/examining-early-child-development-low-income-countries-toolkit-assessment-children-first-five-years-life

Lopez Boo, F. "Lessons from Measuring Child Care Quality in LAC." PowerPoint presentation at SIEF Workshop on ECD Measurement, May 27 2015.

Lopez Boo, F., C. Araujo, and R. Tome. *How Is Child Care Quality Measured? A Toolkit*. Washington, DC: Interamerican Development Bank, 2016. https://publications. iadb.org/bitstream/handle/11319/7432/How-is-child-care-quality-measured.pdf?sequence=4

Young, Mary E., and Linda M. Richardson. *Early Child Development from Measurement to Action: A Priority for Growth and Equity.* Washington, DC: World Bank, 2007. http://documents.worldbank.org/curated/en/2007/01/8341509/early-child-development-measurement-action

Raikes, Abbie. *Early Childhood Care and Education: Addressing Quality in Formal Pre-Primary Learning Environments.* Geneva: UNESCO.

Young, Mary Eming, and Linda M. Richardson. *Early Child Development from Measurement to Action: A Priority for Growth and Equity.* Washington, DC: World Bank, 2007. http://documents.worldbank.org/curated/en/2007/01/8341509/early-child-development-measurement-action

NOTES

[1] An accompanying brief, "Measuring the quality of early learning," provides information on measuring the quality of early childhood care and education settings.

[2] This section is based on various consultations in 2014 and 2015 with MELQO's Technical Advisory Group on Child Development and Early Learning.

[3] Raikes 2014 and Fernald et al. 2009.

[4] This section is based on Fernald et al. 2009 and on Lopez Boo 2015 and Lopez Boo, Araujo, and Tome 2016.

[5] These considerations were compiled from anecdotal experience with pre-piloting the MELQO tool, specifically through conversations with MELQO's Technical Advisory Group for Child Development and Early Learning Outcomes.