

Overview



Measuring Early Learning Quality and Outcomes



Published by the United Nations Educational, Scientific and Cultural Organization (UNESCO), 7, place de Fontenoy, 75352 Paris 07 SP, France, the United Nations Children's Fund (UNICEF), 3 United Nations Plaza, New York, NY 10017, United States, the Center for Universal Education, Brookings Institution, 1775 Massachusetts Ave., NW Washington, DC 20036, United States, and the World Bank, 1818 H Street, NW Washington, DC 20433, United States

© UNESCO, UNICEF, Brookings Institution and the World Bank, 2017

ISBN 9789231002205



This publication is available in Open Access under the Attribution-NonCommercial-ShareAlike 3.0 IGO (CC-BY-NC-SA 3.0 IGO) license (<http://creativecommons.org/licenses/by-nc-sa/3.0/igo/>). By using the content of this publication, the users accept to be bound by the terms of use of the UNESCO Open Access Repository (www.unesco.org/open-access/terms-use-ccbyncsa-en).

The designations employed and the presentation of material throughout this publication do not imply the expression of any opinion whatsoever on the part of UNESCO, UNICEF, Brookings Institution and The World Bank concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

The ideas and opinions expressed in this publication are those of the authors; they are not necessarily those of UNESCO, UNICEF, Brookings Institution and the World Bank and do not commit these Organizations.

Designed by UNESCO

Overview



Measuring Early Learning
Quality and Outcomes

Table of contents

ACKNOWLEDGEMENTS	4
ABBREVIATIONS	6
INTRODUCTION	7
BACKGROUND AND RATIONALE	8
Equity and ECD: why measurement matters	8
The foundations of global ECD measurement	9
Tracking progress towards goals for young children: measurement at scale	9
Brief overview of existing ECD measurement efforts	9
Learning from the larger education measurement dialogue	10
MELQO's innovation	11
BUILDING THE MELQO MODULES	11
What are the MELQO modules?	11
Defining the content of the MODEL and MELE modules	13
Building on existing tools	13
Selecting domains, constructs and items	14
Child development and learning (MODEL) module	14
Early learning environments (MELE) module	15
Establishing technical rigour	16
Reliability	17
Validity	17
Achievements and next steps	18
USING THE MELQO MODULES	18
MELQO in the country context	19
Key considerations for early childhood measurement within the country system	19
Use of MELQO in countries to date	19
Tanzania	20
Mongolia	20
Lao PDR	20
Madagascar	21
Nicaragua	21
Overview of the process of implementing the MELQO modules	21
Planning (2–6 months)	21
Planning phase: discussion topics and roles	23
Adaptation and pre-field-testing (1–2 months)	23
Field-testing and data collection (1–3 months)	23
Analysis and application to policy (1–2 months)	24
Principles for working with governments	25
Building capacity for country-level measurement	25
Protecting the rights of children, teachers and families	26
REFERENCES	27
MODEL Module	29
MELE Module	71

Acknowledgements

This report represents the collaborative work of the Measuring Early Learning Quality and Outcomes (MELQO) core team, technical advisory groups and steering committee, as well as the individuals and organizations consulted in countries where the modules were piloted.

MELQO Core Team

Technical Lead: Abbie Raikes, University of Nebraska (formerly of UNESCO and UNICEF)

UNESCO	<i>Sobhi Tawil Mariana Kitsiona</i>
World Bank	<i>Amanda Devercelli Lucy Bassett</i>
Center for Universal Education, Brookings Institution	<i>Kate Anderson Tamar Manuelyan Atinc Fabiola Lara</i>
UNICEF	<i>Pia Britto Ana Nieto</i>

MODEL and MELE Technical Advisory Group Members, Contributors and Meeting Attendees

Aga Khan Development Network	<i>Caroline Arnold Alison Joyner Sheila Manji</i>	Open Society Foundations	<i>Sarah Klaus</i>
Ambedkar University Delhi	<i>Venita Kaul</i>	Results for Development Institute	<i>Michelle Neuman</i>
Bernard van Leer Foundation	<i>Katelyn Hepworth Joan Lombardi</i>	RTI International	<i>Amber Gove Katherine Merseth Maria Dzula Ritka Michael Willoughby</i>
Brookings Institution	<i>Emily Gustafsson-Wright</i>	San Francisco State University	<i>Linda Platas*</i>
Child Trends	<i>Kelly Maxwell</i>	Save the Children	<i>Lauren Pisani Nitika Tolani</i>
Fundação Carlos Chagas	<i>Eliana Bhering</i>	SEAMEO	<i>Abigail Lanceta</i>
Global Partnership for Education	<i>Jean-Marc Bernard Ian Robert Georges Macpherson Alexandra Solano Rocha</i>	University of North Carolina at Greensboro	<i>Belinda Hardin</i>
Inter-American Development Bank	<i>Florencia Lopez-Boo Aimee Verdisco</i>	UNESCO	<i>Yoshie Kaga</i>
International Step by Step Association	<i>Mihaela Ionescu</i>	UNICEF	<i>Ivelina Borisova Claudia Cappa</i>
Kenya Institute of Education	<i>Hellen Kimathi</i>	UNICEF WCARO	<i>Mariavittoria Ballotta</i>
Mathematica Policy Research	<i>Kim Boller</i>	University of Adelaide	<i>Sally Brinkman</i>
McGill University	<i>Frances Aboud*</i>	University of Hong Kong	<i>Yvonne Belcher</i>
McMaster University	<i>Eric Duku* Magdalena Janus*</i>	University of Hong Kong	<i>Nirmala Rao*</i>
National Association for the Education of Young Children	<i>Stephanie Olmore</i>	University of Oregon	<i>Sylvia-Linan Thompson*</i>
New York University	<i>Hiro Yoshikawa*</i>	University Research Co.	<i>Truphena Choti</i>
		USAID	<i>Marcia Davidson</i>
		World Bank	<i>Sophie Naudeau Rebecca Sayre</i>

* Lead module developers

Steering Committee Members

UNESCO	<i>Maki Hayashikawa Jordan Naidoo</i>
UNESCO Institute for Statistics (UIS)	<i>Albert Motivans Silvia Montoya</i>
World Bank	<i>Luis Benveniste</i>
Center for Universal Education, Brookings Institution	<i>Tamar Manuelyan Atinc</i>
UNICEF	<i>Jo Bourne Pia Britto</i>
Global Partnership for Education	<i>Karen Mundy</i>
World Health Organization	<i>Tarun Dua</i>

Primary responsibility for the report was taken by the technical lead, Abbie Raikes, with assistance from the core team and with inputs from Paul Dizona (University of Nebraska), Leslie Hawley (University of Nebraska), Traci Kutaka (SRI International), Alonzo Sanchez (World Bank) and Rebecca Sayre (World Bank). Rebecca Kraut provided valuable editorial support to the report.

Funding support

Support for MELQO came from the Children's Investment Fund Foundation, Dubai Cares, the Government of Norway, Omidyar Network, and the William and Flora Hewlett Foundation. Special thanks to Carmela Salzano and Janne Kjaersgaard for their contribution to the MELQO Evaluation, through the Children's Investment Fund Foundation.

The views expressed here do not necessarily reflect the opinions or official policies of these institutions. The organizations also recognize that the value they provide is in their absolute commitment to quality, independence, and impact. Activities supported by these donors reflect this commitment and the analysis and recommendations are not determined or influenced by any donation.

Abbreviations

ACEI	Association for Childhood Education International
CLASS	Classroom Assessment Scoring System
ECD	Early childhood development
ECDI	Early Child Development Index (MICS)
ECE	Early childhood education
ECERS	Early Childhood Environment Rating Scale
EDI	Early Development Instrument
EGMA	Early Grade Mathematics Assessment
EGRA	Early Grade Reading Assessment
GAML	Global Alliance for Monitoring Learning
GPE	Global Partnership for Education
HECDI	Holistic Early Childhood Development Index
IRB	Institutional Review Board
ISSA	International Step by Step Association
LMIC	Low- and middle-income country
LMTF	Learning Metrics Task Force
M&E	Monitoring and evaluation
MELE	Measure of Early Learning Environments
MELQO	Measuring Early Learning Quality and Outcomes
MICS	Multiple Indicator Cluster Survey (UNICEF)
MODEL	Measure of Development and Early Learning
MoEVT	Ministry of Education and Vocational Training (Tanzania)
n.d.	No date
NGO	Non-governmental organization
OECD	Organisation for Economic Co-operation and Development
PISA	Programme for International Student Assessment (OECD)
PPE	Pre-primary education
SABER	Systems Approach for Better Education Results (World Bank)
SDG	Sustainable Development Goals
SEAMEO	Southeast Asian Ministers of Education Organization
TAG	Technical Advisory Group
UIS	UNESCO Institute for Statistics
UN	United Nations
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund
USA	United States of America
USAID	United States Agency for International Development
WCARO	West and Central Africa Regional Office (UNICEF)
WHO	World Health Organization

Introduction

In 2015, Member States of the United Nations agreed upon an ambitious set of global Sustainable Development Goals (SDGs). For the first time, early learning and development is featured in the global development agenda via SDG Target 4.2, which calls upon Member States to ensure, by the year 2030, ‘that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education’ (UN, 2015). With this target comes a call to action for the early childhood education community to improve the quality, feasibility and accessibility of population-based measures of early childhood environments and learning outcomes associated with readiness for primary education.

The Measuring Early Learning Quality and Outcomes (MELQO) initiative began in 2014 in anticipation of this new global emphasis on early childhood development (ECD). Led by UNESCO, the World Bank, the Center for Universal Education at the Brookings Institution and UNICEF, the initiative aims to promote feasible, accurate and useful measurement of children’s development and learning at the start of primary school, and of the quality of their pre-primary learning environments. Items are designed for children between the ages of 4 and 6 years. Following the premise that many existing tools include similar items, the leading organizations’ core team worked with a consortium of experts, non-governmental organizations (NGOs) and multilaterals to build upon current measurement tools to create a common set of items organized into modules for measuring: 1) early childhood development and learning, and 2) the quality of pre-primary learning environments.¹ The MELQO core team and experts also collaborated to outline a process for context-specific adaptation of the measurement modules resulting from lessons learned from field-testing in several countries in 2015 and 2016.

The modules are designed to be implemented at scale, with an emphasis on feasibility for low- and middle-income countries (LMICs). A key question addressed by MELQO was the balance between a global tool suitable for use everywhere, and local priorities and goals for children’s development. As a result, MELQO placed strong emphasis on the process of cultural adaptation and alignment with national standards. The modules can be integrated into other measurement tools for young children that may have already been used, or adapted to contain locally generated items with the possibility of using this common set for

regional and global comparisons. In its first phase, MELQO’s long-term goal was to contribute to technically robust measurement tools in a synchronized, integrated manner, which will lead to both national and global improvements in the quality of information on early childhood development and learning. The MELQO consortium has worked to meet this goal in two ways: first, through technical work on developing, field-testing and refining the common items in the modules, and second, by facilitating a platform where experts and users of measurement tools for early childhood can learn from each other’s work.

Many tools have been developed to measure child development and learning and the quality of learning environments for young children, which together represent a wealth of technical ingenuity and commitment to reliable measurement within specific countries or regions. However, there is now greater demand at the global and national levels for population-based tools, especially those that could be used to track trends in equity over time. It is clear that measuring children’s nutrition and health status is not enough to fully address equity because the indicators are narrowly focused on important but not holistic information on children’s development. Nor are enrolment and completion rates sufficient measures of equity in education, as millions of children are enrolling in school and reaching third grade without acquiring even basic literacy and math skills (UNESCO, 2014). Knowing the quality of children’s learning environments and what children are able to do at the start of school is an essential first step for any improvement efforts, as reflected by SDG Target 4.2. The MELQO consortium envisions its measurement tools as a lever for change, providing data to support specific changes in policies and programmes that will improve learning worldwide.

This report is intended as an overview of the MELQO initiative, outlining the overall purpose of the effort, the module development process and rationale, and the content of the modules themselves – MODEL, focused on measuring child development and learning, and MELE, focused on measuring early learning environments. It is organized into three main sections:

1. Background and rationale

This section provides a discussion of why measurement of ECD is important and how the MELQO initiative grew out of the current dialogue about ECD and equity in the SDGs.

¹ For a full description of the existing tools see MODEL and MELE modules and technical manuals.

2. Building the MELQO modules

This section provides an introduction to the MELQO modules, including a description of their purpose and content, and a brief overview of how they were developed.

3. Using the MELQO modules

This section presents the processes for using and adapting MELQO within the country context, including key considerations for integrating MELQO into existing country systems and building capacity for measurement efforts. It

also details the use of MELQO in countries to date, offers guidance for working with governments and touches on critical issues such as protecting children's rights. More detail on the content and adaptation process for each module appears in the technical manuals for MODEL and MELE.

The report is accompanied by two technical manuals which detail the constructs and items for each module at the end of MELQO Phase 1 and describe the process for adapting and using them within countries.

Background and rationale

This section describes the impetus for creating the MELQO modules, including the link between measurement and equity, and the need for an integrated global measurement tool for early childhood development and learning. It also discusses the origins and rationale behind MELQO's approach to supporting early childhood measurement.

Equity and ECD: why measurement matters

The sustainable development agenda places emphasis on the many interrelated factors needed to promote equity in all countries. Indeed, achieving equity is fundamental to reaching each of the agreed-upon goals for sustainable development. In regards to education, SDG Goal 4 states that countries must 'ensure inclusive and equitable quality education and promote lifelong learning opportunities for all', and the targets under this goal specifically emphasize learning and the importance of ensuring that all people are able to acquire the skills and competencies required for full participation in society (UN, 2015).

While assessing learning at all levels of education is critical for tracking progress towards these targets, measuring children's development and learning at the start of school is especially important for equity. Years of research on child development have demonstrated that the rate of young children's neurological development peaks in the early childhood years (Shonkoff and Phillips, 2000). During the first eight years of life, children develop increasingly complex cognitive, linguistic, social-emotional and physical skills, which have long-term implications for learning and school achievement, and also set children on lifelong trajectories towards health and well-being. The rapid rate of growth during this period

means that young children's environments are especially important, and interventions put into place at this time can be more effective than those implemented later in life. If school systems are functioning well, the gaps between advantaged and disadvantaged children should lessen over time. The only way to know whether this is the case is to establish a baseline by measuring children's learning at the start of school. While measuring health and nutrition is critical, these indicators are not enough to fully capture the state of children's learning and development.

Equally important for understanding whether equity is being achieved is measuring the quality of children's learning environments and their experiences in classrooms or other pre-school and early primary-grade settings. In the past decade, many countries have made progress towards increasing access to early childhood education (ECE), including formal and non-formal pre-primary education (PPE), parenting education and support, and infant and toddler programmes. In particular, access to PPE has expanded globally, with the number of children enrolled in pre-school increasing by 46 per cent between 1999 and 2010 (UNESCO, 2012). Yet these efforts to expand ECE have generally not been matched by efforts to improve quality. Evidence from high-, middle- and low-income countries alike demonstrates that even when access goes up, children's outcomes do not always improve (e.g. Wong et al., 2013). At times efforts to increase access may even exacerbate the problem of low-quality ECE: the rapid scaling of PPE in recent years may have amplified underlying issues related to ECE programme quality in some countries, as governments focused on extending services without devoting commensurate attention to building a quality infrastructure to support the expansion (Rao and Sun, 2015). Ensuring equity in education requires measuring not just access to pre-primary educational services but also the quality of the environment in which those services are being delivered.

The foundations of global ECD measurement

When MELQO was initiated, the post-2015 debate on goals, targets and measurement had recently begun, and several global efforts highlighted the need for better data on ECD. First, in January 2013, the World Health Organization (WHO) held a meeting to emphasize the importance of linking health, nutrition and education in early childhood, including measurement. Following from that meeting, WHO, UNESCO and UNICEF started work on a global measurement framework for early childhood, from birth to 8 years of age (Raikes et al., 2014). This birth-to-8 framework is intended to cut across sectors to efficiently prioritize, validate and report on new indicators and measurement systems for ECD. At the same time, the Learning Metrics Task Force (LMTF) brought together representatives from organizations and governments in 118 countries to agree on core elements of learning and measurement from early childhood through secondary schooling. The LMTF identified the assessment of early childhood development and learning around the age of school entry as a globally feasible indicator (LMTF, 2013). MELQO was established in 2014 in order to take action on the recommendations proposed by these two related initiatives. Additionally, development of the Holistic Early Childhood Development Index (HECDI) also identified the need for more coordinated and holistic approaches to ECD measurement (UNESCO, 2014).

Now that the SDGs and corresponding targets and indicators have been agreed upon, MELQO's original goals are still central to promoting high-quality measurement of ECD – but there is now greater opportunity to address the potential and challenges of national measurement, especially in relation to national and regional measurement of progress towards SDG Target 4.2, which is focused on early childhood development and learning. It is also critical to note that ECD contributes to several other goals and targets, including education targets focused on lifelong learning, and goals focused on health, poverty reduction, nutrition and gender equity. Early childhood measurement can help support progress in many areas.

Tracking progress towards goals for young children: measurement at scale

As part of the SDGs, several types of indicators have been proposed. A small set of global indicators has been identified to help support global tracking. A larger set of indicators, referred to as 'thematic' indicators, has also been proposed to provide a more comprehensive view of each target. National and regional indicators will also be developed to build more locally relevant monitoring systems. For SDG Target 4.2, the specific global indicator is the 'percentage of children under age five who are developmentally on track in

health, learning and psychosocial well-being' (UNESCO, 2015). Thematic indicators include ECE participation and access. Most centrally, tracking progress towards this target requires measurement of children's development at scale – across a country's entire population of children. While the definition of 'developmentally on track' is not yet clear for all countries, it begins with measurement of children's development, ideally at multiple time points but at minimum at least once before school begins, and in a manner that promotes equity by including all children and ensuring alignment with national goals and cultural priorities for children's development. Measuring the quality of children's learning environments at scale is also central to Target 4.2, because without a strong emphasis on quality, investments in PPE are not likely to lead to desired improvements in learning and development.

Measurement at scale means producing population-level, ideally representative data on trends in ECD that can be used to draw conclusions about the overall state of children's well-being and the quality of early childhood services. Measurement at scale is *not* intended to identify children with specific disabilities, or to make decisions about specific classrooms, teachers or children. Rather it aims to provide information on groups of children and the quality of programmes overall, which then can be used to inform system-level decision-making about how best to support young children's development and learning. This emphasis on 'group-level' measurement means that measures must be easy to use and applicable across many diverse groups and settings, and ideally should also accurately represent all children in the population through use of representative samples.

As the sustainable development agenda unfolds and indicators are decided upon to promote tracking at the global level, countries will develop plans for monitoring their progress towards SDG targets. As part of this process, measurement systems will also need to be developed, modified and expanded, particularly in LMICs where monitoring of quality and outcomes in early childhood is especially limited (Winthrop and Anderson Simons, 2013).

Brief overview of existing ECD measurement efforts

The idea of measuring ECD at a population level has been articulated by several experts and organizations, beginning with the application of epidemiological models derived from public health to the measurement of child development (Hertzman, 1999) – for example, the Early Development Instrument (EDI), which was among the first and most validated tools for measuring child development. These population-based measures were designed to provide comparable data on children's development and learning that could be used to map growth and development in different domains using a common set of items. 'Population-based measures' are intended for use at the population level, meaning that use of the measures can

provide an overview of disparities in child outcomes by region and group characteristics.

A vision for measuring ECD in a globally comparable manner was outlined by UNICEF in 2007 (Zill and Ziv, 2007) as the first step towards the development of the MICS-ECDI – the Early Child Development Index of UNICEF’s Multiple Indicator Cluster Surveys, which is the largest source of globally comparable data on ECD. Over the past decade, work has proceeded on several fronts to develop globally comparable measures of ECD and learning environments, with the most comprehensive and representative data to date coming from the MICS-ECDI. With data collected through household surveys, the MICS-ECDI offers a broad and globally comparable look at child development for children between the ages of 3 to 4, while currently undertaking an effort to expand the age group to 2-4 years. The WHO is also in the process of creating a developmental scale for children from birth to 3 years of age, which will be an important complement and an opportunity to generate a single measure that covers the whole age range (0-59 months) together with UNICEF. While MICS and WHO measures are focused on a small set of items, MELQO modules provide a more detailed look at child development and learning at the start of formal schooling and aim to identify the effectiveness of children’s learning environments in supporting their development.

Less work has been done to date in creating tools to measure the quality of children’s early learning environments within national samples in LMICs. However, some tools, such as the Early Childhood Environmental Rating Scales (ECERS) and the Classroom Assessment Scoring System (CLASS) have been used extensively within large, representative samples in the United States (e.g. Zaslow et al., 2016).

In addition to the EDI, MICS-ECDI, ECERS and CLASS, several other measures exist to measure ECD and the quality of early learning environments. Some of these measures are intended to provide globally comparable data, while some were developed for use within specific countries or regions; others were designed more for use as programme evaluation tools or in research studies, and may not be feasible to use across an entire population. Existing tools played an important role in laying the groundwork for MELQO.

Learning from the larger education measurement dialogue

As the MELQO initiative was underway, a larger discussion about education measurement in the sustainable development agenda was also evolving. Several themes emerged from this dialogue, which have implications for MELQO:

- First, there is a much stronger focus on measuring learning than ever before. During the last decade, the emphasis of global monitoring was largely focused on monitoring access to education, including enrolment and participation in

primary and secondary education. While educational access and completion continue to be very important focus areas, there is now broad agreement that learning must also be measured (e.g. UN, 2015; LMTF, 2013).

- Second, while measuring quality across countries and cultural contexts is not easy, there is consensus that attention must be placed on the quality of children’s learning environments (UN, 2015).
- Third, countries are increasingly investing in measurement and are showing particular interest in developing national and regional measurement systems that reflect local values and priorities, as opposed to joining international studies (LMTF, 2014).
- Fourth, globally comparable data plays an essential role in providing information on progress towards global education goals, but the advantages of globally comparable data must be balanced with the more nuanced and nationally relevant information that can come from national-level assessments.
- Finally, the tremendous increase in the total number of learning assessments is notable with innovations in measurement taking place in every part of the world.²

Taken together, the increased emphasis on measurement, the desire for national and regional systems, and the large number of new measures signifies a move towards:

1. coordinating among existing measures to facilitate the development of nationally relevant tools and take advantage of information from global, regional and national assessments;
2. sharing information about innovations in measurement across a wide group of stakeholders; and
3. generating analytic approaches across national and regional assessments of learning.

MELQO was designed in response to demand from governments, civil society and researchers for an approach that reflects these shifting priorities. Many efforts are under way that promote information sharing and networking in the space of ECD measurement, including the Early Childhood Development Action Network, the Inter-agency Expert Group on ECD Measurement in the SDGs convened by UNICEF (whose role is to support the revision, testing and validation of the ECDI measure and to support with capacity-building initiatives for countries) and the Global Alliance for Monitoring Learning (GAML) established by the UNESCO Institute for Statistics to develop the standards and methodologies needed to measure learning outcomes in SDG 4 more broadly, including in Target 4.2

² For more information on learning assessments throughout the world, see: <http://www.uis.unesco.org/Education/Pages/learning-outcomes.aspx>

MELQO's innovation

With all of the measures that exist now, how is MELQO's contribution unique? There are four main attributes that distinguish MELQO from previous efforts:

1. MELQO's modules are open-source and available for all to use;
2. MELQO modules look at both children's development and the quality of their learning environments, creating a more holistic picture of influences on early childhood development;
3. MELQO provides an evidence-based framework for measurement to be adapted at the country level;
4. Participation in MELQO builds further capacity by identifying and engaging a range of local stakeholders in the adaptation and implementation process.

Efforts to measure young children's development are currently ongoing in many countries around the world, yet they mostly operate on a small scale, usually in the context of a particular project or impact evaluation. The process of developing a new instrument, usually led by researchers, is costly and time-consuming. Even adapting an instrument from an existing

one is a labour-intensive undertaking, as there are only a few instruments to choose from that have been validated in LMICs. The result tends to be an instrument which is unique to the country of focus yet not sustainable or scalable at the national level, for a number of reasons:

- the cost of hiring and training enumerators is prohibitive and there is not a consistently available funding source;
- the assessment is not designed to feed into the country's information management system(s);
- the length of the assessment may be too long for regular use; and
- the overall design may not be suited for replication at scale without intensive involvement from the original evaluation team.

This process is carried out in countries around the world, with a great deal of time and resources currently being expended on similar, narrowly focused measurement efforts. The MELQO approach is based on the assumption that if research teams had a globally relevant yet locally adaptable instrument and a simple, straightforward process for integrating the assessment into the country's measurement systems, time and resources could be conserved.

Building the MELQO modules

What are the MELQO modules?

Building on the substantial work already underway in many parts of the world, MELQO used existing measures and the expertise of a technical advisory group to create two measurement modules, one aimed at measuring **child development and early learning (MODEL)** and one focused on measuring the quality of **early learning environments (MELE)** in pre-school and the early grades (beginning with classroom-based settings). These modules are intended to provide the backbone for national measurement by outlining sets of items and a process for adaptation that has been tested in many countries and can now be used as the starting point for national measurement; they may also be used to inform global and regional monitoring. It is anticipated that the modules will be improved and refined as their use and research into their psychometric properties increase. The modules are open-source and freely available to all countries, with particular relevance and applicability within LMICs.

MELQO was designed to create an efficient approach to early childhood measurement that would simultaneously promote national-level monitoring and inform global monitoring, by providing scientifically grounded items that can be supplemented and modified to work well for each country. The initiative was designed to efficiently generate locally relevant data on children's learning and development at the start of school as well as their pre-primary learning environments, with specific relevance to national ECD policy. MELQO modules were designed to address the question of whether sufficient comparability could be achieved to inform global monitoring. After feedback from users and analyses of data, there are some elements of the modules that can be considered comparable, and others that are best used only to inform national policy. As outlined in the technical manuals, the constructs of the MELE module may be useful across contexts, while the specific items may differ. Items from the MODEL module are included for all countries because analyses indicated relevance across contexts.

The modules were developed based on evidence gathered to date through many assessments intended to address feasibility in measurement and the science of child development. Based on expert input and on-the-ground testing, the modules were further refined and condensed to be:

- reflective of consensus on important constructs to measure;
- cross-culturally relevant, as established either through input from experts (MELE) or empirical analyses (MODEL); and
- feasible for use in LMICs.

MODEL – which stands for Measure of Development and Early Learning – measures children’s learning and development through two tools, a direct assessment and a teacher/ caregiver survey, designed to assess the basic domains of children’s development at the start of school, including executive function, social–emotional development and pre-academic skills (early mathematics and literacy skills). It is important to note that MODEL is intended to *characterize* child development and learning by describing what children can do, not prescribe standards of what children should be able to do. Further, MODEL outlines fundamental skills but does not set out to capture all elements of children’s development that may be important to measure.

For each domain, MELQO proposes a ‘global core’ set of items, with one instrument that can be used as a direct assessment with children and another that can be used as a parent and/or teacher survey. The items recommended as ‘core’ are those that have been shown to have a reasonable degree of comparability across countries and cultural

contexts (additional technical detail provided in the manuals and below). The instruments are designed to work together: the direct assessment instrument provides information on children’s early learning, while the teacher/caregiver instruments provide insight into children’s behaviour in school and at home, and, for the caregiver instrument, on children’s family backgrounds and home learning environments. It is important to note that the current MODEL module represents an early draft at the end of Phase 1, which is expected to change as more countries pilot the instruments.

MELE – which stands for Measure of Early Learning Environments – includes seven domains for quality in early learning environments and sample items that may be useful in indexing them. This module is structured in a slightly different way than the MODEL module: it outlines key domains of quality learning environments for young children and provides examples of items in existing tools that have been used to measure constructs within each domain. The state of evidence on quality in early learning environments, and the strong cultural influences on what defines ‘good quality’, led to a decision to focus on constructs rather than specific items. Various tools have been developed to describe elements of quality within classrooms and also include questions on parent, teacher and director experiences and support for quality. There is little experience to date in measuring the quality of learning environments at population level within LMICs, and so, as with MODEL, the MELE items are expected to change in response to feedback from countries and validation in the field.

Table 1: MELQO modules

Module	Instruments	Global core items	Countries field-tested to date
Child Development and Learning (MODEL)	<ol style="list-style-type: none"> 1. Direct child observation 2. Parent or teacher survey 	Includes items that are globally comparable in the domains of executive function, social–emotional development, early mathematics skills and early literacy skills. Items on children’s home and family environments are also included.	Bangladesh, Kenya, Kyrgyzstan, Lao People’s Democratic Republic (Lao PDR), Madagascar, Mongolia, Nicaragua, Sudan and Tanzania
Early Learning Environments (MELE)	<ol style="list-style-type: none"> 1. Classroom observation 2. Teacher survey 3. Supervisor survey 4. Parent survey 	Includes seven key domains of quality: play, pedagogy, interactions, environment, parent/community engagement, personnel and inclusion.	Colombia, Kenya, Nicaragua and Tanzania (mainland Tanzania and Zanzibar)

As explained above, the modules are intended for use in representative samples to identify trends in learning and

development in groups of children, and in the quality of learning environments. The modules are not intended to be

used for individual diagnoses or screening. In the context of MELQO, measurement at scale is intended to inform country, regional and global tracking of progress towards proposed goals related to ECD, as well as other goals to which successful ECD can contribute, such as lifelong well-being, poverty reduction, and women's empowerment and financial viability.

In recognition of the profound influence of culture and context on young children, both the MODEL and MELE modules are designed to be further adapted and aligned with curricula, national goals for learning, and cultural contexts. Many constructs that underlie child development are universal and applicable across contexts. However, the items that are most appropriate to measure these constructs might be different in one country versus another. As well, some items may be unique and context-specific, and therefore would be added based on local priorities and goals. The steps for country-specific adaptation are outlined in the technical manual.

Defining the content of the MODEL and MELE modules

Reliable measurement is based on clear concepts of what is to be measured. Defining the core concepts of child development and learning and the quality of learning environments requires integrating science and practice to generate definitions to guide measurement that is both scientifically valid and feasible for measurement at scale.

The MODEL constructs and items are designed to index the skills and competencies that reflect normative development and acquisition of age- and culturally appropriate competencies for children roughly between the ages of 4 and 6 years across a range of domains. The related term 'school readiness' can be considered the set of fundamental skills and competencies that facilitate children's success in school. Items in the MODEL module reflect the latest science on child development and were selected on the basis of their technical rigour across countries (for a full description of how the items were selected, see the technical manuals).

MELE's early learning environments refer to the settings in which children learn. While this can include a range of environments, MELQO begins with a focus on pre-primary and early primary school environments, including community-based, private and public (or part of the formal education system). In the broadest terms, high-quality environments are those which are most effective in promoting the development of children in ways that are culturally valued, adaptive and encouraging of children's developmental potential. In designing the MELE module for measuring the quality of early learning environments, initial constructs were selected based on consensus among academics on key predictors of child

outcomes and further refined through factor analysis after field-testing.

Development of the MELQO tools rests upon constructs of development, learning and quality, concepts that are widely discussed but not easily defined. A vast body of scientific evidence suggests that the development of children all over the world follows the same fundamental progressive pattern, with variations depending on culture and context (Bowman, 1994; Greenfield et al., 2003; Volger et al., 2008). Thus while countries may need to adapt the definitions to their individual contexts, the basic building blocks should be the same, and differences will often be expressed in the way specific skills are measured, rather than a choice of one skill over another. The definitions used here were informed by existing measurement tools, available scientific literature on child development, and global conventions on children's rights.³ The indicators for SDG Target 4.2, focused on ECD, call for the 'percentage of children who are developmentally on track'. Ideally, countries would be able to define 'developmentally on track' using information on children's development that is reflective of universal steps in development but also locally normed. Few countries have the volume or depth of information required to meet this standard. In the absence of such data, MELQO is designed to help countries move forward and generate the necessary information to know whether children are developmentally on track.

Building on existing tools

MELQO modules build on substantial progress made by civil society and multilateral organizations to develop feasible, actionable measures relevant to ECD. Several tools measuring both child development and learning and the quality of early learning environments have been developed for use at the group level and tested within several countries.

The MELQO modules draw upon and complement these tools but reach beyond their scope and differ from them in several important ways. First, the MELQO modules provide a conceptual frame for integrating the measurement of child development and learning with the measurement of early learning environments, to create a tool that can point to specific changes in learning environments and policies which could help promote children's development and learning. Second, the modules emphasize the link between children's development at the start of school and their development in the early primary grades, by identifying the skills and competencies that promote children's development in the first few years of school. Third, MELQO focuses on adaptability and scalability, with low-cost, easy-to-use materials that can be integrated into

³ E.g. the Convention on the Rights of the Child (UN, 1990).

regular citizen-led assessments,⁴ national assessment systems, and monitoring and evaluation (M&E) systems for learning environments. Finally, the MODEL module emphasizes specific areas that may be globally relevant, resulting in a small set of common items that can be used across existing national and regional measures to generate global estimates of child development and learning. For MELE, the development process focused on constructs rather than items, because there is less consensus at present regarding universally applicable elements of quality learning environments.

Selecting domains, constructs and items

Child development and learning (MODEL) module

The development of content for the MODEL module was guided by the following goals:

1. Measure predictors of children's later academic performance, as identified in existing literature;
2. Rely on a broad definition of children's development, with pre-academic skills, social-emotional development and contextual variables all included to ensure identification of factors associated with inequity relevant for children in LMICs; and
3. Allow adaptation to align with national and cultural priorities (following the framework outlined in the technical manual).

Because the module is intended to be adapted based on national priorities and cultural contexts, the recommended set of items primarily reflects the first two goals, with room to add additional items. Guidance on how to adapt the measure to country contexts is also outlined in the technical manuals. While an effort was made to strike a balance between the first two goals, this was not always possible, as reliable evidence on predictors of children's achievement does not exist in many countries. Together, these three goals guided the identification of domains, constructs and items.

In creating the MODEL module, the MELQO team convened a technical advisory group (TAG) composed of experts who had developed tools to measure early childhood development and learning that had been used in multiple countries. A list of TAG members for the MELQO modules is located in the acknowledgements section of this document, along with identification of a small group of lead experts who were instrumental in the design of the measure.

The TAG first identified three domains, or broad areas, as important to measure:

1. **Executive function:** This domain refers to the management of cognitive processes and includes regulation, approaches to learning and other skills that drive learning across a number of areas, as well as fine motor skills. The domain contains some of the strongest drivers of children's academic performance over time, including sustained attention, working memory and the ability to inhibit impulses.
2. **Social-emotional development:** This domain refers to the skills that facilitate children's successful interactions with others, including peers, teachers and family members. These skills and competencies are crucial to children's inclusion and the development of healthy, positive relationships with peers and adults, and play an important role in helping children engage in school over time.
3. **Pre-academic skills (literacy and mathematics):** These two domains comprise the sub-domains of early mathematics and literacy skills. This area is perhaps the best-established in existing literature and includes early skills like counting and letter/sound identification, which are considered fundamental to developing math and literacy skills later in life, and fine motor skills like being able to draw a person, which are also predictive of later development.

The TAG then selected constructs, or specific indicators of development and learning, in each of the domains. These include, for example, working memory (executive function), peer-to-peer interaction (social-emotional development) and counting (pre-academic skills). After identifying the important domains and constructs to measure, the next step was to index existing assessments (both parent/teacher surveys and direct child assessments) that were developed for use at the group level – meaning they were not intended to diagnose individual children – and had been tested in more than one country, including some in LMICs. The resulting spreadsheets of items were then shared with a group of experts who selected 'core' items which, based on their psychometric performance and relevance in diverse contexts, were considered to best represent the most important constructs to measure.

Many existing assessments contain similar items, reflecting a good deal of consensus on what should be measured, with more emphasis and clarity on pre-academic skills than other domains. Across all existing assessments, a total of 65 different early math and literacy items were identified, many of which were slight variations of other items. From this list, a smaller set was established. Fewer items were identified for social-emotional development, with a higher proportion measured through parent/teacher reporting only. For executive function, there were very few items available across

⁴ Citizen-led assessments refer to assessments of learning administered through non-profit organizations and designed to document levels of learning within a given population (Results for Development Institute, 2015).

all existing assessments. Both social–emotional development and executive function are areas that will require further development over time.

It is important to note that how the various domains are manifested in different cultures is assumed to vary, with some constructs and items likely to reveal more variation than others. For example, the manifestation of early social–emotional skills might differ significantly depending on cultural norms (e.g. whether children are expected to resolve conflicts with peers on their own or ask for adult intervention), whereas early academic skills such as counting may vary slightly based on exposure to materials but generally apply to all children as the formative steps in learning mathematics.

Several versions of the items were tested in seven countries, with ongoing evaluation of the feasibility and accuracy of each item. The final version of the module was developed at the end of MELQO Phase 1 based on the results of psychometric analyses (described in the section below on establishing technical rigour).

Early learning environments (MELE) module

Content development for the MELE module followed a similar process, with emphasis on selecting domains, constructs and items that have been shown to predict children's learning across settings, as well as those identified as being critical for protecting children's rights and preserving their well-being. As with the MODEL module, the MELQO team convened a TAG consisting of experts who had developed and implemented measures of quality in early childhood settings, as well as experts on early childhood systems more generally. The acknowledgements section of this document contains a list of TAG members for the MELQO modules, along with the technical experts who were especially instrumental in shaping the measure.

Research indicates that the quality of children's learning environments is impacted by several important elements of the learning setting – including the presence of learning materials within classrooms, teacher–child interactions, and the presence of a clear, usable curriculum – as well as by broader policy elements, such as support for teachers' professional development and the provision of adequate funding for pre-schools. The MELE module is therefore designed to account for the classroom environment as well as the level of support and engagement from parents, communities and programme directors.

Drawing from a meeting of experts on quality in early childhood settings hosted by the International Step by Step Association (ISSA) in September 2014, the TAG started construction of the MELE module by identifying seven domains as important to measure. These domains were

informed by several existing frameworks, including the ISSA's Principles of Quality Pedagogy (ISSA, 2010) and the Association for Childhood Education International's Global Guidelines (ACEI, n.d.), as well as perspectives from several practitioners who have developed and used scales in a range of settings. The domains were selected to be relevant in all countries, with the acknowledgement that how quality is demonstrated within the domains may vary from country to country as well as within countries. Slight modifications were made to the proposed domains based on results from field-testing, leading to the final set of domains and their accompanying constructs described below:

- 1. Environment and physical setting:** The physical space is safe, clean and promotes good health practices. The learning environment provides children and adults with a sense of well-being and community and offers frequent opportunities for interaction. A variety of culturally relevant and meaningful learning materials are available, including visual displays, books, art supplies, musical instruments, etc.
- 2. Family and community engagement:** Programmes share information, promote positive relationships and create opportunities for parent and community engagement. Families and the local community are actively involved in planning, decision-making and action to improve early childhood care and education.
- 3. Personnel:** The teachers and staff have been trained in ECD and pedagogy and are adequately compensated. There are enough trained staff to maintain appropriate teacher/child ratios for the age groups in the classroom. Teachers receive regular, effective professional development with ongoing opportunities for reflection and skill development.
- 4. Interactions:** Children experience daily interaction with teachers and school staff who are nurturing, emotionally supportive, trained in pedagogy and ECD, and attuned to children's individual needs.
- 5. Inclusiveness:** All children and families have access to high-quality ECE services. Teachers speak the home language of the majority of students. Teachers are trained in providing ECE to children with disabilities and special needs. Teachers foster age-appropriate development, positive social interactions and play between all children in the classroom.
- 6. Pedagogy:** Curriculum content addresses children's physical, social–emotional, linguistic and cognitive development needs and stimulates early literacy and numeracy skills. Child-centred teaching encourages initiative, curiosity, persistence, attentiveness, cooperation, participation and active engagement. Children engage in age-appropriate play, activities and routines. The necessary policies required to support good pedagogy – regarding

human resources, fiscal management, evaluation and quality improvement – are also in place.

- 7. Play:** Children have access to play materials and opportunities to play during the school day, with time for all children to explore and engage in play with peers.

After constructs were identified for each domain, items were selected from existing scales, much in the same way that items were identified for MODEL.⁵ While there was much agreement on the overall constructs, the selection of items was more difficult for the MELE module than for the MODEL module, for multiple reasons. First, although there are many valuable studies in the quality of pre-primary learning environments, there are fewer large-scale studies on quality than on child development and learning. Second, how quality is defined and understood within various cultures appears to be quite varied. A first step is to describe what settings look like in a range of countries, and then to determine if there are in fact items that may be relevant across all settings. Overall, quality measurement is in an earlier phase of development than the measurement of ECD and learning, emphasizing the importance of approaching this effort in a systematic way. MELQO hopes that by summarizing the work to date, it will be possible to objectively assess the similarities and differences between various countries, which in turn can help inform questions about the feasibility of global measurement.

Field-testing of the MELE module in Colombia, Kenya, Nicaragua and Tanzania (both mainland Tanzania and Zanzibar) helped clarify how to simplify the seven domains, which were difficult to measure all at once. Data was collected on the items themselves and on the contextual issues that arose in data collection, such as the time required to train enumerators and administer the module, and the difficulty in defining key terms like ‘positive interaction’ in a manner that was easily understood and culturally relevant across settings. After a factor analysis of the data was performed, the TAG proposed a new format for the MELE module, which emphasizes four main elements of classroom settings – play, physical environment, pedagogy/instruction and interactions – and three elements of the broader setting: personnel (including access to professional development), inclusion and parent/community engagement. This format served as the basis for the revision of the MELE module.

Establishing technical rigour

Reaching acceptable standards of technical strength is a high priority for MELQO and an essential responsibility of making a measure strong enough to recommend to governments and to use for tracking progress towards international goals. Any recommendation made to countries on what and how to measure should reflect the most robust and technically sound conclusions on measurement. Yet at the same time, the various challenges involved in measuring child development and learning and the quality of early learning environments across different settings mean that a large investment in module development and validation must be made before the modules can be considered complete. Despite the time and resources required, fully establishing the technical strength of the modules is essential for producing the results desired from MELQO and, perhaps more importantly, for ensuring results are equitable and provide all children with the ability to demonstrate their skills and competencies. Meeting full standards for test development requires extensive documentation of psychometric properties (e.g. Downing and Haladyna, 2006).

The process of validating a set of items proposed for the MELQO modules for use at the national level has been initiated, but the full validation of the modules within national contexts should be part of long-term, national-level plans to integrate child development and quality measurement into national systems. Standards for valid, reliable tests of children’s abilities and the quality of their learning environments have been developed over decades of research on test development, and reaching these standards requires a substantial investment of time and resources. Further, when measures are adapted and applied to new settings, validity must be established again. To date, many of the measures used to assess child development and learning and the quality of learning environments have met some but not all of these standards. Countries interested in measurement should therefore work towards full validation of the scales within each country, with a long-term plan to establish validity and reliability.

A brief summary of the types of information needed to document psychometric properties appears below. **Reliability** refers to the degree to which an assessment tool produces stable and consistent results, while **validity** is the extent to which the test measures what it is supposed to measure. Tests can be reliable without being valid – reliability is necessary but not sufficient for establishing validity. Basic requirements for validity and reliability include the following:

⁵ A list of all scales included as part of the MELQO development process can be found in the MODEL and MELE Technical Manuals.

Reliability

- **Test–retest:** The test produces the same results when administered at different times.
- **Inter-rater:** Different administrators of the assessment are able to use the assessment to gain similar scores.
- **Internal:** The items that measure the same construct are related to one another (inter-item correlations or Cronbach's alpha).

Validity

- **Predictive:** There is an implicit assumption in measures of school readiness that they are valuable because they indicate how well children will perform in school over time; likewise, it is often assumed that quality measures are useful because they will demonstrate how likely children are to gain the desired skills during their time in PPE and beyond. For both quality and child development and learning, it is important to have evidence that the measures predict children's future skills and development.
- **Construct:** The assessments measure what they say they measure and not other constructs – for example, items meant to reflect children's language abilities through expressive language are not measuring children's self-confidence.

- **Cross-cultural:** A key question for this project is the extent to which items work in similar ways when used in larger populations and across different cultures.
- **Concurrent:** Scores on the assessment are correlated with scores on more established and fully validated measures, or participant demographic characteristics (e.g. older children have better scores than younger ones), indicating that the new measure is at least as valid in measuring the underlying construct of interest as other measures, and appropriately reflects demographics.

The table below outlines the results of preliminary testing of each module. A full report with results of data analyses to date will be available in late 2017. The goals for technical development were different for each module and reflected both the overall clarity in the field about how measures should be designed, and the extent to which field-testing sites provided reliable data for analyses. For MODEL, because emphasis was placed on identifying previously used items and testing them in a range of countries to establish basic psychometric properties, more data were available and it was possible to build complex statistical models to inform the final recommendation for the module. For MELE, emphasis was placed on identifying a set of items and a process for defining locally relevant quality measurement. Because there were fewer field-testing sites for MELE, it was not possible to conduct the same type of analyses on both modules.

Table 2: Psychometric properties established

Psychometric properties	Definition	MODEL	MELE
Construct validity	Items accurately index key constructs.	Review of literature and expert input on the most valuable and relevant items.	Review of literature and expert input on the most valuable and relevant items.
Concurrent validity	Items correlate well with other established sets of items.	For concurrent validity, parent/teacher reports were validated against direct assessment items. Analytic techniques including factor analyses were used to determine how items were related to one another.	Analytic techniques including factor analyses were used to determine how items were related to one another.
Inter-rater reliability	Items can be reliably collected and do not vary from one assessment to the next.	Inter-rater reliability established for the direct assessments and teacher/parent reports.	Inter-rater reliability established for observational measures.
Internal consistency	Items within a scale measure the same underlying construct.	Cronbach's alphas calculated for each scale; modifications made to direct assessment and parent/teacher scales based on results.	Cronbach's alphas calculated for scales within each country; results vary by country.
Cross-cultural relevance	Items are relevant across contexts.	Analytic models to test item relevance and meaning across countries were developed, leading to small set of items with evidence of cross-cultural relevance.	Expert input requested to determine cultural relevance and applicability of constructs and items for observational measures.
Predictive validity	Items reliably predict children's development in the future.	Not yet established.	Not yet established.

Achievements and next steps

MELQO has been able to achieve the first steps in developing modules that can be used at a population level in a range of countries, and has made progress both in identifying the scope of work required to support countries in measurement, and in further validating the tools. Achievements include the establishment of a set of constructs and a small set of items to measure child development and learning (MODEL module) that were deemed reliable and conceptually relevant to children in a range of settings and showed signs of psychometric strength across the four countries included in analyses. The MELQO team has also established a common set of constructs to measure the quality of children's early learning environments (MELE module), along with examples of items and a proposed scale that can be further adapted based on national context.

Further steps to continue the process of refining the MELQO modules over the next few years include:

- establishment of predictive validity, both globally and within countries, to determine whether the items on both the MODEL and MELE modules reliably predict children's learning in the primary school years; and
- further development of a module for social-emotional development in particular, which may be more sensitive to cultural expectations than other developmental domains.

Because MELQO's modules are designed to respond to ongoing use and adaptation, MELQO will also engage in 1) documentation of innovative approaches to measurement, to help inform the process of adapting the modules to national contexts; and 2) documentation of results from the modules, to improve items and describe children's development and learning and the quality of learning environments across a range of settings. The contribution of data from many sites over time will support improvement of the modules by providing feedback on how the items are working in different countries. The MELQO team expects that the modules will be updated and improved in response to the data collected.

Using the MELQO modules

MELQO's MODEL and MELE modules are designed to establish a baseline of skills and competencies for groups of children and the quality of their learning environments, which could then be used to identify gaps between groups of children reflecting inequities (such as disadvantages linked to family income, cultural background or geographic location), and potentially to evaluate programmes, if the modules are deemed consistent with the programme model and sensitive enough to detect programme effects. Specific uses for each module are outlined below.

- **MODEL:** This module can be used to identify areas of strength and weakness for groups of children in classrooms or schools (for example, what proportion of children per school has mastered foundational math concepts or what proportion has working memory challenges). Taken together at a regional or country level, these results may be used to establish how well curricula are working in different regions and for different developmental areas. This in turn will inform approaches to curricula at the policy level, allow reasonable expectations to be set for what children are able to accomplish by the end of the school year, and help identify areas where additional support is needed. Scores on the measures can be used to track changes within a region or country over time, or to provide broad-level

comparison between countries on a core set of items, when the same set of core items is used.

- **MELE:** This module is intended to help inform national monitoring systems, by identifying items or areas that could be monitored on an ongoing basis in addition to the health and safety standards that often comprise national monitoring systems. The results from the MELE module can be used to track the overall quality of young children's learning environments and to identify the areas in which specific actions are needed, such as providing teachers with more training and support for interacting with young children, offering teachers professional development opportunities in a particular domain, or ensuring that classrooms have access to high-quality, age-appropriate learning materials for all children.

In most countries, measurement at scale for early childhood development and learning is not fully achievable through school-based measurement alone. Because many young children are not in school, it is necessary to ensure household surveys are also used in order to provide a full accounting of all children in the population. Similarly, measurement of the quality of early learning environments should include both formal and informal settings, such as community-based pre-schools, as well as parent and teacher input on quality.

To capture national and sub-national trends in children's preparedness for formal schooling and the quality of children's learning environments over time, countries must develop a measurement and sampling plan that establishes a regular cycle of administration and data analysis. This is best achieved by integrating the MELQO modules into a country's current measurement system(s). An essential first step, therefore, is to adapt the modules to align with the content and goals of national curricula and regulations governing learning environments.

Guidance on how to begin the country adaptation process for MELQO is presented in the section below. Detailed information on adapting specific items to country contexts can be found in the technical manuals for each module.

MELQO in the country context

The MELQO modules and process are designed to help countries gather information to inform policy decisions and meet national priorities for ECD. While MELQO unfolded in close partnership with the World Bank, UNICEF and other multilateral partners, ministries and other national stakeholders can undertake the MELQO process with as little or as much external support as preferred. It is critical that governments and local stakeholders lead the effort from the very beginning to help define the goals of measurement; guide the process of adaptation to reflect a particular country's context and provide policy-relevant information for that context; and decide on the end uses of the data. Perhaps most importantly, measurement requires a dedicated team with the capacity to walk through each of the steps needed to adapt the measures, implement them, and use the data, especially on an ongoing basis. This section is designed to inform governments in their efforts to adapt and implement MELQO. A subsection also offers preliminary guidance for NGOs, multilateral agencies and others working with governments on incorporating MELQO modules into national monitoring systems.

The MELQO modules are intended for adaption and use across countries to support the design and implementation of locally relevant measurement over time. Many efforts that measure young children's development or quality of learning environments are small-scale, usually operating in the context of a particular project or impact evaluation. Adapting an instrument from an existing one can be labour-intensive and costly, leading to narrowly focused measurement efforts. An instrument that is globally relevant yet locally adaptable can ideally be integrated into a country's M&E system, or used over time to measure national-level trends in child development or the quality of learning environments.

Key considerations for early childhood measurement within the country system

The MELQO modules are meant to be used beyond the evaluation of a specific project; rather, they were designed to be feasible for ongoing use in a country to gather information at a systems level. For this to be achieved, it is critical to consider from the start how the modules will be used within the existing system, in order to align or integrate the assessment efforts into a national monitoring system. Key questions for this process include:

- Is there a unit in the government that can be involved in the first MELQO effort and has the capacity (and mandate) to continue regular measurements?
- Is there an information system in the country with which MELQO could integrate (for example, indicators of pre-school classroom quality in school census efforts, or national education management information systems)?
- Are there other data collection efforts to which MELQO could link (e.g. education surveys, assessments such as EGRA or EGMA, household surveys, etc.)?
- Are data collectors available and sustainable? These could be 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, etc.); local university students; or others.

Use of MELQO in countries to date

The MELQO modules can serve as a starting point for longer-term investment in ECD measurement, beginning with national-level discussions among all stakeholders. The approach to national adaptation and use of the MELQO modules has varied by country. As of December 2016, the MODEL module has been field-tested in Bangladesh, Cambodia, Kenya, Kyrgyzstan, Lao PDR, Madagascar, Mongolia, Nicaragua, Sudan and Tanzania. In each instance, the MODEL module was adapted to contribute to different measurement goals. For example, some countries have used the direct assessment along with the teacher/caregiver report, while others have used different combinations of these, sometimes with significant adaptation. Various versions of the MELE module have also been field-tested in Colombia, Kenya, Nicaragua and Tanzania (mainland Tanzania and Zanzibar). In Colombia, for example, many changes were made through consultations with experts and stakeholders, resulting in a tool that differed substantially from the module tested in East Africa. In Tanzania, a consultation and institutional assessment were undertaken to understand the current policies, curricula, standards and M&E systems for PPE. The information was used to adapt the MODEL and MELE modules to meet the needs of the Tanzanian context. A similar process took place in Nicaragua.

Several lessons on using the modules emerged during Phase 1 of the MELQO initiative. First, it was found that the MODEL and MELE modules may be best suited for use in countries where there are existing standards for early development and learning and the quality of pre-primary environments, which can then serve as the backbone for adapting the measures to meet national standards. The standards ideally should have been developed through a consultative process that engaged a range of stakeholders. Additionally, political will from the government agency responsible for ECE, as well as the ability to provide human and financial resources, are enabling factors for this process. In countries where the government not only expressed enthusiasm for using the modules but also had existing standards, devoted resources to adaptation, and viewed the process as an internal capacity-building exercise, the process tended to be quicker and the outcome more effective. In these countries, governments leveraged internal resources, including assigning a government focal point to coordinate the work, recruiting university students as data collectors, and translating the materials themselves.

In countries where there was interest but less commitment from the government to provide resources, where fewer local experts were identified and engaged, and where standards were emerging but not yet in place, the process was more complex and the overall timeline was longer. In these countries there was a greater reliance on international consultants and experts which lengthened timelines and increased costs. Countries planning to use the modules should examine the existing political will, capacity and overall readiness for undertaking a measurement effort. Lower levels of political will and capacity do not necessarily mean the modules should not be attempted. However, it may be useful to reduce the scale of the project, for example by conducting a study in several districts rather than a national study, or by using teacher and parent interviews instead of direct child and classroom observations. Additionally, it could be helpful to build in a longer planning process which includes getting national agreement on standards and buy-in before beginning training and data collection. Finally, active engagement of local researchers and experts who can help with the adaptation of the measures is also important, both for the initial testing and also for the long-term sustainability of the measurement effort.

While the exact process will vary by country, the following examples from Tanzania, Mongolia, Lao PDR, Madagascar and Nicaragua illustrate five different approaches to implementing the MODEL and MELE modules.

Tanzania

The Government of Tanzania recently changed its education policy to require one year of compulsory pre-primary education. MELQO therefore presented an opportunity for

the country as it plans to operationalize this new policy. In July 2015 the government, with leadership from the Ministry of Education and Vocational Training (MoEVT), convened a MELQO launch workshop to bring together the MELQO global team and international and national experts to plan for piloting the modules in the country. At the workshop, the preliminary institutional assessment was presented in order to inform international and national experts about the country's current status in pre-primary programming and policies. After the launch, further consultations were held to finalize the institutional assessment, and the MoEVT Permanent Secretary appointed a MELQO Task Force to coordinate MELQO activities in Tanzania. MoEVT also convened a one-day adaptation workshop with national experts and practitioners to review the draft global modules, consider how they align with the new draft pre-primary policy, and make necessary adjustments according to the context of Tanzanian classrooms. The MODEL module was field-tested with 200 children and further adapted and refined based on feasibility and relevance. The MELE module was also field-tested, and areas for further refinement were identified. A nationally representative study, including both mainland Tanzania and Zanzibar, is planned for in 2017.

Mongolia

The Mongolian government was interested in assessing the quality of pre-schools and children's development, with a specific interest in outcomes and how they vary by socio-economic status. After reviewing different possible tools, the government decided that the MODEL module seemed the most relevant, given its simplicity and ability to capture a range of levels of child outcomes. An adaptation workshop was organized with the participation of the Ministry of Education, local language experts, local ECE experts (e.g. academics, practitioners involved in pre-school teacher training, etc.), and representatives from UNICEF and the World Bank. Overall, modifications were not focused on creating alignment with existing standards or curriculum but instead were aimed at answering the specific question posed by the government (how outcomes vary by socio-economic status). Adaptation of the direct assessment was minimal (an additional shape – an oval – was added to the shape identification item to make it more challenging). Adaptations to the parent questionnaire focused on socio-economic status and the home environment, while removing almost all references to what parents thought their children could do. Results were used to inform policy decisions regarding the quality of pre-primary education and the differences in access, quality and child development outcomes for various regions within the country.

Lao PDR

In Lao PDR, the MODEL module was used to evaluate the impact of an early education and development intervention on a sample of 6,000 children aged 2 to 7 years. The process of adaptation on both the parent report and the direct

assessment took place via workshops with various members of the government, including the director of the early childhood unit, the manager of the training and curriculum unit, a representative from the M&E unit, a local paediatrician, the manager of the public health unit from the Department of Health, and the director/manager of school health, as well as some trained early childhood teachers. Numerous modifications were needed to adapt the modules to the Lao context and language.

Madagascar

Madagascar was able to test the MODEL module of the MELQO instrument as part of an ongoing World Bank evaluation of pre-schools in the country. While the government was open to both the direct assessment and the teacher/caregiver report, given time and funding constraints it chose to use only the teacher/caregiver report. Adaptation began with one expert consultant translating the items and adapting them for cultural relevance. This version was then reviewed by pre-school experts in the Ministry of Education, who made additional revisions. Representatives from the Ministry of Education as well as directors of private and Catholic pre-schools participated in the adaptation and training process, which was a fairly iterative process that also involved pilot-testing the modules. A firm administered the module and collected data from parents and caregivers. The Ministry of Education used the content and results from the MELQO implementation to inform the design of a new pre-school curriculum.

Nicaragua

Nicaragua expressed interest in expanding their educational measurement system to include pre-school education. By April 2016, a core technical group comprised of Ministry of Education officials and local researchers had reviewed Spanish-translated versions of items from the MODEL module and the Colombian adaptation of the MELE module, with the intent to create alignment with the existing national pre-school policy and curriculum. Further adaptation of the MODEL module was completed through an in-country adaptation workshop. The workshop was led by one international expert consultant who reviewed the group's efforts and made further revisions with input from the group and a comprehensive analysis of key curricular and quality standards. The group then participated in a mini-pilot and made additional modifications based on findings from this exercise. The same process was undertaken for the MELE module, and both modules were renamed to reflect government plans for long-term integration. Implementation of the modules took place in summer 2016. The core technical group supervised data collection, and university students administered the modules. The Ministry of Education plans to use results and findings from collected data to inform design and planning for the country's pre-school measurement system, expected to be launched in 2017.

Overview of the process of implementing the MELQO modules

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 and linking to existing systems; 2) adaptation and pre-field-testing; 3) field-testing and data collection; and 4) analysis and application to policy. Detailed descriptions of each stage are included in the technical manuals, with the caveat that these descriptions are intended only as guidance, as each country has followed a unique process to date.

Planning (2–6 months)

Many implementations of MELQO modules have originated through expressions of interest from ministries of education. There are several steps that can provide information on whether MELQO is an appropriate instrument for the purposes indicated by agencies, and where capacity for implementation can be found. 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 lead agency refers to the institution taking on the majority of the leadership for MELQO, such as a ministry or a multi-lateral organization working in partnership with a ministry. The technical assistance personnel refer to the measurement experts, training team, and those responsible for data collection and analyses. These two teams should meet together to discuss the purpose for MELQO and define roles and responsibilities, with a primary priority placed on defining the main research questions to be addressed through MELQO. The process of defining key questions for MELQO is also typically done in consultation with government and non-governmental actors, including teachers, development partners and national and international experts. As part of an institutional assessment, mapping out the existing curricula and standards, recent policies and reform efforts, and the existing or upcoming educational assessments can provide important input for the research questions.

Uses and misuses of the instrument and results should also be discussed. Examples of uses include population measurement for the purposes of: monitoring and tracking where children are in the development of the skills and abilities measured; aligning existing or proposed standards; and programme evaluation. Examples of misuses include individual-level diagnoses and teacher evaluation.

Figure 1: Four phases of MELQO implementation



Planning phase: discussion topics and roles

AGENCY PROVIDES INFORMATION ON:	TECHNICAL ASSISTANCE PROVIDES INFORMATION ON:	JOINT DISCUSSIONS TOPICS:
<ul style="list-style-type: none"> • Purpose and goals for assessment • Information about in-country expertise in: <ul style="list-style-type: none"> - child development assessment - research design - data collection - data analysis • Other instruments proposed for concurrent use 	<ul style="list-style-type: none"> • Overview of MELQO development and design • Overview of uses/misuses of instrument • Overview of MELQO constructs and examples of items • Overview of assessment process 	<ul style="list-style-type: none"> • Appropriate modules/instruments for purposes outlined by agency • Mapping MELQO constructs onto standards and curriculum • Language(s) of assessment (based on purpose and goal) • Country capacity and need for outside technical assistance • Appropriate level of disaggregation and appropriate sampling frame • Contextual variables that may confound the results (and the possibility of collecting data to reduce these confounds)

It is also important to map out the resources available from the government and development partners (donors) for implementing the MELQO modules. This will help determine the number of children and classrooms (sample size) the initial implementation is able to cover, and the frequency with which the data are collected.

An overview discussion of the assessment process – including adaptation, training, field implementation, data collection, data confidentiality, data analysis, results reporting and policy implications – should also be included. Processes to optimize reliable assessment and collection of accurate data, including date of birth and representative samples, should be discussed as well.

Several countries using the MELQO modules have found it useful to appoint a focal person or task force within the government to oversee the work. This person or group is the liaison between the high-level government officials authorizing the work and the experts and international agencies involved in resourcing the initiative. The initial tasks include proposing a study design including the framework for sample selection, developing a project timeline and budget, identifying a mechanism for data collection, and obtaining ethical approval to assess children as needed. This phase also includes deciding on a name for the initiative, which could be the MELQO name or a name more relevant to the country. Some countries have found it advantageous to use a formal launch of the initiative or other external communication efforts to inform the public and other departments in the government about the process and aims of the initiative.

Following these meetings, if MELQO is still deemed appropriate to the needs of the agency, implementation plans can begin. Preparation tasks for the adaptation workshop should be assigned. Prior to adaptation, the MELQO instruments need to be translated and back-translated and checked by appropriate personnel. Note that the MELQO instruments have been translated into several languages and beginning with those versions may be more efficient. If the modules are not being administered in one of the available languages, they should be translated into the language(s) of administration and then translated back to verify that no changes were introduced to the initial meaning of the items.

Adaptation and pre-field-testing (1–2 months)

The adaptation process refers to the process of ensuring that the modules are aligned with national goals and cultural expectations for young children and early learning environments. MELQO recommends ‘adaptation workshops’ to provide a forum for discussion of the modules. Such workshops should include key stakeholders in the country, including parents and teachers, who are familiar with the language and culture. This should be the process through which core items are adapted to the country/cultural/

language context and any new country-specific items are added. Country-specific items should be added through joint decisions based on the purpose and goals of the assessment, the feasibility of assessing specific constructs, funding, and capacity. In order to insure that the new items measure what they purport to measure, they must be translated and back-translated, field-tested and piloted for reliability and validity. Detailed information on adapting both the MODEL and MELE modules appear in the technical manuals.

The end results of the adaptation workshop are:

- 1) instruments that have been accurately translated into the language(s) of assessment, maintaining the integrity of the items, 2) instruments that include additional items deemed necessary for the goals of the assessment, 3) agreement on the entirety of the instruments as necessary and sufficient to measure those constructs in pursuit of the purposes of the assessment, and 4) establishment of sufficient reliability and validity of the instruments.

Deciding on the language of administration is an important discussion, particularly for the MODEL module. This will depend on the language of instruction in the classrooms being assessed as well as the home language of the children. In Tanzania, for example, the direct assessment was administered in the language of instruction of the classroom, the teacher interview was given to the teachers in the language in which they teach, and the parent interview was administered in the language the respondent preferred. The most recent versions of the MODEL and MELE modules are available in English, with previous versions available in Kiswahili, Arabic, French and Spanish. If the modules are not being administered in one of the available languages, they should be translated into the language(s) of administration and then translated back to verify that no changes were introduced to the initial meaning of the items.

Field-testing and data collection (1–3 months)

The purpose of field-testing is to ensure that the items assess what they purport to test. For MODEL, field-testing involves an iterative process of testing the items with children in order to refine the items so that they obtain the information required. After being adapted, the modules should be tested on a small sample of children to identify any major problems with the items. This can be done by the government task force, international experts or a combination of both. To help in the refining process, it is important to capture contextual information about how the items are working by taking notes on the issues that arise when administering each item, and then to modify the modules to address these issues before moving to field-testing. For MELE, field-testing is focused on making sure that assessors can be reliably trained in the administering the items, and that the items capture quality in a range of settings.

During the field-testing phase, expertise to support the process can come either from an expert on the MODEL and MELE modules (usually from outside the country) or through an expert researcher within the country, who then can work with the national MELQO team to train enumerators to the standard of inter-rater reliability identified in the study design. This training should include an interactive presentation of the items to answer any questions from enumerators, as well as a field component where the enumerators assess real children and classrooms and receive feedback. For some countries, a key component of the training is a brief overview of the principles of ECD. Because collecting data on young children is a relatively new activity in many countries, there are often a limited number of experienced data collectors with knowledge of ECD, even from within the education sector.

To inform national planning and policy and to track trends in equity, representative samples can be very useful. Representative samples provide a comprehensive picture of all children in a country, and allow the data to be analysed for differences between groups. Obtaining a representative sample requires work with a sampling expert to determine a sampling frame that allows disaggregation of data by gender, regions, languages or other factors that may exert influence on either child development and learning or the quality of learning environments. Creation of a sampling frame should be undertaken with members of the research team or statistical offices within the government, to ensure that data are available and that the sample size is adequate to compare between groups.

Analysis and application to policy (1-2 months)

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. There are existing guides and publications on how to analyse and report on national assessment results (see, for example, Greaney and Kellaghan, 2008). 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.

At the time of writing, data from the MODEL module has been used in several countries to begin to inform policy, with more countries planning to use the data in the future. Some potential policy applications for MELQO include:

- **Better understanding young children's developmental status, in order to inform the quality and content of early childhood programmes**

Available evidence indicates that monitoring of ECE quality is relatively limited, particularly in LMICs. Measurement at scale can help inform national, regional and global tracking of progress towards goals and provide information on the quality of programmes overall. This information can be used to inform decision-making about how best to support young children's development and learning. For example, the Ministries of Education in Madagascar and Mongolia used the content and results from their assessment (using the adapted MODEL module) to inform the design of a new pre-school curriculum.

- **Informing funding allocation decisions**

Resources are scarce and governments and funders are interested in allocating resources to interventions that deliver the most impact. Documenting the quality of early learning and the baseline level of children's development can be a powerful tool to advocate for more resources, especially for the schools and programmes most in need. Once interventions are operating, measuring to demonstrate impact and results can help make the case for continued or increased investment.

- **Demonstrating commitment to measuring outcomes and results-based funding as a means to attract investment**

Interest in assessments and in measuring results has increased in recent years, in part driven by 'impact investing' and a push for greater emphasis on M&E from governments and funders. Merely demonstrating increased enrolment in programmes is not an adequate outcome measure; increasingly, the impact of programmes (in this case, the learning and developmental outcomes of children) is the outcome of interest. The Global Partnership for Education's strategic plan for 2016–2020 describes a results-based financing scheme in which the release of 30 per cent of each implementation grant is 'contingent upon countries achieving agreed-upon results in equity, learning and system efficiency.' (GPE, 2016).

To achieve the goals of identifying clear connections to policy and increasing the efficiency of measurement, it is critical that governments lead the effort from the very beginning to help define the goals of measurement, decide on the end uses of the data and guide the process of adaptation to reflect the particular country context. Securing strong government buy-in and participation may not be easy in every country, given the range of competing policy priorities and a tendency for ministries of education to focus resources on the primary and secondary years of the education cycle (GPE, 2015).

Principles for working with governments

This section offers specific guidance to non-governmental agencies working with governments to implement the MELQO modules.

1. Identify policy-relevant entry points.

The exact strategy for initiating the MELQO process in a country will depend on the specific pre-primary landscape and political environment. It is important to understand a government's existing institutional arrangements and capacity, as well as its ongoing or planned policies and programmes for the pre-primary sector. Due to the multisectoral nature of early childhood and the fact that modalities for service delivery, governance and quality assurance vary widely by country, an institutional analysis can help identify the exact entry point for a particular country. Analysing a country's existing pre-primary policy and planning landscape and its existing systems for M&E and quality assurance (including standards, curricula and teacher training programmes) can help identify specific strengths to be leveraged and gaps to be addressed by MELQO. The World Bank's Systems Approach for Better Education Results (SABER) ECD module and student assessment, as well as the OECD's Capacity Needs Analysis methodology within the PISA for Development project, may serve as useful tools during this process (see World Bank, 2013; Clarke, 2012; OECD, 2015). Working with governments from the beginning to understand the specific policy priorities and challenges in the country's pre-primary system will also help outline key research questions. Discussing what key questions the government wants to answer at the outset will ensure that the study design is able to provide relevant information to inform future decision-making.

2. Encourage government ownership from the beginning.

To ensure national ownership and sustainability, launching MELQO in a country should be a government-driven process. Seeking endorsement from a minister, permanent secretary, commissioner or other high-level leader can help bring necessary attention to the importance of measuring child development and learning and the quality of early learning environments, which in many countries might be a relatively new concept. Governments may convene a wide group of stakeholders (e.g. policy-makers, policy planners, academics, civil society representatives, parents, teachers, etc.) to discuss the needs of the country's pre-primary system and consider how MELQO data can be used to address some of those needs. A multistakeholder and participatory process helps the government take the lead in building national buy-in to prioritize early learning and ECD within its education system, ensure alignment with national standards and curricula, and

help present results in a way that is relevant to the country and useful for decision-making.

3. Engage technical lead(s) in government.

While high-level government endorsement will be critical for initial momentum, having a strong technical lead or team to carry work forward is just as (if not more) important. An officially appointed government task force, working group or delegate will keep the government connected throughout the MELQO adaptation and implementation process, ensuring that national needs are considered and government procedures and systems are appropriately navigated. The technical lead or leads will be invaluable for a wide range of activities, including obtaining the necessary background data for sampling; convening the right group of stakeholders; leading decisions on module adaptation; facilitating access to classrooms, teachers and students; assisting in ethical and research clearances; and synthesizing and disseminating findings to the people who need them.

Building capacity for country-level measurement

In many low-income countries, the only data collected in early childhood are related to enrolment in formal pre-primary programmes and possibly the age of the child, although in countries where birth registration is not universal, even age can be hard to capture. Some countries have information on the qualifications of teachers and on types of provision (i.e. public, private, community-based, faith-based, etc.). In many countries, an inspector visits a pre-primary classroom once per year, or even less often, and measures aspects such as square meters per child and the number of toilets. This level of quality assurance and monitoring does not lend itself to a more in-depth evaluation process, with a focus on building towards improvement.

For MELQO to be used in a sustainable way, there is a need to build capacity at the country level, including emphasizing integration of the modules into the ongoing M&E systems for PPE. Building the capacity of local researchers and a government team is also important, both for increasing the efficiency of scaling MELQO, and for creating capacity for high-quality, sustainable early childhood programmes. This is likely to be a long-term process that will take time and resources to address.

In Tanzania, for example, national stakeholders have suggested that for sustainable integration of MELQO, capacity will need to be built at all levels. At the national level, technical teams and policy planners will need to understand how to digest the results and use them for future policy and programme design. At the local level, school inspectors, district education officers and even parents will need to be sensitized to what 'good quality' looks like, which in the case of Tanzania may differ from their current perspectives. Teachers will also need to be

oriented on how MELQO findings can help them improve the quality of their classrooms and cater to the needs and levels of development of their students.

In Mongolia, part of the success of the MELQO implementation was attributed to local researcher participation in adaptation and training workshops, during which their capacity was further strengthened. The government also expressed interest in building capacity at the departmental level; while this was not feasible during the first round of MELQO, it is a priority for the future. One of the factors that led to the success of the survey firm was that supervisors participated in the adaptation workshop and thereby were exposed to the MELQO modules and the reasons for modification early in the process. Thus they were better able to train and supervise enumerators during implementation.

These and other implementations have revealed several lessons for building capacity for country-level ECD measurement. First, it is critical to think through the timeline from the beginning, taking into account school year dates, availability of key people (e.g. the expert leading the adaptation), time to contract a data collection firm, and so on. Data collected at the start of the school year will lead to different conclusions than data collected at the end of the year, and decisions on the timeline should be made within the context of the overall policy goals for measurement.

Second, to enhance sustainability, local researchers should be engaged in each country beginning with adaptation and field-testing, in order to help with training, adapt items and play a role in using data for improvement. The roles and responsibilities for local researchers are outlined in the technical manuals for MODEL and MELE.

Third, plans for how the data will be used should include an emphasis on both policy and practice. Several innovative projects have demonstrated that data can lead to better practices in classrooms, and more work is needed to explore how data can best leverage change within a system. For agencies working with governments, significant time should be dedicated to securing country buy-in and, during the adaptation process, to ensuring items are culturally appropriate and consistent with national goals and priorities.

Protecting the rights of children, teachers and families

The design and implementation of ECD measurement should place children's rights at the heart of all decisions. Risks to children can arise through a number of routes, including subjecting children to tests that are not appropriate for their age or skills; failing to obtain parental consent for children's participation; and, perhaps most critically, using the data to penalize children or their families, such as by making decisions about children's abilities to learn on the basis of measurement results. This includes using the information to exclude children from entering primary school, which could be a violation of their right to education (UN, 1990).

Before any data collection takes place, protection of the rights of all participants, including parents, teachers and children, should be ensured through the appropriate mechanisms in each country. In the USA, all researchers participating in research must obtain approval from an Institutional Review Board (IRB). One issue that surfaces in ethics reviews is how the data will be used. If there is any risk to a child's well-being, proper protection and/or modification to study design should take place. This includes using results from the MELQO modules for decision-making about individual children, teachers or schools, which is considered an inappropriate use of MELQO results. At present, the most responsible use of data from the MODEL module outlined in this report is tracking at the group level, which means that no information about the performance of individual children should be reported. As countries develop systems for identifying and serving children with special needs (including children with disabilities and especially talented children), tests can be further developed to help identify these children and ensure they receive the appropriate services.

References

- Association for Childhood Education International (ACEI). n.d. *Global Guidelines for Early Childhood Education and Care in the 21st Century*. Washington, DC, Author. <http://acei.org/global-guidelines>
- Bowman, B. T. 1994. *Cultural Diversity and Academic Achievement. Urban Education Program. Urban Monograph Series*. Oak Brook, Ill., North Central Regional Educational Lab. <http://eric.ed.gov/?id=ED382757>
- Clarke, M. 2012. *What Matters Most for Student Assessment Systems: A Framework Paper*. SABER Working Paper Series, No. 1. Washington, DC, World Bank. http://wbfiles.worldbank.org/documents/hdn/ed/saber/supporting_doc/Background/SAS/Framework_SABER-Student_Assessment.pdf
- Downing, S. M. and Haladyna, T. M. (eds). 2006. *Handbook of Test Development*. Mahwah, NJ, Lawrence Erlbaum Associates.
- Global Partnership for Education (GPE). 2015. Chapter 3: Domestic and external financing for education. *Results for Learning Report 2014/15: Basic Education at Risk*. Washington, DC, Author. <http://www.globalpartnership.org/content/results-learning-report-2014-15>
- Global Partnership for Education (GPE). 2016. *GPE 2020*. Washington, DC, Author. <http://www.globalpartnership.org/content/gpe-2020-strategic-plan>
- Greaney, V. and Kellaghan, T. (eds). 2008. *Assessing National Achievement Levels in Education*. National Assessments of Educational Achievement, Vol. 1. Washington, DC, World Bank. <http://documents.worldbank.org/curated/en/393021468141552365/Assessing-national-achievement-levels-in-education>
- Greenfield, P. M., Keller, H., Fuligni, A. and Maynard, A. 2003. Cultural pathways through universal development. *Annual Review in Psychology*, Vol. 54, pp. 461–90. doi:10.1146/annurev.psych.54.101601.145221
- Hertzman, C. 1999. Population health and human development. D. P. Keating and C. Hertzman (eds), *Developmental Health and the Wealth of Nations: Social, Biological, and Educational Dynamics*. New York, Guilford Press, pp. 21–40.
- International Step by Step Association (ISSA). 2010. *Competent Educators of the 21st Century: Principles of Quality Pedagogy*. Amsterdam, Author. http://www.issa.nl/sites/default/files/Quality-Principles-final-WEB_1.pdf
- Learning Metrics Task Force (LMTF). 2013. *Toward Universal Learning: Recommendations from the Learning Metrics Task Force*. Montreal, PQ/Washington, DC, UNESCO Institute for Statistics/Brookings Institution. <http://www.brookings.edu/research/reports/2013/09/learning-metrics-task-force-universal-learning>
- Learning Metrics Task Force (LMTF). 2014. *Toward Universal Learning: Implementing Assessment to Improve Learning*. Montreal, PQ/Washington, DC, UNESCO Institute for Statistics/Brookings Institution. <http://www.brookings.edu/research/reports/2014/07/implementing-assessment-improve-learning>
- OECD. 2015. *PISA for Development. Capacity Needs Analysis: Zambia*. Paris, Author. https://www.oecd.org/pisa/aboutpisa/NEW_Pisa%20for%20Development_Zambia_FINAL_revised.pdf
- Raikes, A., Britto, P. R. and Dua, T. 2014. *A Measurement Framework for Early Childhood: Birth to 8 Years of Age*. Discussion Paper. Washington, DC, National Academy of Sciences, Institute of Medicine. <https://nam.edu/wp-content/uploads/2015/06/EarlyChildhoodFramework.pdf>
- Rao, N. and Sun, J. 2015. Quality early childhood care and education in low-resource level countries in Asia. P. T. M. Marope and Y. Kaga (eds), *Investing Against Evidence: The Global State of Early Childhood Care and Education*. Paris, UNESCO, pp. 211–30. <http://unesdoc.unesco.org/images/0023/002335/233558E.pdf>
- Results for Development Institute. 2015. *Bringing Learning to Light: The Role of Citizen-led Assessments in Shifting the Education Agenda*. Washington, DC, Author. http://www.r4d.org/sites/resultsfordevelopment.org/files/resources/Bringing%20Learning%20to%20Light_English.pdf

- Shonkoff, J., & Phillips, D., and the National Research Council and Institutes of Medicine (2000). *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Committee on Integrating the Science of Early Childhood Development. Jack P. Shonkoff and Deborah A. Phillips, eds. Board on Children, Youth, and Families, Commission on Behavioral and Social Sciences and Education. Washington, D.C.: National Academy Press.
- UNESCO. 2012. *EFA Global Monitoring Report 2012. Youth and Skills: Putting Education to Work*. Paris, Author. <http://unesdoc.unesco.org/images/0021/002180/218003e.pdf>
- UNESCO. 2014. *EFA Global Monitoring Report 2013/4. Teaching and Learning: Achieving Quality for All*. Paris, Author. <http://unesdoc.unesco.org/images/0022/002256/225660e.pdf>
- UNESCO. 2015. *Education 2030: Incheon Declaration and Framework for Action*. Paris, Author. http://www.unesco.org/new/fileadmin/MULTIMEDIA/HQ/ED/ED_new/pdf/FFA-ENG-27Oct15.pdf
- United Nations. 1990. *Convention on the Rights of the Child*. Geneva, Switzerland, UN Office of the High Commissioner for Human Rights. <http://www.ohchr.org/Documents/ProfessionalInterest/crc.pdf>
- United Nations. 2015. *Transforming Our World: The 2030 Agenda for Sustainable Development*. New York, Author. <https://sustainabledevelopment.un.org/post2015/transformingourworld/publication>
- Volger, P., Crivello, G. and Woodhead, M. 2008. *Early Childhood Transitions Research: A Review of Concepts, Theory, and Practice*. Working Papers in Early Childhood Development, No. 48. The Hague, Netherlands, Bernard van Leer Foundation.
- Winthrop, R. and Anderson Simons, K. 2013. Can international large-scale assessments inform a global learning goal? Insights from the Learning Metrics Task Force. *Research in Comparative and International Education*, Vol. 8, No. 3., pp. 279–95. <http://rci.sagepub.com/content/8/3/279.full.pdf+html>
- Wong, H. L., Luo, R., Zhang, L. and Rozelle, S. 2013. The impact of vouchers on preschool attendance and elementary school readiness: a randomized controlled trial in rural China. *Economics of Education Review*, Vol. 35, pp. 53–65. http://iis-db.stanford.edu/pubs/23803/Impact_of_Vouchers.pdf
- World Bank. 2013. *What Matters Most for Early Childhood Development: A Framework Paper*. SABER Working Paper Series, No. 5. Washington, DC, Author. http://wbfiles.worldbank.org/documents/hdn/ed/saber/supporting_doc/Background/ECD/Framework_SABER_ECD.pdf
- Zaslow B., et al., 2016. "Quality Thresholds, Features, and Dosage in Early Care and Education: Introduction and Literature Review" <http://onlinelibrary.wiley.com/doi/10.1111/mono.12236/abstract>
- Zill, N. and Ziv, Y. 2007. *Toward a Global Indicator of Early Child Development: Summary Report*. Prepared for UNICEF ECD Unit. http://www.issa.nl/newsletter/11_07_docs/Toward_A_Global_ECD_Indicator.pdf (Accessed 13 June 2016.)

MODEL Module

MELQO Measure of Child Development and Learning

Technical Manual

Table of contents

ABBREVIATIONS	32
INTRODUCTION	33
Background and rationale	33
Approach and methodology	34
DOMAIN 1: EXECUTIVE FUNCTION	35
Domain overview	35
Constructs to be measured	36
Working memory	36
Inhibitory control	36
Assessment items	36
DOMAIN 2: SOCIAL-EMOTIONAL DEVELOPMENT	37
Domain overview	37
Constructs to be measured	39
Self-regulation	39
Social cognition	39
Social competence	39
Emotional well-being	39
Assessment items	40
DOMAIN 3: EARLY MATHEMATICS SKILLS	40
Domain overview	40
Constructs to be measured	41
Verbal counting	41
Set production	41
Mental addition	41
Numeral identification	42
Spatial sense	42
Measurement vocabulary	42
Assessment items	42
DOMAIN 4: EARLY LITERACY SKILLS	43
Domain overview	43
Constructs and items	43
Alphabet knowledge	43
Phonological awareness	44
Expressive vocabulary	44
Listening comprehension	44
Assessment items	44
HEALTH AND FAMILY BACKGROUND	44
ADAPTING AND USING THE TOOL	45
Planning	45
Adaptation	47
Field-testing and full data collection	48
Analysing results	50
APPLICATION TO POLICY	50
CONCLUSION	51
REFERENCES	52
APPENDICES	60
Appendix 1: Domains, constructs and items for MODEL	60
Appendix 2: Example of national alignment worksheets	63
Appendix 3: Item adaptation notes	68

Abbreviations

ASER	Annual Status of Education Report
DA	Direct assessment
DIBELS	Dynamic Indicators of Basic Early Literacy Skills
ECD	Early childhood development
ECE	Early childhood education
EGRA	Early Grades Reading Assessment
GPE	Global Partnership for Education
HTKS	Head, toes, knees and shoulders
IDELA	International Development and Early Learning Assessment
IOM	Institute of Medicine
IQ	Intelligence quotient
IRT	Item response theory
LMIC	Low- and middle-income country
M&E	Monitoring and evaluation
MELQO	Measuring Early Learning Quality and Outcomes
MICS	Multiple Indicator Cluster Survey (UNICEF)
MODEL	Measure of Development and Early Learning (MELQO)
NELP	National Early Literacy Panel
NGO	Non-governmental organization
NRC	National Research Council
PALS	Phonological Awareness Literacy Screening
PIAT	Peabody Individual Achievement Test
PTSD	Post-traumatic stress disorder
SDG	Sustainable Development Goal
TCR	Teacher/caregiver report
UK	United Kingdom
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund
USA	United States of America
USAID	United States Agency for International Development
WISC	Wechsler Intelligence Scale for Children

Introduction

This technical manual outlines the content and items for the MELQO child development and learning module, MODEL (Measure of Development and Early Learning). The module includes parent interview, teacher interview and direct assessment instruments, along with recommendations for measurement of contextual information. The MODEL instruments were designed to serve as a common set of items that can be modified or complemented by other items to reflect national priorities. The MODEL module reflects a set of commonly used items that can be seen as the starting point for national-level assessment of child development and learning for children between 4 and 6 years of age.

The MODEL module contains a set of items that were selected to provide a starting point for national-level measurement, and that are also able to inform regional and global monitoring. The items were selected to represent key areas of children's development at the start of school; because they are predictive of later development; and are feasible to collect on a large scale. Results should not be used to make decisions about individual children. Results can be used to inform policy development and practice related to pre-primary and early primary education, such as informing curricular development and professional development, and identifying groups of children in need of greater support.

Background and rationale

Before describing the content of the MODEL module, it is helpful to explore the context of the Sustainable Development Goals (SDGs). Goal 4 of SDGs focuses on education, and Target 4.2 specifically concerns ECD. Target 4.2 was generated in response to policy-makers' questions about school readiness and whether children are 'developmentally on track' when school begins. In keeping with the intent behind the target, measurement of Target 4.2 should begin with a shared understanding of what it means for children to be developmentally on track. The study of how children develop, which elements of development are universal versus culturally specific, and how early development influences later development, is now decades old. The scientific and theoretical literature on child development have notable implications for population-based measurement, in some ways underscoring both the importance and the difficulty of building measures that are accurate, reliable and useful in influencing policy.

At the highest level, 'developmentally on track' means that children are developing the skills and competencies that will allow them to participate successfully in their environments

and reach their developmental potential, both at present and by building the groundwork for lifelong development. While the idea of developmentally on track is intuitive to many parents, teachers, caregivers and policy-makers, the nature of child development also presents complexities that are critical to understand when deciding how best to measure.

At a basic level, there is substantial agreement that children's development at the start of school should include many domains or areas of development, such as cognition, early mathematics, language, literacy and social-emotional development. Many assessments have items from each of these areas. As well, some of children's early skills are relevant across domains (NRC and IOM, 2000). For example, self-regulation, or the ability to focus attention and behaviour, is hypothesized to be relevant across all domains because it is so central to what children learn and experience. Self-regulation is understood to play an especially central role in school readiness (Eisenberg et al., 2010). Language development also has a strong influence on many other domains of development, including mathematics, literacy and social interactions. Social interactions are especially sensitive to expressive language because children with better language skills can communicate better with peers. Expressive language in particular plays a role across multiple domains, including social-emotional development, literacy and cognitive development. Finally, some areas of development, such as executive function, seem to facilitate the acquisition of new skills and knowledge, and thus show strong relationships with many domains of learning from early childhood onward (Blair and Razza, 2007). Scientific findings therefore suggest that (1) the measurement of ECD should include all domains of development, rather than focusing on early academic skills alone; (2) that domains will be interconnected; and (3) that for some domains the effects of ECD may be apparent several years into the future.

While there is good science backing the basic developmental processes and solid understanding of some of the mechanisms by which environment affects them, the concept of 'healthy development' will have some elements that are consistent for all children and some elements that are based on the culture or context in which children live. What is 'normative' in any one culture reflects scientifically based patterns of development and cultural influences on children's development. Defining typical development is ideally established through the creation of normal distributions of children's development and skill acquisition in various parts of the world, with ages at which percentages of children are typically demonstrating specific skills. For example, a recent project by the World Health Organization

to analyse developmental milestones in various parts of the world for children between birth and age 3 found that some developmental steps, like sitting and standing in infancy, seem to emerge on similar timetables. As children grow older, environmental influences – including poor health and nutrition, home environments and exposure to school – have a greater effect on development, in some domains more so than others, and the range of ages at which children will develop skills within the ‘normative’ range widens. Cultural and contextual influences remain important influences on the scope and timing of developmental achievements throughout childhood.

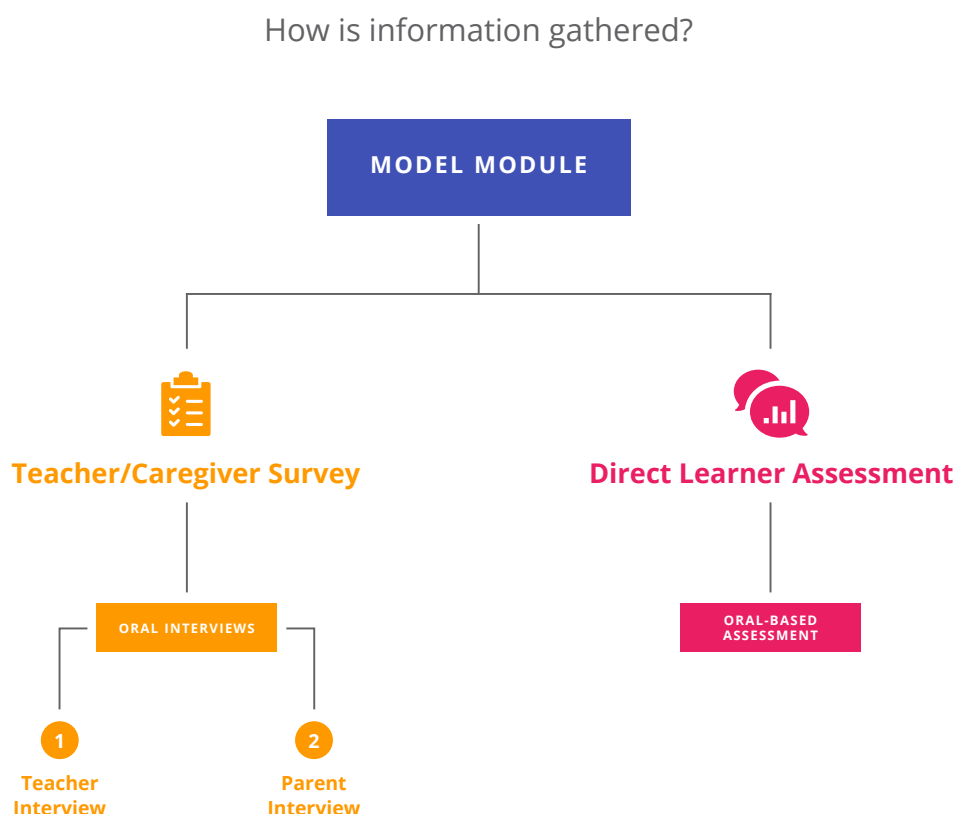
Approach and methodology

The MELQO modules were designed to present a framework that can serve as a starting point for national measurement of ECD. The framework is intended to be further adapted to align with national priorities and goals for young children’s development, and to allow measurement through the method of data collection that is the most expedient and useful for countries. As described in the MELQO Overview document, the items contained in the MELQO instruments were selected by reviewing existing measures, soliciting

feedback from experts and testing the items in a range of countries. An overview of other measures can be found in the upcoming World Bank toolkit for research and evaluation measurement (World Bank, 2016).

There are several tools available to measure young children’s development and learning, MELQO items are just one example; other assessments may use different items to measure the same domains. In surveying existing assessments to construct the MODEL module, the MELQO team identified items they found to be relevant to children across countries in which they worked. Items were selected based on their importance for children’s learning and development, evidence suggesting relevance and applicability across settings, and feasibility to collect at scale. Further, the MELQO team deliberately chose to measure skills that are (1) actionable (e.g. can be taught in the classroom), (2) predictive of later achievement, and (3) related to common curricular goals. This set of items, identified in Appendix 1, can be considered a ‘common module’ that may be useful across many countries. Sources for each item are documented in the Item Source Tables (in Excel format), and full descriptions of each tool are available in the Measures Summaries. Both of these documents are available upon request from the MELQO team.

Figure 1: MODEL Instruments



Measurement of child development and learning can take place through at least three approaches: direct assessment of children by a trained observer who administers a battery of items to each child; parent reports of children's development; and teacher reports of children's development. The first type is referred to as direct assessment (DA), while the second and third types are referred to collectively as teacher/caregiver report (TCR). Each type of assessment can provide reliable information on children's learning and development. There are advantages and drawbacks to each approach: direct assessment, for example, provides specific information on young children's academic skills, while parent or teacher reports provide information on social-emotional development, as parents and teachers have more opportunities to become familiar with children's social skills. The MELQO team prepared the MODEL module using all three methods, so that countries can choose whichever method

seems most appropriate to their aims. The MELQO team strongly recommends that contextual information on each child is also collected, especially information on gender, age, family environments and parents' education, and if possible, children's mother tongue. This information is very valuable in analysing results and using the information to inform policy and practices in early childhood.

The sections that follow describe the domains of development covered by the MODEL module – executive function, social-emotional development, early mathematics skills and early literacy skills; outline the constructs selected to be measured for each domain; and discuss the items developed to measure each construct. It also includes a discussion of health and family background data. The manual ends with guidelines for adaptation and use within country settings, along with information on how to analyse and use results.

Domain 1: Executive function

Domain overview

Executive function refers to a group of skills that promote children's learning across a range of areas, including working memory, self-regulation and mental flexibility. Executive function skills reflect an underlying set of abilities that help children and adults focus on what matters, update ideas to reflect new information, shift their attention and control their impulses. These skills are relevant for people of all ages and are sometimes called neurocognitive skills, stemming from neurological functioning. Children experience rapid growth in executive function in the early childhood years, which sets the stage for continuing development throughout childhood and into adulthood.

Executive function skills are strongly affected by children's early environments. Early caregiving is an important predictor of how well children develop executive function. Children who experience high-quality parenting – in particular, support for exploring in the first few years of life – develop better executive function skills later in life (Bernier et al., 2010). In addition, early health and nutrition play a critical role in executive function. Poor nutrition early in life is especially problematic for developing executive function skills, with effects of early undernutrition on cognition that are apparent through adulthood (e.g. Waber et al., 2014).

There is now a strong body of evidence suggesting that neurocognitive and regulatory abilities have strong and

reliable effects on learning throughout the school years and possibly beyond. Executive function skills are implicated in early learning and are predictive of later mathematics and literacy achievement (Berlin and Bohlin, 2002; Blair and Peters, 2003; Blair and Razza, 2007; Bull et al., 2008; Bull and Scerif, 2001; Duncan et al., 2007; Gathercole et al., 2003; Matthews et al., 2009; Welsh et al., 2010). Evidence from high-income countries shows that children with better executive function skills in the pre-school years learn more math and reading skills in the first years of primary school (Bull et al., 2008; Blair, 2002). Executive function skills may be especially important for keeping children engaged in school over time, and could help untangle the findings indicating that children in the same classrooms often have very different levels of learning (e.g. Glewwe et al., 2014). In sum, measuring executive function is beneficial not only because the skills are predictive, but also because these skills can be remediated in the classroom, and measurement can lead to more effective classroom interventions (Diamond et al., 2007; Whitehurst and Lonigan, 2001).

Overall, executive function may be one of the most critical areas to measure and is likely relevant to children across a range of cultures. However, the manifestation of these skills is not necessarily the same across contexts. Children develop executive function skills with practice, and expectations are higher in some cultures than in others. For example, research has demonstrated that inhibition and self-regulation are

more quickly developed in some Asian countries than in Western countries (Oh and Lewis, 2008). At present, our understanding of why such differences arise in development is somewhat limited. MELQO tools use items that have been shown to work well in a range of cultures, but it is important to engage in discussion with local stakeholders on the items, and pilot-test items in each context to ensure relevance across cultures and countries. It is also essential to continue reviewing tools and updating items as needed based on findings from local research.

Constructs to be measured

Executive function represents a set of cognitive and self-regulation skills needed for goal-directed behaviour (Gioia et al., 2000; Zelazo and Müller, 2002). The three main elements, or constructs, associated with executive function are working memory, inhibitory control and flexible switching (Liew, 2012). Working memory is defined as the ability to store and mentally manipulate memory contents (Baddeley and Hitch, 1974; Davidson et al., 2006). Inhibitory control refers to the ability to suppress the processing of irrelevant yet dominant stimuli or responses (Nigg, 2000). Flexible switching describes a child's ability to adapt responses according to changing demands (Cragg and Nation, 2009).

These three constructs are related to one another but represent distinct skills that can be measured separately. All three have been studied extensively with pre-school populations. Evidence suggests that working memory and inhibitory control are more reliably linked to school achievement than flexible shifting. Among existing batteries of items tested in low- and middle-income countries (LMICs), tasks that require children to demonstrate working memory and inhibit responses seem to have the most reliable associations with later development. Because of this, MELQO has included items to measure only two constructs of executive function: working memory and inhibitory control. These constructs are described in the sections below.

Working memory

Working memory is positively related to general intellectual functioning, mathematics and reading comprehension (Monette et al., 2011; Passolunghi et al., 2007; Swanson and Jerman, 2007; Lesaux et al., 2007). Issues with working memory are associated with difficulty following multistep directions, forgetfulness and inattention (Bignell and Cain, 2007; Engle et al., 1991; Gathercole et al., 2006). Verbal working memory is a weak but significant predictor of attentional ability in everyday activities, such as remembering long sentences and phone numbers (Groth-Marnat and Baker, 2003). Working memory is similarly important for learning,

and may be closely related to inhibitory control. In young children (2 to 6 years old), measures of working memory and inhibitory control are strongly related to one another and can be assessed using the same tasks (Wiebe et al., 2008).

Inhibitory control

Inhibitory control, along with flexible switching, has been identified as one of the first executive function skills to emerge, developing quickly during the pre-school years (Diamond and Doar, 1989; Diamond et al., 2002). Inhibitory control skills help children suppress inappropriate behaviours and thoughts that may distract them from instruction (Alexander et al., 1993; Blair, 2002; Sáez et al. 2012). Further, these skills are essential for cognition, educational attainment and social functioning (Blair, 2002; Espy et al., 2004). Measured prior to kindergarten, inhibitory control predicts mathematics skills in kindergarten; when measured in kindergarten, it predicts mathematics and early literacy skills in primary school (Blair and Razza, 2007).

Assessment items

Working memory and inhibitory control can both be assessed with low-cost and easy to administer tasks. In the first phases of pre-field-testing, the direct assessment portion of the MODEL module contained three items that measure executive function: head, toes, knees and shoulders (HTKS); forward digit span; and backward digit span.

HTKS is a measure of both working memory and inhibitory control. The task was conceptualized by Ponitz et al. (2008) as a measure of inhibitory control, as well as working memory and attention focusing. At the pre-kindergarten level, it predicts letter knowledge, vocabulary and emergent mathematics; when administered in kindergarten, it predicts student sound awareness and letter/word identification (Matthews et al., 2009). In a study conducted in China, South Korea, Taiwan and the United States with 3- to 6-year-old children, HTKS scores predicted children's academic achievement in all samples. Specifically, HTKS scores significantly predicted early mathematics in all four samples, early literacy in all three samples where it was measured (China, South Korea and the USA), and early vocabulary in two out of three of the samples where it was measured (Taiwan and the USA, but not South Korea) (Wanless et al., 2011). A study conducted in Germany and Iceland (von Suchodoletz et al., 2013) found that German children who scored higher on HTKS also scored higher on academic outcome measures (the vocabulary, reading and mathematics subtests of the German version of the Kaufman Assessment Battery for Children, 2006). In Iceland HTKS was related to phonological awareness and single-word reading.

The authors speculated that discrepancies could be due to differences in the children's ages.

Forward digit span measures short-term auditory memory, sequencing and simple verbal expression (Hale et al., 2002). It is part of the Wechsler Intelligence Scale for Children (WISC) assessment and is often administered with backward digit span (Wechsler, 2003). Of the two, forward digit span is thought to be an easier task and may be more appropriate for young children. Two longitudinal studies have examined the role of short-term memory in learning. Bull et al. (2008) found that short-term memory predicts reading and math skills at least a year into the future. Stipek and Valentino (2015) also found a significant predictive relationship between short-term memory and children's initial scores in both math and reading comprehension, and a link between short-term memory and growth in mathematics skills. Savage et al. (2005) found that short-term memory is associated with reading comprehension but not reading accuracy.

Backward digit span measures short-term verbal memory and working memory (Dehn, 2011; Hale et al., 2002). It has been found to predict both general intelligence and academic achievement. Developmental studies examining 5- to 9-year-old children show that measures of working memory (e.g. counting and backward digit recall) typically account

for about 11.6 per cent of the variance in performance on non-verbal measures of fluid intelligence (e.g. Raven's Colored Progressive Matrices) (Engel de Abreu et al., 2010). They were also found to predict the scores of older children on three subtests of the Peabody Individual Achievement Test (PIAT): reading recognition, mathematics and reading comprehension (Dunn and Markwardt, 1970).

Including both forward and backward digit span is important because the memory processes involved in forward recall of both digits and letters are distinctly different from those involved in backward recall (e.g. Reynolds, 1997). Participants typically achieve higher scores when recalling items in their original (forward) order than when recalling them in reverse (backward) order (e.g. Li and Lewandowsky, 1995; St Clair-Thompson, 2010), and the measures differentially predict attention, executive function and scores on behaviour rating measures (Hale et al., 2002). Similar assessments include the digit span subtest from the WISC-IV, which consists of both forward and backward tests.

In the TCR, executive function is measured using interview questions addressing children's abilities to focus, control their behaviour and stay engaged in difficult tasks (see Appendix 1 for specific items).

Domain 2: Social–emotional development

Domain overview

Social–emotional development refers to children's social behaviour and emotional welfare. It consists of several interrelated elements of children's functioning, including self-regulation, social cognition (or understanding) and its impact on prosocial behaviour, social competence, emotional health and well-being, and approaches to learning. At its core, social–emotional development is the process of learning what is culturally and socially appropriate, and then behaving in a manner that allows one to develop strong relationships with others and handle emotions in positive ways. This domain overlaps with executive function, as some elements of executive function also influence social–emotional development.

Social–emotional development takes place within the context of children's relationships with others, beginning early in life. Children who have secure, positive relationships with caregivers develop better social–emotional skills than those

who do not. Through bonding with caregivers, children learn how to understand feelings, maintain close connections even during conflict, and learn social and cultural expectations for social situations. These experiences are central for healthy development and serve as the basis for the quality of relationships throughout childhood and into adulthood (Thompson, 2006).

Exposure to excessive stress – such as that caused by poverty, caregivers with mental health problems, or disruptions in relationships with primary caregivers through separation – leads to biological changes in young children, which in turn increase their susceptibility for poor mental and physical health throughout childhood and into adulthood. Children who have been exposed to mental illnesses like depression or who have experienced very harsh, strict parenting show lifelong negative consequences, with greater stress levels and a greater likelihood of mental illness and school failure (Middlebrooks and Audage, 2008; Felitti, 2009). Of course, all young children experience times

when they are not able to control their emotions or respond positively to others. What differentiates children with strong skills from those without is the frequency of their difficulties – children who consistently have trouble responding well, or who are frequently sad or withdrawn, might be considered to have less developed social–emotional skills than children who encounter such difficulties less often. Children with strong social–emotional skills are part of relationships that can be characterized by supportiveness and intimacy. Additionally, these children are able to manage disagreement and conflict with others, as well as generate socially acceptable strategies to resolve social problems.

Social–emotional development forms the basis for children’s social skills and emotional health throughout life. A decades-long longitudinal study of children in New Zealand demonstrated that social–emotional development has as powerful an impact on later health and wealth as the participants’ socio-economic status (Moffitt et al., 2011). Social–emotional development in early childhood is, therefore, increasingly understood as an important predictor of both school engagement and achievement as well as well-being later in life. Self-regulation in particular is viewed as central to school readiness in the USA, as it affects children’s abilities to get along well with peers and teachers, as well as the ability to engage productively in classroom routines (Liew, 2012). Social competence helps children navigate relationships with peers and teachers, and is linked to children’s successful engagement in school learning activities.

Social–emotional development is also affected by other developmental domains, particularly linguistic development. Language skills help children to better understand and navigate their social worlds. Children with more developed linguistic skills are better able to communicate their feelings, wishes and goals for social interactions (Cutting and Dunn, 1999; Morgan et al., 2010), and can also ask more pointed questions about their own emotions as well as the emotions of others (Beck et al., 2012). Exposure to certain types of language can influence how and when children develop certain social–emotional skills: for instance, Dunn and Kendrick (1982) provide evidence that the birth of a sibling is associated with an increase in family talk about the mental states of others – particularly the feelings and wants of the newborn sibling. The authors, as well as White et al. (2014), suggest that such talk may facilitate children’s prosocial behaviour with their siblings and with peers.

Although existing evidence comes mostly from high-income countries, the importance of optimal early social–emotional development is widely recognized (Wachs et al., 2009). However, the power to actually detect highly predictive patterns is currently limited to those children who demonstrate the highest deficits – only very low social–emotional skills in early childhood have been shown to

correlate reliably with certain behaviours and skills later in life, whereas mid-range or high social–emotional skills are less indicative of future outcomes in existing research.

Also, while social–emotional development is undoubtedly important for children in all cultures, what is considered ‘normal’ or culturally appropriate is likely to be different from one place to the next. For example, relationships and self-regulation are relevant to development for all children, regardless of culture; stress like depression has been shown to have negative effects on children in a range of cultures (Wachs et al., 2009); and children’s social–emotional skills are enhanced by pre-school environments in many countries (Baker-Henningham, 2013). However, cultural expectations for how relationships are formed, what appropriate social interactions look like, and how children should regulate their emotions means that these constructs may not look the same in all cultures. This means that children’s answers to some items will likely take on different meanings depending on the cultural context, and that parents and teachers may interpret items differently as well. For that reason, it is essential to pre-field-test social–emotional items by discussing them with key stakeholders; test items with samples of children; assure alignment with any local research on social–emotional development; and ideally improve items over time.

By its very nature of reflecting a child’s abilities to interact with others, measurement of social–emotional development can be challenging and imprecise. Each of the three main modes of measurement – direct assessment through interviews and tasks, direct observation, and teacher/caregiver report – has advantages and disadvantages, and there is no one method that fully captures the range of skills and competencies that comprise social–emotional development in young children. Measuring social–emotional development poses unique challenges that are not as evident in other domains. For example, when measuring counting, researchers can be reasonably confident that if a child can count to ten in a direct assessment task, they are able to count out five objects if asked to do so by a parent or a friend. But if a child can describe what to do in a conflict with a friend – a common direct assessment task – it is not as certain that they will be able to use the same skill in a similar situation with any of their peers or siblings. Other issues arise when using direct observation or TCR. Reliable observation of children’s behaviour in groups requires extensive training and habituation, while teacher and caregiver reports can contain positive social bias, with respondents presenting a more competent picture of the child than may actually be the case. Direct assessments of children’s social–emotional skills are also heavily dependent on children’s language abilities. Direct assessments often indicate that children with more developed language skills are also more advanced in their social–emotional development; while it may be true

that children with better language skills also have better social skills, such overlaps make results more difficult to interpret and harder to use in improving practice and policy.

In light of these challenges, it is important to affirm that social-emotional development *can* be measured, and that existing measurements, though not comprehensive, provide valuable insights into children's development.

Constructs to be measured

Recent conceptualizations of children's social-emotional development in the pre-school years have focused on two key areas that are central to children's development: competence in building and maintaining relationships with others; and self-regulation, or the ability to successfully manage emotional states (Thompson, 2015). Using these two key principles as a starting point for measurement, there are a number of ways to describe the constructs that can be used to develop items to measure social-emotional development. To provide some categorization, MELQO has identified four constructs of social-emotional development: self-regulation, social cognition, social competence and emotional well-being. These constructs are discussed in the sections below. It is important to note that the social cognition items were tested in early versions of MODEL, but were removed after finding that children's responses seemed sensitive to cultural expectations for discussion and expression of emotions, as described below.

Self-regulation

Self-regulation refers to the ability to control emotions and social behaviour in the interest of engagement and participation in both social interactions and independent work. It encompasses the regulation of one's own emotions both in social contexts and non-social contexts (such as delayed gratification). This construct is also strongly related to executive function, specifically inhibitory control (Liew, 2012).

Social cognition

Social cognition captures children's abilities to think about and comprehend social relationships with others, recognize the feelings of others, and, if required, take actions that are meant to make others feel better. Social cognition encompasses empathy – children's abilities to read others' emotions and respond appropriately – as well as prosocial behaviour that includes helping others who may be in distress. Understanding the feelings of oneself and others includes the ability to (1) comprehend basic emotions (e.g. happiness, sadness and anger) and how these emotions are expressed, as well as their antecedents, causes and consequences; (2) recognize

that emotions are complex (e.g. two individuals can feel two different emotions in response to the same event); and (3) distinguish rules for how to display basic and more complex emotions (e.g. shame or guilt). Accurate interpretation of the feelings of others provides important information about social situations and what constitutes appropriate responses. Indeed, emotional understanding in pre-school and elementary school-aged children is associated with positive peer status, lack of negative peer status or perceptions of rejection, prosocial behaviours and acts, and adult ratings of social competence (Denham et al., 2002; Denham et al., 1990; Denham et al., 1997; Garner, 1996; Izard et al., 2001; Miller et al., 2005; see also Trentacosta and Fine, 2010).

As part of the MODEL development process, the MELQO team tested items on how children think about social relationships through direct assessment, by asking children to describe how they would respond or feel within a given social situation. Results indicated that responses to these items were dependent on children's language and cultural expectations for how and when to discuss emotions, as well as social-emotional skills. In high-income countries, a wealth of social cognition tasks has been devised and tested (e.g. Denham et al., 2010), demonstrating that children who can talk more effectively about social interactions also show better social behaviour. Translating such measures into a global set of core items is challenging because the available measures have largely been tested only in high-income countries. Cultural influences on the social expression of emotions could have profound effects on children's responses – in some settings, for example, children may not be used to talking about feelings, even if they do understand them, and may say things that have no relation to their actual skills, because answering questions about emotions is not typical. In sum, existing measures of social cognition may be useful in LMICs, but more information is needed before it can be assumed that such measures are valid in diverse contexts. Consequently, items related to this construct were removed from the MODEL module.

Social competence

Social competence, or how well children develop and maintain relationships with peers and adults, is a construct that includes the ability to coexist and interact with others in a competent manner – essentially getting along with other children and adults, and being part of a social group. Social competence is an important developmental skill that children start learning from the moment they are born and which forms the foundation for a social human being.

Emotional well-being

Emotional well-being refers to aspects of optimal mental health that, if not evident in a child at this early stage, could

predict more serious mental health problems, such as anxiety – a potential sign of post-traumatic stress disorder (PTSD); sadness, a potential precursor of depression; or the inability to control aggressive impulses – a precursor of oppositional and conduct disorders. These issues could also be signs of other problems in the child's life, such as exposure to violence or a lack of attachment figures. It is understood that other aspects of children's development contribute to their mental health (such as high-quality parenting and safe environments), which are measured by other parts of the MELQO tools. This section covers those aspects of child behaviour that, if reported, could indicate deficits in mental health. Mental health diagnoses for young children have been found to be fairly stable and persistent from 3 to 6 years of age (Bufferd et al., 2012). The existence of chronic anxiety and aggression in childhood, even when not formally diagnosed, can predict the presence of a diagnosable mental disorder later in life (Fryers and Brugha, 2013).

Assessment items

All of the social-emotional development constructs are measured through questions on the TCR (see Appendix 1 for specific items).

Before it was removed from the module, the construct of social cognition was measured through direct assessment, including a task to measure children's emotional self-knowledge (i.e. their ability to understand their own feelings), and a 'perspective-taking' task. Cognitive perspective-taking is operationally defined as the ability to imagine how things are experienced from another point of view (Taylor, 1988), while affective perspective-taking is operationally defined as the ability to understand the feelings of another by taking that person's point of view (Harris et al., 1989). Perspective-taking has been positively associated with prosocial behaviour (for a review see Holmgren et al., 1998).

Domain 3: Early mathematics skills

Domain overview

Mathematics skills refer to the ability to think using mathematical concepts (Nunes and Bryant, 1996), which provide powerful tools for describing and understanding the world around us (Butterworth, 1996; NRC, 2009). In early childhood these concepts generally include numbers and operations, geometry (including spatial thinking) and measurement. Children develop mathematical skills from their interactions with more knowledgeable caregivers, teachers and peers, as well as construct their own understanding of foundational mathematics through interacting with their environments. Talking about mathematics requires a special language, which young children can only gain through interactions with those who already have acquired this knowledge.

Mathematical learning begins in the home and in the community. Early informal mathematical skills – such as counting fingers, sharing toys equally, setting the table with enough plates and cups for everyone, and learning that a mother goat and two baby goats make three goats – serve as the foundation for later formal mathematical learning that children acquire in school. Pre-school curricula build on these early skills and provide instruction in numeracy, informal addition and subtraction, classification, geometry and measurement.

Across the globe there has been a rising interest in understanding the contributions of early mathematics skills to later academic

achievement and economic well-being. Increased policy interest in early mathematics instruction is due in large part to a growing recognition that participation in the modern world requires competence in mathematics, and that early mathematical skills provide the foundation for later learning (Baroody et al., 2006; Clements and Sarama, 2014; NRC, 2001). Without these early skills, children's later acquisition of more complex skills becomes problematic. A lack of early experiences that support mathematics development may lead to lower mathematical skill acquisition as well as lower overall educational attainment (Geary, 2000). All later mathematics are built upon early foundational mathematics, without which later mathematical learning may be difficult or impossible.

In a meta-analysis of six longitudinal studies from three countries, researchers concluded that early mathematics skills at kindergarten entry were the strongest predictor of academic achievement in Grades 3 and 5, compared to early reading, social-emotional and attention skills (Duncan et al., 2007). Further investigations into the effects of persistent low mathematics scores on later achievement showed that mathematics skills are the greatest predictor of college attendance and a major contributor to whether a student graduates from high school (Duncan and Magnuson, 2011). Research on cognitive skills, including mathematics, shows a causal relationship between these skills and later individual economic well-being and broader economic growth. These results have been found in

low-income and high-income countries alike (Hanushek and Woessmann; 2008; Jolliffe, 1998).

While fostering early mathematics skills is universally important, it is particularly vital for LMICs to do so, both to improve the economic well-being of its citizenry and to help boost national economic growth and development overall.

Constructs to be measured

Most of the constructs measured by the MODEL module belong to the subdomain of 'numbers and operations'. This subdomain includes important knowledge foundational to almost all mathematics. While not every child gains mathematical skills in the same order, there is a certain degree of mathematical knowledge that must be acquired before more sophisticated skills can be developed. This includes knowledge of the counting word sequence and understanding of: cardinality and properties of sets; magnitude (e.g. being able to identify the larger of two sets); and early addition and subtraction (e.g. knowing that adding items to a set makes it larger and taking items away makes it smaller). Because these are foundational skills, they are often measured in assessments of children's early mathematical abilities and have been found to be predictive of later mathematical development and achievement. The MODEL module measures three constructs of these early numeracy skills: verbal counting, set production (production of a specified number of objects), and mental addition (adding two small quantities). As children become more proficient in these activities, more advanced numeracy skills such as numeral identification become important, as they lead to using the symbols so necessary to later mathematics. For this reason, the MODEL module also measures the construct of numeral identification.

Outside of the numbers and operations subdomain but equally important to early mathematics development are spatial sense and measurement skills. These skills begin early when children describe and manipulate their environments (e.g. 'My shoes are under the chair', 'I am taller than my brother', and 'These three sticks make a triangle!') and engage in conversations that help them navigate the world and get their needs met. Recent research has begun to support not only the importance of spatial and measurement language in mathematics but also its place as a predictor of later mathematical understanding (Arcavi, 2003; Clements et al., 2004; Gunderson et al., 2012; Pruden et al., 2011). Because of this, the MODEL module also measures the constructs of spatial sense and measurement vocabulary.

The six mathematics constructs measured by the MODEL module – verbal counting, set production, mental addition, numeral identification, spatial sense and measurement vocabulary – are described in the sections below.

Verbal counting

Although very early numeracy does not require knowledge of the number words or sequence (e.g. comparing two groups of objects and stating which group has more objects), later mathematical activities are predicated on this knowledge. The counting words are cultural artefacts, which must be learned through interactions with more knowledgeable others such as peers and caregivers. Most known languages possess these words, although there are exceptions (see Gordon, 2004; Pica et al., 2004). The ability to count fluently and accurately, along with other number-sense skills, is predictive of later mathematics achievement (Jordan et al., 2009). Verbal counting measures children's knowledge of the names and order of the counting words; it does not, however, measure whether children can apply this knowledge to counting sets.

Set production

While knowledge of the counting sequence is important, it does not signify knowledge of a corresponding quantity – in other words, just because a child can count to ten does not mean he or she can count ten objects. The accurate counting of objects requires five subskills: (1) knowledge of the counting word sequence; (2) an understanding of one-to-one correspondence between each object and number word; (3) knowledge that the last number word in the count is the cardinal value of the set (i.e. numerosity); (4) knowledge that objects can be counted in any order; and (5) an understanding that any collection of objects can be counted (Gelman and Gallistel, 1978). The task of producing a smaller set from a larger set tests many of these skills. In order to produce the correct numerosity, children must know the counting sequence, use one-to-one correspondence between each counting word and object, stop counting when the correct numerosity is attained, and not be constrained by counting objects in any particular order.

The ability to count items accurately is a fundamental mathematical skill and is predictive of later mathematical skill development (Clements and Sarama, 2014). Many early mathematics assessments measure cardinal value in one form or another.

Mental addition

Older research suggested that early mathematical development progressed linearly from knowing number words to counting to addition. Recently, evidence has been mounting that even very young children (3 years old) appear to have some limited arithmetic competence, primarily with non-verbal addition and tasks (Huttenlocher et al., 1994). Young children encounter addition with small set sizes frequently – for example, the sharing of toys or resources can

bring up such problems (e.g. pooling a set of marbles for ‘fair’ redistribution). Developmental studies illustrate that children generally adopt more sophisticated strategies for counting as they develop. For instance, counting on from a set size of four stones to add two more stones is fairly straightforward, whereas counting on becomes more problematic when the first set is nine. In this last example, counting all might be a preferred strategy. Early arithmetic activities provide young children with the opportunity to develop a rudimentary understanding of the rules and concepts that inform later number-sense understanding (Bisanz et al., 2005).

Almost all early mathematics assessments include mental addition tasks. These tasks – whether phrased as a word problem (e.g. ‘If Father has four cows and he is given two more, how many cows will Father have altogether?’) or with number words (e.g. ‘Four plus two equals...?’) – generally require young children to use working memory to solve the problem (Fuchs and Fuchs, 2007). The cognitive load is lessened by the availability of manipulatives (e.g. fingers, stones, blocks), and even more so if the manipulatives are related to the word problem (i.e. ‘If you have four stones and I give you two more, how many stones will you have altogether?’) (Clements and Sarama, 2014).

Numeral identification

Proficiency in the use of numbers is exemplified through counting, operations-based fact retrieval (e.g. addition and multiplication), and accurate computation (Jordan et al., 2009). At the core of these competencies is the identification of numerals and knowledge of their corresponding numerosities (Geary, 2000). Young children encounter numerals in their everyday lives through exposure to digital and analogue clocks, apartment numbers, numbers on buildings, prices for items in stores or markets, license plates and the like. In the vast majority of languages, numerals do not resemble the numerosities they represent but are instead abstract symbols with no obvious connection to the numerosity (Chrisomalis, 2010). Therefore children must commit to memory the numeral and its lexical equivalent (the number word), and eventually the numerosity to which the numeral refers. The ability to identify numerals grows over the early childhood years and has been shown to be predictive of later mathematics abilities (Chard et al., 2005).

Spatial sense

Considerable evidence suggests that long before children gain the skills to describe their environments, they are equipped with a skeletal (but unconscious) physical reasoning framework that helps them make sense of how objects and force interact in the world (Baillargeon and Carey, 2012). Development of spatial language varies across cultures and languages: while most children have a significant grasp of spatial relationships

upon which language acquisition can be based (Sarama and Clements, 2009), there is a great deal of cross-linguistic variation in the semantics and syntax of spatial language. A simple example is the difference between the prepositions applicable to door handle. In English, a door handle is *on* a door; in Finnish, a door handle is *in* a door. This variation implies that there is no one-to-one universal preposition-to-position correlation in spatial language, making the acquisition of these spatial terms complicated (Bowerman, 1996). Caregivers and teachers use a variety of prepositions when interacting with young children. Children in turn learn to use these terms accurately in conversations with adults and peers.

Spatial skills are highly related to achievement in math and science, and appear to be foundational to some aspects of mathematics (Casey et al., 1995; Clements and Sarama, 2014; Mix and Cheng, 2011). Children’s early spatial skills predict performance on numerical arithmetic tasks at age 8 (Gunderson et al., 2012). A recent study of spatial skills, as measured by a transformation task, explained a significant portion of the variance in mathematics performance in pre-schoolers (Verdine et al., 2014). Training in spatial tasks can also improve performance in mathematics, specifically missing term equations (Cheng and Mix, 2014).

Measurement vocabulary

Research on measurement suggests that measurement-related procedures can serve as cognitive tools that later help children organize the way they reason in mathematical problem-solving. Measurement activities themselves can help children create a framework for thinking about problem-solving that calls attention to crucial aspects of the provided information (Miller, 1989). Yet long before children learn to formally measure with rulers or scales, they use measurement vocabulary in their interactions with caregivers and peers (e.g. ‘I am taller than you are’). Caregivers and early childhood teachers reinforce understanding of these concepts through a variety of interactions such as requests (‘Please hand me the big spoon’), declarative statements (‘That cow is huge!’), and questions (‘Which do you think is longer?’). Children’s vocabulary has long been a predictor of later academic achievement (Hart and Risley, 2003), and the use of measurement vocabulary relating to size from toddlerhood to pre-school is predictive of later spatial problem-solving tasks (Pruden et al., 2011). Quite a few early childhood assessments contain measurement vocabulary items.

Assessment items

In the MODEL module, constructs of mathematics are measured through direct assessment and questions on the TCR (see Appendix 1 for specific items).

Domain 4: Early literacy skills

Domain overview

Literacy refers to the ability to read, write and use language proficiently, and encompasses the skills necessary to complete these activities. Although literacy instruction has been the focus of school reform, children begin to develop the skills they need to learn to read and write long before they enter school. In the years before formal schooling, children acquire an implicit understanding of how language works – which sounds are acceptable, which words have meaning, what word order is comprehensible to others in their community, and so on. Through these interactions and exposure to rhymes and songs, children develop phonological awareness. During this period they also acquire several thousand vocabulary words, knowledge of letter shapes and names, and an understanding of the grammar and discourse rules of the language(s) they speak (Snow et al., 2005). Further, they begin to relate oral language to written language when they read signs in their environment and when they begin to experiment with writing. Therefore, by the time they enter school, many children understand that there is a relationship between oral and written language and have some knowledge of reading and writing (Graves, 1983). The development of these early language and literacy skills is similar across languages.

The literacy skills that children develop prior to school entry – such as oral language, phonological awareness, reading of logos and concepts of print – are sometimes referred to as pre-literacy or emergent reading skills. MELQO adopts a more expansive view, seeing literacy development as a continuum that begins at birth and continues until children are proficient readers and writers. This perspective recognizes that although there are some prerequisite skills children must acquire, literacy development is not linear. Also, while children develop language naturally by interacting with members of their community, learning to read and write requires experiences with print and instruction.

The importance of learning to read is undisputed. Ensuring that all children learn to read is a goal shared by educators worldwide because they understand that reading serves as a gateway through which children acquire knowledge in writing, math and content-area classes. Meta-analyses and longitudinal studies provide evidence of the foundational skills and knowledge that children need to develop literacy. The National Early Literacy Panel (NELP) conducted a meta-analysis of approximately 300 studies from high-income countries that included data about the predictive relationship

between skills measured in pre-school or kindergarten and later literacy skills (word decoding, reading comprehension, or spelling for children learning to read in an alphabetic language). The results indicate that children's skills related to print knowledge, phonological processing and oral language are independent predictors of later literacy development (NELP, 2008). Further, several specific skills – including alphabet knowledge and phonological awareness – not only correlated with later literacy but also maintained their predictive power even when the role of other variables, such as age, IQ or socio-economic status were accounted for. This indicates that they may be suitable in contexts where children's experiences prior to school entry vary greatly.

Another meta-analysis of literacy development in high-income countries, conducted by Scarborough (1998), examined the findings of 61 studies that used a variety of measures administered in kindergarten to predict reading achievement in first or second grade. Results indicate that the best predictors of reading skills were measures of print-specific knowledge, including letter identification and phonological awareness. Perceptual skills, motor skills and speech perception were less predictive of later reading outcomes. The stability of the findings over time indicates that the antecedent skills that underlie the acquisition of reading are developed early and can be assessed before school entry (e.g. Lonigan et al., 2000; Storch and Whitehurst, 2002). However, in countries where children might not interact with print prior to formal schooling, these measures may be less predictive, since many children score zero even on measures of the most basic literacy skills.

Constructs and items

Since reading is essential for learning, it is important to use measures that are predictive of later reading skills. In research, three skill sets have consistently emerged as the strongest predictors of reading: print knowledge, phonological awareness and oral language (Lonigan, 2006; Lonigan et al., 2008; Whitehurst and Lonigan, 2001). MELQO has selected four constructs that reflect these three skill sets: alphabet knowledge, phonological awareness, expressive vocabulary and listening comprehension.

Alphabet knowledge

Alphabet knowledge includes letter naming and letter sound knowledge. Letter identification is the strongest predictor among reading readiness skills of later literacy achievements,

such as decoding, spelling and reading comprehension skills (Burgess and Lonigan, 1998; McBride-Chang, 1999; Scarborough, 1998; Schatschneider et al., 2004; Wagner et al., 1994; National Early Literacy Panel, 2008). Although some children develop alphabet knowledge incidentally, other will require explicit instruction. Explicit instruction for students with poor alphabet knowledge is essential since these students are likely to encounter difficulty learning to read (O'Connor and Jenkins, 1999; Torppa et al., 2006). Alphabet knowledge is a constrained skill – meaning it is finite and therefore relatively easy to teach. However, children also need to learn how to use that knowledge to read and spell.

Knowledge of both letter names and sounds is assessed in the early grades; the decision to assess one over the other is usually based on the curriculum. Tests that assess letter name knowledge include the Dynamic Indicators of Basic Early Literacy Skills (DIBELS) (Kaminski and Good, 1996); the International Development and Early Learning Assessment (IDELA) (Pisani et al., 2015); the Uwezo Annual Learning Assessment; and the Annual Status of Education Report (ASER). Measures that assess both letter names and sounds include the Phonological Awareness Literacy Screening (PALS) PreK Assessment, and the Early Grade Reading Assessment (EGRA) (RTI International, 2016).

Phonological awareness

Phonological awareness is one of the most researched early literacy skills. It is consistently found to predict the acquisition of later word-reading skills in every language in which it has been studied (Bradley and Bryant, 1983; Roth et al., 2002; Wagner and Torgesen, 1987; Ziegler and Goswami, 2005). The hierarchy of skills is also consistent across languages (e.g. Alcock et al., 2010; Cho and McBride-Chang, 2005; de Jong and van der Leij, 1999; Holopainen et al., 2001; Tolchinsky et al., 2012). For example, Anthony et al. (2003) reported that 2- to 6-year-old children's phonological awareness develops

gradually, with word-level skills developing first, followed by syllable-level skills, then onset/rime-level skills and lastly phoneme-level skills. Phoneme awareness usually develops when children are taught to read (Alcock et al., 2010; Ziegler and Goswami, 2005). Once children learn to read, the amount of information gained from phonological awareness measures decreases (Torgesen, 1999; Wagner et al., 1997).

Criterion-based tests of phonological awareness provide information on children's acquisition of a particular phonological skill. While most norm-referenced tests specify that the child should be 5 years of age or older, there are many skills that can be assessed reliably in younger children.

Expressive vocabulary

Expressive vocabulary refers to the vocabulary a child can produce orally. In five studies, expressive vocabulary measures accounted for a considerable amount of variance in subsequent reading skills (NELP, 2008).

Listening comprehension

Language comprehension – both listening and reading – requires a number of general language abilities. Students must decode words, access meaning, integrate words within sentences and integrate information across sentences. Listening comprehension is moderately related to decoding and reading comprehension (NELP, 2008) and to general reading achievement (Stanovich et al., 1984).

Assessment items

In the MODEL module, expressive vocabulary, alphabet knowledge and listening comprehension are all measured through both direct assessment and questions on the TCR. Phonological awareness is measured only through direct assessment. See Appendix 1 for specific items.

Health and family background

In addition to items measuring child development and learning in the four domains described above, MODEL also includes items focused on children's health status and family context. This information is important for understanding children's development holistically and in particular for identifying the role of family background in shaping children's experiences.

There are several reasons why MODEL includes information beyond children's development and learning. Family environments, health and nutrition status exert very powerful influences on children's development and learning at the start of school (e.g. Walker et al., 2011). While access to pre-primary education plays an important role in preparing children for school, participation in pre-primary education does not mitigate the effects of children's experiences prior to school, as

children's experiences beginning in infancy exert influence on school readiness years later (e.g. Mistry et al., 2010).

Health status: Children with poor health status have more difficulties in the early school years and beyond, when compared with children with good health status at the start of school (Hair et al., 2006; Janus and Duku, 2007; Currie, 2009). Vaccination status, or the number of children who have received up-to-date vaccinations, has been significantly associated with child cognitive development (Bloom et al., 2011). Links between early nutrition status and later cognitive development are very well-established (Engle et al., 2007; Walker et al., 2011), but since early nutrition status requires careful anthropomorphic measurement and is more predictive prior to age 2, MODEL does not include measures of underweight. Instead, MODEL questions focus on children's overall health status, including vaccinations and how frequently children's activities are impaired due to sickness or poor health. Many of these questions also appear in the Multiple Indicator Cluster Survey (MICS) carried out by Governments with support from UNICEF.

Family and home environments: Family and home environments, especially the degree to which children

experience nurturing and stimulating care, is likely one of the most powerful and consistent predictors of children's development (Lancet, 2016). MODEL includes questions on the degree of stimulation that children experience, including access to books and toys, adequate care and typical interactions with parents. Like health items, many of these items also appear in the MICS. Further, parental education levels and income are reliable influences on children's development, and collecting information on these variables is important for tracking equity.

It is recommended that each administration of MODEL include questions on health status and family/home environments for two main reasons: first, as described above, these questions are critical for understanding influences on children's development; and second, these items help establish construct validity, or the evidence needed to determine that MODEL is accurately indexing children's development and learning. Presence of reliable relationships between health, family and home environments and children's development and learning provides validation that the MODEL items are working as intended.

Adapting and using the tool

This section offers specific guidance on working with partners to implement the MODEL module. Because MELQO was primarily intended to support governments in measurement efforts, the explanations below are targeted more towards government ministries than other partners, although the points here are relevant to working with all partners on MELQO. A more complete description of adapting and using both modules appears in the MELQO Overview. Below the general process of MODEL adaptation is outlined; more detailed guidance on specific items that must be adapted for each language can be found in Appendix 3. Adaptation takes place prior to the final construction of the tool for field-testing and data collection. A complete guide to training appears in the Training Manual.

Figure 2 outlines the four main phases of the MELQO implementation process, which applies to both MELQO modules. Each phase is described in detail in the sections that follow.

Planning

Defining policy questions: Planning for use of the MODEL modules should begin with clarification of policy questions that will be informed by the measurement results. Examples of research questions that MODEL can help inform include the following:

- What are the characteristics of children's development and learning at the end of pre-primary or the start of primary school? How do children's skills align with curricula and expectations for learning?
- How do children's skills vary by region, gender or other factors associated with inequity? How does family background influence children's development and learning?
- What are parents' and teachers' perceptions of children's development and learning?

Over time, MODEL is intended to help inform ongoing monitoring and evaluation (M&E) of pre-primary environments by serving as a starting point for developing monitoring tools focused on the most critical elements of children's learning environments. The planning process can

Figure 2: Phases of MELQO Implementation

FOUR PHASES OF MELQO IMPLEMENTATION



be viewed as a first step in building an effective system for monitoring quality over time.

Team structure and skills required: Large-scale data collection and analysis requires a team of individuals with varying levels of expertise and experience with the local population. Ideally, teams should be comprised of individuals at the country level with individuals at the global level serving in an advisory role.

Country-level team members: Depending on roles and responsibilities, the majority of positions should be filled by individuals from the local country. Team members at the country level are best situated to understand and address the needs of the local population. Within each country, MELQO recommends the creation of senior-level positions to assume leadership roles for particular aspects of the data collection and management processes. The section on leadership roles and responsibilities below outlines the recommended leadership structure.

Global-level team members: MODEL is intended to promote sustainable measurement of children's development and learning by government ministries and/or local leaders in ECD. However, depending on the context, some positions require expertise that may not be available locally within the country. In these cases, members of a global advisory team can provide guidance and oversight on specific areas. Advisory team members may include individuals with select technical expertise in areas such as child development, survey research methodology and psychometrics. Depending on the country, some of the global-level functions may be filled within the country, either by the ministry or through a local university, while others may be filled through partnerships with universities outside the country or through connections with multilateral agencies.

Leadership roles and responsibilities:

- **Project Lead:**
 - Overall project manager; responsible for coordinating all technical, training and data components; coordinates with local liaison in the development of country-specific scale components, adaptations and alignment needs
- **Technical and Data Analytic Lead:**
 - Responsible for overseeing the technical quality of the assessment (e.g. sampling design, data analysis)
 - Psychometrician, to assist in determining whether the items are appropriate in overall level of difficulty and to conduct data analyses
- **Training Lead:**
 - Responsible for conducting trainings for classroom observers; in charge of maintenance and fidelity of observer standards
- **Data Collection Lead:**
 - Responsible for coordinating all aspects of data collection; maintains data quality; coordinates with technical lead on data analysis
- **Local Stakeholder and Policy Liaison:**
 - Responsible for working with local parents, teachers, school directors and policy-makers to ensure local voices are reflected in the instruments; coordinates with the technical lead on alignment with national standards

The leadership structure outlined above serves as a starting point for countries to consider when conducting global measurement. In addition to the leadership team, multiple individuals focused on specific tasks and goals are needed within each team in order to conduct quality measurement. With the aid of the global advisory team in certain technical areas and/or in the early phases of implementation, this type of model provides the structure to facilitate countries in taking ownership of their measurement needs and capabilities.

Adaptation

Adaptation is the process by which MODEL core items are adapted to the country/cultural/language context and any new country-specific items are added to improve overall alignment with national standards and priorities for children's development. In order for the core MODEL items to maintain their validity and reliability, adaptations must provide the same sub-construct measurement as the original item. This means that core item adaptation does not include replacement by related but not identical items. New items should be added only if they align with the goals and purpose of the assessment. Country-specific items are added through joint decisions based on the purpose and goals of the assessment, feasibility of assessing specific constructs, funding, and capacity. To ensure that the new items measure what they purport to measure, they must be field-tested and piloted for reliability and validity.

Participants in the adaptation workshop should include experts in: child development, the country's education system, measurement, data collection, assessment language(s), linguistics and curriculum. Where in-country experts are not available, outside expertise should be sought. Technical assistance from personnel who are well-trained on the MELQO instruments and have extensive implementation knowledge is essential. These individuals may also serve in some of the roles previously mentioned (e.g. experts in child development).

Before the adaptation workshop, analyses should be conducted to determine the alignment of the items to national standards. An example of how these analyses have been conducted in other countries appears in Appendix 2. The core items in the module should be translated and back-translated to serve as a starting point.

During the adaptation workshop, the following topics should be addressed:

- Determination of cultural/language/curriculum appropriateness of items on both the DA and the TCR
- Identification and prioritization of country-specific items to develop
- Adaptation of literacy items (choice of letters for letter identification and common initial sounds for initial sound identification)
- Adaptation of mathematics items (type of objects to be counted, word problem context)
- Adaptation of fonts to match curriculum or common letter/numeral displays (numbers and letters in the assessment should look like numbers and letters that children are familiar with)

For MODEL, workshop agendas have begun with a discussion of the domains of child development, followed by presentation of the instruments and the criteria for how the items were selected. Next, the items are selected by reviewing existing curriculum and service delivery standards to identify areas of alignment and non-alignment. Where domains exist in the curriculum or standards and not in the MODEL module, additional items are discussed. Where domains exist in the MODEL module but not in the curriculum and standards, the stakeholder group should decide whether collecting this information would be useful (for example, data could be used to make a case for including these domains in the next national curriculum review). Finally, the group comes to a consensus on the initial set of adapted items for the preliminary field test (see below).

Detailed instructions on adapting items appear in Appendix 3.

By the end of the adaptation workshop, next steps should be agreed upon by the group. Depending on how much work is required to add and/or adapt items, time may not allow for all goals to be reached within the scope of the workshop, so it is important to establish a process for achieving the following: (1) instruments that have been accurately translated into the language(s) of assessment, maintaining the integrity of the items and following the recommendations for modifications outlined in Appendix 3; (2) instruments that include additional items deemed necessary for the goals of the assessment; and (3) agreement

on the entirety of the items in the instruments as necessary and sufficient to measure the constructs in pursuit of the purposes of the assessment. Additionally, it is recommended that the team develop a document outlining the key questions answered by the modules and a preliminary plan for how the results will be used.

Field-testing and full data collection

The purpose of field-testing is to ensure that the items assess what they purport to test. Field-testing involves an iterative process of testing the modules with children in order to refine items so that they obtain the information required. **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. **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.

While conducting pre-field-testing and field-testing, it is important to capture contextual information about how the items are working to help in the refining process, by taking notes on the issues that arise when administering each item, and then to modify the modules to address these issues before moving on to field-testing. This process should ideally be led by the lead researcher/expert or by the government partner. Finally, the team should also keep detailed documentation of the adaptation process, noting which items were modified and why, as well as how they performed in the pre-field test.

A key element of data collection is ensuring **highly trained data collectors**. At the start of the field-testing phase, an expert on the MODEL modules or a local lead researcher should work with the national MELQO team to train enumerators to the standard of inter-rater reliability identified in the study design. This training should include an interactive presentation of the items to answer any questions from enumerators, as well as a field component where the enumerators assess real children and classrooms and receive feedback. For some countries, a key component of the training is a brief overview of the principles of ECD. Because collecting data on young children is a relatively new activity in many countries, there are often a limited number of experienced data collectors with knowledge of ECD, even from within the education sector. A detailed guide to training data collectors is available in the Training Manual.

Table 1: Stages of MODEL Data Collection

STAGES OF MODEL DATA COLLECTION

STAGE	PRE-FIELD-TESTING	FIELD-TESTING	FULL DATA COLLECTION
Purpose	<p>To establish that the newly translated and adapted items are understandable and feasible. The results will not be analysed for assessment of child development.</p> <p>Note: If a country has more than one official language of instruction, this stage and all subsequent stages need to be carried out in each of the languages to be used in the final implementation.</p>	<p>To establish a distribution of scores on the scales of the instrument(s) to allow for psychometric analyses and recommendations for revisions prior to final data collection.</p>	<p>To address the goals of ECD measurement as per country requirements.</p>
Method	<p>The newly translated and adapted items are implemented with children, parents and teachers. Notes are kept on the timing of the testing, difficulties encountered and potential ambiguities in wording.</p> <p>Note: If new, country-specific items are being considered, they need to be included in testing from this stage onwards.</p>	<p>At the start of this phase, a full team of data collectors are trained. The prototype instrument(s) are implemented in a rigorous manner by trained assessors. Care is taken to ensure accurate recording and scoring of items. A subsample of children/parents/teachers are tested/interviewed twice within 2 weeks for test-retest reliability; if feasible, a subsample of children are tested by 2 assessors, and a subsample of parents/teachers are interviewed by 2 assessors for inter-rater reliability.</p>	<p>Implementation of the final version of the instruments in the agreed-upon design (e.g. samples from rural and urban regions, private and public schools, over a period of time, etc.) by the trained assessors with the full rigour of testing and interviewing. Data are entered into databases and analysed to address the project questions.</p>
Sample	<p>Convenience sample for each of the tests being adapted: children for DA, parents for parent TCR, and teachers for teacher TCR.</p> <p>A minimum of 10 children and 5 parent/teacher respondents is required.</p>	<p>The sample size will be dictated by the country's needs for full-scale implementation (to ensure that the groups of interests are represented in the pilot). Clustering of children in classrooms (per teacher) needs to be considered.</p> <p>At a minimum 5 children per classroom in 20 classrooms (100 children, 100 parents, 20 teachers) is required.</p>	<p>Sample is based on the goals identified by the country partners; size is determined by the scope of the project.</p>
Outcome	<p>Responses and administration notes are closely examined to identify problematic items, which may have to be reviewed and potentially revised/removed.</p> <p>The final outcome is the instrument prototype ready for pilot testing.</p>	<p>The results of pilot testing will be analysed for psychometric properties of the items (internal consistency, association with age and gender, test-retest and inter-rater reliability). Items that do not perform well will be identified, and potential reasons for this explored (e.g. bad translation, inadequate administration, etc). Recommendations for final revision of the instruments will be made as needed. The final instruments will then be back-translated for review by the MELQO team. Once the continuity of constructs is confirmed, the instruments will be ready for full implementation.</p>	<p>A data set reflecting children's development and early learning in a country, within the framework of the country requirements. Psychometric properties will be re-examined, and group comparisons will be carried out as planned. Additional validity will be established through association with sociodemographic variables and, if possible, regional characteristics.</p>

Table 1 summarizes the stages of data collection from pre-field testing to large-scale data collection.

Analysing results

Once data have been collected, scores can be calculated indicating overall scores for each child. **These scores are not intended to define which children are developmentally on track or ready for school.** Instead, scores should only be used to examine results in the aggregate and to understand influences on children's development and learning, differences among groups, and the extent to which children's learning aligns with the government standards in place for pre-primary and early primary grades. In addition to routine steps of data cleaning and conducting basic descriptive analyses, the following steps are recommended to ensure the MODEL module is functioning well:

After field-testing:

1. Review descriptive statistics to determine if any items show limited variance. This is a first step in determining whether any of the items are clearly too easy or too hard for the majority of children within the population. Also, view descriptive data to determine if there are any large blocks of missing data and take steps to determine if that data is missing at random. Missing data could also indicate that an item is too hard or easy for the population.
2. Conduct item response theory (IRT) analyses to determine the relative ease or difficulty of the items when taken across the sample. Triangulate the information provided by the IRT model: item difficulty, item discrimination, item characteristics curves, and reliability levels.
3. Remove or revise items that are: (1) redundant, (2) do not reliably distinguish high- and low-performing children from the rest of the sample, or (3) are too difficult or too easy.

After full data collection:

1. Conduct item-level analyses again, using the methods outlined above, to determine the overall patterns of item difficulty.
2. Create total scores from the instruments administered, using the suggested MODEL Syntax as a starting point, and report group differences according to the groups representatively sampled.
3. Establish construct validity by examining relationships between child and family characteristics and child outcomes. For example, children's ages should show a strong and consistent relationship with scores, especially on the direct assessment, as children will show more knowledge and gain higher scores as they grow. Children's scores will also likely be correlated with family income and/or parental education levels.
4. Identify policy implications, by examining differences in scores by region, child gender, or other factors of interest to policy-makers. Scores can also be used to inform policy-makers of the degree of alignment between children's abilities and national standards for pre-primary and early primary grades.

Examples of field-testing reports and recommendations for modifications to items are available for review. A full report on all psychometric properties and results of analyses will be available in late 2017. Please note that in field-testing, the MODEL direct assessment instruments were best described by one factor structure with no reliable subscales. For that reason, the MODEL syntax describes overall scores only, and does not include any syntax to create subscales. Because items may be modified based on context, and additional items may be added to align with national standards, factor structures will change from one country to the next. Factor analyses should be conducted on data from each country. Examples of syntax to calculate overall scores is available based on the version of the assessment used in Tanzania in 2015.

Application to policy

The first phase of MELQO was focused on creating feasible, reliable modules for national adaptation. As outlined in the MELQO Overview, the modules are intended to be used to inform policies and practices to improve child development and learning, by informing curricular standards; helping to identify areas of equity and inequity in child development and learning; and over time, providing information to

teachers, parents and other stakeholders on developmentally appropriate standards for young children. **It is important to note that MODEL results are not intended to be used in high-stakes decision-making about children's readiness for school, as the items are not standardized for each context or intended to be interpreted as a signal of individual children's 'school readiness'.**

Some potential policy applications for MELQO include:

- **Better understanding young children's developmental status, in order to inform the quality and content of early childhood programmes**

Available evidence indicates that monitoring of early childhood education (ECE) quality is relatively limited, particularly in LMICs. Measurement at scale can help inform national, regional and global tracking of progress towards goals and provide information on the quality of programmes overall. This information can be used to inform decision-making about how best to support young children's development and learning. For example, the Ministries of Education in Madagascar and Mongolia used the content and results from their assessment (using the adapted MODEL module) to inform the design of a new pre-school curriculum.

- **Informing funding allocation decisions**

Resources are scarce and governments and funders are interested in allocating resources to interventions that deliver the most impact. Documenting the quality of early learning and the baseline level of children's development can be a powerful tool to advocate for more resources, especially for the schools and programmes most in need. Once interventions are operating, measuring to demonstrate

impact and results can help make the case for continued or increased investment.

- **Demonstrating commitment to outcome measurement and results-based funding as a means to attract investment**

Interest in assessments and in measuring results has increased in recent years, in part driven by 'impact investing' and a push for greater emphasis on M&E from governments and funders. Merely demonstrating increased enrolment in programmes is not an adequate outcome measure; increasingly, the impact of programmes (in this case, the learning and developmental outcomes of children) is the outcome of interest. The Global Partnership for Education (GPE) strategic plan for 2016–2020 describes a results-based financing scheme in which the release of 30 per cent of each implementation grant is 'contingent upon countries achieving agreed-upon results in equity, learning and system efficiency' (GPE, 2016).

To date, MELQO modules have been initiated by several countries to assist in the process of policy planning and programmatic improvements. Details about how each country used the modules is provided in the MELQO Overview. As modules are updated and used in additional countries, more information will be available on how results can be used to inform policy.

Conclusion

Measuring early childhood development and learning provides critical information on the overall status of young children, and on the factors associated with inequity. This information should play a key role in helping to shape policies and practices for young children, perhaps most centrally in ensuring that children have access to services that will help address inequity in early childhood experiences.

In its first phase, MODEL was able to identify a set of items with both scientific backing and policy relevance in a number of countries. This guide is intended to provide preliminary information on how to use the MODEL module, based on

experience to date. This guide will be further updated and expanded as the module is tested in other contexts. To date, MODEL has been useful in outlining policy priorities for pre-primary education and in highlighting the importance of investments in young children. The module has not yet been fully validated, as outlined in the MELQO Overview. Users of the MODEL module should make plans for full validation of the instruments in their countries' contexts, ideally including the establishment of both construct and predictive validity. As more countries use the MODEL module, updated information on the technical properties, including reliability and validity, will be available.

References

- Alcock, K. J., Ngorosho, D., Deus, C. and Jukes, M. C. 2010. We don't have language at our house: disentangling the relationship between phonological awareness, schooling, and literacy. *British Journal of Educational Psychology*, Vol. 80, No. 1, pp. 55–76.
- Alexander, K. L., Entwisle, D. R. and Dauber, S. L. 1993. First grade classroom behavior: its short- and long-term consequences for school performance. *Child Development*, Vol. 64, No. 3, pp. 801–14.
- Anthony, J. L., Lonigan, C. J., Driscoll, K., Phillips, B. M. and Burgess, S. R. 2003. Phonological sensitivity: a quasi-parallel progression of word structure units and cognitive operations. *Reading Research Quarterly*, Vol. 38, No. 4, pp. 470–87.
- Arcavi, A. 2003. The role of visual representations in the learning of mathematics. *Educational Studies in Mathematics*, Vol. 52, No. 3, pp. 215–41.
- Baddeley, A. D. and Hitch, G. 1974. Working memory. *The Psychology of Learning and Motivation*, Vol. 8, pp. 47–89.
- Baillargeon, R. and Carey, S. 2012. Core cognition and beyond: the acquisition of physical and numerical knowledge. S. M. Pauen (ed.), *Early Childhood Development and Later Outcome*. Cambridge, UK, Cambridge University Press, pp. 33–65.
- Baker-Henningham, H. 2013. The role of early childhood education programmes in the promotion of child and adolescent mental health in low-and middle-income countries. *International Journal of Epidemiology*, Vol. 43, No. 2, pp. 407–33.
- Baroody, A. J., Lai, M. L. and Mix, K. S. 2006. The development of young children's early number and operation sense and its implications for early childhood education. B. Spodek and O. N. Saracho (eds), *Handbook of Research on the Education of Young Children*, 2nd edn. Mahwah, NJ, Lawrence Erlbaum Associates Publishers, pp. 187–221.
- Beck, L., Kumschick, I. R., Eid, M. and Klann-Delius, G. 2012. Relationship between language competence and emotional competence in middle childhood. *Emotion*, Vol. 12, No. 3, pp. 503–14.
- Berlin, L. and Bohlin, G. 2002. Response inhibition, hyperactivity, and conduct problems among preschool children. *Journal of Clinical Child and Adolescent Psychology*, Vol. 31, No. 2, pp. 242–51.
- Bernier, A., Carlson, S. M. and Whipple, N. 2010. From external regulation to self-regulation: early parenting precursors of young children's executive functioning. *Child Development*, Vol. 81, No. 1, pp. 326–39.
- Bignell, S. and Cain, K. 2007. Pragmatic aspects of communication and language comprehension in groups of children differentiated by teacher ratings of inattention and hyperactivity. *British Journal of Developmental Psychology*, Vol. 25, No. 4, pp. 499–512.
- Bisanz, J., Sherman, J. L., Rasmussen, C. and Ho, E. 2005. Development of arithmetic skills and knowledge in preschool children. J. I. Campbell (ed.), *Handbook of Mathematical Cognition*. New York, Psychology Press, pp. 143–62.
- Blair, C. 2002. School readiness: integrating cognition and emotion in a neurobiological conceptualization of children's functioning at school entry. *American Psychologist*, Vol. 57, No. 2, pp. 111–27.
- Blair, C. and Peters, R. 2003. Physiological and neurocognitive correlates of adaptive behavior in preschool among children in Head Start. *Developmental Neuropsychology*, Vol. 24, No. 1, pp. 479–97.
- Blair, C. and Razza, R. P. 2007. Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child Development*, Vol. 78, No. 2, pp. 647–63.
- Bloom, D. E., Canning, D. and Seiguer, E. 2011. *The Effect of Vaccination on Children's Physical and Cognitive Development in the Philippines*. Program on the Global Demography of Aging (PGDA) Working Paper No. 69. Boston, Mass., Harvard School of Public Health.

- Bowerman, M. 1996. The origins of children's spatial semantic categories: cognitive versus linguistic determinants. J. J. Gumperz and S. C. Levinson (eds), *Rethinking Linguistic Relativity*. Cambridge, UK, Cambridge University Press, pp. 145–76.
- Bradley, L. and Bryant, P. E. 1983. Categorizing sounds and learning to read: a causal connection. *Nature*, Vol. 301, No. 5899, pp. 419–21.
- Bufferd, S. J., Dougherty, L. R., Carlson, G. A., Rose, S. and Klein, D. N. 2012. Psychiatric disorders in preschoolers: continuity from ages 3 to 6. *American Journal of Psychiatry*, Vol. 169, No. 11, pp. 1157–64.
- Bull, R., Espy, K. A. and Wiebe, S. 2008. Short-term memory, working memory, and executive functioning in preschoolers: longitudinal predictors of mathematical achievement at age 7 years. *Developmental Neuropsychology*, Vol. 33, No. 3, pp. 205–28.
- Bull, R. and Scerif, G. 2001. Executive functioning as a predictor of children's mathematical ability: inhibition, shirting, and working memory. *Developmental Neuropsychology*, Vol. 19, No. 3 pp. 273–93.
- Burgess, S. R. and Lonigan, C. J. 1998. Bidirectional relations of phonological sensitivity and prereading abilities: evidence from a preschool sample. *Journal of Experimental Child Psychology*, Vol. 70, No. 2, pp. 117–41.
- Butterworth, B. (ed.) 1996. *Mathematical Cognition: Volume 1*. Hove, UK, Psychology Press.
- Casey, M. B., Nuttall, R., Pezaris, E. and Benbow, C. P. 1995. The influence of spatial ability on gender differences in mathematics college entrance test scores across diverse samples. *Developmental Psychology*, Vol. 31, No. 4, pp. 697–705.
- Chard, D. J., Clarke, B., Baker, S., Otterstedt, J., Braun, D. and Katz, R. 2005. Using measures of number sense to screen for difficulties in mathematics: preliminary findings. *Assessment for Effective Intervention*, Vol. 30, No. 2, pp. 3–14.
- Cheng, Y. L. and Mix, K. S. 2014. Spatial training improves children's mathematics ability. *Journal of Cognition and Development*, Vol. 15, No. 1, pp. 2–11.
- Cho, J. R. and McBride-Chang, C. 2005. Levels of phonological awareness in Korean and English: a 1-year longitudinal study. *Journal of Educational Psychology*, Vol. 97, No. 4, pp. 564–71.
- Chrisomalis, S. 2010. *Numerical Notation: A Comparative History*. Cambridge, UK, Cambridge University Press.
- Clements, D. H. and Sarama, J. 2014. *Learning and Teaching Early Math: The Learning Trajectories Approach*, 2nd edn. New York, Routledge.
- Clements, D. H., Wilson, D. C. and Sarama, J. 2004. Young children's composition of geometric figures: a learning trajectory. *Mathematical Thinking and Learning*, Vol. 6, No. 2, pp. 163–184.
- Cragg, L. and Nation, K. 2009. Shifting development in mid-childhood: the influence of between-task interference. *Developmental Psychology*, Vol. 45, No. 5, pp. 1465–79.
- Currie, J. 2009. Healthy, wealthy, and wise: socioeconomic status, poor health in childhood, and human capital development. *Journal of Economic Literature*, Vol. 47, No. 1, pp. 87–122.
- Cutting, A. L. and Dunn, J. 1999. Theory of mind, emotion understanding, language, and family background: individual differences and interrelations. *Child Development*, Vol. 70, No. 4, pp. 853–65.
- Davidson, M. C., Amso, D., Anderson, L. C. and Diamond, A. 2006. Development of cognitive control and executive functions from 4 to 13 years: evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, Vol. 44, No. 11, pp. 2037–78.
- Dehn, M. J. 2011. *Working Memory and Academic Learning: Assessment and Intervention*. Hoboken, NJ, John Wiley & Sons.
- de Jong, P. F. and van der Leij, A. 1999. Specific contributions of phonological abilities to early reading acquisition: results from a Dutch latent variable longitudinal study. *Journal of Educational Psychology*, Vol. 91, No. 3, pp. 450–76.
- Denham, S. A., Caverly, S., Schmidt, M., Blair, K., DeMulder, E., Caal, S. ... Mason, T. 2002. Preschool understanding of emotions: contributions to classroom anger and aggression. *Journal of Child Psychology and Psychiatry*, Vol. 43, No. 7, pp. 901–16.

- Denham, S. A., Ji, P. and Hamre, B. 2010. *Compendium of Preschool Through Elementary School Social-Emotional Learning and Associated Assessment Measures*. Chicago, University of Illinois at Chicago/CASEL.
- Denham, S. A., McKinley, M., Couchoud, E. A. and Holt, R. 1990. Emotional and behavioral predictors of preschool peer ratings. *Child Development*, Vol. 61, No. 4, pp. 1145–52.
- Denham, S. A., Mitchell-Copeland, J., Strandberg, K., Auerbach, S. and Blair, K. 1997. Parental contributions to preschoolers' emotional competence: direct and indirect effects. *Motivation and Emotion*, Vol. 21, No. 1, pp. 65–86.
- Diamond, A., Barnett, W. S., Thomas, J. and Munro, S. 2007. Preschool program improves cognitive control. *Science*, Vol. 318, No. 5855, pp. 1387–88.
- Diamond, A. and Doar, B. 1989. The performance of human infants on a measure of frontal cortex function, the delayed response task. *Developmental Psychobiology*, Vol. 22, No. 3, pp. 271–94.
- Diamond, A., Kirkham, N. and Amso, D. 2002. Conditions under which young children can hold two rules in mind and inhibit a prepotent response. *Developmental Psychology*, Vol. 38, No. 3, pp. 352–62.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P. ... Japel, C. 2007. School readiness and later achievement. *Developmental Psychology*, Vol. 43, No. 6, pp. 1428–46.
- Duncan, G. J. and Magnuson, K. 2011. The nature and impact of early achievement skills, attention skills, and behavior problems. G. J. Duncan and R. J. Murnane (eds), *Whither Opportunity? Rising Inequality, Schools, and Children's Life Chances*. New York, Russell Sage Foundation, pp. 47–70.
- Dunn, J., and Kendrick, C. 1982. *Siblings: Love, Envy, and Understanding*. Cambridge, Mass, Harvard University Press.
- Dunn, L. M. and Markwardt, F.C. 1970. *Peabody Individual Achievement Test*. Circle Pines, Minn., American Guidance Services.
- Eisenberg, N., Valiente, C. and Eggum, N. D. 2010. Self-regulation and school readiness. *Early Education and Development*, Vol. 21, No. 5, pp. 681–98.
- Engel de Abreu, P. M. J., Conway, A. R. and Gathercole, S. E. 2010. Working memory and fluid intelligence in young children. *Intelligence*, Vol. 38, No. 6, pp. 552–61.
- Engle, P. L., Black, M. M., Behrman, J. R., Cabral de Mello, M., Gertler, P. J., Kapiriri, L. ... International Child Development Steering Group. 2007. Strategies to avoid the loss of developmental potential in more than 200 million children in the developing world. *The Lancet*, Vol. 369, No. 9557, pp. 229–42.
- Engle, R. W., Carullo, J. J. and Collins, K. W. 1991. Individual differences in working memory for comprehension and following directions. *Journal of Educational Research*, Vol. 84, No. 5, pp. 253–62.
- Espy, K., McDiarmid, M., Kwik, M., Stalets, M., Hamby, A. and Senn, T. 2004. The contribution of executive function to emergent mathematic skills in preschool children. *Developmental Neuropsychology*, Vol. 26, No. 1, pp. 465–86.
- Felitti, V. J. 2009. Adverse childhood experiences and adult health. *Academic Pediatrics*, Vol. 9, No. 3, pp. 131–32.
- Fryers, T. and Brugha, T. 2013. Childhood determinants of adult psychiatric disorder. *Clinical Practice and Epidemiology in Mental Health*, Vol. 9, pp. 1–50.
- Fuchs, L. S. and Fuchs, D. 2007. A model for implementing responsiveness to intervention. *Teaching Exceptional Children*, Vol. 39, No. 5, pp. 14–20.
- Garner, P. W. 1996. The relations of emotional role taking, affective/moral attributions, and emotional display rule knowledge to low-income school-age children's social competence. *Journal of Applied Developmental Psychology*, Vol. 17, No. 1, pp. 19–36.
- Gathercole, S. E., Brown, L. and Pickering, S. J. 2003. Working memory assessments at school entry as longitudinal predictors of National Curriculum attainment levels. *Educational and Child Psychology*, Vol. 20, No. 3, pp. 109–22.

- Gathercole, S. E., Lamont, E. and Alloway, T. P. 2006. Working memory in the classroom. S. J. Pickering (ed.), *Working Memory in Education*. Burlington, Mass., Academic Press, pp. 219–40.
- Geary, D. C. 2000. From infancy to adulthood: the development of numerical abilities. *European Child and Adolescent Psychiatry*, Vol. 9, No. 2, pp. 511–16.
- Gelman, R. and Gallistel, C. 1978. *The Child's Understanding of Number*. Cambridge, Mass., Harvard University Press.
- Gioia, G. A., Isquith, P. K., Guy, S. C. and Kenworthy, L. 2000. Test review behavior rating inventory of executive function. *Child Neuropsychology*, Vol. 6, No. 3, pp. 235–8.
- Glewwe, P., Krutikova, S. and Rolleston, C. 2014. *Do Schools Reinforce or Reduce Learning Gaps Between Advantaged and Disadvantaged Students? Evidence from Vietnam and Peru*. Young Lives Working Paper No. 133. Oxford, UK, Young Lives, Oxford Department of International Development (ODID), University of Oxford. https://assets.publishing.service.gov.uk/media/57a089ed40f0b652dd00048c/wp133_learning-gaps.pdf
- Global Partnership for Education (GPE). 2016. *GPE 2020*. Washington, DC, Author. <http://www.globalpartnership.org/content/gpe-2020-strategic-plan>
- Gordon, P. 2004. Numerical cognition without words: evidence from Amazonia. *Science*, Vol. 306, No. 5695, pp. 496–99.
- Graves, D. H. 1983. *Writing: Teachers and Children at Work*. Exeter, NH, Heinemann Educational Books.
- Groth-Marnat, G. and Baker, S. 2003. Digit span as a measure of everyday attention: a study of ecological validity. *Perceptual and Motor Skills*, Vol. 97, No. 3, pp. 1209–18.
- Gunderson, E. A., Ramirez, G., Beilock, S. L. and Levine, S. C. 2012. The relation between spatial skill and early number knowledge: the role of the linear number line. *Developmental Psychology*, Vol. 48, No. 5, pp. 1229–41.
- Hair, E., Halle, T., Terry-Humen, E., Lavelle, B. and Calkins, J. 2006. Children's school readiness in the ECLS-K: predictions to academic, health, and social outcomes in first grade. *Early Childhood Research Quarterly*, Vol. 21, No. 4, pp. 431–54.
- Hale, J. B., Hoepfner, J. B. and Fiorello, C. A. 2002. Analyzing digit span components for assessment of attention processes. *Journal of Psychoeducational Assessment*, Vol. 20, No. 2, pp. 128–43.
- Hanushek, E. A. and Woessmann, L. 2008. The role of cognitive skills in economic development. *Journal of Economic Literature*, Vol. 46, No. 3, pp. 607–68.
- Harris, P. L., Johnson, C. N., Hutton, D., Andrews, G. and Cooke, T. 1989. Young children's theory of mind and emotion. *Cognition and Emotion*, Vol. 3, No. 4, pp. 379–400.
- Hart, B. and Risley, T. R. 2003. The early catastrophe: the 30 million word gap by age 3. *American Educator*, Vol. 27, No. 1, pp. 4–9.
- Holmgren, R. A., Eisenberg, N. and Fabes, R. A. 1998. The relations of children's situational empathy-related emotions to dispositional prosocial behaviour. *International Journal of Behavioral Development*, Vol. 22, No. 1, pp. 169–93.
- Holopainen, L., Ahonen, T. and Lyytinen, H. 2001. Predicting delay in reading achievement in a highly transparent language. *Journal of Learning Disabilities*, Vol. 34, No. 5, pp. 401–13.
- Huttenlocher, J., Jordan, N. C. and Levine, S. C. 1994. A mental model for early arithmetic. *Journal of Experimental Psychology: General*, Vol. 123, No. 3, pp. 284–96.
- Izard, C., Fine, S., Schultz, D., Mostow, A., Ackerman, B. and Youngstrom, E. 2001. Emotion knowledge as a predictor of social behavior and academic competence in children at risk. *Psychological Science*, Vol. 12, No. 1, pp. 18–23.
- Janus, M. and Duku, E. 2007. The school entry gap: socioeconomic, family, and health factors associated with children's school readiness to learn. *Early Education and Development*, Vol. 18, No. 3, pp. 375–403.
- Jolliffe, D. 1998. Skills, schooling, and household income in Ghana. *The World Bank Economic Review*, Vol. 12, No. 1, pp. 81–104.

- Jordan, N. C., Kaplan, D., Ramineni, C. and Locuniak, M. N. 2009. Early math matters: kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, Vol. 45, No. 3, pp. 850–67.
- Kaminski, R. A. and Good, R. H., III. 1996. Toward a technology for assessing basic early literacy skills. *School Psychology Review*, Vol. 25, No. 2, pp. 215–27.
- Lancet., . Year. Title. *Journal Title*, Vol. #, No. #, pp. #-#. Manuscript under preparation.
- Lesaux, N. K., Rupp, A. A. and Siegel, L. S. 2007. Growth in reading skills of children from diverse linguistic backgrounds: findings from a 5-year longitudinal study. *Journal of Educational Psychology*, Vol. 99, No. 4, pp. 821–34.
- Li, S. C. and Lewandowsky, S. 1995. Forward and backward recall: different retrieval processes. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, Vol. 21, No. 4, pp. 837–47.
- Liew, J. 2012. Effortful control, executive functions, and education: bringing self-regulatory and social-emotional competencies to the table. *Child Development Perspectives*, Vol. 6, No. 2, pp. 105–11.
- Lonigan, C. J. 2006. Development, assessment, and promotion of pre-literacy skills. *Early Education and Development*, Vol. 17, No. 1, pp. 91–114.
- Lonigan, C. J., Burgess, S. R. and Anthony, J. L. 2000. Development of emergent literacy and early reading skill in preschool children: evidence from a latent variable longitudinal study. *Developmental Psychology*, Vol. 36, No. 5, pp. 596–613.
- Lonigan, C. J., Schatschneider, C., Westberg, L. and NELP. 2008. Identification of children's skills and abilities linked to later outcomes in reading, writing, and spelling. *Developing Early Literacy: Report of the National Early Literacy Panel*. Washington, DC, National Institute for Literacy, pp. 55–106.
- Matthews, J. S., Ponitz, C. C. and Morrison, F. J. 2009. Early gender differences in self-regulation and academic achievement. *Journal of Educational Psychology*, Vol. 101, No. 3, pp. 689–704.
- McBride-Chang, C. 1999. The ABCs of the ABCs: the development of letter-name and letter-sound knowledge. *Merrill-Palmer Quarterly*, Vol. 45, No. 2, pp. 285–308.
- Middlebrooks, J. S. and Audage, N. C. 2008. *The Effects of Childhood Stress on Health Across the Lifespan*. Atlanta, Ga., Centers for Disease Control and Prevention, National Center for Injury Prevention and Control.
- Miller, A. L., Gouley, K. K., Seifer, R., Zakriski, A., Eguia, M. and Vergnani, M. 2005. Emotion knowledge skills in low-income elementary school children: associations with social status and peer experiences. *Social Development*, Vol. 14, No. 4, pp. 637–51.
- Miller, K. F. 1989. Measurement as a tool for thought: the role of measuring procedures in children's understanding of quantitative invariance. *Developmental Psychology*, Vol. 25, No. 4, pp. 589–600.
- Mistry, R. S., Benner, A., Biesanz, J., Clark, S. L. and Howes, C. 2010. Family and social risk, and parental investments during the early childhood years as predictors of low-income children's school readiness outcomes. *Early Childhood Research Quarterly*, Vol. 25, No. 4, pp. 432–49.
- Mix, K. S. and Cheng, Y. L. 2011. The relation between space and math: developmental and educational implications. *Advances in Child Development and Behavior*, Vol. 42, pp. 197–243.
- Moffitt, T. E., Arseneault, L., Belsky, D., Dickson, N., Hancox, R. J., Harrington, H. ... Sears, M. R. 2011. A gradient of childhood self-control predicts health, wealth, and public safety. *Proceedings of the National Academy of Sciences*, Vol. 108, No. 7, pp. 2693–98.
- Monette, S., Bigras, M. and Guay, M. C. 2011. The role of the executive functions in school achievement at the end of Grade 1. *Journal of Experimental Child Psychology*, Vol. 109, No. 2, pp. 158–73.
- Morgan, J. K., Izard, C. E. and King, K. A. 2010. Construct validity of the Emotion Matching Task: preliminary evidence for convergent and criterion validity of a new emotion knowledge measure for young children. *Social Development*, Vol. 19, No. 1, pp. 52–70.

- National Early Literacy Panel (NELP). 2008. *Developing Early Literacy: Report of the National Early Literacy Panel*. Washington, DC, National Institute for Literacy. <http://lincs.ed.gov/publications/pdf/NELPReport09.pdf>
- National Research Council (NRC). 2001. *Adding It Up: Helping Children Learn Mathematics*. Washington, DC, National Academies Press.
- National Research Council (NRC). 2009. *Mathematics Learning in Early Childhood: Paths Toward Excellence and Equity*. Washington, DC, National Academies Press.
- National Research Council (NRC) and Institute of Medicine (IOM). 2000. *From Neurons to Neighborhoods: The Science of Early Childhood Development*. Washington, DC, National Academies Press.
- Nigg, J. T. 2000. On inhibition/disinhibition in developmental psychopathology: views from cognitive and personality psychology and a working inhibition taxonomy. *Psychological bulletin*, Vol. 126, No. 2, pp. 220–46.
- Nunes, T. and Bryant, P. 1996. *Children Doing Mathematics*. Hoboken, NJ, Wiley-Blackwell.
- O'Connor, R. E. and Jenkins, J. R. 1999. Prediction of reading disabilities in kindergarten and first grade. *Scientific Studies of Reading*, Vol. 3, No. 2, pp. 159–97.
- Oh, S. and Lewis, C. 2008. Korean preschoolers' advanced inhibitory control and its relation to other executive skills and mental state understanding. *Child Development*, Vol. 79, No. 1, pp. 80–99.
- Passolunghi, M. C., Vercelloni, B. and Schadee, H. 2007. The precursors of mathematics learning: working memory, phonological ability and numerical competence. *Cognitive Development*, Vol. 22, No. 2, pp. 165–84.
- Pica, P., Lemer, C., Izard, V. and Dehaene, S. 2004. Exact and approximate arithmetic in an Amazonian indigene group. *Science*, Vol. 306, No. 5695, pp. 499–503.
- Pisani, L., Borisova, I. and Dowd, A. J. 2015. *International Development and Early Learning Assessment Technical Working Paper*. Fairfield, Conn., Save the Children. http://www.savethechildren.org/atf/cf/%7B9def2ebe-10ae-432c-9bd0-df91d2eba74a%7D/IDELA%20TECHNICAL%20WORKING%20PAPER_V4.PDF
- Ponitz, C. C., McClelland, M. M., Jewkes, A. M., Connor, C. M., Farris, C. L. and Morrison, F. J. 2008. Touch your toes! Developing a direct measure of behavioral regulation in early childhood. *Early Childhood Research Quarterly*, Vol. 23, No. 2, pp. 141–58.
- Pruden, S. M., Levine, S. C. and Huttenlocher, J. 2011. Children's spatial thinking: does talk about the spatial world matter? *Developmental Science*, Vol. 14, No. 6, pp. 1417–30.
- Reynolds, C. R. 1997. Forward and backward memory span should not be combined for clinical analysis. *Archives of Clinical Neuropsychology*, Vol. 12, No. 1, pp. 29–40.
- Roth, F. P., Speece, D. L. and Cooper, D. H. 2002. A longitudinal analysis of the connection between oral language and early reading. *Journal of Educational Research*, Vol. 95, No. 5, pp. 259– 72.
- RTI International. 2016. *Early Grade Reading Assessment (EGRA) Toolkit*, 2nd edn. Washington, DC, United States Agency for International Development (USAID).
- Sáez, L., Folsom, J. S. and Al Otaiba, S. 2012. Relations among student attention behaviors, teacher practices, and beginning word reading skill. *Journal of Learning Disabilities*, Vol. 45, No. 5, pp. 418–32.
- Sarama, J. and Clements, D. H. 2009. *Early Childhood Mathematics Education Research: Learning Trajectories for Young Children*. New York, Routledge.
- Savage, R. S., Frederickson, N., Goodwin, R., Patni, U., Smith, N. and Tiersley, L. 2005. Relationships among rapid digit naming, phonological processing, motor automaticity, and speech perception in poor, average, and good readers and spellers. *Journal of Learning Disabilities*, Vol. 38, No. 1, pp. 12–28.



- Scarborough, H. 1998. Predicting the future achievement of second graders with reading disabilities: contribution of phonemic awareness, verbal memory, rapid naming, and IQ. *Annals of Dyslexia*, Vol. 48, pp. 115–36.
- Schatschneider, C., Fletcher, J. M., Francis, D. J., Carlson, C. D. and Foorman, B. R. 2004. Kindergarten prediction of reading skills: a longitudinal comparative analysis. *Journal of Educational Psychology*, Vol. 96, No. 2, pp. 265–82.
- Snow, C. E., Griffin, P. and Burns, M. S. 2005. *Knowledge to Support the Teaching of Reading: Preparing Teachers for a Changing World*. San Francisco, Calif., Jossey-Bass Education.
- St Clair-Thompson, H. L. 2010. Backwards digit recall: a measure of short-term memory or working memory? *European Journal of Cognitive Psychology*, Vol. 22, No. 2, pp. 286–96.
- Stanovich, K. E., Cunningham, A. E. and Feeman, D. J. 1984. Intelligence, cognitive skills, and early reading progress. *Reading Research Quarterly*, Vol. 19, No. 3, pp. 278–303.
- Stipek, D. and Valentino, R. A. 2015. Early childhood memory and attention as predictors of academic growth trajectories. *Journal of Educational Psychology*, Vol. 107, No. 3, pp. 771–88.
- Storch, S. A. and Whitehurst, G. J. 2002. Oral languages and code-related precursors to reading: evidence from a longitudinal structural model. *Developmental Psychology*, Vol. 38, No. 6, pp. 934–47.
- Swanson, H. L. and Jerman, O. 2007. The influence of working memory on reading growth in subgroups of children with reading disabilities. *Journal of Experimental Child Psychology*, Vol. 96, No. 4, pp. 249–83.
- Taylor, M. 1988. Conceptual perspective taking: children's ability to distinguish what they know from what they see. *Child Development*, Vol. 59, No. 3, pp. 703–18.
- Thompson, R. A. 2006. The development of the person: social understanding, relationships, conscience, self. N. Eisenberg (ed.), *Handbook of Child Psychology*, Vol. 3: *Social, Emotional, and Personality Development*. Hoboken, NJ, John Wiley & Sons, pp. 24–98.
- Thompson, R. A. 2015. Relationships, regulation, and early development. M. E. Lamb (ed.), *Handbook of Child Psychology and Developmental Science*, Vol. 3: *Socioemotional Processes*. Hoboken, NJ, John Wiley & Sons, pp. 201–46.
- Tolchinsky, L., Levin, I., Aram, D. and McBride-Chang, C. 2012. Building literacy in alphabetic, abjad and morphosyllabic systems. *Reading and Writing: An Interdisciplinary Journal*, Vol. 25, No. 7, pp. 1573–98.
- Torgesen, J. K. 1999. Phonologically based reading disabilities: toward a coherent theory of one kind of learning disability. R. J. Sternberg and L. Spear-Swirling (eds), *Perspectives on Learning Disabilities: Biological, Cognitive, Contextual*. Boulder, Colo., Westview Press, pp. 106–35.
- Torppa, M., Poikkeus, A. M., Laakso, M. L., Eklund, K. and Lyytinen, H. 2006. Predicting delayed letter knowledge development and its relation to grade 1 reading achievement among children with and without familial risk for dyslexia. *Developmental Psychology*, Vol. 42, No. 6, pp. 1128–42.
- Trentacosta, C. J. and Fine, S. E. 2010. Emotion knowledge, social competence, and behavior problems in childhood and adolescence: a meta-analytic review. *Social Development*, Vol. 19, No. 1, pp. 1–29.
- Verdine, B. N., Irwin, C. M., Golinkoff, R. M. and Hirsh-Pasek, K. 2014. Contributions of executive function and spatial skills to preschool mathematics achievement. *Journal of Experimental Child Psychology*, Vol. 126, pp. 37–51.
- von Suchodoletz, A., Gestsdottir, S., Wanless, S. B., McClelland, M. M., Birgisdottir, F., Gunzenhauser, C. and Ragnarsdottir, H. 2013. Behavioral self-regulation and relations to emergent academic skills among children in Germany and Iceland. *Early Childhood Research Quarterly*, Vol. 28, No. 1, pp. 62–73.
- Waber, D. P., Bryce, C. P., Girard, J. M., Zichlin, M., Fitzmaurice, G. M. and Galler, J. R. 2014. Impaired IQ and academic skills in adults who experienced moderate to severe infantile malnutrition: a 40-year study. *Nutritional Neuroscience*, Vol. 17, No. 2, pp. 58–64.

- Wachs, T. D., Black, M. M. and Engle, P. L. 2009. Maternal depression: a global threat to children's health, development, and behavior and to human rights. *Child Development Perspectives*, Vol. 3, No. 1, pp. 51–59.
- Wagner, R. K. and Torgesen, J. K. 1987. The nature of phonological processing its causal role in the acquisition of reading skills. *Psychological Bulletin*, Vol. 101, pp. 192–212.
- Wagner, R. K., Torgesen, J. K. and Rashotte, C. A. 1994. Development of reading-related phonological processing abilities: new evidence of bidirectional causality from a latent variable longitudinal study. *Developmental Psychology*, Vol. 30, No. 1, pp. 73–87.
- Wagner, R. K., Torgesen, J. K., Rashotte, C. A., Hecht, S. A., Barker, T. A., Burgess, S. R. ... Garon, T. 1997. Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: a 5-year longitudinal study. *Developmental Psychology*, Vol. 33, No. 3, pp. 468–79.
- Walker, S. P., Wachs, T. D., Grantham-McGregor, S., Black, M. M., Nelson, C. A., Huffman, S. L. ... Gardner, J. M. M. 2011. Inequality in early childhood: risk and protective factors for early child development. *The Lancet*, Vol. 378, No. 9799, pp.1325–38.
- Wanless, S. B., McClelland, M. M., Acock, A. C., Ponitz, C. C., Son, S., Lan, X. ... Li, S. 2011. Measuring behavioral regulation in four societies. *Psychological Assessments*, Vol. 23, No. 2, pp. 364–78.
- Wechsler, D. 2003. *Wechsler Intelligence Scale for Children – Fourth Edition (WISC-IV)*. San Antonio, Tex., PsychCorp/Pearson.
- Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L. and Nelson, K. E. 2010. The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology*, Vol. 102, No. 1, pp. 43–53.
- White, N., Ensor, R., Marks, A., Jacobs, L. and Hughes, C. 2014. 'It's Mine!' Does sharing with siblings at age 3 predict sharing with siblings, friends, and unfamiliar peers at age 6? *Early Education and Development*, Vol. 25, No. 2, pp. 185–201.
- Whitehurst, G. J., and Lonigan, C. J. 2001. Emergent literacy: development from prereaders to readers. S. B. Neuman and D. K. Dickinson (eds), *Handbook of Early Literacy Research*. New York, Guilford, pp. 11–29.
- Wiebe, S. A., Espy, K. A. and Charak, D. 2008. Using confirmatory factor analysis to understand executive control in preschool children: I. latent structure. *Developmental Psychology*, Vol. 44, No. 2, pp. 575–87.
- World Bank. 2016. *TITLE*. Washington, DC, Author. Manuscript under preparation
- Zelazo, P. D. and Müller, U. 2002. Executive function in typical and atypical development. U. Goswami (ed.), *Blackwell Handbook of Childhood Cognitive Development*. Malden, Mass., Blackwell Publishers Ltd, pp. 445–69.
- Ziegler, J.C. and Goswami, U. 2005. Reading acquisition, developmental dyslexia, and skilled reading across languages: a psycholinguistic grain size theory. *Psychological Bulletin*, Vol. 131, No. 1, pp. 3–29.

Appendices

Appendix 1: Domains, constructs and items for MODEL

DOMAINS, CONSTRUCTS AND ITEMS FOR MODEL

DOMAIN	CONSTRUCT	TEACHER CAREGIVER REPORT (TCR) ITEM	DIRECT ASSESSMENT (DA) ITEM
 Literacy	Motivation	Would you say <i>(name)</i> is interested in reading (inquisitive/curious about the meaning of printed material)?	--
	Expressive language	Can <i>(name)</i> communicate his/her own needs/what s/he wants in a way understandable to adults and peers?	Can you name some things you can eat that you buy at the market? Can you tell me the names of some animals you know?
		--	I want you to tell me the beginning sound of <i>(familiar word)</i> .
	Alphabet knowledge	Can <i>(name)</i> identify at least 10 letters of the alphabet?	Can you write your name here?
	Receptive language	Can <i>(name)</i> understand on first try what is being said to him/her?	I am going to tell you a story. After the story, I will ask you some questions about what happened.
		Can <i>(name)</i> write his/her own name?	Can you write your name here?
 Mathematics (continued on next page)	Numbers and operations	How high can <i>(name)</i> count?	How high can you count?
		Can <i>(name)</i> count 10 objects?	Please give me 3 stones. Now please give me 6 stones. Now please give me 14 stones.
		--	Which number is greater, 3 or 5? 8 or 6? 4 or 7?
		Can <i>(name)</i> identify written numerals up to 5? If yes ask 10. If yes, then ask 20.	I will point to a number and I'd like you to tell me which one it is.
		Can <i>(name)</i> add three and two together?	If you have three balls, and I give you two more, how many will you have?
	Measurement	Does <i>(name)</i> know that a cow is taller than a cat?	--
		Can <i>(name)</i> identify: The largest/smallest/longest/shortest of three objects?	--
		Does <i>(name)</i> know that a goat weighs more than a mouse?	--
		Does <i>(name)</i> understand the concepts of: today/yesterday/tomorrow?	--

NOTE: This table displays items that are conceptually linked to more than one construct. In the actual modules, however, the items are not administered twice.


DOMAINS, CONSTRUCTS AND ITEMS FOR MODEL

DOMAIN	CONSTRUCT	TEACHER CAREGIVER REPORT (TCR) ITEM	DIRECT ASSESSMENT (DA) ITEM
Mathematics (continued)	Spatial relations	Can (<i>name</i>) name shapes like circles, triangles and squares?	Point to the picture with the ball on/under/in front of/next to the chair. (4 items)
		Can (<i>name</i>) complete a five-piece puzzle?	Look at these pieces. Now look at these shapes. If you put the pieces together, they will make one of these shapes. Point to the shape that the pieces make.
 Social-Emotional Development (continued on next page)	Self-regulation	How often does (<i>name</i>) plan ahead?	--
		How often does (<i>name</i>) stop an activity when told to do so?	--
		How often does (<i>name</i>) rudely intrude on others?	--
		Is (<i>name</i>) over-active?	--
		When you take your child somewhere new, does (<i>name</i>) seek to explore the new environment freely?	--
		Does (<i>name</i>) adjust easily to transitions? (for example starting pre-primary school for the first time, or going to visit relatives in another house)	--
		Does (<i>name</i>) settle down after periods of exciting activity?	--
		Is (<i>name</i>) unable to sit still?	--
		When interacting with others, for example, sharing food, does (<i>name</i>) show self-control?	--
	Social understanding/ pro-social behaviour	Does (<i>name</i>) show consideration of other people's feelings?	--
		Does (<i>name</i>) offer to help someone who seems to need help?	--
		Does (<i>name</i>) share with his/her peers?	--
 Social-Emotional Development (continued)	Social competence	Does (<i>name</i>) get along with other children s/he plays with, such as siblings or family members?	--
		Does (<i>name</i>) have difficulty taking turns when playing together with others?	--
		Does (<i>name</i>) accept responsibility for his/her actions?	--
	Emotional well-being	Would you say (<i>name</i>) kicks, bites or hits other children or adults?	--
		Is (<i>name</i>) upset when left by parents/guardians?	--
		Would you say that (<i>name</i>) is often sad or unhappy?	--
		Is (<i>name</i>) easily distracted? (i.e. how often does his/her concentration wander?)	--
		How often does (<i>name</i>) keep working at something until s/he is finished?	--
		How often does (<i>name</i>) have difficulties doing things that s/he does not like?	--
 Executive Function	Working memory	When asked to do several things, how often does (<i>name</i>) remember all the instructions?	Please do the opposite of what I say: When I say touch your head, you touch your toes instead.
		--	I am going to say a list of numbers, one after another. After you hear the numbers, I want you to repeat them after me in the same order.
	Inhibition ¹	Is (<i>name</i>) easily distracted? (i.e. how often does his/her concentration wander?)	I am going to say a list of numbers, one after another. After you hear the numbers, I want you to repeat them after me backwards.
		How often does (<i>name</i>) rudely intrude on others?	--
		Is (<i>name</i>) over-active?	--
	Fine motor	Can (<i>name</i>) write his/her own name?	Can you write your name here?
		--	Here is a square. Can you draw a square like this one?

¹ Overlaps with 'Self-regulation' and 'Emotional well-being' (Social-Emotional Development).

NOTE: This table displays items that are conceptually linked to more than one construct. In the actual modules, however, the items are not administered twice.

DOMAINS, CONSTRUCTS AND ITEMS FOR MODEL

DOMAIN	CONSTRUCT	TEACHER CAREGIVER REPORT (TCR) ITEM	DIRECT ASSESSMENT (DA) ITEM
 Contextual Information	Socio-economic conditions ²	Does your household have: <i>[do not ask if already known/visible]</i> Does anyone in your household have:	--
	Parent education	What is the highest level of school you attended? (child's mother)	--
		What is the highest level of school attended by (name)'s father?	--
		If caregiver is not mother or father, what is the highest level of school attended by caregiver?	--
	Household composition	How many other children live in (name)'s household? (list by age)	--
		How many adults (including yourself and excluding the children) live in this home?	--
		How are adults in the household related to the child?	--
	Participation in early learning	Does (name) attend any organized learning or early childhood education programme, such as a private or government facility, including kindergarten or community child care? <i>[do not ask if already known]</i> If yes: Within the last seven days, about how many hours did (name) attend?	--
	Home learning environment/parent involvement	How many children's books or picture books do you have for (name) at home? (enter number)	--
		Does (name) attend any organized learning or early childhood education programme, such as a private or government facility, including kindergarten or community child care? <i>[do not ask if already known]</i>	--
		Do you believe that in order to bring up, raise or educate a child properly, the child needs to be physically punished?	--
	Neglect	Sometimes adults taking care of children have to leave the house to go shopping, wash clothes, or for other reasons and have to leave young children. On how many days in the past week was (name): • left alone for more than an hour? • left in the care of another child (someone less than 10 years old) for more than an hour?	--
	Child health	Is (name) often sick?	--
		Is (name) often tired?	--
		Does (name) have any special needs or difficulties? If yes, please explain.	
		Has (name) ever received any vaccinations to prevent him/her from getting diseases, including vaccinations received in a campaign or immunization day?	--

NOTE: This table displays items that are conceptually linked to more than one construct. In the actual modules, however, the items are not administered twice.

Appendix 2: Example of national alignment worksheets

The tables below provide sample worksheets from a workshop to align MELQO tools with national pre-primary curricula and standards. This workshop was designed to engage local stakeholders in the process of fully aligning MODEL items with national standards.

General suggestion: Ensure all overarching competencies are covered. Try to have good representation of items for each competency/domain.

USING MATHEMATICAL CONCEPTS			
COMPETENCY	ACTIVITY PERFORMED BY CHILD	ALIGNMENT WITH MELQO	WORKSHOP DISCUSSION AND AGREEMENTS
Identifying things found in the school environment	Identifying objects in the environment	No	Sorting is the most important to assess with MELQO because sorting captures many of the other competencies (suggest to add measurement vocabulary where child points to objects according to size).
	Identifying primary colours	No	
	Identifying shapes	No	
	Sorting things according to their characteristics	#1 Measurement Vocabulary: Pointing to objects according to size	
	Playing games involving comparing and differentiating things		
Developing the concept of time	Identifying different times of the day	No	This is not a priority for now (children already know this and it is difficult to measure).
	Dramatizing activities done during different times of the day		
	Identifying days of the week		
	Differentiating days of the week		
Mastering measurement skills	Comparing things according to their characteristics	#1 Measurement Vocabulary: Pointing to objects according to size	<ul style="list-style-type: none">• This is important to measure and include in MELQO. Recommend to bring back item on measurement vocabulary – pointing to objects according to size.• If the item is too easy, could consider including more difficult comparison words (heavy/light, for example).

USING MATHEMATICAL CONCEPTS

Developing number concept	Reading numbers 1–10	#8 Number Identification	<ul style="list-style-type: none"> • There is already a lot included here, could potentially remove some. • Could consider capturing sorting and measurement with same item ('Give me all of the big sticks'). • Verbal counting is crucial for Tanzania, but it is already covered in other ways.
	Performing simple activities of counting numbers	#2 Verbal Counting: Counting as high as possible #8 Number Identification: Identification of numbers	
	Counting numbers using objects	#3 Producing a Set: Give specified number of objects	
	Matching objects with number	#3 Producing a Set: Give specified number of objects	
	Performing actions involving addition and subtraction of things	#5 Mental Addition: Simple mental addition and strategies	
	Identifying number shapes by using different things	No	
	Moulding/constructing number shapes	No	
	Tracing number shapes	See Pre-writing skills (copying shapes)	
	Writing numbers step by step	See Pre-writing skills (name writing)	
Additional MELQO items		#4 Number Comparison: Comparing number magnitude #6 Spatial Vocabulary: Understanding positional terms #7 Mental Transformation: Geometry; Problem-solving	

COMMUNICATION			
COMPETENCY	ACTIVITY PERFORMED BY CHILD	ALIGNMENT WITH MELQO	WORKSHOP DISCUSSION AND AGREEMENTS
Listening	Listening to songs, conversations and short stories	#16 Listening Comprehension Story: Listen to story and answer questions (listen to 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	No	
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 also recognizes it is difficult to find items that measure if children are pronouncing words correctly (there are different regional accents).
	Identifying different riddles	No	
	Singing different songs	No	
	Telling different stories	No	
	Explaining daily activities	No	
	Expressing things he/she likes/dislikes	No (with links to #12 Expressive Vocabulary)	
Mastering pre-reading skills	Reading pictures	No	<ul style="list-style-type: none"> It may be difficult for some children to name things from their environment if they are from poor families and 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	No	
	Identifying different names/words in his/her environment	No	
	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	
	Recognizing sounds of vowels and consonants; recognizing and identifying first sounds	#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)	
	Doing exercises of using drawing and writing instruments	#21 Copying: Copy shapes and lines (hold writing instruments; draw lines)	
	Doing exercise of drawing lines from left to right and from top to bottom		
	Moulding/constructing shapes of vowels	No	
	Tracing vowel shapes; writing vowels step by step	#17 Name Writing: Write own name (in any way the child knows)	
	Moulding/constructing consonant shapes	No	
	Tracing consonant shapes; writing consonants step by step	#17 Name Writing: Write own name (in any way the child knows)	

RELATING TO EACH OTHER			
COMPETENCY	ACTIVITY PERFORMED BY CHILD	ALIGNMENT WITH MELQO	WORKSHOP DISCUSSION AND AGREEMENTS
Caring for one another	Introducing himself or herself	No	<ul style="list-style-type: none"> Figure out how to improve the item with girl crying (not just crying), maybe could show a photograph of a sad child (rather than drawing) In Tanzania, children do not know the names of adults (go by 'Baba' or 'Mama'). But this is still important to assess. Item to include: Children should know their own full name, their address, and where they go to school.
	Cooperating with others in different activities – problem-solving; turn-taking; sharing; getting along	#9 Perspective-Taking/Empathy: Identifying emotions and suggesting appropriate emotional caring responses (cooperate with others) #10 Understanding Feelings: Appropriate naming of feelings (cooperate with others) #9 Perspective-Taking/Empathy: Identifying emotions and suggesting appropriate emotional caring responses (demonstrate good conduct)	
Respecting one another	Greeting one another according to age	No	
	Demonstrating good conduct activities (manners)	#9 Perspective-Taking/Empathy: Identifying emotions and suggesting appropriate emotional caring responses (demonstrate good conduct)	
	Identifying appropriate clothes according to the environment	No	
Self-control	Performing activities with independence	No	<ul style="list-style-type: none"> Teacher interview could assess whether child cares for her own things. Existing executive function tasks already work well.
	Using acceptable ways of controlling emotions	See Executive Function items	
	Performing different activities related to caring for things	No	

DOMAINS IN CURRICULUM NOT CURRENTLY COVERED IN MELQO			
COMPETENCY	ACTIVITY PERFORMED BY CHILD	ALIGNMENT WITH MELQO	WORKSHOP DISCUSSION AND AGREEMENTS
Caring for health	Identifying external parts of the body and their functions	No	Add item on personal hygiene – hand-washing (when should you wash your hands, with what?).
	Personal hygiene	No	
	Taking care of clothes	No	
	Taking care of eating utensils	No	
	Identifying nutritious foods	No	
	Identifying diseases	No	
Caring for the environment	Identifying things in the environment	#11 General Knowledge: knowing how things grow	<ul style="list-style-type: none"> Add item on nutrition – what is healthy vs. unhealthy food? Group suggests that children identify unsafe areas/things. If it is possible to find generalizable unsafe things, this could work (but may be difficult to identify something for the entire country).
	Cleaning the environment	No	
	Taking precautions	No	
Mastering artistic skills	Mastering creative arts involving use of hands	#17 Name Writing: Write own name (in any way the child knows) #21 Copying: Copy shapes and lines (hold writing instruments; draw lines)	Drawing shapes is already included in math concepts, but group suggests that drawing a picture be included in the instrument.
	Mastering creative arts involving use of the whole body	No	
	Mastering creative arts involving use of sounds	No	

Appendix 3: Item adaptation notes

The sections below offer guidance for adapting literacy and mathematics items on the DA as well as items on the TCR.

Adaptation of literacy items

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 of items

- Initial sound identification
 - a. The target sounds should be among the most frequent from the letter/syllable frequency list. In many languages, vowels are the most common but this will probably work best with consonants. Although continuous sounds (sound you can hold as you say them such as /mmmmm/ or /sssss/) are easier than stop sounds (/t/ or /b/), the structure of the language should guide the selection from among the most frequent sounds.
 - b. When modelling continuous sounds, do not hold the sound (/m/ not /mmmmmmmm/).
 - c. Use words that begin with continuous and stop sounds.
- Letter name knowledge
 - a. Use the frequency list to identify the easiest and hardest letters appropriate for beginning readers (in alphasyllabic languages children would not get to some graphemes until second grade, so the letters should be drawn from those they are expected to learn in first grade).
 - b. Randomly list the easiest letters in the first column and the most challenging in the second column.
 - c. For languages that read from right to left, reverse the columns (put the easiest letters in the right column and most challenging in the left column, and present to the child in that order).

- d. Stop rule is at five incorrect consecutively.

- Listening comprehension

The story for the listening comprehension task is below. The approximate length of the story should be maintained. 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

- a. The two animals can be changed to two animals that are common in the country.
- b. The story opening can be changed to what is commonly used in the language.

The Mouse and the Cat

Once upon a time there was a fat cat. He always wore a red hat. Once when he was sleeping, a small mouse came silently and stole the hat. The cat woke up to see his hat gone, got very angry and started chasing the mouse. After a while, the mouse was trapped under a table and could not find any way to escape. So the mouse cried to the cat, 'Please don't eat me cat. If you spare my life I will return your hat.' So, after getting back his hat the cat said, 'Never touch my hat again' and he went back to sleep in a happy mood.

Adaptation of mathematics 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.

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.

Adaptation guidelines for the TCR

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

MELE Module

MELQO Measure of Early Learning Environments

Technical Manual

Table of Contents

ABBREVIATIONS	74
INTRODUCTION	75
Background and rationale	76
Why is measuring quality important?	76
How much quality is 'enough'?	76
Why invest in adapting and using MELE?	77
Approach and methodology	77
Overview of the MELE module	77
Development process	79
Domain selection	80
Format selection	80
MELE Instruments	82
KEY DOMAINS OF QUALITY	83
Environments and materials	83
Interactions	83
Pedagogy	85
Family and community engagement	85
Inclusion	86
Play	86
Personnel	86
USING AND ADAPTING THE TOOL	87
Planning	87
Adaptation	88
Field-testing, data collection and validation	90
Issues to consider	90
Next steps in development	92
APPLICATION TO POLICY	92
Improving quality in classroom settings	92
The role of government policies in improving quality	93
CONCLUSION	94
Considerations for the future	95
REFERENCES	96

Abbreviations

ACEI	Association for Childhood Education International
AKF	Aga Khan Foundation
CGECCD	Consultative Group on Early Childhood Care and Development
CLASS	Classroom Assessment Scoring System
CRC	Convention on the Rights of the Child
ECD	Early childhood development
ECE	Early childhood education
ECEQAS	Early Childhood Education Quality Assessment Scale
ECERS	Early Childhood Environment Rating Scale
ECERS-E	Early Childhood Environment Rating Scale – Extension
ECERS-R	Early Childhood Environment Rating Scale – Revised
EFA	Education for All (UNESCO)
EIU	Economist Intelligence Unit
ELDS	Early Learning and Development Standards (UNICEF)
GGA	Global Guidelines Assessment (ACEI)
HECDI	Holistic Early Childhood Development Index
ISSA	International Step by Step Association
LMIC	Low- and middle-income country
MDG	Millennium Development Goal (UN)
MELE	Measure of Early Learning Environments (MELQO)
MELQO	Measuring Early Learning Quality and Outcomes
MODEL	Measure of Development and Early Learning (MELQO)
NGO	Non-governmental organization
OECD	Organisation for Economic Co-operation and Development
PPE	Pre-primary education
SABER	System Analyses to Better Education Results (World Bank)
SDG	Sustainable Development Goal (UN)
SEAMEO	Southeast Asian Ministers of Education Organization
STEPP	Survey of Teachers in Pre-Primary Education (UNESCO)
TECERS	Tamil Nadu ECERS
TIPPS	Teacher Instructional Practices and Processes System
UIS	UNESCO Institute of Statistics
UK	United Kingdom
UN	United Nations
UNESCO	United Nations Educational, Scientific and Cultural Organization
UNICEF	United Nations Children's Fund
USA	United States of America
USAID	United States Agency for International Development
WCARO	West and Central Africa Regional Office (UNICEF)

Introduction

Given the rapid expansion in enrolment in early childhood education (ECE) worldwide and increases in investment, it has become ever more urgent to support implementation of high-quality ECE (Britto et al., 2011; Myers, 2006). Policy-makers, practitioners and administrators need reliable and valid information about early learning environments to support ECE programme development, establishment and improvement.

The MELE module is a set of constructs that are important for quality in pre-primary settings. The module is intended to serve as a starting point for national adaptation, to develop measurement tools that are locally relevant and useful for improving policies and practices in early childhood education. Results from the tools can be used to inform policies, professional development, and classroom practices for young children.

ECE has a profound impact on children's development and their ability to reach their full potential later in life. The recently adopted Sustainable Development Goals (SDGs) acknowledge the primacy of ECE in protecting children and supporting their development holistically, stating in Goal 4, Target 4.2, that countries should 'ensure that all girls and boys have access to quality early childhood development, care and pre-primary education so that they are ready for primary education' by 2030. This SDG target on ECD maintains the emphasis on quality that began with UNESCO's Education for All (EFA) Goal 1 (UNESCO, 2000). Several international agreements, including the Millennium Development Goals (MDGs), the EFA goals and the Convention on the Rights of the Child (CRC) have similarly affirmed the importance of quality in ECE.

In the past decade, many countries have made progress towards increasing access to ECE, including formal and non-formal pre-primary education (PPE), parenting education and support, and infant and toddler programmes. In particular, access to PPE has expanded globally, with the number of children enrolled in pre-school increasing by 64 per cent between 1999 and 2014 (UNESCO, 2015). Yet these efforts to expand ECE have generally not been matched by efforts to improve quality. Evidence from high-, middle- and low-income countries alike demonstrates that even when access goes up, children's outcomes do not always improve (e.g. Wong et al., 2013). At times efforts to increase access may even exacerbate the problem of low-quality ECE. The rapid scaling of PPE in recent years may have amplified underlying issues related to ECE programme quality in some countries, as governments focus on extending services without commensurate attention to building a quality infrastructure to support the expansion).

A basic level of quality in the learning environment – comprising aspects such as the tone and content of teacher–child interactions, the range of play materials available, the safety and atmosphere of the physical space, the attention paid to health and nutrition, and the level of parent engagement – is critical to ensuring children receive the maximum benefit from ECE and to preserving children's rights. Children's rights are upheld when they are in environments that respect and nurture their natural curiosity, value supportive interactions with adults, and empower their growing abilities to make their own decisions. Attention to quality is particularly important to safeguard the rights of children at risk of social exclusion due to poverty, ethnic or cultural background, gender or disability. Quality can be seen as an essential link between policy commitments and actual results: children will only benefit from increased access to ECE if the services being provided meet core standards for quality.

As a platform for measurement, the MELQO initiative is designed to help countries advance quality measurement by outlining the issues and approaches to measuring quality (please see the MELQO Overview for more a more detailed description). At its start, MELQO addressed two key questions: (1) **how can quality be defined in a culturally and contextually relevant and measurable way;** and (2) **how can results be used to promote improvement through integration into ongoing national monitoring systems?** MELQO's efforts are focused on pre-primary settings only, as a starting point for measurement that should ideally include all of children's learning environments, including homes and informal pre-school settings.

This guide outlines questions and conclusions regarding measurement of quality that arose throughout the MELQO process. It presents key constructs that have been shown to contribute to quality in pre-primary settings in previous studies, explains why these constructs are important, and presents items that can be used as a starting point for capturing those constructs.

A central question for MELQO was how to draw upon the work done to date in developing global modules and processes to support national-level measurement. In measuring quality, there is a delicate balance to strike between a 'global' approach to using existing measures, in order to increase the efficiency of measurement and leverage global attention; and a 'local' approach to creating new measures that may be more contextually and culturally relevant. During the MELQO process, this issue was examined from multiple angles. At this stage, MELQO is able

to identify domains of quality that could be relevant across countries, but items may vary from one country to the next. This document outlines these domains, offers suggestions on issues to consider in the process of measuring quality, and provides a small item bank.

Background and rationale

Why is measuring quality important?

Quality in ECE is an issue in almost every country. While many countries may have model ECE programmes available for some children, a range of sources suggests that quality in ECE settings at scale is a grave concern in most countries (Economist Intelligence Unit, 2012; OECD, 2012, 2015). This includes high-income countries and is especially true for children at risk of social exclusion. Difficulty in providing global access to high-quality, pre-primary ECE is due to the relatively low importance often placed on ECE in general, which results in a lack of adequate funding, appropriately trained teachers, space and materials. Even when ECE access is universal, most countries appear to have variations in quality in pre-primary settings. For instance, research from Sweden based on observational measures suggests an overall high quality of care when compared to Korea, yet lower-quality pre-schools can be found in Sweden as well (Sheridan et al., 2009; Sheridan and Samuelsson, 2001). ECE in Belgium is of lower quality for children of low-income families, and low-income children in the United States also tend to receive lower-quality ECE (Vandenbroeck et al., 2008; Vandell and Wolf, 2000). Some countries may have more consistently low-quality ECE than others. Studies conducted in Turkey, China, India, Ethiopia, Peru and Korea indicate that, on average, quality is low in ECE settings (Kalkan and Akman, 2009; Rao, Sun, Zhou and Zhang, 2012; Sheridan et al., 2009).¹

Quality ECE is essential for promoting equity and social inclusion. Families and communities worldwide share a desire to support their children's development. But despite this dedication, millions of children do not have access to the basic necessities for healthy development, due to family poverty; disease; ethnic, gender or cultural discrimination; or natural and human-induced disasters. Inadequate attention to young children's health, nutrition, and social-emotional and cognitive well-being impedes their development. Early intervention is essential, as developmental delays reflective of social inequities appear before birth and become greater and more intractable as time passes. ECE can help address the gaps in development that arise among children who face considerable challenges in early childhood. It can also promote school attendance, achievement and completion

for children who are likely to struggle academically due to family poverty, low levels of parental education, geographic isolation, disability or other reasons (Walker et al., 2011; Learning Metrics Task Force, 2012). But in order to fully support children's development, ECE must be affordable and accessible, especially to the most vulnerable children. Current research reveals large gaps in access to pre-primary ECE based on poverty, ethnic status, gender and the presence of disabilities, clearly demonstrating that equitable, quality ECE systems are not yet the global norm (UNESCO, 2012). Moreover, vulnerable children in both high- and low-income countries tend to receive the lowest-quality care (Helburn, 1995; Streuli, 2012). Access to ECE is necessary but not sufficient for supporting young children's development. For example, enrolment in PPE alone does not describe the extent to which children are experiencing learning environments that support their development. Increasing access to low-quality programmes for children and families who already face disadvantage widens inequities. Holistic information on ECE quality that includes a broad spectrum of quantitative and qualitative data from multiple sources, contexts and caregivers holds great promise as a monitoring and improvement tool to reduce societal inequity and exclusion.

How much quality is 'enough'?

Quality can best be understood as a continuum, from very low-quality ECE with conditions that are harmful to children's development to very high-quality ECE with conditions that offer young children considerable support. To ensure that all young children reach their full developmental potential, a long-term commitment to stimulating, nurturing environments with trained, high-quality ECE teachers is needed.

For healthy development, children need sustained access throughout the early childhood years to supportive, nurturing environments that protect their health and ensure good nutrition. These environments must also provide a high degree of emotional care and cognitive stimulation obtained through activities and interaction with objects, peers and adults that give children the opportunity to test out their ideas and experiment with new ones. For this reason, any number of activities or objects can support children's learning, provided that the child also has an opportunity to experiment with new ideas. To determine what degree of ECE quality is needed, it is important to emphasize the intersection between quality and access. *Both are critical to supporting young children's development.* Cross-country research shows that children who experience higher-quality ECE are positively influenced both concurrently and in the future, when compared to children who do not have access *and* when compared to children who attend ECE of lower observed quality (e.g., Rao et al., 2015; Burger, 2010). Yet limited access, even to very high-quality programmes, is unlikely to

¹ These studies offer cross-country comparisons because all used the Early Childhood Environment Rating Scale (ECERS) (Harms et al., 2005).

have lasting effects on development. While some access is better than none, the benefits ECE are greater when there is continuous access to higher quality ECE for at least one to two years. In particular, children who face significant hurdles to healthy development – such as malnutrition, low parental education, poor health, disabilities, or family or community conflict – can be supported by regular attendance to quality ECE programmes over an extended period of time.

But what is the threshold for quality? For many years, research from high-income, Anglo-Saxon countries has supported the view that ECE settings must meet certain quality standards to have positive impacts on children's development, and could even be harmful if some standards are not reached (Helburn, 1995). While those findings are an appropriate and useful guide for high-income countries that already have high rates of access, recent research suggests that for countries with very limited access to ECE, exposing children to environments that have some positive elements but do not meet all of the criteria for 'quality' can also contribute to development (e.g. Rao, Sun, Pearson et al., 2012). It is critical to note that children who attend ECE programmes with trained staff and adequate materials show better results. As countries work to make high-quality ECE available to all, each step along the way can lead to benefits for young children. Recent studies suggest that supporting community-based pre-schools and other low-cost interventions may be worthwhile, especially as countries build systems to provide universal access to PPE.

MELQO and its quality module, MELE (Measure of Early Learning Environments), are especially relevant in light of the renewed effort to focus on quality as part of the SDGs. Available indicators for measuring quality globally are currently limited to the few data sources that are reliably collected across a number of countries, including teacher–child ratios and the percentage of teachers who have undergone formal training (UNESCO, 2014). MELQO's work on this module demonstrates that more comprehensive data on quality are both possible and desirable to obtain. As countries consider plans for implementing the Education 2030 agenda, an emphasis on quality will be essential to reaching goals for ECE.

Why invest in adapting and using MELE?

Ensuring good quality learning environments for all children requires a systematic way to collect regular, accurate information on what is happening in early childhood classrooms. Monitoring and assessment can play an important role in identifying strengths and areas for improvement. A strong body of evidence from high-income countries demonstrates that monitoring in ECE settings is essential for maintaining quality over time (OECD, 2012, 2015). Assessments and monitoring of quality can be used to provide teachers with specific feedback on how to improve their performance, identify areas for improvement in teacher training, provide

information to policy-makers on the overall functioning of the ECE system, and help identify what policy and funding changes are needed. Decisions about how to assess quality should be made with the goal of improvement in mind – specifically, what kind of information is most likely to be useful to ECE professionals and stakeholders, and how well the assessment reflects the local values and priorities for quality. Indicators intended to measure country, regional or global progress towards supporting children's holistic development, including the quality of ECE settings, should reflect stakeholders' visions of the kind of world in which they would like to live (R. Meyers, personal communication, May 2014).

Approach and methodology

Overview of the MELE module

MELQO's quality module and guidance on how to use it can help provide a national picture of the quality of PPE – for example, through a probability or representative sample of PPE settings. The information resulting from the MELE module can inform quality improvement efforts of the following kinds:

- Improvement of curriculum and materials.
- Professional development and in-service training of teachers, supervisors and trainers.
- Feedback and reflection sessions among teachers.
- Pre-service training for pre-primary teachers as well as for supervisors, trainers and ministry of education staff.
- Creation of monitoring tools that focus on the most critical elements of quality, such as developmentally appropriate activities and teacher–child interactions.

MELE is designed to assess group-based care in community centres, schools and kindergartens, for children from age 3 to primary school entry. The measure is meant to be relevant in both low-income and middle-income country contexts, and in both rural and urban ECE settings. **It is not meant to be used for 'high-stakes' purposes** – for example, making decisions concerning hiring and firing teachers, or funding or defunding programmes.

This guide is divided into two main sections:

1. Domains included in MELE module

Measurement begins with a locally relevant definition of 'quality' in early childhood settings. MELE outlines seven constructs and explains why they are critical to serve as a starting point for local adaptation processes. A definition of ECE quality that is collaboratively developed with input from multiple groups of stakeholders, including parents,

professionals and government officials, forms the foundation for quality measurement and improvement. This definition, which should reflect local standards for quality and the science of child development, serves as the bedrock for establishing goals for quality, outlining what quality measures should contain and identifying elements to be targeted in improvement efforts.

While this document is focused on the quality module, the MELQO suite of modules is intended to promote measurement of both child development and learning, and quality of learning environments. The modules refer to the overall tool or measure that can be used as a starting point for measurement, and the tools refer to the specific sets of items that are administered either to children, parents or teachers and school directors. The table below describes the content of the MODEL and MELE modules:

Table 1: MELQO modules

Module	Instruments	Global core items	Countries field-tested to date
Measure of Development and Early Learning (MODEL)	<ol style="list-style-type: none"> 1. Direct child observation 2. Parent or teacher survey 	Includes items that are globally comparable in the domains of executive function, social-emotional development, early mathematics skills and early literacy skills. Items on children's home and family environments are also included.	Bangladesh, Kenya, Lao People's Democratic Republic, Madagascar, Mongolia, Nicaragua, Sudan, Kyrgyzstan and Tanzania
Measure of Early Learning Environments (MELE)	<ol style="list-style-type: none"> 1. Classroom observation 2. Teacher survey 3. Supervisor survey 4. Parent survey 	Includes seven key domains of quality: play, pedagogy, interactions, environment, parent/community engagement, personnel and inclusion.	Colombia, Kenya, Nicaragua, Peru, Uganda and Tanzania (mainland Tanzania and Zanzibar)

2. Adapting and using the modules

- Learning about quality through measurement, or the systematic collection and analyses of data from children's learning environments, is an important step in creating plans for improvement. With well-designed measures that reflect the goals for ECE, regular monitoring of ECE settings can increase quality by providing information on strengths and areas for improvement.
- This section covers four steps of adapting and using the modules: (1) planning for use of MELE modules; (2) adapting to the national context; (3) field-testing and validation; and (4) application to policy. After measuring quality, many countries may find that ECE quality is not consistently high enough to support young children's development. Improving ECE quality can focus on elements such as revising the curriculum, teacher training and coaching, establishing regulations for ratios and teacher qualification, and increasing investments in classroom supplies and materials.

Throughout these steps, the following principles should guide efforts:

- Use science to determine what matters most for young children's development, by focusing on research to help clarify the most critical elements of quality.

- While building a measure to help learn about quality is important, the emphasis should be placed on 'mentoring' to help providers improve, rather than 'monitoring,' which is important but limited in its ability to promote improvement.
- Quality cannot be assumed solely from the outcomes that children achieve. For a variety of reasons, equating 'quality' with 'outcomes' produces misleading information and, most importantly, often fails to provide enough information on where to focus improvement efforts. Quality should be measured in partnership with measurement of learning and development.
- Quality is the result of many factors, including policies, monitoring and evaluation systems, teacher qualifications and supports, and characteristics of classrooms. Ideally, some indication of these factors would be included in quality measurement.

It is important to note that the scope of this review is constrained by the limited range of current research. Investment in local research is greatly needed. In an attempt to compensate for the lack of globally relevant data, this paper emphasizes research and analyses that address ECE quality in more than one country. Although most current quality rating scales were developed in the USA or the United Kingdom, this review focuses on studies conducted in more than one

country, with an emphasis on studies conducted in low- and middle-income countries (LMICs). Also, while ECE can refer to services delivered through a variety of mechanisms – including home visiting, parent education, and school- or community-based pre-primary settings – the majority of cross-cultural research to date was conducted in formal pre-schools.² Therefore, the scope of the recommendations herein is limited to ECE quality in formal, pre-primary settings designed for children 3 to 6 years old – namely state- or privately funded pre-schools, run either in association with local schools or separately.

Development process

Quality in ECE should be defined and measured by its effectiveness in supporting young children's development in a manner desired by children themselves, their families and caregivers, and their communities and countries (UNESCO, 2007). Two widely consulted UNESCO publications – the *Dakar Framework for Action* (UNESCO, 2000), which established the EFA goals, and *Learning: The Treasure Within*, a report by the International Commission on Education for the Twenty-First Century (1996) – define 'quality' in education in large part by the outcomes achieved by students. Similarly, the Convention on the Rights of the Child outlines children's rights to holistic support of their development, as defined by the developmental outcomes they achieve (United Nations, 1990). MELQO proposes that countries define quality in ECE as the extent to which policies, programmes and classroom settings support the holistic development of all children, particularly those at risk of social exclusion, with specific goals established collaboratively by parents, teachers, communities and other stakeholders.

Measuring child outcomes is important. But looking at child outcomes alone provides limited information as to what should be changed in children's learning environments, and therefore may be less immediately applicable to education policy and programming than measurement of quality in learning environments. An accurate assessment of quality should ideally be complemented with a measure children's development, evaluated holistically and in culturally appropriate ways, but it should focus on the 'health' of the overall ECE system – the extent to which teachers, programmes, schools, the larger community and the government function and cooperate in ways that support children and are consistent with local and country priorities (Britto et al., 2011).

Points of consensus and disagreement arose during the MELQO process. There is agreement among all stakeholders that quality in ECE can be reliably measured. While there is

not a single universal definition of quality that should drive all measurement, we can draw upon what is currently known about the quality of environments that promote development and learning, to generate locally relevant definitions of quality. There is also not just one type of 'quality' learning environment. In ECE, there is no conclusive evidence that one specific type of setting – connected with schools, community-based, public or private – consistently promotes quality and child development in all countries. Instead, it seems that high-quality ECE can be delivered in a range of pre-primary settings. A number of different types of programmes have been found to benefit young children, and variation in quality among settings depends on many factors, including country context. In Bangladesh, for example, home-based pre-schools can be just as effective as school-based ones in preparing children for school (Aboud and Hossain, 2011). Research has found that in Turkey, public, private and institution-based pre-schools do not differ reliably in quality (Kalkan & Akman, 2009), whereas school-based pre-schools in Cambodia generally offer better quality ECE than home-based, although children benefited from attending both (Rao, Sun, Pearson et al., 2012).

MELQO's module and the content of this manual reflect the points of consensus among experts in quality measurement. There is a fair amount of agreement among ECE professionals on the core elements of quality, which can serve as the groundwork for community-based adaptations. In general, high-quality ECE in pre-primary settings is characterized by attention to certain key 'domains' affecting children's experience, an overall focus on child development, an emphasis on language and literacy, and the presence of well-trained teachers. These domains are described below, along with examples of items that have been tested to date as part of the MELE effort. However, there was not consensus on one set of items with relevance across all settings – thus the MELQO recommendation was to outline a small bank of items along with guidelines for local adaptation.

There is also agreement on the importance of developmentally appropriate settings. A key element of quality across cultures is an appropriate emphasis on children's development in all areas, not only early academic skills. With ECE-specific research lacking in many countries, policy-makers can look to the substantial body of evidence that describes how children learn and develop to inform local definitions of quality. Studies spanning high- and low-income countries clearly show that stimulating environments, emotionally supportive relationships, and access to health care and good nutrition support ECD around the world (Engle et al., 2007; Grantham-McGregor et al., 2007; Walker et al., 2011). High-quality ECE settings are specifically designed to meet young children's developmental needs for play, social stimulation, creative expression and active engagement with their environments. ECE should not prematurely push

² Some research on the quality of home visiting is available, e.g. Alderman, 2011, and Gomby et al., 1999.

children to acquire academic skills that are developmentally beyond their reach, through rote memorization. Rather, in quality ECE settings, children develop age-appropriate skills in early literacy and numeracy – for example, learning letter sounds, looking at books, hearing stories read aloud and sorting objects by colour, shape and physical properties. Research demonstrates that placing young children in classrooms designed for children even one to two years older, without adequate adjustment of the curriculum and teaching practices to address the range of developmental needs, slows rather than accelerates their academic achievement over time (Rao, 2012). **Therefore, a central component of quality ECE is a match between the classroom setting, teaching practices and children’s developmental needs.**

The domain selection process and key characteristics of the MELE module are outlined in the sections below. Before describing details of the module’s content and structure, it is helpful to review the desirable characteristics of tools to measure quality. First, quality measurement should meet basic psychometric properties, such as construct and content validity (more detail is provided in the MELQO Overview), and reliability between people administering the scales. While some scales, such as the ECERS-R, have often been adapted for use within LMICs, evidence demonstrates both success (e.g. Hadeed, 2014) and failure (e.g. Gordon et al., 2013) in replicating the scales within new populations. Second, recent large-scale studies suggest that there are elements of children’s environments that are more critical to measure than others – in particular, teacher–child interaction and support for children’s learning have been found to be most important for children’s later development (Zaslow, et al., 2016). Finally, research scales are often designed for use among highly trained observers, who may or may not be readily available in some settings. The MELE module was developed to overcome these challenges, and the question of how best to align the module with cultural and national expectations should be addressed with these issues in mind.

Domain selection

The MELQO project brought together stakeholders to address questions of measuring quality in learning environments. As outlined in the MELQO Overview document, the MELE group included stakeholders, experts and others who had developed and used a number of quality measures and/or had supported pre-schools in various countries. This group identified important domains for quality in pre-school centres. Following the recommendations of this group, the MELE module was developed to represent dimensions, or constructs, of a broad conceptual framework for the measurement of quality (Britto et al., 2011), with sample items for measuring each construct.

Several rating scales have been developed to measure ECE quality in formal pre-primary settings, and many have been validated by showing relationships to children’s learning and development (see UNICEF, 2012, for an overview of some several scales). Many of these scales overlap, suggesting a certain degree of international agreement on the key indicators of quality in ECE classrooms. Among the scales currently available, there is a good deal of consensus on the following basic elements: an indication of the provision of materials and physical setting; the type and emotional tone of teacher–child interactions; children’s engagement in classroom activities; adequate play time and space; and attention to health and safety.³ Some of the source measures for the MELE module include:

- the Early Childhood Environment Rating Scale-Extension (ECERS-E) (Sylva et al., 2003);
- the International Step by Step Association’s Principles of Quality Pedagogy (ISSA, 2010);
- the ACEI Global Guidelines Assessment (GGA) (ACEI, 2011);
- the Early Childhood Education Quality Assessment Scale (ECEQAS) designed for India (Kaul et al., 2012)
- the Tamil Nadu ECERS (TECERS) as modified for use in India and Cambodia (Isely, 2001);
- the Classroom Assessment Scoring System (CLASS) (Pianta, Le Paro and Hamre, 2008); and
- the Teacher Instructional Practices and Processes System (TIPPS) (Seidman et al., 2014).

Each measure has strengths in particular domains. For example, the CLASS and TIPPS concentrate on interactions, the ECERS-E concentrates on literacy, numeracy and science, and the TECERS has many appropriate items on the safety of the physical environment. The MELE team relied on measures that had been modified for use in LMICs, and where local assistants had been trained to use the measures with good reliability.

Format selection

Observations of teacher–child interactions and ECE settings, teacher and parent surveys, and evaluations of child outcomes are all approaches that can be used to assess quality (OECD, 2012, 2015). Assessments of quality in classrooms can rely on rating scales, which are administered by trained observers across programmes, or on documentation of children’s learning and development through portfolios and other means. Considerations and issues for each type of measurement are outlined briefly below.

³ The scales tend to differ from one another in the extent to which physical settings and range of materials are emphasized, due to differences in types of classroom settings.

Observational scales. Observational scales refer to scales that use a trained, outside observer to identify key elements of quality during a typical day. Many quality scales have been developed to use trained observers who visit classrooms on a typical school day to obtain information about classroom interactions and routines. Observation is considered one of the most reliable ways to collect data on ECE, mainly because when observers are trained well, they are able to capture the most critical elements of learning environments in an objective manner. At present, ECE experts argue that reliable estimates of quality can only be obtained through direct observation, because indicators like class size and teacher–child ratios are inconsistent predictors of quality, especially in countries with little research on ECE quality in the local context (Cryer et al., 1999). In many countries, attention is now turning to observational measures of quality that are more likely to capture the most critical elements of children's experiences, teacher–child interactions and dialogue.

Quality assessments from teachers and parents. Feedback from teachers, including self-reflection or self-evaluation, can be very useful in measuring ECE quality and identifying where to focus improvement efforts. Teacher feedback and perceptions of work environments are perhaps ideally solicited through detailed reflection on classroom practices but can also take place through surveys, which have been proposed as one route towards obtaining national-level information on quality. The UNESCO Survey of Teachers in Pre-Primary Education (STEPP) project, for example, is focused on obtaining teacher feedback on experiences in classrooms, along with information on professional development and support.

However, some caution is needed in approaches relying on teachers' responses to survey questions, as teachers' perceptions of quality may vary widely depending on context. Research suggests that teachers tend to rate quality differently depending on whether the programme is functioning well or not. Contrary to expectations, staff in low-functioning programmes have been found to rate their overall quality as higher than staff in high-functioning programmes (Sheridan et al., 2009). Teachers and other staff members may not have a common understanding of what defines quality, especially if there has been no effort to define it collaboratively. Despite this potential limitation, engaging and listening to teachers should be considered an important element of assessing quality. It might be better to simply ask teachers what and how they conduct various activities rather than asking them to evaluate the quality of their programme.

Soliciting feedback from parent and families is also important, and careful consideration should be given to how best to engage parents. In spite of general agreement that parents should be engaged in conversations about defining and measuring ECE quality, there is little research available on parents' engagement in ECE settings outside the USA.

Cultural differences may influence parental perceptions of some elements of ECE quality, particularly around the degree of autonomy or independence encouraged by the school and the expectations for moral teaching. On the other hand, there may also be notable areas of agreement among parents of different cultural backgrounds, such as emphasis on the importance of teacher quality, curriculum, school practice and environment (Yamamoto and Li, 2012). Parent engagement and involvement has been shown to increase primary school quality in Africa (Duflo Dupas & Kremer, 2009), but there is no analogous research on ECE. Overall, parent reports of ECE quality tend to be very weakly associated with observed quality and vary by education level, with more educated parents viewing ECE quality more critically than less educated parents (Torquati et al., 2011). In the USA, parents have been shown to perceive ECE programmes more positively than trained observers, probably because the trained observers have the advantage of comparing multiple programmes whereas parents are likely to observe only the programme their child attends (Cryer & Burchinal, 1997). In sum, while it is important to engage parents as one of several voices in determining and monitoring ECE quality, parent perception is an important piece of complementing information from objective observers and surveys to provide a multidimensional view of quality.

Globally comparable indicators. Some indicators collected globally by the UNESCO Institute for Statistics (UIS) include regulatory data such as teacher–child ratios, the percentage of trained teachers and funding for PPE. While undoubtedly useful, the degree of correlation between these indicators and children's experiences in ECE settings is likely to vary depending on context. Although data on teacher–child ratios, teacher qualifications and funding levels are easier to obtain and compare across countries, such data are also significantly less likely to provide valid assessments of the actual quality of ECE services and can therefore lead to inaccurate comparisons (Hustedt and Barnett, 2010; Tobin, et al., 2009). Some reports have attempted to infer ECE quality based on country-level policies, regulations and access levels as reported by governments, using this information to conclude that some countries have higher-quality ECE than others (EIU, 2012). These types of structural indicators are limited at best, and relying on them to infer process quality at the classroom level is not recommended.

Nevertheless, quality in ECE settings such as pre-schools should be conceptualized within the larger framework of the overall infrastructure that sets regulations, provides funding and trains teachers (Britto et al., 2011), and developing a systematic way to summarize policy influences related to quality can be very valuable. Several international organizations have recently addressed questions about policies in ECE, providing sources of information on important elements and indicators of the policy context. The OECD,

for example, has identified five strategic policy levers for quality, including establishing quality goals and regulations; designing and implementing curriculum standards; improving staff qualifications, training and workforce conditions; engaging parents and communities; and advancing data collection, research and monitoring (OECD, 2012). The World Bank's System Analyses for Better Education Results (SABER) ECD framework outlines three policy goals for effective early childhood systems, including but not limited to PPE (Neuman and Devercelli, 2013):

1. Establishing an enabling environment for ECD, including the development of policies guaranteeing sustained access, funding and quality assurance for ECE settings;
2. Implementing widely and ensuring access to health, education and nutrition, especially for the most vulnerable children; and
3. Learning about quality in a systematic way through regular monitoring, and assuring quality through the development of data systems to track child development, policies that set learning standards and teacher qualifications, and ongoing evaluations to ensure compliance.

To date there is a limited body of evidence explicitly examining the link between quality as defined by globally

comparable indicators and quality in ECE settings in countries around the world. Local research and country-level experiences are essential to identifying key quality indicators and determining which policy approaches are most beneficial for improving ECE quality and supporting child development.

MELE Instruments

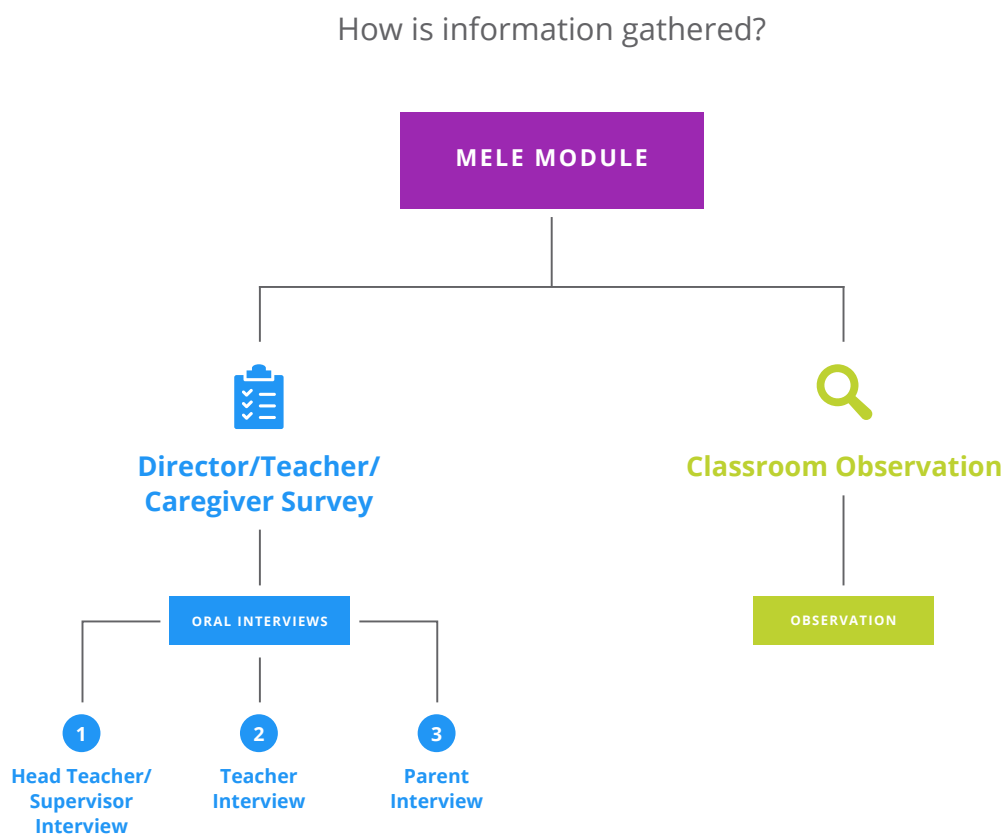
In light of these considerations, MELE includes an observational instrument along with surveys for teachers, parents and directors. The module consists of four instruments or tools:

1. Observation tool (for use in the classroom or programme itself during at least a half day of activities)
2. Teacher interview
3. Supervisor interview
4. Short parent or caregiver survey (to gauge experiences of family engagement)

Figure 1 displays the different instruments for gathering information through the MELE module.

The items in these tools should be considered a starting point for national adaptation, and users are expected to remove, add or modify items as needed.

Figure 1: MELE Instruments



Key domains of quality

A growing body of research demonstrates that certain elements or domains influence children's learning and development in ECE settings across countries. The following domains were first identified at a meeting of ECD experts and stakeholders in Leiden, Netherlands in September 2014, and were then modified based on experiences with the MELE module, as outlined in the MELQO Overview.

- Environments and materials
- Interactions
- Pedagogy
- Family and community engagement
- Inclusion
- Play
- Personnel

Each domain is outlined in detail in the sections below, with information on why it is important. Examples of items for each domain can be found in Figure 2.

Across all domains, emphasis should be placed on finding culturally relevant ways to address quality. Domains are expected to be defined somewhat differently in different countries, based on cultural expectations and context.

Also, while all of the domains are important, some may correlate more closely with certain aspects of child development and learning. In general, children's activities and language environments tend to be more strongly related to cognitive development than other factors, and the immediate experiences of children in the classroom context, which occur in interaction with teachers, teachers' assistants and other children, are important influences on development and learning. Although the immediate setting may vary quite substantially – from a classroom with four walls to a more informal space that could be either indoors or outdoors – the quality of interactions and pedagogy are perhaps the most proximal and powerful influence on children in PPE.

Environments and materials

Environments and materials refer to the physical environment, or the physical space and safety of the classroom, including access to clean water, toilets and space for each child. To meet quality standards, the physical space should be safe, clean and promote good health practices. The learning environment

should provide children and adults with a sense of well-being and community, and offer frequent opportunities for interaction. A variety of culturally relevant and meaningful learning materials should be available, including visual displays, books, art supplies, etc. Programmes should provide access to clean drinking water, nutritious meals and adequate sanitation facilities. Standards for physical environments are often included in government regulations on pre-school settings.

The physical environments of schools have long been recognized as central to children's learning (Bernard, 2012). Early learning environments are ideally designed to support children's learning, by providing safe, clean spaces for learning, play and socializing. For children to reach learning goals, learning materials should be provided as well, including access to toys, indoor and outdoor spaces, and writing and math materials. Addressing safety issues is essential for ensuring that children are not harmed during school and for protecting children's rights. Aspects of the immediate physical environment, including safety, access to sanitation and improved water have direct health benefits children. The programme's procedures for child protection and for emergencies or disasters can make a difference in situations where children's welfare is endangered. These aspects of structural quality are obtained largely through the teacher-survey component of this instrument.

Interactions

Interactions refer to the type and quality of interactions between teachers and children, and between children and their peers. Children should experience daily interaction with teachers and school staff who are nurturing, emotionally supportive, trained in pedagogy and ECD, and attuned to children's individual needs.

Interaction between teachers and children is perhaps the most critical element of ECE quality, regardless of the specific curriculum used (UNESCO, 2007; Britto et al., 2011). In the USA and the UK, where much of the research on young children's learning processes has been conducted, it has been proposed that teacher-child interactions are at the heart of the teaching and learning process (e.g. Siraj-Blatchford and Wong, 1999; Mashburn et al., 2008). In pre-primary settings, teachers play a fundamental role in supporting young children's development. Teachers' qualifications, and specifically their knowledge of ECD, translate into developmentally appropriate teaching practices and interactions with children, with many opportunities for children to fully participate in classroom

CONSTRUCTS AND SAMPLE ITEMS FOR MELE

Figure 2: MELE Constructs and Sample Items

CONSTRUCT	CONSTRUCT	CONSTRUCT	CONSTRUCT	CONSTRUCT	CONSTRUCT	CONSTRUCT	CONSTRUCT
	Play		Pedagogy		Interactions		Environment
Emphasis of the programme on creating opportunities for all children to explore and engage in free play and group play; the presence of adequate toys and spaces to play.	Approaches that teachers take in teaching children, including individualized and/or group lessons and opportunities for dialogue, and in supporting a successful transition to primary school independent work.	Type and quality of interactions between teachers and children, and between children and their peers.	Physical space and safety of the classroom, including access to clean water and toilets, and adequate space for each child.	Experiences of teachers and directors in training, years of service, compensation, supervision and mentoring.	Extent to which parents and community members are encouraged and able to engage in children's education.	Extent to which the classroom is able to support participation for all children, which may include gender, learning needs, and cultural, ethnic and linguistic accommodations.	
<ul style="list-style-type: none"> Do all children have time for play during the school day? Do all children get an opportunity to use toys during play? Do all children have access to sufficient, varied and challenging materials, such as blocks, books and coloured pencils? 	<ul style="list-style-type: none"> Has an age-appropriate curriculum or set of guidelines been developed outlining competencies and lesson plans? Do children use objects to learn mathematics; for example, do teachers encourage children to use objects for numerical exploration like sorting, counting and operations? Do teachers introduce new vocabulary by reading storybooks to children daily? Are children learning to perform new skills independently? 	<ul style="list-style-type: none"> Do teachers discipline and maintain order without being excessively negative? How often do teachers smile or verbally praise children? Do teachers patiently coach children who struggle to learn a new concept? Do teachers encourage children's questions and respond to them with sentences of explanation? 	<ul style="list-style-type: none"> Is there clean drinking water available for the children? Are toilets available for both boys and girls? Are there safety hazards? Is indoor and outdoor space sufficient for play? Is there enough space for all children to sit and room for play? 	<ul style="list-style-type: none"> How many years have you been a teacher overall? Do you receive support from your supervisor, through in-class observations and professional development? During the last 12 months, how often have you been observed in your classroom teaching as a part of supervision, monitoring or training? How useful was the feedback you received from supervisors and from peers? 	<ul style="list-style-type: none"> Do parents have regular meetings with teachers to discuss children's learning and development? Do community members who are not parents (in the neighbourhood or village) participate in making decisions about the programme? How often does your pre-primary programme provide group sessions on parenting or home visits? 	<ul style="list-style-type: none"> Does the programme show evidence of encouraging enrolment and participation of all ethnic, linguistic, religious and gender groups? Are children with disabilities included in the programme? Does the programme include a focus on mother-tongue instruction? 	

dialogue (Mellor and Chan, 2002; Sheridan, 2007; Dahlberg et al., 2005). The nature and importance of teacher–child interactions in LMICs has not yet received as much research attention as in high-income countries, but existing data suggest that teachers are just as critical for young children in these contexts, if not more so (Aboud, 2006; AKF, 2010).

A definition of high-quality ECE teaching, based on developmental and pedagogical theories, includes giving children the support and structure needed to creatively experiment with objects and artwork, offering emotional support, and avoiding punitive interactions (Sheridan, 2007; AKF, 2010). This research may be consistent with the observation that ECE settings around the world with free-choice activities for children tend to result in better outcomes over time (Montie et al., 2006). Despite potential benefits, however, a child-centred approach – which includes children’s choice of activities, encouragement of initiative, and learning through play – may not be culturally acceptable in all settings, nor are all teachers trained with the pedagogical perspectives or skills required to handle young children’s participation in decision-making. Again, the limited research in diverse cultural and economic contexts is problematic. The question of whether child-centered approaches predict children’s development in all contexts, and the extent to which they are culturally influenced, must be further investigated, but evidence to date suggests that they are critical to learning for all children.

Pedagogy

Pedagogy refers to the approaches that teachers take in teaching children new skills, including individualized and/or group lessons and opportunities for dialogue, and in supporting a successful transition to primary school, which is characterized by more independent work. Child-centred teaching should encourage initiative, curiosity, persistence, attentiveness, cooperation, participation and active engagement. Children should engage in age-appropriate play, activities and routines. Curriculum content should address children’s physical, social–emotional, language and cognitive development needs and stimulate early literacy and numeracy skills.

Pedagogy, or the instructional practices used in the classroom, has strong implications for children’s learning. Research shows that high-quality language and literacy environments in the mother tongue may have a strong positive impact on young children (Ball, 2011). Studies from low-income countries, such as Bangladesh, have demonstrated that rich language and literacy experiences are critical for developing expressive vocabulary as a foundation for learning to read (Opel et al., 2009). Similarly studies from high-income countries have demonstrated

that rich language and literacy environments are especially critical for later literacy development. Teachers who build and expand on children’s ideas, allow them to engage in dialogue, and expose them to books, songs and stories help promote reading skills several years into the future (Storch and Whitehurst, 2002; Lonigan et al., 2000). An emphasis on literacy is thought to be particularly crucial for children who have few opportunities for reading at home. It is also important that children have the opportunity to learn in their native language. In areas with a great deal of linguistic diversity, investment in the challenging task of finding and training teachers who are fluent in children’s native tongues could have enormous benefits for children’s language and literacy development.

Teachers’ scaffolding of numeracy and abstract reasoning in the classroom can lead to higher levels of early numeracy as well as executive function, attention and language skills. Numeracy skills consist of far more than counting skills. Although number recognition and counting are important foundations of later quantitative reasoning skills, just as important is exposure to notions of magnitude, seriation, spatial skills (e.g. concepts of over, under and behind) and shape recognition. Children’s use of manipulatives to learn properties and operations of numbers, rather than rote memorization, is expected to link to numeracy and math skills. Play-based activities with objects can build these skills – for example, by having children identify simple geometric shapes and then combine them into more complicated shapes of the child’s own devising. Children’s use of manipulatives to solve math problems stimulates reasoning that goes beyond what is involved when children copy teachers’ actions. As with language and conversation, these activities encourage children’s participation. Young children’s thinking, reasoning, affect and responses to new shapes, comparisons and concepts can be actively facilitated by the teacher so that a ‘math lesson’ in the classroom moves beyond rote instruction (Opel, Zaman, Khanom & Aboud, 2012).

Family and community engagement

Family and community engagement refers to the extent to which parents and community members are encouraged and able to engage in children’s education. Programmes should share information, promote positive relationships, and create opportunities for parent and community engagement. Families and the local community should be actively involved in planning, decision-making and action to improve ECE.

Family engagement in children’s education, beginning with ECE, is one way that ECE programmes may create positive benefits for children over time (e.g. Duflo, Dupas & Kremer, 2009). By engaging parents in children’s learning, ECE programmes build parents’ awareness of the importance

of education and a deeper understanding of how best to support it, which in turn has been shown to have substantial, long-term impacts on children's educational outcomes and well-being (e.g. Reynolds, Ou, & Topitzes, 2004; Campbell, et al., 2002). Family engagement takes place through a range of activities, including parent–teacher conferences, opportunities for parents to engage in children's classroom activities, and information-sharing with parents about children's activities and how to support learning at home. Parent engagement has been shown to be especially important for families without a strong sense of connection to schools, who may not have had positive experiences with their own education (Halgunseth, 2009).

Family engagement includes programmes' relations with families and the wider community, and can also include opportunities for parents to engage in governance of pre-primary settings through participation on committees, volunteering in classrooms, and meeting regularly with teachers and directors. This type of family engagement in PPE can vary widely depending on the cultural and policy context. In some settings, families are not allowed on the pre-school grounds; in others, there is active provision of workshops and visits to classrooms, with explicit efforts to translate and transfer some learning activities conducted in the classroom to the home to strengthen learning across the child's entire day.

Inclusion

Inclusion refers to the extent to which the classroom is able to support participation for all children, which may include addressing diversity in gender; learning needs; and cultural, ethnic, religious and linguistic backgrounds. All children and families should have access to high-quality ECE services. Teachers should speak the home language of the majority of students. Teachers should be trained in providing ECE to children with disabilities and special needs, and should foster age-appropriate development, positive social interactions and play among all children in the classroom.

Meeting goals for equity in PPE requires an emphasis on including and supporting the learning of all children, including children with special needs, children representing diverse cultural or family backgrounds, and children living in all parts of the country. Children with special needs are often among those who benefit most from quality early childhood programmes, and importantly, *all* children have been found to benefit from inclusive settings (CGECCD, 2016). Linguistic inclusion is also important: children who learn in mother-tongue pre-primary settings benefit more from instruction and learn faster over time (Ball, 2011). Inclusive settings are also central for protecting children's rights by ensuring access for all children (UNESCO, 2015).

Play

Play refers to the provision of opportunities for all children to explore and engage in free play and group play, as well as their access to adequate play/learning materials and spaces to play. Children should be given ample opportunity to explore and enjoy learning through play, with time for pretend play and interactions with peers.

Play has been identified by many as a critical contributor to children's development, by providing children with stimulating social interactions and encouraging imagination. Autonomy, or the ability of children to spend some time each day choosing their own activities, is developmentally appropriate and associated with children's learning. A study of pre-school quality in Bangladesh, for example, showed that daily access to a range of learning stations with varied materials, and the inclusion of free play and group time, were strongly associated with cognitive development (Aboud, 2006). In another multicountry study, independent activities were associated with higher cognitive development scores at age 7 across several countries, especially when compared with activities focused on sharing social experiences in a group setting (Montie et al., 2006). Increases in the variety of equipment and materials available in ECE settings have also been associated with higher scores in math at age 7 (Montie et al., 2006).

Free play provides an important opportunity for children to actively participate in their own learning. The provision of different materials – print, objects (e.g. human-made objects or naturally occurring materials), and visual and art materials – can provide a 'menu' of activities for free play periods. Free play is not conducted to teach children about specific concepts (e.g. colours or shapes) but rather allows children to experiment with concepts like colours and shapes to construct something new. For this reason, it is not the specific play area that is linked to child development so much as the number, variety and challenge of the materials manipulated by a child to attain his or her goal. This may explain why independent studies have found a relationship between child outcomes and free play activities and materials (Aboud, 2006; Aboud and Hossain, 2011; Montie et al., 2006). Free play also encourages peer interaction and so promotes cooperation, social–emotional skills, and language. It can be supplemented (but not replaced) by teacher-led group participation in art, song, rhymes and cooperative games. These activities also instil appreciation of cultural forms of expression, as well as self-regulation and gross- and fine-motor skills.

Personnel

Personnel refers to the experiences of staff (teachers and directors) in training, years of service, compensation, supervision and mentoring through professional

development. Teachers should be well-supported in improving their practices throughout their tenure, including regular participation in professional development, adequate training and recognition of their work. Directors should be engaged with teachers and effective in supporting teachers to improve their practices.

The nature and intensity of teachers' professional development and training (both pre-service and in-service) have been shown to influence the quality of their interactions with children. In pre-service training, different early childhood programmes have different requirements for formality, intensity and content of pre-service training. Typically, the more informal or community-based models have less intensive pre-service training, with some programmes providing as little as two or three days of training. More formal PPE qualifications are usually associated with much more prolonged and frequent training sessions, up to and including those associated with formal degrees. Specialized and more intensive training in ECD in particular has been linked more strongly to setting-level quality in PPE than has generalized training in education. In-service training often consists of periodic didactic workshops; however, recent evidence

from multiple LMICs suggests that on-site observation of teachers' pedagogy, instruction and interactions, coupled with supportive feedback in a trusting mentoring relationship, can improve classroom setting quality (Opel et al., 2009; Yoshikawa et al., 2015).

The nature of leadership of the ECE centre – including directors' training in the early childhood field, and the provision of supervision that includes mentoring and not simply oversight of structural factors such as safety or health and sanitation – can also influence the quality of what children experience in the classroom. The MELE module therefore includes questions about both training and supervision.

Overall, while there is not as much research from diverse settings as desirable, it is clear that teachers are central to children's ECE experiences. Quality teaching that provides children with stimulating, language-rich and emotionally supportive environments seems likely to promote learning. There is now an opportunity to use the existing research base to test these assumptions more fully, and to identify effective, feasible and culturally relevant approaches to teacher training and support.

Using and adapting the tool

The MELE planning process is centred around creating a scientifically based, feasible and locally relevant tool for measuring quality. Achieving these goals requires careful planning to bring together important stakeholders, evaluate items, train data collectors and anticipate how results can best be used to inform policy and practice. Outlined below are the steps to guide countries through the process of creating their own MELE tool. An overview of the general structure and process for adaptation appears in the MELQO Overview as well this section.

Figure 3 outlines the four main phases of the MELQO implementation process, which applies to both MELQO modules. Each phase is described in detail in the sections that follow.

At the end of the adaptation process, countries will have a set of items that reflect national priorities and goals. These items are intended to help inform national monitoring systems, by identifying areas that could be monitored on an ongoing basis in addition to the health and safety standards that often comprise national monitoring systems. The results from the MELE tools can be used to track the overall quality of young children's learning environments and to identify the areas in which specific actions are needed, such as providing

teachers with more training and support for interacting with young children, offering teachers professional development opportunities on a particular developmental domain, or ensuring that classrooms have access to high-quality, age-appropriate learning materials for all children.

Planning

Defining policy questions: Planning for use of the MELE module should begin with clarification of policy questions that will be informed by the measurement results. Examples of research questions that MELE can help inform include the following:

- What are the characteristics of pre-primary classrooms? What types of activities are typically underway, and how are children spending their time during the day?
- What are the characteristics of teachers and directors? How much training and support do teachers typically receive?
- What are parents' views of early learning environments?

Over time, MELE is intended to help inform on-going monitoring and evaluation of pre-primary environments by serving as a starting point for developing monitoring tools focused on the most critical elements of children's learning environments. The planning process can be viewed as a first step in building an effective system for monitoring quality over time.

Team structure and skills required: Large-scale data collection and analysis requires a team of individuals with varying levels of expertise and experience with the local population. Ideally, teams should be comprised of individuals at the country level with individuals at the global level serving in an advisory role.

Country-level team members: Depending on roles and responsibilities, the majority of positions should be filled by individuals from the local country. Team members at the country level are best situated to understand and address the needs of the local population. Within each country, MELQO recommends the creation of senior-level positions to assume leadership roles for particular aspects of the data collection and management processes. The section on leadership roles and responsibilities below outlines the recommended leadership structure.

Global-level team members: Some positions require expertise that may not be available within the local country. In these cases, members of a global advisory team can provide guidance and oversight on specific areas. Advisory team members may include individuals with select technical expertise in areas such as child development, survey research methodology and psychometrics.

Leadership roles and responsibilities

- Project Lead:
 - Overall project manager; responsible for coordinating all technical, training and data components; coordinates with local liaison in the development of country-specific scale components, adaptations and alignment needs
- Technical Lead:
 - Responsible for overseeing the technical quality of the assessment (e.g. sampling design, data analysis)
- Training Lead:
 - Responsible for conducting trainings for classroom observers; in charge of maintenance and fidelity of observer standards
- Data Collection Lead:
 - Responsible for coordinating all aspects of data collection; maintains data quality; coordinates with technical lead on data analysis
- Local Stakeholder and Policy Liaison:

- Responsible for working with local parents, teachers, school directors and policy-makers to ensure local voices are reflected in the instruments; coordinates with technical lead on the alignment with national standards

The leadership structure outlined above serves as a starting point for countries to consider when conducting global measurement. In addition to the leadership team, multiple individuals focused on specific tasks and goals are needed within each team in order to conduct quality measurement. With the aid of the global advisory team in certain technical areas and/or in the early phases of implementation, this type of model provides the structure to facilitate countries in taking ownership of their measurement needs and capabilities.

Adaptation

The adaptation process calls for examination of the seven MELE constructs with local experts and stakeholders to design a scale that is appropriate and useful for each setting. This process involves reviewing the constructs, discussing relevant items, and deciding upon a set of items that stakeholders view as accurately representing the goals of early learning settings. As noted in the description of the constructs, the specific manifestation of each construct may vary based on the setting. A successful adaptation process results in a tool that is able to index the range of quality within a given country (meaning that some programmes are considered higher quality than others) and with a range of constructs included (meaning that the quality rating is based on several elements of early learning settings).

It is recommended that the adaptation process take place in a workshop focused specifically on discussing the constructs and generating a list of items that seem well-aligned to local settings. National standards for children's learning and development and the quality of early learning environments are an important part of this process. In advance of the adaptation workshop, it is useful to summarize the standards and conduct a preliminary mapping to determine which items are most closely aligned to the national standards. Existing tools developed for use in other countries as part of the MELE process can be used as a guide. A PowerPoint deck has been created to serve as a starting point for this process, which is available upon request.

In MELQO experiences to date, the adaptation workshop has yielded important feedback from country stakeholders, which then can be used to create a draft version of the tool for field-testing. Expertise required to create a workable version of the scale for testing may exist within local universities, or it may be useful to partner with a global measurement expert to ensure that the items are workable to use for reliable training and data collection. The adapted version of the module

Figure 3: Phases of MELQO Implementation

FOUR PHASES OF MELQO IMPLEMENTATION



also requires a training manual so that observers and data collectors can be reliably trained.

Field-testing, data collection and validation

The adaptation process will result in a tool that is specific to each context – likely with some items that have been used repeatedly across countries, and some that have been revised or newly developed. As a result, it is necessary to field-test the items before moving to large-scale data collection.

Field-testing and piloting may require more time and attention for the classroom observation tool than for the parent/teacher/director questionnaires. Special emphasis on the classroom observation instrument is required because good data are dependent upon **reliably trained assessors**. To ensure consistent interpretation and use of scores from the observation instrument, standardized training needs to be provided for all classroom observers. This training needs to provide observers with opportunities to gain exposure to the tool and complete practice exercises with guided feedback, and should be planned over the course of at least two to three days, to give observers plenty of time to practice making observations and reflecting on their ratings. Training materials need to provide clear definitions for the discrete behaviours included in each of the domains and items. Comprehensive training combined with precise definitions for the behaviours observers are asked to rate will lead to less ambiguity in observers' scores. An instrument with clear, well-defined items and descriptions, facilitates the ability for classroom observers to better complete their task and subsequently raises the level of data quality. Although observation scores will always be subject to some degree of error, error can be minimized by providing clear instructions and standardized training opportunities for observers.

Table 1 summarizes the stages of data collection from pre-field testing to large-scale data collection.

Scores from any scale intending to measure quality should ideally be validated to determine their usefulness both within and across countries and contexts. Validation includes testing the items to make sure they accurately and reliably predict quality. Because ECE quality should be determined by its effectiveness in reaching agreed-upon goals for children's development, countries should ensure that any quality indicators they use are in fact reliably related to young children's development, particularly for those children at risk of social exclusion. The items on ECERS, for example, have been shown to reliably predict children's cognitive development in Bangladesh, Cambodia, China, India, Sweden, the UK, the USA and Viet Nam, as described by the studies referenced in this paper, although recent evidence also

questions the extent to which all items are important for children's development (e.g. Zaslow, 2016).

Thus while the existing work on MELE provides a very sound starting point for countries to begin quality measurement, local validation of the instruments to ensure their applicability within the national context should be considered a high priority. This validation process should include soliciting feedback from teachers and parents on the proposed constructs and their relevance to goals for children, as well as determining whether children who attend the settings considered higher-quality actually do benefit more from PPE than children in lower-quality settings (this is called 'predictive validity').

Designing a validation study requires careful planning. Possible designs include following children over the course of a school year, to see if children in the settings rated as higher-quality gain more over the course of the year; and measuring quality and child development at the end of the pre-primary years and then following children into early primary school to see if children in higher-quality ECE environments benefit in early primary grades. Various designs can be used to assess the validity of the items in quality tools; the most critical point is that an approach must be envisioned to ensure the tools work well.

Issues to consider

To date, there have been several updates and revisions to the MELE tools, each reflecting the local goals and priorities of each setting. Although these updates have built upon each other to develop the current module, there are still several considerations that need to be addressed as the tools continue to be modified for use within specific settings.

One of the main aspects stakeholders need to address is the creation of concise, specific definitions for each of the domains and items included on the MELE module for each country. For example, interpretations of positive or negative interactions may reflect cultural standards for how people interact with one another, and so should be carefully field-tested to be sure that observers are clear on how to define them. Another example relates to the item on 'free play.' The term free play can mean multiple things depending on the context, so a specific definition needs to be provided. Observers rating classroom environments need to have a strong understanding of what aspects do and do not constitute 'free' play compared to another type of play, such as 'gross motor' play or school recess play. It is important that stakeholders provide specific directions and examples so there is no ambiguity in what terms mean across the different items.

Additionally, stakeholders need to consider the clarity and consistency of the descriptions provided within each

Table 1: Stages of MELE Data Collection

STAGES OF MELE DATA COLLECTION

STAGE	PRE-FIELD-TESTING	FIELD-TESTING	FULL DATA COLLECTION
Purpose	To establish that the newly translated and adapted items are understandable and feasible. Results should not be analysed for assessment of quality in early learning settings.	To establish a distribution of scores on the scales of the instrument(s) to allow for psychometric analyses and recommendations for revisions prior to final data collection.	To address the goals of the measurement of quality in learning environments as per country requirements.
Method	<p>The newly translated and adapted items are implemented with classroom observers, teachers, directors and parents. Notes are kept on the timing of the testing, difficulties encountered and potential ambiguities in wording.</p> <p>Note: If new, country-specific items are being considered, they need to be included in testing from this stage onwards.</p>	The prototype instrument(s) are implemented in a rigorous manner by trained assessors. Care should be taken to ensure accurate recording and scoring of items. A subsample of classroom observations are conducted twice within 2 weeks for test-retest reliability; if feasible, a subsample of classroom observations are tested by 2 assessors, and a subsample of parents/teachers/directors interviewed by 2 assessors for inter-rater reliability.	Implementation of the final version of the instruments in the agreed-upon design (e.g. samples from rural and urban regions, private and public schools, over a period of time, etc.) by the trained assessors with the full rigour of testing and interviewing. Data are entered into databases and analysed to address the project questions.
Sample	<p>Convenience sample for each of the tests being adapted.</p> <p>A minimum of 10 classroom observations is recommended, along with surveys of teachers, parents and directors.</p>	<p>The sample size will be dictated by the country's needs for full-scale implementation (to ensure that the groups of interest are represented in the pilot).</p> <p>At a minimum 20 classrooms and 20 teacher/parent/director surveys are recommended.</p>	Sample is based on the goals identified by the country partners; size is determined by the scope of the project.
Outcome	<p>Responses and administration notes should be closely examined to identify problematic items, which may have to be reviewed and potentially revised/removed.</p> <p>The final outcome is the instrument prototype ready for pilot testing.</p>	The results of pilot testing will be analysed for psychometric properties of the items (internal consistency, test-retest and inter-rater reliability). Items that do not perform well will be identified, and potential reasons for this explored (e.g. bad translation, inadequate administration, etc). Recommendations for final revision of the instruments will be made as needed. The final instruments will then be back-translated for review by the MELQO team. Once the continuity of constructs is confirmed, the instruments will be ready for full implementation.	A data set reflecting the quality of early learning environments in a country, within the framework of the country requirements. Psychometric properties will be re-examined, and group comparisons will be carried out as planned. Additional validity will be established through association with sociodemographic variables and, if possible, regional characteristics.

of the items. Some of the aspects within the descriptions that need to be evaluated include the terminology and progression across levels (low, medium, high). Terms used within the item descriptions such as ‘vocabulary’, ‘dialogue’, ‘discussion’, ‘discipline’, ‘interaction’, and ‘time’ need to be well-defined for classroom observers – likely to a level of detail that goes beyond what highly knowledgeable early childhood practitioners may assume, as observers may not have extensive training in ECD. In addition to definitions, examples should be added to help classroom observers better understand what these terms include or do not include. For example, several of the item descriptions reference the number of occurrences observed. However, more information is needed for this type of description, because ‘1 time’ could describe a single instance (e.g. child sings a song) or a collection of activities within a single lesson (e.g. multiple songs during a single large activity). Stakeholders may need

to place specific emphasis on defining terms to provide clear guidance for classroom observers. Creating workable scales for items, or defining what ‘high’ or ‘low’ looks like in various contexts, is an important piece of making the scales work well, and should be a focus within field-testing.

Next steps in development

The current versions of the MELE module have been crafted using existing measures and therefore provide a good starting point for measurement. Several versions of the tools have been developed to date, and more will be generated as the modules are used in more countries. This process will yield ongoing information on the applicability of various items and domains to different countries, and it should be assumed that the modules will change over time.

Application to policy

After assessing quality, most countries will find ECE settings in need of support for improvement. For monitoring to be successful in improving quality, results from quality assessments should be used in a transparent, supportive manner for maximum impact on quality. Countries must ensure that the data are used to provide useful feedback to teachers and programmes, with an emphasis on improvement rather than accountability or punitive actions. Delivering quality ECE services is challenging in many ways, and measurement is only useful if providers are also given the opportunities, support and means to improve. This is only possible through investment in visiting and learning about ECE settings directly. It is necessary to review documents and materials specific to each ECE programme, conduct observations, and talk directly with teachers, children and parents to determine the extent to which programmes are reaching their stated goals for each child who attends. Such in-depth, individualized assessments of quality require considerable investment but are likely to yield reliable and valuable results.

The following subsections describe strategies for improving both process and structural quality for ECE in a variety of contexts. Effective process quality improvement requires measurement, feedback to programmes on their results, and support for improvement efforts. On the structural level, national policies and regulations and community engagement can help improve system quality in ECE.

Improving quality in classroom settings

Targets for process quality improvement may include expanding and improving the physical space, providing classroom materials, integrating nutrition and health programmes, improving management of ECE programmes, increasing teacher qualifications and training, and strengthening connections with parents and community members. These targets tend to be interconnected: for instance, providing classrooms with appropriately designed learning materials, including culturally relevant picture storybooks, can be an important step in improving quality. However, it is equally important to ensure that teachers know how best to use the materials, and moreover that children are given regular access. Simply providing or improving classroom materials may not be enough; teacher training and changes to the curriculum or activity schedule might be necessary to enable children to benefit from these investments.

Several sources have articulated the tremendous need for additional investments in teacher education and training, to address the issue of untrained and/or unmotivated staff and to ensure that early childhood educators use age-appropriate methods (e.g. Orkin Yadete & Woodhead, 2012). Well-trained teachers are particularly important for the early years of child development and for children who face barriers to education. A study of pre-schools in ten countries, for example, found that teachers with more training or education correlated with children having higher language scores at age 7 (Montie et al., 2006). While there is consensus that well-trained teachers

are essential to high-quality ECE, evidence is lacking on which method of teacher training is most effective for quality improvement – in-service, pre-service or some combination of both. There is a general assumption that more education equals better teaching (e.g. EIU, 2012), and some countries have mandated higher degrees for all teachers. Yet some USA-based evidence suggests that in-service training and professional development may be more important than formal education (such as a university degree) in promoting good teaching and ensuring quality over time (Early et al., 2006; UNESCO, 2012). Research also shows that teacher training, whether in-service or pre-service, must combine theory with practice. Targeted training programmes that integrate information about child development with coaching within classrooms have been shown to be more effective than education about child development alone (e.g. Early et al., 2006). Training conducted in formal settings such as universities, with little or no follow-up once teachers enter or return to the classroom, has typically not had strong impacts on ECE quality, especially among new teachers or those without previous professional development experience (Sheridan et al., 2009). Training in implementing curricula can be useful, but this type of training alone is not enough to support teachers in cultivating the types of high-quality interactions with children that will promote their development. Instead, professional development opportunities such as intensive coaching and support for teachers are especially important for improving quality, and can lead to significant changes in classroom practices (Yoshikawa et al., 2015). Pre-service training should be considered a necessary first step in developing teachers, which then should be complemented by on-going coaching and professional development once teachers enter the classroom.

Ultimately, governments' decisions about the level of teacher education and training needed are likely to depend on the specific wages, recruitment procedures and overall education level of each country. That said, there are some conclusions that seem consistent across countries: (1) training and ongoing professional development are important for improving quality, and (2) this training should focus specifically on young children's development. Using upper primary school teacher training methods is not likely to be appropriate for ECE teachers in any country, as the developmental needs and approaches to learning are different for pre-primary children and children in the early grades (Rao et al., 2011).

There are additional strategies for improving programmes, without waiting for changes in teacher training. Those who are currently teaching and managing ECE settings should have an open discussion about items where they scored lower than hoped. Because descriptions of higher-quality settings are available as part of the module itself, the current description for a setting can be compared with the desired description. Strategies to reach that desired description

can be generated and the best ones selected to work on. Such strategies may require, for instance, re-thinking the curriculum, examining how it is translated into interesting lesson plans, or promoting the use of hands-on learning materials by all children.

The role of government policies in improving quality

National policies are an essential building block for ECE systems. National policies affect what happens in classrooms in at least two ways: first, by determining local regulations, resources, staffing, standards and curricula in ECE settings; and second, by ensuring consistency across ECE settings that aligns with the expectations and requirements of families, communities and primary schools, which is especially important for settings that serve children at risk of exclusion. Ensuring consistency across ECE settings does not imply rigidity; instead, common standards are a reflection of what stakeholders deem to be valuable enough to ratify as country- or programme-level policy. In addition to governments, communities often take on great responsibility for ECE, and in some countries the primary source of ECE is community-organized and run.

Government regulations, quality standards and priorities matter: research consistently demonstrates that structural elements influence the quality of ECE classrooms across a range of countries, and even between states in a federated system like the USA (Rao and Li, 2009; Mashburn et al., 2008). More specifically, the prevailing social and political context, the extent and enforcement of regulations, and funding – especially teachers' wages – have all been shown to affect ECE quality (Ahnert and Lamb, 2000; Pianta et al., 2005; Cryer et al., 1999). Changes in these factors have also been associated with changes in observed ECE quality, further underscoring the consensus among experts that policies are important for quality (Deynoot-Schaub and Riksen-Walraven, 2005).

While policies are clearly important for ECE, more research is needed to determine which specific regulations, standards and funding will lead to the most notable gains in quality. The importance of teacher–child ratios, for example, is still open to debate. In the USA and some parts of Europe, consensus has generally been that lower teacher–child ratios promote child development (OECD, 2012; Economist Intelligence Unit, 2012). Some research, however, has questioned this assumption, as the impact of ratios on child development has not been clearly documented outside of Western contexts (e.g., Tobin et al., 2009). Some studies in LMICs have shown that teachers with more qualifications *and* lower ratios led to higher-quality care, but it is not yet clear whether class sizes and ratios influence quality similarly in all countries (e.g. Rao et al., 2003). It is reasonable to assume that the extremely high ratios and

group sizes noted in severely underfunded programmes (such as one teacher for 40 or 50 children) are not ideal and are unlikely to promote child development. However, no specific ratio – of 1:20 or 1:10, for example – has been established through research as a ‘cut-off point’ that countries should adhere to. Whether such a ratio is needed in all settings requires more investigation and must also be weighed against the realities of resource allocation for ECE.

The establishment of learning standards to guide classroom practices has also been explored as a mechanism for improving quality. Defining what children should learn in ECE settings can be an effective route for improving quality, by focusing attention on goals for learning and development. UNICEF’s Early Learning and Development Standards (ELDS), for example, define learning goals for children from birth to age 8 using a participatory process that includes teachers, government representatives, ECD stakeholders and parents, and have been adapted for use in several countries, particularly in East Asia (e.g. Miyahara and Meyers, 2008; Kagan, 2012). With standards in place that outline what children should learn at each age, ECE programmes can orient their curriculum and pedagogy accordingly. Some countries have also used the ELDS as the backbone for national systems to collect data on children’s development. However, while reliable data on children’s development have been noted as an important characteristic of quality ECD systems (e.g. the World Bank’s SABER-ECD), it is important to keep in mind that

data on outcomes do not serve as a substitute for information on quality in ECE settings.

In many countries standards may exist with little or no support for implementing them across all ECE settings. As noted above, focusing on support for implementation of quality practices may be the most promising mechanism for improving quality. Efforts to promote quality should reflect a balance between revised standards and support for implementation, through mechanisms like professional development and systems for accreditation.

In reviewing the state of efforts to improve ECE quality, countries may find the following questions useful:

- Which governmental agency or organization has responsibility for quality improvement in ECE? Do the various ministries responsible for monitoring, curriculum development and teacher training communicate effectively with each other?
- What resources are available for quality improvement?
- What ECE pre-service and in-service training is available to all ECE providers, not only those in formal settings? How many providers have been reached through current efforts and what additional efforts are needed?
- Have service and learning standards been considered as routes for improving quality?
- How can communities be engaged in improving quality? What resources and supports would be most useful to community leaders?

Conclusion

Less than half of the world’s children currently receive PPE, and in low-income countries this figure drops to 15 per cent (UNESCO, 2012). Clearly, much progress is still needed to expand access to ECE. Yet as ECE access expands, attention also needs to be paid to quality. Children have a right to quality ECE, and countries have a responsibility to ensure that each child has access not just to ECE services but to *high-quality* learning environments that promote their holistic development. The MELE module is intended to assist countries in fulfilling that responsibility, in order to safeguard children’s rights and make the most of governments’ investments in early childhood.

Based on the review of available evidence on ECE quality in pre-primary settings, the MELE team has drawn the following broad conclusions:

1. Young children learn through relationships with others. They learn best when their physical, social, emotional and cognitive needs are all addressed, and when expectations for their development are age-appropriate and reflect their needs for play and exploration.
2. Definitions of quality in pre-primary settings should reflect local values and holistic expectations for children’s growth and development, and should take into account teacher–child interactions, available materials, parent involvement, and the needs of children at risk of social exclusion.
3. While more research in diverse settings is needed, certain core elements of early learning environments have been shown to promote children’s learning across cultures:
 - Settings with a variety of learning materials to promote early academic and social skills

- Settings with opportunities for self-initiated learning as well as whole group activities and time for play
 - Teachers who have experienced several years of full-time learning about effective teaching for young children prior to entering classrooms, and who regularly participate in professional development once they are teaching
 - An emphasis on language, with many opportunities for interaction with teachers and peers
4. Quality can be reliably measured using a range of methods, and choices about how to assess quality should be aligned with local values and plans for improvement.
 5. Quality can be improved. One strategy for improving quality in pre-primary settings is to recruit teachers with high levels of formal education and offer ongoing coaching and professional development once teachers are in classrooms. In countries where teachers do not receive formal education prior to teaching, investments in professional development may be even more important for ensuring quality. Discussion with all personnel about how to implement an age-appropriate curriculum with interesting lesson plans and well-used learning materials is also recommended.

Considerations for the future

International and expert guidance and norms can provide useful insights into best practices for improving ECE quality. However, country-level adaptation and implementation is essential. Each country must begin the process of addressing ECE quality with attention to local contexts, values and goals for young children's development, in particular for those children at risk of exclusion due to poverty; gender, ethnic or cultural background; or the presence of disabilities (Dahlberg et al., 2005). Definitions of quality reflect the cultural, social and political context in which children live. Parents, communities and governments all influence the specific goals and structure of ECE programmes and should have a voice in deciding what constitutes quality (Tobin et al., 2009; UNESCO, 2007; Dahlberg et al., 2005). Some ECE experts have proposed that quality must be locally defined and cannot be conceptualized across cultures, or only by identifying common, broadly defined elements like accessibility and availability (e.g. Dahlberg et al., 2005; Tobin et al., 2009; Vandenbroeck et al., 2012). Other experts have focused on adapting more specific, classroom-related quality concepts to local contexts and have demonstrated reliable relationships between ECE quality and child development outcomes (e.g. Rao & Sun, 2012). Both schools of thought not only offer

important perspectives on ECE quality that countries may find useful in developing their own approaches, but also highlight the importance of engaging in a transparent and widely accessible process to determine goals.

To achieve the goal of improving ECE quality globally, more country-level and cross-cultural research is needed, particularly in LMICs. In general, reliable data on ECE quality are lacking, and most of the research to date has focused on pre-primary classrooms in the USA and a few other countries (although a notable body of research has been developed in East Asia, e.g. Rao & Sun, 2012). Only a small number of LMICs have invested in conducting representative research on ECE quality, and very few studies have cross-cultural relevance. This lack of information impedes realistic appraisal of ECE quality and makes it difficult to plan strategically to improve quality. Community and country-based efforts are essential to ensuring the creation and sustainability of locally relevant approaches. Looking across available data and their limitations, the MELQO team emphasizes the need for countries to build capacity to conduct local research that fully captures the values and cultural and political contexts that influence ECE quality.

In particular, a key challenge for the ECE community is to determine the applicability of local findings to children in all countries. While conclusions drawn from existing research and practice are useful in shaping high-level thoughts on ECE quality, additional research is needed to address whether the themes identified by research to date are in fact critical for children's development globally. Some constructs, such as the importance of engaging in self-initiated activities, for example, may not be culturally appropriate in all settings, nor necessary for children's development, as children in some countries seem to develop normally in more authoritarian classrooms (Tobin et al., 2009). Systematic, cross-cultural investigation is needed to distinguish universally applicable concepts of ECE quality from those that are contextually dependent. Countries are encouraged to consider the evidence on ECE quality in light of cultural and contextual factors, to engage in locally designed, culturally relevant research on the relationship between quality and child development, and to work with local researchers to assess the extent to which development milestones are met for all children including those at risk of exclusion. It is also important to share the findings of this research widely. There are currently few designated locations online where up-to-date, reliable information on ECE quality can be found. It is vital to ensure that local and regional experiences in quality assessment and improvement are translated and shared globally, in order to improve our understanding of which aspects of quality are universal across countries and cultures.

References

- Aboud, F. E. 2006. Evaluation of an early childhood preschool program in rural Bangladesh. *Early Childhood Research Quarterly*, Vol. 21, No. 1, pp. 46–60.
- Aboud, F. E. and Hossain, K. 2011. The impact of preprimary school on primary school achievement in Bangladesh. *Early Childhood Research Quarterly*, Vol. 26, No. 2, pp. 237–46.
- Aga Khan Foundation (AKF). 2010. *Improving Learning Achievement in Early Primary in Low-Income Countries: A Review of the Research*. Geneva, Switzerland, Author. http://www.akdn.org/publications/2010_ecd_learning_paper.pdf
- Ahnert, L. and Lamb, M. E. 2000. Infant-care provider attachments in contrasting German child care settings II: individual-oriented care after German reunification. *Infant Behavior and Development*, Vol. 23, No. 2, pp. 211–22.
- Alderman, H. (Ed.). 2011. No small matter: The impact of poverty, shocks, and human capital investments in early childhood development. World Bank Publications.
- Association for Childhood Education International (ACEI). 2011. *ACEI Global Guidelines Assessment (GGA)*, 3rd edn. Washington, DC, Author. <http://www.acei.org/sites/default/files/global-guidelines/GGAenglish.pdf>
- Ball, J. 2010. Educational equity for children from diverse backgrounds: Mother tongue-based bilingual or multilingual education in the early years. Literature review. In Presentation to UNESCO International Symposium: Translation and Cultural Mediation, Paris.
- Bernard, J. 2012. A place to learn: Lessons from research on learning environments. UNESCO Institute for Statistics, Montreal, Quebec.
- Burger, K. 2010. How does early childhood care and education affect cognitive development? An international review of the effects of early interventions for children from different social backgrounds. *Early childhood research quarterly*, Vol. 25, No. 2, pp. 140–165.
- Britto, P. R., Boller, K., Yoshikawa, H. 2011. Quality of early childhood development programs in global contexts: rationale for investment, conceptual framework and implications for equity. *Social Policy Report*, Vol. 25, No. 2. Ann Arbor, Mich., Society for Research in Child Development. <http://files.eric.ed.gov/fulltext/ED519240.pdf>
- Campbell, F. A., Ramey, C. T., Pungello, E., Sparling, J., & Miller-Johnson, S. 2002. Early childhood education: Young adult outcomes from the Abecedarian Project. *Applied Developmental Science*, Vol. 6, No. 1, pp 42–57.
- Consultative Group on Early Childhood Care and Development (CGECCD). 2016. *Global Report on Equity and Early Childhood*. Leiden, Netherlands, ISSA. <http://www.ecdgroup.justinluke.us/wp-content/uploads/2015/12/CGGlobal-FullReport-English-R2-WEB-LowRes.pdf>
- Cryer, D., & Burchinal, M. 1997. Parents as child care consumers. *Early Childhood Research Quarterly*, Vol. 12, No. 1, pp 35–58.
- Cryer, D., Tietze, W., Burchinal, M., Leal, T. and Palacios, J. 1999. Predicting process quality from structural quality in preschool programs: a cross-country comparison. *Early Childhood Research Quarterly*, Vol. 14, No. 3, pp. 339–61.
- Dahlberg, G., Moss, P. and Pence, A. 2005. *Beyond Quality in Early Childhood Education and Care: Postmodern Perspectives*, rev. edn. Abingdon, UK, Taylor & Francis.
- Deynoot-Schaub, M. J. and Riksen-Walraven, J. M. 2005. Child care under pressure: the quality of Dutch centers in 1995 and in 2001. *Journal of Genetic Psychology*, Vol. 166, No. 3, pp. 280–96.
- Duflo, E., Dupas, P., & Kremer, M. 2009. Additional resources versus organizational changes in education: Experimental evidence from Kenya. Unpublished manuscript. Abdul Latif Jameel Poverty Action Lab (JPAL), Cambridge, Mass.: Massachusetts Institute of Technology.

- Early, D. M., Bryant, D. M., Pianta, R. C., Clifford, R. M., Burchinal, M. R., Ritchie, S. ... Barbarin, O. 2006. Are teachers' education, major, and credentials related to classroom quality and children's academic gains in pre-kindergarten? *Early Childhood Research Quarterly*, Vol. 21, No. 2, pp. 174–95.
- Engle, P. L., Black, M. M., Behrman, J. R., Cabral de Mello, M., Gertler, P. J., Kapiriri, L. ... International Child Development Steering Committee. 2007. Strategies to avoid the loss of developmental potential among over 200 million children in the developing world. *The Lancet*, Vol. 369, No. 9557, pp. 229–42.
- Gomby, D. S., Culross, P. L. and Behrman, R.E. 1999. Home visiting: recent program evaluations – analysis and recommendations. *Future of Children*, Vol. 9, No. 1, pp. 4–26. http://futureofchildren.org/futureofchildren/publications/docs/09_01_Analysis.pdf
- Gordon, R. A., Fujimoto, K., Kaestner, R., Korenman, S., & Abner, K. 2013. An assessment of the validity of the ECERS-R with implications for measures of child care quality and relations to child development. *Developmental Psychology*, Vol. 49, No. 1, pp. 146.
- Grantham-McGregor, S., Cheung, Y. B., Cueto, S., Glewwe, P., Richer, L., Strupp, B. and the International Child Development Steering Group. 2007. Developmental potential in the first 5 years for children in developing countries. *The Lancet*, Vol. 369, No. 9555, pp. 60–70.
- Hadeed, J. 2014. Reliability and validity of the Early Childhood Environment Rating Scale, Revised Edition, ECERS-R in Arabic. *Early Child Development and Care*, Vol. 184, No. 6, pp. 819-842.
- Halgunseth, L. 2009. Family engagement, diverse families, and early childhood education programs: An integrated review of the literature. *YC Young Children*, Vol. 64, No. 5, pp. 56.
- Harms, T., Clifford, R. M. and Cryer, D. 2005. *Early Childhood Environment Rating Scale*, rev. edn. New York, Teachers College Press. <http://ers.fpg.unc.edu/early-childhood-environment-rating-scale-ecers-r>
- Helburn, S. W. (ed.). 1995. *Cost, Quality, and Child Outcomes in Child Care Centers: Public Report*. Denver, Colo., Cost, Quality and Child Outcomes Study, Economics Department, University of Colorado at Denver.
- Hustedt, J. T. and Barnett W. S. 2010. Issues of access and program quality. P. Peterson, E. Baker and B. McGaw (eds), *International Encyclopedia of Education*, Vol. 2. Oxford, UK, Elsevier, pp. 110–19. <http://nieer.org/pdf/issues-of-access-and-program-quality.pdf>
- International Commission on Education for the Twenty-First Century. 1996. *Learning: The Treasure Within*. Paris, UNESCO. http://www.unesco.org/education/pdf/15_62.pdf
- International Step by Step Association (ISSA). *Competent Educators of the 21st Century: Principles of Quality Pedagogy*. Leiden, Netherlands, Author. http://www.issa.nl/docs_pdfs/Quality-Principles-final-WEB.pdf
- Isely, B. J. 2001. Tamil Nadu Early Childhood Environment Rating Scale. Chennai, India, M.S. Swaminathan Research Foundation.
- Kagan, S. L. 2012 Early Learning and Development Standards An Elixir for Early Childhood Systems Reform In S. L. Kagan & K. Kauerz (Eds.), *Early Childhood Systems Transforming Early Learning* pp. 55-70. New York, NY: Teachers College Press.
- Kalkan, E. and Akman, B. 2009. Examining preschools' quality in terms of physical conditions. *Procedia Social and Behavioral Sciences*, Vol. 1, No. 1, pp. 1573–77.
- LMTF (Learning Metrics Task Force). 2013. *Toward Universal Learning: Recommendations from the Learning Metrics Task Force*. Montreal and Washington, D. C.: UNESCO Institute for Statistics and Center for Universal Education at the Brookings Institution.
- Lonigan, C. J., Burgess, S. R. and Anthony, J. L. 2000. Development of emergent literacy and early reading skills in preschool children: evidence from a latent- variable longitudinal study. *Developmental Psychology*, Vol. 36, No. 5, pp. 596–613.
- Mashburn, A. J., Pianta, R. C., Hamre, B. K., Downer, J. T., Barbarin, O. A., Bryant, D. ... Howes, C. 2008. Measures of classroom quality in prekindergarten and children's development of academic, language, and social skills. *Child Development*, Vol. 79, No. 3, pp. 732–49.

- Mellor, E. J. and Chan, L. K. S. 2002. Conclusion: contexts, issues, developments, trends and challenges. L. K. S. Chan and E. J. Mellor (eds), *International Developments in Early Childhood Services*. New York, Peter Lang, pp. 253–70.
- Miyahara, J., & Meyers, C. 2008. Early learning and development standards in East Asia and the Pacific: Experiences from eight countries. *International Journal of Early Childhood*, Vol. 40, No. 2, pp. 17–31.
- Montie, J. E., Zongping, X. and Schweinhart, L. J. 2006. Preschool experience in 10 countries: cognitive and language performance at age 7. *Early Childhood Research Quarterly*, Vol. 21, No. 3, pp. 313–31.
- Myers, R. G. 2006. *Quality in Program of Early Childhood Care and Education (ECCE)*. Background paper prepared for the Education for All Global Monitoring Report 2007. Paris, UNESCO. <http://unesdoc.unesco.org/images/0014/001474/147473e.pdf>
- Neuman, M. J. and Devercelli, A. E. 2013. *What Matters Most for Early Childhood Development: A Framework Paper*. SABER Working Paper Series, No. 5. Washington, DC, World Bank. http://wbfiles.worldbank.org/documents/hdn/ed/saber/supporting_doc/Background/ECD/Framework_SABER-ECD.pdf
- Opel, A., Ameer, S. and Aboud, F. E. 2009. The effect of preschool dialogic reading on vocabulary among rural Bangladeshi children. *International Journal of Educational Research*, Vol. 48, No. 1, pp. 12–20.
- Opel, A., Zaman, S. S., Khanom, F. and Aboud, F. E. 2012. Evaluation of a mathematics program for preprimary children in rural Bangladesh. *International Journal of Educational Development*, Vol. 32, No. 1, pp. 104–10.
- Organisation for Economic Co-operation and Development (OECD). 2012. *Starting Strong III: A Quality Toolbox for Early Childhood Education and Care*. Paris, Author. <http://www.oecd.org/edu/school/startingstrongiii-aqualitytoolboxforearlychildhoodeducationandcare.htm>
- Organisation for Economic Co-operation and Development (OECD) 2015. *Starting Strong IV: Monitoring Quality in Early Childhood Education and Care*, OECD Publishing, Paris. DOI: <http://dx.doi.org/10.1787/9789264233515-en>
- Orkin, K., Yadete, W. A., & Woodhead, M. 2012. Delivering Quality Early Learning in Low-resource Settings: Progress and Challenges in Ethiopia.
- Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D. and Barbarin, O. 2005. Features of pre-kindergarten programs, classrooms, and teachers: do they predict observed classroom quality and child–teacher interactions? *Applied Developmental Science*, Vol. 9, No. 3, pp. 144–59.
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. 2008. Classroom Assessment Scoring System: Manual Pre-K. Education Review// Reseñas Educativas.
- Rao, N., Koong, M., Kwong, M. and Wong, M. 2003. Predictors of preschool process quality in a Chinese context. *Early Childhood Research Quarterly*, Vol. 18, No. 3, pp. 331–50.
- Rao, N., & Li, H. 2009. Quality matters: Early childhood education policy in Hong Kong. *Early Child Development and Care*, Vol. 179, No. 3, pp. 233–245.
- Rao, N. and Sun, J. 2012. Quality early childhood care and education in low resource level countries in Asia. Chapter prepared for *Early Childhood Care and Education: Building the Wealth of Nations*. UNESCO, Paris.
- Rao, N. and Sun, J. 2015. Quality early childhood care and education in low-resource level countries in Asia. P. T. M. Marope and Y. Kaga (eds), *Investing Against Evidence: The Global State of Early Childhood Care and Education*. Paris, UNESCO, pp. 211–30. <http://unesdoc.unesco.org/images/0023/002335/233558E.pdf>
- Rao, N., Sun, J., Chen, E. E. and Ip, P. 2015. Effectiveness of interventions to promote early childhood development in developing countries: a systematic review and meta-analysis. *Hong Kong Journal of Paediatrics*. In press.
- Rao, N., Sun, J., Pearson, V., Pearson, E., Liu, H., Conostas, M. A., Engle, P. L. 2012. Is something better than nothing? An evaluation of early childhood programs in Cambodia. *Child Development*, Vol. 83, No. 3, pp. 864–76.
- Rao, N., Sun, J., Zhou, J. and Zhang L. 2012. Early achievement in rural China: the role of preschool experience. *Early Childhood Research Quarterly*, Vol. 27, pp. 66–76.

- Reynolds, A. J., Ou, S. R., & Topitzes, J. W. 2004. Paths of effects of early childhood intervention on educational attainment and delinquency: A confirmatory analysis of the Chicago Child Parent Centers. *Child development*, Vol. 75, No. 5, pp. 1299-1328.
- Sheridan, S. 2007. Dimensions of pedagogical quality in preschool. *International Journal of Early Years Education*, Vol. 15, No. 2, pp. 197–217.
- Sheridan S., Giota, J., Han, Y. M. and Kwon, J. Y. 2009. A cross-cultural study of preschool quality in South Korea and Sweden: ECERS evaluations. *Early Childhood Research Quarterly*, Vol. 24, No. 2, pp. 142–56.
- Sheridan S. and Samuelsson, I. P. 2001. Children's conceptions of participation and influence in pre-school: a perspective on pedagogical quality. *Contemporary Issues in Early Childhood*, Vol. 2, No. 2, pp. 169–89.
- Siraj-Blatchford, I. and Wong, Y. L. 1999. Defining and evaluating 'quality' early childhood education in an international context: dilemmas and possibilities. *Early Years: An International Research Journal*, Vol. 20, No. 1: pp. 7–18.
- Storch, S. A. and Whitehurst, G. J. 2002. Oral language and code-related precursors to reading: evidence from a longitudinal structural model. *Developmental Psychology*, Vol. 38, No. 6, pp. 934–47.
- Streuli, N. 2012. *Early Childhood Care and Education in Peru: Evidence from Young Lives*. Young Lives Policy Brief 18. Oxford, UK, Young Lives. <http://www.younglives.org.uk/sites/www.younglives.org.uk/files/YL-PolicyBrief-18.pdf>
- Sylvia, K., Siraj-Blatchford, I., Taggart, B. 2003. Assessing Quality in the Early Years: Early Childhood Environment Rating Scales – Extension (ECERS-E): Four Curricular Subscales. London, Trentham.
- Tobin, J., Hsueh, Y. and Karasawa, M. 2009. *Preschool in Three Cultures Revisited: China, Japan, and the United States*. Chicago, University of Chicago Press.
- Torquati, J. C., Raikes, H. H., Huddleston-Casas, C. A., Bovaird, J. A., & Harris, B. A. 2011. Family income, parent education, and perceived constraints as predictors of observed program quality and parent rated program quality. *Early Childhood Research Quarterly*, Vol. 26, No. 4, pp. 453-464.
- UNESCO. 2000. *The Dakar Framework for Action. Education for All: Meeting our Collective Commitments*. Adopted by the World Education Forum, Dakar, Senegal, 26–28 April 2000. Paris, Author. <http://unesdoc.unesco.org/images/0012/001211/121147e.pdf>
- UNESCO. 2007. *EFA Global Monitoring Report 2007. Strong Foundations: Early Childhood Care and Education*. Paris, Author. <http://unesdoc.unesco.org/images/0014/001477/147794e.pdf>
- UNESCO. 2012. *EFA Global Monitoring Report 2012. Youth and Skills: Putting Education to Work*. Paris, Author. <http://unesdoc.unesco.org/images/0021/002180/218003e.pdf>
- UNESCO. 2014. *Holistic Early Childhood Development Index (HECDI) Framework: A Technical Guide*. Paris, Author.
- UNESCO. 2015. *EFA Global Monitoring Report 2015 Education for All*. Paris, Author. <http://unesdoc.unesco.org/images/0023/002322/232205e.pdf>
- UNICEF. 2012. *A Framework and Tool Box for Monitoring and Improving Quality*. Draft. New York, Author. http://www.unicef.org/ceecis/ECD_Framework_PART_II_june3.pdf
- Unit, E. I., & Britain, G. 2012. *Starting well: Benchmarking early education across the world*. Economist Intelligence Unit.
- United Nations. 1990. *Convention on the Rights of the Child*. Geneva, Switzerland, UN Office of the High Commissioner for Human Rights. <http://www.ohchr.org/Documents/ProfessionalInterest/crc.pdf>
- Vandell, D. L. and Wolfe, B. 2000. *Child Care Quality: Does It Matter and Does It Need to be Improved?* Washington, DC, Office of the Assistant Secretary for Planning and Evaluation, US Department of Health and Human Services. <http://aspe.hhs.gov/hsp/ccquality00/>

- Vandenbroeck, M., De Visscher, S., Van Nuffel, K., & Ferla, J. 2008. Mothers' search for infant child care: The dynamic relationship between availability and desirability in a continental European welfare state. *Early Childhood Research Quarterly*, Vol. 23, No. 2, pp. 245-258.
- Vandenbroeck, M., Roets, G., & Roose, R. 2012. Why the evidence-based paradigm in early childhood education and care is anything but evident. *European Early Childhood Education Research Journal*, Vol. 20, No. 4, pp. 537-552.
- Walker, S., Wachs, T. D., Grantham-McGregor, S., Black, M. M., Nelson, C. A., Huffman, S. L. ... Richter, L. 2011. Inequality in early childhood: risk and protective factors for early child development. *The Lancet*, Vol. 378, No. 9799, pp. 1325-38.
- Wong, H. L., Luo, R., Zhang, L. and Rozelle, S. 2013. The impact of vouchers on preschool attendance and elementary school readiness: a randomized controlled trial in rural China. *Economics of Education Review*, Vol. 35, pp. 53-65.
- Yamamoto, Y. and Li, J. 2012. What makes a high-quality preschool? Similarities and differences between Chinese immigrant and European American parents' views. *Early Childhood Research Quarterly*, Vol. 27, No. 2, pp. 306-15.
- Yoshikawa, H., Leyva, D., Snow, C. E., Treviño, E., Barata, M., Weiland, C., ... & Arbour, M. C. 2015. Experimental impacts of a teacher professional development program in Chile on preschool classroom quality and child outcomes. *Developmental psychology*, Vol. 51, No. 3, pp. 309.
- Zaslow, M., Anderson, R., Redd, Z., Wessel, J., Daneri, P., Green, K., ... & Martinez Beck, I. 2016. Quality Thresholds, Features, And Dosage In Early Care And Education: Introduction And Literature Review. *Monographs of the Society for Research in Child Development*, Vol. 81, No. 2, pp. 7-26.



The importance of early childhood development and learning for later school achievement, health and well-being is well established. The role of early childhood development and learning in reaching global development goals in education, health, nutrition, protection, poverty reduction and gender equity is now clearly recognized in the 2030 Agenda for Sustainable Development. The measurement of early childhood development and learning and reliable data are key for sustaining policy focus and increased funding. The Measuring Early Learning and Quality Outcomes (MELQO) initiative is a response to the current limited information about the state of young children's development and learning experiences prior to primary school. As a joint effort of UNESCO, UNICEF, the World Bank and Brookings Institution, the MELQO initiative aims to address the need for population-based measurement at the country level and to produce measurement tools for both child development and learning and quality of learning environments. This report strengthens the rationale for the development of an integrated global measurement tool for early childhood development and learning. It outlines the methodological approach used to develop the MELQO modules, the different stages of the adaptation process, and examines the relevance of assessment results for policy change.

