A Toolkit for Measuring Early Childhood Development in Low- and Middle-Income Countries

LIA C. H. FERNALD
ELIZABETH PRADO
PATRICIA KARIGER
ABBIE RAIKES

PREPARED FOR THE STRATEGIC IMPACT EVALUATION FUND, THE WORLD BANK
Early Child Development (ECD) Globally

• Estimated 250 million children at risk for not fulfilling their developmental potential

• 193 members of the United Nations have adopted Sustainable Development Goal (SDG) 4.2 to “ensure that all girls and boys have access to quality early childhood development (ECD) so that they are ready for primary education”

→ Increasing demand for ECD assessments
Purpose of the Toolkit

• Provide a resource
  o To researchers, evaluators, and program personnel from varied backgrounds and disciplines
  o For assessing ECD in low- and middle-income countries

• Produce reliable, actionable data on child development

• To inform policy and practice
Why Measure Early Childhood Development?
Three Primary Purposes

<table>
<thead>
<tr>
<th></th>
<th>GLOBAL OR NATIONAL POPULATION MONITORING</th>
<th>PROGRAM EVALUATION</th>
<th>HYPOTHESIS-DRIVEN OR EXPLORATORY RESEARCH</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td><strong>Goal</strong>: Detecting broad trends in child development to inform policy</td>
<td><strong>Goal</strong>: Demonstrating impacts of specific programs or policies</td>
<td><strong>Goal</strong>: Exploring a range of impacts on child development in line with theory and existing understanding of neural mechanisms</td>
</tr>
</tbody>
</table>

Screening tools can also be used to identify children who may need further testing, diagnosis, and treatment.
Domains of ECD: What Can Be Measured?

• Cognitive Skills
• Language Skills
• Motor Skills
• Executive Function/Self-Regulation/Effortful Control
• Temperament
• Social-Emotional Skills
• Personal-Social/Adaptive Skills
• Pre- and Early-Academic Skills
• Approaches to Learning
Where Standard Tests Do No Exist

**TRANSLATION**
Translating items, with no alteration of concepts or pictures used in original test

**ADAPTATION**
Translating items and then changing words or pictures to reflect cultural differences

**EXPANSION**
Adding items to top or bottom of scale or adding items to represent context-specific constructs

**INNOVATION**
Developing a new test or new test items and new methods to examine constructs
How to Decide?

- **Purpose of Assessment**
- **Underlying Construct/Ability**
- **Method**
- **Implementation**

**Data:**
- Item Score
- Subscale Scores
- Composite Scores
- Risk/Delay

Conclusions depend on how well the score reflects the underlying construct/ability.
Bayley Motor Scales in Germany and Cameroon

Pattern Reproduction in Zambia

Participants may perform poorly because they are unfamiliar with the materials

Mental Rotation in Indonesia

Participants may perform poorly because they don’t understand the instructions.

The Ideal Test Does Not Yet Exist!

Ideal Traits of an ECD Assessment

• **Ideal 2:** The test is appropriate, interpretable, and has high reliability and validity in all contexts and cultures.
• **Ideal 4:** The test is easy to administer.
• **Ideal 5:** The test can be administered quickly and at low cost.
• **Ideal 8:** The test is a good indicator of future success.
• **Ideal 9:** The neural mechanisms underlying test performance are well-understood.
• **Ideal 10:** The impact of health, nutrition, and environmental factors on the test score is well-understood.

➔ Prioritize the most important criteria for your purposes then select a test that meets those criteria
STEP 1: Define purpose of assessment
1. Population/Global or National Monitoring
2. Program Evaluation
3. Hypothesis-Driven or Exploratory Research
4. Screening a child for referral

STEP 2: Determine physiology or behavior
- Physiology
  Identify biological mechanisms of effects on behavior to inform more effective interventions.
- Behavior
  More clearly links to child function in everyday life.

STEP 3: Determine mode of assessment
- Autonomic nervous system
- Brain structure
- Brain function
- Direct child assessment
- Parent/teacher report
- Naturalistic or structured observation

STEP 4: Determine type of assessment
- Screening: Binary score indicating risk of delay
- Ability: Continuous score indicating child’s skill level
- Screening: Binary score indicating risk of delay
- Ability: Continuous score indicating child’s skill level

STEP 5: Determine which assessment to use (examples below)
- Heart rate, respiratory rate, stress (e.g., cortisol, galvanic skin response)
- Structural MRI
- Functional MRI, fNIRS, ERP
- RNDA, GMCD
- MDAT, KDI, BSID, NEPSY, WISC, KABC
- ASQ, PEDS, TQQ
- DMC, MacArthur-Bates CDI
- IEA’s Child Coding System
What’s New:

**ECD Measurement Inventory**

Information listed:
- Domains assessed
- Age range
- Method of administration
- Purpose of assessment
- Origin and locations of use
- Logistics
- Cost

Tools can be filtered based on project-specific criteria
What is New: Expanded Tools and Ages

35 new tools for children age 5-8 years

71 new tools for children age 0-5 years

Previous Toolkit reviewed:
41 assessment tools for children age 0-5 years
Domains Assessed of 147 Tools

- Cognitive
- Language
- Motor
- Socio-Emotional/Temperament
- Attention/Executive Function
- Personal-Social/Adaptive
- Academic/Pre-Academic
- Approaches to Learning
Method of Administration of 147 Tools

- Child Assessment/Self-Report: 80%
- Caregiver/Teacher Report: 20%
- Computer-Administered: 0%
Purpose of Assessment of 147 Tools

Percentage of Tools

- Screening Test: 20%
- Ability Test: 90%
- Population-level: 10%
Example: How to Use the *ECD Measurement Inventory*

A research group wants to assess executive function in children age 2 to 4 years.
What’s New: Step-by-Step Guide to Test Adaptation

• Form a Panel of local professionals who inform the process.
• If necessary, conduct preliminary interviews or focus groups.
• Produce an accurate translation.
• Conduct an iterative series of pilot tests.
• Keep track of changes
Adaptation Example: Visual Search Test in Indonesia

Original Version:

Adapted Version:
Guide to Analyze Pilot Data

• Analyze the pilot data to check
  o The percentage of missing item scores
  o Item variability
  o Expected age-related associations
  o Associations with other variables expected to be related

• Evaluate test-retest reliability
What’s New: Tips for Training and Quality Implementation

• Train more testers than you need and hire the top performers.

• Require testers to pass knowledge-based and practice-based evaluations before being certified to administer tests.

• Require testers to achieve inter-rater agreement above 80 or 90 percent.
What’s New:
Review of Tools Using Rapidly Developing Technologies

• Functional Near Infrared Spectroscopy (fNIRS)
• Event-Related Potentials (ERP)
• Eye-tracking
• Accelerometers
• Language Environment Analysis (LENA™)
• Computerized Testing
ERP and fNIRS

Figure 1 | A photo timeline of the setup of the NIRS equipment and first infant to take part in the study at the MRC Field Station in Keneba.

Wearable Devices

Accelerometers: Physical Activity

Language Environment Analysis (LENA™)
Computerized Tests
What’s New:
Review of Tools to Measure the Home Environment and Preschool Quality

• Home Observation for the Measurement of the Environment (HOME) Inventory
  ○ Other related tools
• Parent-Child Interactions
• Measures of Early Learning Environments
What’s New: Review of Predictive Validity of ECD Domains to Forecast Future Performance

• Continuity and plasticity
What to Measure to Forecast Future Performance

<table>
<thead>
<tr>
<th>Early Life Predictor</th>
<th>Outcome in Later Life</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pre-academic and cognitive skills, such as language, general knowledge, and executive function</td>
<td>Academic achievement</td>
</tr>
<tr>
<td>Early development of social-emotional skills and self-regulation</td>
<td>Social-emotional and behavioral function</td>
</tr>
</tbody>
</table>
Key Recommendations 1-3

• Recommendation 1. *Decide on the type of outcome measure that is appropriate.*

• Recommendation 2. *Consider the cultural context and how it may affect children’s development and school readiness.*

• Recommendation 3. *Collect and evaluate pilot data to assess the properties of adapted tests.*
Key Recommendations 4-6

Recommendation 4. Look for national-level tests where possible and use parent or teacher report when large-scale assessments are needed.

Recommendation 5. To assess an indicator of future success, assess children at age three to five years.

Recommendation 6. Include assessments of home and early learning environments.
Recommendation 7. If possible, rely upon multiple measures of children’s development.

Recommendation 8. If possible, use computerized tests (administered by laptop, tablet, or smartphone).

Recommendation 9. For program evaluations, assess characteristics of the child that the intervention is intending to affect and dimensions of a child’s development that you expect to be affected at the target age.

Recommendation 10. In program evaluations, include the same assessments at both baseline and endline.
A Toolkit for Measuring Early Childhood Development in Low- and Middle-Income Countries

Lia C. H. Fernald, Elizabeth Prado, Patricia Kariger, Abbie Raikes

WWW.WORLDBANK.ORG/SIEF