Population-level assessments for 0-3

Maureen Black, Ph.D.
University of Maryland School of Medicine
RTI International
on behalf of GSED Team
Background

- Early childhood development: priority for many countries
- Reporting on childhood development: Sustainable Development Goals

- 147 measures of child development
- Lack of validated **metric** or **measure** at population level for children 0-3
Two Objectives

• Describe a global metric of early childhood development:
  • D-Score (Development Score)
  • DAZ (age-adjusted Z-score)

• Describe a global measure of childhood development for children 0-3:
  • Global Scale of Early Development (GSED)
Global Scale of Early Development (GSED)

- **CREDI**: Caregiver Reported Early Development Instrument
- **IYCD**: Infant & Young Child Development
- **D-Score**: Developmental Score

Global Child Development Group
Metric of Early Childhood Development

- Underlying latent construct, invariant across countries
- Predictable sequence representing domains of motor, language, cognition, personal-social development
- Interval scale (e.g., centimeters: metric for height)
- Interpretable
- Development Score (D-score)
- DAZ (age-adjusted z-score): allows comparability across ages

DAZ: a metric for child development, as HAZ (height-for-age z-score): metric for height.

Shonkoff & Phillips, 2000; Sameroff, 2009
Example of a D-score reference chart

Data from existing longitudinal studies

16 cohorts in 11 countries
> 36,000 children, 1,339 items

• Africa
  – Ethiopia
  – Madagascar
  – South Africa
• Americas
  – Brazil (2)
  – Chile (2)
  – Colombia (2)
  – Ecuador
  – Jamaica (2)
• Asia
  – Bangladesh
  – China
• Europe:
  – The Netherlands (2)

• Birth cohorts, instrument validation studies, intervention evaluations

Time 1 Measures for ages 0-48 months
• Bayley I, II, III
• Griffiths
• Denver
• Dutch Scale
• Battelle
• Barrera Moncada
• Others
Item level coded (pass =1, no =0)

Time 2 Later measures age 5-18 years
• WPPSI/WAIS
• Ravens
• PPVT
• Others
Instruments use similar items

Mapped equivalent items across different instruments to form

Equate Groups

<table>
<thead>
<tr>
<th>Bayley 3 Item Description</th>
<th>Griffiths Item Description</th>
<th>Mapping Score</th>
<th>Denver Item Description</th>
<th>Mapping Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>Child imitates at least four different repetitive consonant-vowel combinations</td>
<td>Babbled phrases: 4+ syllables</td>
<td>excellent</td>
<td>Child repeats the same syllable 3 or more times, e.g. &quot;Dadada&quot; &quot;Gagaga&quot;?</td>
<td>moderate</td>
</tr>
<tr>
<td>Child uses at least two different words appropriately</td>
<td>Says 2 clear words</td>
<td>very good</td>
<td>Says 2 words</td>
<td>very good</td>
</tr>
<tr>
<td>Child correctly names at least four colors</td>
<td>Knows 6+ colors</td>
<td>moderate</td>
<td>Child names color of 4 blocks</td>
<td>excellent</td>
</tr>
</tbody>
</table>
Example of a successful equate group

Outfit 0.64 (−10.26)
Infit 0.74 (−18.19)

ACTIVE

<table>
<thead>
<tr>
<th>Study</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangladesh</td>
<td>use a short sentence</td>
</tr>
<tr>
<td>Chile 1</td>
<td>Says sentences with 2 words</td>
</tr>
<tr>
<td>Columbia 2</td>
<td>Uses word combinations</td>
</tr>
<tr>
<td>Jamaica 1</td>
<td>Combine Words</td>
</tr>
<tr>
<td>Jamaica 2</td>
<td>sentence of 2 words</td>
</tr>
<tr>
<td>Netherlands 1</td>
<td>Uses a two-word utterance</td>
</tr>
<tr>
<td>Netherlands 2</td>
<td>EXP26</td>
</tr>
</tbody>
</table>
Equate group – variable item difficulty

<table>
<thead>
<tr>
<th>Item Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>n43</td>
<td>Throws ball without falling down</td>
</tr>
<tr>
<td>ge2_14</td>
<td>(Eye and hand) Can throw a ball</td>
</tr>
<tr>
<td>dg23</td>
<td>Throw Ball Overhand</td>
</tr>
<tr>
<td>b3g44</td>
<td>Throws ball</td>
</tr>
<tr>
<td>b2p64</td>
<td>Throws ball</td>
</tr>
<tr>
<td>229</td>
<td>GM44</td>
</tr>
</tbody>
</table>

Outfit 1.45(6.31)
Infit 0.94(-3.37)
Model estimation

- **Rasch Model**: Probability of passing item function of difference between child ability and item difficulty
- **Equate groups**: same-skill items from different instruments same difficulty level. Used to connect instruments to common scale
- **Iterative approach**: i) active equate groups, ii) cut points for acceptable fit to the model, iii) subject matter experts
- **Final model**: retained items in active equate groups and individual items with infit and outfit statistics < 1
Distribution of D-score by age & cohort
Association of DAZ and original measures in children < 48 months with IQ at 4.5 to 9 years

- Predictive validity comparable to predictive validity of original instruments
- Fewer items!
Association of DAZ and original measures in children < 48 months with IQ at 9.5 to 18 years
Summary

• D-score (fewer items than original assessment): feasibility of a relatively **short instrument** while maintaining validity.

• **Interval-scale property** of the D-score: quantitative comparisons across ages.

• Estimation of **D-scores for pre-existing studies**, enabling external validation.
GSED Criteria

- Reliable
- Valid (construct, content, convergent, discriminant, predictive) across different contexts (e.g. culture, income)
- Relatively easy to administer with few materials
- Items culturally neutral or easily modifiable for cultural differences
- Interpretable by local personnel, including policymakers
GSED Criteria

• 2 instruments:
  • Population: Caregiver report, surveys
  • Program Evaluation: Direct & caregiver report
• Not designed for individual screening
• Tablet administration
• Administered by local personnel with training & supervision
• Adaptive testing
• Open access, minimal cost (materials)
Early Child Development Measurement

Individual Measures of Child Development
N=147

Global Scale of Early Development (GSED, D-score)

WHO, BMGF
Early Child Development Interventions

- Intensive
- Individual
- Community (Programs)
- Universal (Policies)

- Children with special needs
- At risk community (high rate of poverty)
- Population Equity
Measurement Precision

- High Precision
- Modest Precision
- Modest Precision

Individual Measures of Child Development
N=147

Global Scale of Early Development (GSED, D-score)

WHO, BMGF
Global Scale of Early Development (GSED)

• Population (short, caregiver-report)
  • Guide country-level planning & investment
  • Monitor country-level progress

• Programmatic (long, direct & caregiver)
  • Plan community programming
  • Monitor & evaluate community programming

• Psychosocial (caregiver-report)

• Individual
  • Identify individual child needs & plan intervention
  • Evaluate individual child progress
Process (replicated)

- Formation and harmonization of item bank
- Determine item difficulty
- Iterative statistical/subject matter expert matching and selection of items
- Evaluation using Item Response Theory:
  - 1PL (Rasch): items in equate group constrained to equal difficulty
  - 2PL: item difficulty not constrained
- 1 PL Rasch model retained, yields D-score
Process

• Organized items and instructions into instruments
• Close collaboration with Advisory Group
• Align with other initiatives
• Field sites selected (Bangladesh, Pakistan, Tanzania) and field test planned:
  • Feasibility (timing, training, fidelity, etc.)
  • Psychometrics (reliability and validity, including 6-month predictive validity)
  • Acceptance
• Link with policy and SDGs
Issues and Next Steps

• Integration of psychosocial variables?
• Implementation/utilization within & across countries?
• Extension beyond age 3?
• Use of tablet administration for children?
• Incorporate other components of child development, i.e., self-regulation, executive functioning?
• Develop standards based on “healthy” communities as per the growth standards based on the Multi-country Growth Reference Study?
Thank You!!

Donors

Bill & Melinda Gates Foundation
Grand Challenges Canada
Bernard van Leer Foundation
Hilton Foundation