Assessing Cognitive Processes at the Core of Reading Comprehension: Pre-Kindergarten through Third Grade for the General English Speaking Population

Presented at the 2012 Annual Meeting of the National Council on Measurement in Education, April 12-16, Vancouver, British Columbia

Goals of Our Presentation

To provide:

• An overview of the conceptual framework behind LARRC
  — Research questions
  — Sample Design
  — Preliminary sample demographics
  — Measures
  — Data collection methodologies and cross-site quality control
• Summarize and review challenges for our first year of RFU

Series of Connected Studies

• Study 1: Research study on the cognitive, language, and environmental processes underlying reading comprehension and its development across pre-K through Grade 3
• Study 2: Define and develop 2 instantiations of a language-based reading comprehension intervention for Pre-K through Grade 3
• Study 3: Efficacy trials of our intervention through a randomized control trial

All studies include a complementary ELL component. Our presentation is on the General (non-ELL) sample.
Study 1 Focus

• Cross-sectional and longitudinal
  – Identify grade-related differences in the relative contribution of lower- and higher-level language skills to children’s listening and reading comprehension in PK through G3
  – Determine the nature of the prospective relationship between children’s lower- and higher-level language skills in the early grades and reading comprehension in G3
  – Which child- and environmental-level attributes are positive, prospective predictors of reading comprehension; do these moderate the relations between language skills and reading comprehension?

Multisite School and Teacher Selection

• 69 schools/centers recruited across 4 data collection sites
  – 29 were Pre-school only schools/centers, 4 had both PK and K, and 4 were used for K only
  – 32 schools were used for range of grades PK to G3
  – 73% suburban; 22% urban schools/centers
• 265 teachers
  – 98.5% of recruited teachers are female
  – 70% of PK teachers reported (only half-day programs) at their center/school (22.5% both half and full-day programs)
  – Only 2.2% of K teachers reported half-day programs; 2.2% both.
  – For Primary teachers: 12.3% have no college degree; 32.2% have Associates; 39.5% have Bachelors degree
  – For the Elementary teachers (including K), 34% have BA/BS; 58% with a Masters

Data Collection Methodology

• Initial screeners for Parents were used to exclude children with severe/profound developmental disability (ELL children assessed in ELL study)
• Teacher Background Questionnaire and Parent Background Questionnaire for capturing classroom, home environment characteristics
• Annual assessments each Spring, following children up to G3
• Assessments divided into 9 blocks of 25-35 minutes each
  – Total Assessment time: approximately 4.5 hours
• Classroom observations: for the 400 children in the PK entry sample, classroom observations are collected longitudinally until G3
  – CLOP, CLASS, ISI

Study 1 Sample per Site, and Total

Chart showing number of children assessed/followed up per year, per site.

<table>
<thead>
<tr>
<th>GRADE</th>
<th>PK</th>
<th>K</th>
<th>G1</th>
<th>G2</th>
<th>G3</th>
<th>N (Total)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Year 1</td>
<td>PK = 100</td>
<td>K = 30</td>
<td>G1 = 30</td>
<td>G2 = 30</td>
<td>G3 = 30</td>
<td>230</td>
</tr>
<tr>
<td>Year 2</td>
<td>K = 100</td>
<td>G1 = 30</td>
<td>G2 = 30</td>
<td>G3 = 30</td>
<td>190</td>
<td>760</td>
</tr>
<tr>
<td>Year 3</td>
<td>G1 = 100</td>
<td>G2 = 30</td>
<td>G3 = 30</td>
<td>160</td>
<td>640</td>
<td></td>
</tr>
<tr>
<td>Year 4</td>
<td>G2 = 100</td>
<td>G3 = 30</td>
<td>130</td>
<td>520</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Year 5</td>
<td>G3 = 100</td>
<td>100</td>
<td>480</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Sampling Methods for Child Sample

• Schools selected for willingness to work with researchers; many already had an established collaboration
• Sites looked for diversity in school/center size, ethnic/racial composition, poverty rates (FRL)
• After initial principal meeting:
  – All teachers invited; info sent home to parents of consenting teachers
  – Goal was 30 K-G3 children per grade per site, and 100 PK per site
  – random sample taken from those families who returned consents
• Primary and secondary samples (K-G3: 30 primary, 20 secondary)
Age (years) on September 1, 2010

<table>
<thead>
<tr>
<th>Grade</th>
<th>Pre-K</th>
<th>K</th>
<th>G1</th>
<th>G2</th>
<th>G3</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean</td>
<td>4.55</td>
<td>5.47</td>
<td>6.56</td>
<td>7.53</td>
<td>8.58</td>
</tr>
<tr>
<td>SD</td>
<td>.36</td>
<td>.40</td>
<td>.34</td>
<td>.34</td>
<td>.38</td>
</tr>
</tbody>
</table>

Race/Ethnicity and Language at Home

- 78% of families reported speaking primarily English at home
  - Rest indicated Spanish, Chinese, Amharic, Vietnamese, other (or missing)

9.5% of students in sample are Hispanic
88.2% White/Caucasian
7.5% Black
5.0% Asian
2.5% other

SES and Free/Reduced Lunch

- 38% of families indicated making more than $85K per year
- 77% of families have 2 or more children
- 70% reported both parents living with child
- 15.6% of sampled children receive free or reduced lunch

Parent Ratings: SWAN, ODD

Parent ratings on SWAN: Attention, Impulse Control, and Oppositional Defiance Scale

- Lower score implies poor attention, strong hyperactivity or impulse, and stronger oppositional defiance, relative to other children.

Match with General US Population

- Our General English sample is racially and ethnically diverse, and includes children with IEPs (9.6%) and from families below poverty level (15.6% receiving free or reduced lunch).
- Schools across the four data-collection sites capture between-school differences representing a range of schools in the US.

Measures Selection

- Three guiding principles for selections
  - Multiple constructs, enable use as latent variables
    - Important for reading comprehension
    - Need for different modalities of measurement
    - Need for variability in task demand.
  - Battery developmentally appropriate in content and length, still able to address multidimensional study aims
  - Measures skills that span pre-reading to conventional reading
Measures: Experimental Measures

- Generalizable versus Experimental Measures
- Limited availability of quality assessments of certain study constructs, specifically higher-level language skills
  - Detecting inconsistent information in short narratives
  - Tests of knowledge of narrative structure

Measures: Psychometric Properties

- Process
  - Reliability
  - Validity
  - Reliability assessments for 70% of measures done to date
  - Process is on-going

Measures: Organization/Administration

- Codebook
  - For each measure, a codebook entry was created to provide detailed records of the measures' psychometric qualities, as well as any modifications made to the measure for project purposes.
- Training of assessors
  - Assessors-in-training undergo comprehensive measurement training and in-lab observations to ensure consistent field assessor training, measurement administration, and fidelity.
- Scoring protocols
  - For those measures scored during administration, assessors are trained in scoring procedures as part of the training protocols.

Cross-Site Data Collection: TeleForm

- Teleform
  - After data is collected at each site, it is entered into a general database via Teleform© scans.
  - In order to capture high quality data, all hard-copies of data being entered into the database via Teleform© are thoroughly checked to eliminate any errors prior to scanning.
Cross-Site Data Collection

<table>
<thead>
<tr>
<th>Task</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Teleform Scan Verification</td>
<td>Identify errors on scans (e.g., missing values)</td>
</tr>
<tr>
<td>Data Audit</td>
<td>Identify random selection of 10 assessment forms (2 children per grade) for audit</td>
</tr>
<tr>
<td></td>
<td>• Compare item level data in database to hard copy</td>
</tr>
<tr>
<td></td>
<td>• Compare total scores in database to hard copy</td>
</tr>
<tr>
<td></td>
<td>• Compare computed/derived standard scores in assessment manuals to hard copy</td>
</tr>
<tr>
<td></td>
<td>• Compare any hand-written data (e.g., those with 'other' options and constrained number variables such as birth date) on assessment scans to database</td>
</tr>
<tr>
<td>Data Check</td>
<td>Check item-level data and sum scores against plausible ranges</td>
</tr>
<tr>
<td></td>
<td>• Identify missing data (between and within measures)</td>
</tr>
<tr>
<td></td>
<td>• Verify ID's (child, teacher) and check for duplicates</td>
</tr>
</tbody>
</table>

Summary

- Challenges and Considerations
  - Measure selection versus time
  - Data cleaning etc., across site
  - Post scoring and time required (training and scoring)
  - Primary/secondary sample
  - Communications across site
  - Incentives for children to keep them engaged in completing the lengthy battery
  - Treasure maps, encouragement to continue
  - Teacher videotaping concerns
  - Verify our sample characteristics for teachers/families who consented versus those who did not (ongoing)

Contacts

- For more information about this presentation, contact:
  - Ann A. O'Connell (OSU) aoconnell@ehe.osu.edu
  - Jill Pimentoni (OSU) jpimentoni@ehe.osu.edu
- Or, visit our LARRC website: http://LARRC.ehe.osu.edu

Thank you!