The Effects of the Tennessee Voluntary Pre-Kindergarten Program: Initial Results

Mark Lipsey  Dale Farran
Kerry Hofer  Carol Bilbrey  Nianbo Dong
Peabody Research Institute, Vanderbilt University

Society for Research on Educational Effectiveness, March 2011
In partnership with the Tennessee Dept. of Education, Division of School Readiness and Early Learning
Bobbi Lussier, Assistant Commissioner
Connie Casha, Director of Early Childhood Programs
Robert Taylor, Consultant

With the invaluable assistance of our fine research team:
Patricia Abelson          Kurt Scheib
Richard Feldser          Ilknur Sekmen
Janie Hughart            Travis Wimsett
And many more …

Supported by a grant from the U.S. Department of Education, Institute of Education Science (R305E090009)
Tennessee invests over $85 million a year in its statewide voluntary Pre-K program.

934 state-funded Pre-K classrooms serve 18,000+ economically disadvantaged children across all 95 Tennessee counties.

Support for Pre-K is based on the belief that high quality Pre-K:
- Improves at-risk children’s readiness for kindergarten
- Improves achievement test scores and decreases retention rates, special education placements, and drop outs
- In adulthood, increases employment rates and earnings, and reduces welfare needs and criminal behavior
Many Tennessee legislators question the value of Pre-K, some calling it “expensive babysitting,” and whether it warrants funding given the current budget constraints. The evidence for long-term Pre-K effects is from small intensive programs that are not typical of state programs. The national Head Start study provides little support for the view that large scale preschool programs are especially effective for boosting academic performance. No research using a randomized control trial design has studied the benefits of typical public Pre-K programs and their effects beyond the beginning of kindergarten.
Evaluating the Effectiveness of the TN Voluntary Pre-K Program

- 5-year study funded by the Institute of Education Sciences beginning in 2009.
- Conducted by the Peabody Research Institute (PRI) at Vanderbilt in partnership with the TN Dept. of Education Division of School Readiness and Early Learning.
- Objective: To conduct a scientifically rigorous evaluation of the effectiveness of the Tennessee Pre-K program for improving children's readiness for kindergarten and their achievement in later grades.
Two Components of the Study

- Randomized control trial comparing children who attended Pre-K with those on the waiting list who were unable to be admitted
  - All randomly assigned children to be tracked in the state Education Information System through third grade TCAP achievement tests
  - Intensive Substudy: Children assessed at the beginning and end of the pre-k year, and the end of K, 1st, and 2nd grade.

- Regression-discontinuity design for end of Pre-K effects with a statewide representative sample of classrooms

- Currently available and reported here:
  - Results of the Intensive Substudy for the first of two randomization cohorts
  - Results of the regression-discontinuity analysis for the classrooms in the middle TN region (one of 4 being sampled)
Phase 1 RCT Intensive Substudy

- Phase 1 in 2009-10 school year (Phase 2 underway for 2010-11 school year)
- Randomized admissions in 23 schools in 14 TN school districts
- 907 children in full randomization (will be tracked in state EIS database)
- 303 consented children with assessment data, 73 no pre-k controls and 230 pre-k
Intensive Substudy Sample

- Mean age, 4.4 yrs; 56% girls, 44% boys
- 57% white, 23% African-American, 20% Hispanic
- 31% language other than English in home
- Median parent education: High school/GED
- No Pre-K control childcare alternatives:
  - 11% Head Start
  - 22% Private childcare center
  - 51% Home with parent or other
  - 16% Unknown
**Intensive Substudy Outcome Measures**

**Woodcock Johnson III Scales**
- **Literacy**: Letter-Word Identification, Spelling
- **Language**: Picture Vocabulary, Oral Comprehension
- **Math**: Applied Problems, Quantitative Concepts
- **Overall WJ Composite**: Mean W score across all WJ scales (supported by factor analysis)

**Kindergarten Teacher Ratings**
- Cooper-Farran Work-Related & Social scales
- Academic Child Behavior Record: School Readiness, Liking for School, & Behavior Problems
Complications, Part 1

- Indirect process required to approach parents for consent to assess children resulted in low and variable consent rate
  - 37% overall across the 23 schools
  - Ranged from 7.5% to 82.6%
  - Note: Changed procedure for 2010-11 has produced higher overall rate for Cohort 2

- Response: Use the consent rate and its interaction with treatment conditions as a covariate in the analysis
Complications, Part 2

- **Some baseline differences between conditions**
  - Pre-K children: more African-Americans, lower mean parent education
  - Pretest averaged 9 days later for no Pre-K control children; pre-post test interval average of 6.5 days shorter
  - No significant ($p<.10$) differences on age, gender, proportion Hispanic, or parent report of weekend TV, home literacy support, or whether language other than English spoken in home (correlates of WJ outcomes)

- **Response:** Use all these variables as covariates in the analysis to ensure no associated bias
Complications Continued: Pretest Differences Despite Prior Covariates

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Control Mean (N=73)</th>
<th>Pre-K Mean (N=230)</th>
<th>Difference as an Effect Size</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Composite</td>
<td>391.0</td>
<td>396.0</td>
<td>.30*</td>
<td>.000</td>
</tr>
<tr>
<td>Letter-Word ID</td>
<td>313.6</td>
<td>320.6</td>
<td>.28*</td>
<td>.003</td>
</tr>
<tr>
<td>Spelling</td>
<td>350.1</td>
<td>355.9</td>
<td>.23*</td>
<td>.008</td>
</tr>
<tr>
<td>Picture Vocabulary</td>
<td>449.0</td>
<td>454.9</td>
<td>.27*</td>
<td>.002</td>
</tr>
<tr>
<td>Oral Comprehension</td>
<td>442.5</td>
<td>442.8</td>
<td>.02</td>
<td>.583</td>
</tr>
<tr>
<td>Applied Problems</td>
<td>384.4</td>
<td>393.9</td>
<td>.39*</td>
<td>.000</td>
</tr>
<tr>
<td>Quantitative Concepts</td>
<td>406.3</td>
<td>408.2</td>
<td>.13</td>
<td>.071</td>
</tr>
</tbody>
</table>

- **Response**: Use the respective pretest as a covariate in each analysis.
Analysis

- W-scores from WJIII scales
- Teacher rating scales
- Multilevel with children nested within schools/randomized lists (blocking factor)
- Covariates for all analyses: age, gender, race/ethnicity, home literacy, weekend TV, parent education, home language, response rate, response rate x condition
- Additional covariates for WJ outcomes: pretest, lag to pretest, pretest-posttest interval
Pre-Post Gain for WJ Composite Scale with Initial Differences Controlled

-0.2
-0.4
-0.6
-0.8
-1.0
-1.2

0.0
0.2
0.4
0.6
0.8
1.0
1.2

Pre-K
No Pre-K

.35 SD
(50% increase)

.69 SD

Pretest

Posttest

Standard Deviations
### Summary: Pre-Post Gains and Effect Sizes for WJIII Scales

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Control Gain in SD Units</th>
<th>Pre-K Gain in SD Units</th>
<th>Difference (Effect Size)</th>
<th>% Improvement for Pre-K</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Overall Composite</strong></td>
<td>.69</td>
<td>1.03</td>
<td>.35*</td>
<td>50%</td>
</tr>
<tr>
<td><strong>Literacy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letter-Word ID</td>
<td>.51</td>
<td>1.00</td>
<td>.49*</td>
<td>96%</td>
</tr>
<tr>
<td>Spelling</td>
<td>.70</td>
<td>1.01</td>
<td>.31*</td>
<td>45%</td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture Vocabulary</td>
<td>.18</td>
<td>.43</td>
<td>.25*</td>
<td>141%</td>
</tr>
<tr>
<td>Oral Comprehension</td>
<td>.28</td>
<td>.58</td>
<td>.31*</td>
<td>110%</td>
</tr>
<tr>
<td><strong>Math</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Applied Problems</td>
<td>.69</td>
<td>.91</td>
<td>.22*</td>
<td>32%</td>
</tr>
<tr>
<td>Quantitative Concepts</td>
<td>.64</td>
<td>1.05</td>
<td>.41*</td>
<td>63%</td>
</tr>
</tbody>
</table>

Note: All the differences between the Control and Pre-K children are statistically significant, $p<.05$. 
### Effect Sizes for Teacher Ratings

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Effect Size</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cooper-Farran Work Related</td>
<td>.29*</td>
<td>.027</td>
</tr>
<tr>
<td>Cooper-Farran Social Behavior</td>
<td>-.04</td>
<td>.791</td>
</tr>
<tr>
<td>ACBR Readiness for Kindergarten</td>
<td>.36*</td>
<td>.006</td>
</tr>
<tr>
<td>ACBR Likes School</td>
<td>-.09</td>
<td>.653</td>
</tr>
<tr>
<td>ACBR Behavior Problems No/Yes</td>
<td>.06</td>
<td>.468</td>
</tr>
<tr>
<td>ACBR Number of Behavior Problems</td>
<td>.00</td>
<td>.939</td>
</tr>
</tbody>
</table>

Note: Based on ratings by 19 teachers at 19 of the 23 schools and 203 children rated (133 T and 70 C).
Regression-Discontinuity Substudy: Middle Tennessee Region

- 36 schools in 17 middle Tennessee school districts
- 682 children who attended Pre-K during the 2009-10 school year
- 676 children who were below the age cutoff and attended Pre-K during the 2010-11 school year
- All children assessed early in the fall of 2010
Regression-Discontinuity Sample

- **Mean age at time of assessment**
  - Beginning of Pre-K control sample: 4.4 yrs
  - Beginning of K treatment sample: 5.4 yrs
- **50% boys; 50% girls**
- **52% white, 35% African-American, 13% Hispanic**
- **12% native language other than English**
- **Urban and rural schools**
RD Outcome Measures

Woodcock Johnson III Scales

- **Literacy**
  - Letter-Word Identification
  - Spelling

- **Language**
  - Picture Vocabulary
  - Oral Comprehension

- **Math**
  - Applied Problems
  - Quantitative Concepts

- **Overall WJ Composite**
  - Mean W score across all WJ scales (supported by factor analysis)
Pre-K Age Cutoff RDD: Timing of Outcome Measures

Year 1 (2009-10)  Year 2 (2010-11)

Treatment: First cohort (before cutoff)  Control: Second cohort (after cutoff)

Pre-K (T)  Kindergarten
No Pre-K (C)  Pre-K

Administer Tests
Entry into Pre-K Selected by Birthday

- **Born before October 1**: Completed pre-K; tested at beginning of K

- **Born after October 1**: No Pre-K yet; tested at beginning of pre-K year
Analysis

- W-scores from WJIII scales
- Multilevel with children nested within schools/classrooms
- Covariates to account for functional form: age (centered at cutpoint), age x condition, age squared, age cubed
- Additional covariates: gender, race/ethnicity, non-English native language, school start to test lag
- ±3 months around cutpoint (conservative analysis)
  - 170 control cases
  - 157 treatment cases
RD Effect Size Relative to Pre-K to K Gain for WJ Composite Scale

-0.2
0.0
0.2
0.4
0.6
0.8
1.0
1.2
Pre-K
Kindergarten

Total Year-to-Year
Estimate w/o Pre-K

.40 SD
.61 SD Effect Size
(152% increase)
## RDD Effect Sizes for WJIII Scales (± 3-months around birth date cutoff)

<table>
<thead>
<tr>
<th>Outcome Measure</th>
<th>Est. 1-yr Gain w/o Pre-K in SD Units</th>
<th>1-year Gain with Pre-K in SD Units</th>
<th>Pre-K Effect Size Estimate</th>
<th>% Improvement for Pre-K</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Composite</td>
<td>.40</td>
<td>1.01</td>
<td>.61**</td>
<td>152%</td>
</tr>
<tr>
<td><strong>Literacy</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Letter-Word ID</td>
<td>.35</td>
<td>1.01</td>
<td>.66**</td>
<td>190%</td>
</tr>
<tr>
<td>Spelling</td>
<td>.44</td>
<td>1.13</td>
<td>.69**</td>
<td>159%</td>
</tr>
<tr>
<td><strong>Language</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Picture Vocabulary</td>
<td>.16</td>
<td>.46</td>
<td>.30</td>
<td>193%</td>
</tr>
<tr>
<td>Oral Comprehension</td>
<td>.08</td>
<td>.45</td>
<td>.37*</td>
<td>456% (!)</td>
</tr>
<tr>
<td><strong>Math</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Applied Problems</td>
<td>.42</td>
<td>.76</td>
<td>.34*</td>
<td>81%</td>
</tr>
<tr>
<td>Quantitative Concepts</td>
<td>.40</td>
<td>.78</td>
<td>.39*</td>
<td>97%</td>
</tr>
</tbody>
</table>

* p<.10, ** p<.05.
Conclusions and Final Thoughts

- Policy level research is remarkably difficult to do with rigor
  - Schools prize their independence and are reluctant to change procedures for the sake of research
- Close collaboration with the TN DOE was essential for obtaining the RCT and RDD samples
- Multiple substudies and the ability to ‘triangulate’ results helps compensate for the weaknesses of each individual study
More Concluding Thoughts

- **Important findings so far**
  - Strong effects demonstrated for TN state-funded Pre-K compared to what is otherwise available in the community
  - Same pattern of effects found in both the RCT and the RDD samples lends credibility to the conclusion
  - Teacher ratings corroborate the effect and expand it to include important learning dispositions

- **Future results still to come on this project**
  - Effect estimates from the second cohort RCT sample
  - Grade retention, special education placement, and 3rd grade achievement test outcomes for the full RCT sample
  - Effect estimates from the RDD in the remaining 3 regions of the state, eventually including 140 Pre-K classrooms
Thanks!

Contact information:

mark.lipsey@vanderbilt.edu