
Deborah L. Linebarger, Ph.D.
Children’s Media Lab
Annenberg School for Communication
University of Pennsylvania
April 7, 2009
We would like to thank the talented and dedicated efforts of the staff and students in Mississippi and Pennsylvania who helped with this project including Cathy Grace, Stacey Callender, Nikki McCelleis, Jessamine Huffman, Beverly Willis, and Katie McMenamin.

In addition, we would like to thank the children, families, teachers, directors, and staff at the child care centers where we were fortunate enough to work. Without their time, energy, and enthusiasm, this project would not have been completed.

For additional information, please contact:

Dr. Deborah L. Linebarger  
Director, Children’s Media Lab  
Annenberg School for Communication  
University of Pennsylvania  
3620 Walnut Street  
Philadelphia, PA 19104  
215.898.7855 (lab)  
215.898.2024 (fax)  
Email: rtl@asc.upenn.edu
Introduction

Substantial evidence documents the success of the program *Between the Lions* in helping young children acquire early literacy skills, particularly those young children who may be at-risk for reading failure due to economic disadvantage. Young children who watch this program have demonstrated consistent gains across alphabet knowledge, phonological awareness, phonemic awareness, and fluency through simple exposure (Linebarger, Kosanic, Greenwood, Doku, 2004; Uchikoshi, 2006) and exposure combined with classroom materials (Linebarger, 2006; Prince, Grace, Linebarger, Atkinson, & Huffman, 2001). The study presented here extends the research on the use of *Between the Lions*, supplemental classroom materials, and a mentoring program to a new sample of teachers, classrooms, and children.

- **Hypothesis 1:** The treatment group would outperform the control group at the post-test.
- **Hypothesis 2:** The maintenance group would outperform the control group at the post-test.
- **Hypothesis 3:** There should be no or, at most, a slight difference between the treatment group and the maintenance group.

Intervention

Child care classrooms were categorized into three groups: Treatment, Maintenance, and Control.

- **9 Treatment classrooms** were encouraged to use Between the Lions in their classrooms and each received a set of the Between the Lions Preschool Literacy Initiative classroom materials as well as training and mentoring described below, beginning in November 2007 following the collection of pre-test data. For both teachers and students, this was their first year of participation using Between the Lions.

- **9 Maintenance classrooms** had previously been Treatment classrooms, so these teachers had received the Between the Lions classroom materials, training, and mentoring in a previous school year. The children in these classrooms had not previously been exposed to the intervention. The teachers were encouraged to continue using Between the Lions in their classrooms once the pre-test data was collected. Three of the original maintenance classrooms were not included in this study because in two of the classrooms the original teachers left and the new teachers had not previously been in Treatment classrooms that received training/mentoring. A third Maintenance classroom was pulled from the project because the teacher stopped using the Between the Lions materials and was replaced by another classroom that had not been a Treatment classroom the previous year.

- **13 Control classrooms** were not provided with any Between the Lions materials, training, or mentoring, but could opt to receive the materials and training following the collection of post-test data. They could also be put on the list to be entered into the random selection process to become Treatment classrooms the following year.
Classroom Materials

Treatment and Maintenance classrooms received the following materials:

- **Between the Lions Preschool Literacy Initiative Lesson Plans**: A series of 30 theme-based weekly lessons, organized into 5 units, with ideas and strategies for conducting daily literacy activities. Lessons include ideas for activity centers along with whole- and small-group activities. Additional lesson components include a weekly planner, suggestions for setting up the classroom, A Family Letter, theme-related songs and poems, and a recommended booklist. The lessons follow a scope and sequence designed to address all key early literacy skills and are aligned with state preschool standards and Head Start frameworks.

- **DVDs**: Each unit is accompanied by a DVD with six Between the Lions episodes edited specifically for preschool use, for a total of 30 episodes plus bonus tracks.

- **Books**: Each lesson includes at least two accompanying books; the diverse collection of 61 trade books includes folktales, contemporary stories, rhyming books, alphabet books, concept books, and nonfiction.

- **Song and Poem Charts**: Illustrated song and poem charts linked to the lessons aid in teaching children about concepts of print and the sounds of spoken language.

- **Additional Classroom Resources**: A bin of additional materials includes letter cards, word cards, picture cards, story figures for a Velcro board, magnetic letters, lion puppets, and other basic supplies for use with the lessons.

Training and Mentoring

Immediately after the pre-testing is completed, teachers in the Treatment classrooms participate in five hours of training provided by mentors from Mississippi Public Television. The training provides an overview of the Between the Lions curriculum materials and strategies for setting up their classrooms to encourage literacy. Teachers also gain an understanding of how the mentors will work with them during the intervention. Teachers in the Maintenance classrooms are also invited to attend this training.

Teachers in Treatment classrooms receive two three-hour visits from a mentor every week, for 16 weeks, for a total of 96 hours of mentoring. Teachers in their first year as a Maintenance classroom receive two three-hour visits per month for 16 weeks, for a total of 24 hours of mentoring. Teachers in their second year as a Maintenance classroom receive one three-hour visit per month for 16 weeks, for a total of 12 hours of mentoring. The mentors guide teachers in setting up and organizing their classrooms, model Between the Lions lessons, and provide feedback on how teachers carry out these lessons.

After 16 weeks, the mentors do not return to the classrooms until post-testing has been completed. At the point of post-testing, most teachers have completed 17 to 20 of the 30 lessons. As a result, they have not covered all the letters, and have not devoted as much focus on initial sound fluency or on blending beginning and ending sounds and words, which are the focus of later units. After post-testing, the Treatment and Maintenance classrooms continue using the Between the Lions lessons.
Method

Participants

Table 1 provides detailed information about teacher and child participants.

Teachers and Child Care Centers

Thirty-one teachers and classrooms in 23 different child care centers were recruited to participate in this study. Nearly equal numbers of classrooms were matched across the three conditions using teacher education, classroom age, percentage of subsidies, census data on percent of families living in poverty by zip code, and center size. While there was a systematic attempt to match classrooms across conditions, the results presented in Table 1 indicate that there were differences at the start of the study associated with centers and classrooms.

Dropped Classrooms. Two teachers originally in the Maintenance condition left their respective centers and were not replaced with teachers who were trained to use the intervention materials. A third teacher who was assigned to the control condition was at a center where other classrooms were in the treatment or maintenance groups. There was substantial evidence that this teacher implemented the intervention despite a request not to do so.

Dropped Centers. Two child care centers were dropped from the original sample due to the teacher changes described above.

Children

The original sample consisted of 319 children attending preschools and child care centers in Mississippi. Of this total, 23 children who were in the three classrooms described above were dropped from the final analyses; therefore, the final sample consisted of 296 children (Mean Age = 59.93 months, SD = 16.87 months). Matching by classroom characteristics resulted in 95 children in 9 treatment classrooms, 90 children in 9 maintenance classrooms, and 113 children in 13 control classrooms. Just over half of the children were boys (i.e., 53.3%). Children participated in the assessments at their child care centers. All 21 centers served children who were predominantly from economically disadvantaged backgrounds. No other demographic information was collected.
Table 1. Characteristics of the Sample

<table>
<thead>
<tr>
<th>Attribute</th>
<th>Description</th>
<th>All</th>
<th>Treatment</th>
<th>Maintenance</th>
<th>Control</th>
<th>Initial Group Differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>Original Total</td>
<td>Entire Sample</td>
<td>319</td>
<td>95</td>
<td>113</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>Revised Total¹</td>
<td></td>
<td>296</td>
<td>95</td>
<td>90</td>
<td>111</td>
<td></td>
</tr>
<tr>
<td>Center Characteristics</td>
<td>Total Number</td>
<td>23</td>
<td>7</td>
<td>9</td>
<td>7</td>
<td></td>
</tr>
<tr>
<td></td>
<td>% Subsidy</td>
<td>69.0%</td>
<td>67.9%</td>
<td>86.5%</td>
<td>55.8%</td>
<td>F (2, 293) = 67.19***</td>
</tr>
<tr>
<td>Classroom Characteristics</td>
<td>Total Classrooms</td>
<td>31</td>
<td>9</td>
<td>9</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FOURS</td>
<td>156</td>
<td>53</td>
<td>30</td>
<td>73</td>
<td>F (2, 153) = 0.26, ns</td>
</tr>
<tr>
<td>Mean Age (months)</td>
<td>54.1</td>
<td></td>
<td>54.3</td>
<td>54.4</td>
<td>53.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>FIVES</td>
<td>140</td>
<td>58</td>
<td>65</td>
<td>17</td>
<td>F (2, 137) = 5.86**</td>
</tr>
<tr>
<td>Mean Age (months)</td>
<td>65.6</td>
<td></td>
<td>64.4</td>
<td>66.8</td>
<td>65.4</td>
<td></td>
</tr>
<tr>
<td>Teacher Characteristics</td>
<td>Education</td>
<td>49.3%</td>
<td>57.9%</td>
<td>35.9%</td>
<td>60.4%</td>
<td>X² = 69.77***</td>
</tr>
<tr>
<td></td>
<td>29.1% CDA/AA</td>
<td></td>
<td>13.7% CDA/AA</td>
<td>52.4% CDA/AA</td>
<td>17.1% CDA/AA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>19.9% BA</td>
<td></td>
<td>23.2% BA</td>
<td>13.3% BA</td>
<td>22.5% BA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>1.7% MA</td>
<td></td>
<td>5.3% MA</td>
<td>0 MA</td>
<td>0 MA</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BTL Lessons Completed</td>
<td>13.18</td>
<td>18.35</td>
<td>23.98</td>
<td>0</td>
<td>F (2, 293) = 1075.48***</td>
</tr>
<tr>
<td>Child Characteristics</td>
<td>Gender</td>
<td>53.4%</td>
<td>52.6% boys</td>
<td>47.8% boys</td>
<td>58.6% boys</td>
<td>X² = 2.35, ns</td>
</tr>
<tr>
<td></td>
<td>Child Age</td>
<td>59.6 months</td>
<td>62.8</td>
<td>56.1</td>
<td>59.6</td>
<td>F (2, 293) = 25.97***</td>
</tr>
</tbody>
</table>

¹ Three classrooms and 23 children were dropped from the analyses due to either teacher changes or implementation of the curriculum in a control classroom.
Measures

Measures were selected or developed to assess targeted skills supported through the *Between the Lions Literacy Initiative* and to reflect the key early literacy skills as described by Neuman and Roskos (2005). These skills are language development, letter knowledge, phonemic awareness, and print conventions. Normative, or standardized, measures tapped into each of these domains using multiple indices.

**Demographic Information**

**Children and Families**

Children’s gender and dates of birth were recorded.

**Teachers and Child Care Centers**

Teachers provided information regarding their years of education and any degrees they possessed. Centers indicated the percentage of children and families at each center who received subsidies.

**Indicators of the Classroom Environment**

The classroom literacy environment was examined using the *Early Literacy and Language Classroom Observation Tool (ELLCO)*. The ELLCO measured literacy and language practices and materials in early childhood classrooms across 4 components: the General Classroom Observation and Teacher Interview; the Literacy Environment; the Language, Literacy, and Curriculum Assessment; and the Literacy Activities Rating Scale.

1. **General Classroom Observation and Teacher Interview**: measures organization, contents, technology, and classroom climate and management
2. **Literacy Environment**: measures availability, content, and diversity of reading, writing, and listening materials.
3. **Language, Literacy, and Curriculum Assessment**: measures reading and writing instruction, oral language use, cultural sensitivity, and assessment approaches
4. **Literacy Activities Rating Scale**: measures how many times and for how long nine literacy behaviors occurred in two categories, Book Reading and Writing

**Indicators of Language Development**

**IGDI Picture Naming Task**

Generalized vocabulary knowledge was evaluated using the *Picture Naming Task*, a tool that measured children’s expressive language knowledge (PNT, Missall & McConnell, 2004). The PNT is an Individual Growth and Development Indicator (IGDI) used to track preschoolers’ vocabulary acquisition on a
regular basis over time. Children were presented with images of objects familiar to preschoolers one at a time and asked to name the pictures as fast as possible for one minute. Categories of objects used included animals, food, people, household objects, games and sports materials, vehicles, tools, and clothing. Psychometric properties for this measure were adequate. Specifically, alternate forms reliability ranged between .44 and .78 while test-retest reliability over a two-week period was .69. Concurrent validity estimates with the Peabody Picture Vocabulary Test – 3rd Edition (Dunn & Dunn, 2000) and with the Preschool Language Scale – 3 (Zimmerman, Steiner, & Pond, 1992) were adequate, .53 to .79. The PNT was also sensitive to developmental status and growth over time. Children identified 21.4 pictures at the pretest (SD = 6.7). Benchmarking norms were provided by the authors: scores at 59 months averaged 16.97 for typically developing children; 16.51 for children from low income backgrounds; and 14.13 for children with identified disabilities (Missall & McConnell, 2004).

**Indicators of Letter Knowledge**

**PALS-PreK - Alphabet Knowledge**

The PALS PreK Alphabet Knowledge Task (Invernizzi, Sullivan, & Meier, 2002) was used to evaluate alphabet letter knowledge. The developers of the PALS included three different tasks that tapped into various components of letter knowledge: 1) identification of the 26 Upper Case letters; 2) identification of the 26 Lower Case letters; and 3) identification of the sounds associated with 23 letters and 3 digraphs. Children are first presented all 26 Upper Case letters in a random order. To be eligible to proceed to the second task, identification of all 26 Lower Case letters, the child must correctly identify 16 Upper Case letters. To be eligible to proceed from Lower Case letters to Letter Sounds, the child must correctly identify 9 Lower Case letters. Psychometrics are adequate with reported reliabilities ranging from .74 to .94.

With this task, we derived three types of scores: 1) the number of letters or sounds a child could correctly identify; 2) the number of children in each viewing group who were able to identify any Lower Case letters or Letter Sounds (i.e., only children who reached a certain cut-off were able to proceed to Lower Case letters and Letter Sounds); and 3) fluency scores (i.e., the number of seconds it took to identify one letter or sound).

1. **Number of Letters or Sounds Correctly Identified.** The total number of upper case, lower case, and letter sounds were recorded.
2. **Identification of Any Lower Case Names or Letter Sounds.** Children were presented with these tasks if they were able to 1) identify 16 or more Upper Case letters and 2) 9 or more Lower Case letters.
3. **Fluency Scores.** Children’s performance on each of the 3 subscales (i.e., Upper Case, Lower Case, Letter Sounds) was timed. Then, the number of letters or sounds accurately identified was divided by the number of seconds it took the child to complete each task. This produced a letter or sound identification per second rate. All children were administered the Upper Case task; therefore, all children had a fluency score associated with Upper Case Letter Knowledge. Only those children eligible to complete the Lower Case Letter Knowledge and the Letter Sounds tasks were included in those analyses.
Indicators of Phonemic Awareness

**IGDI Initial Sounds Fluency**

The DIBELS Initial Sound Fluency task is an individually administered and timed measure of children’s ability to recognize and produce the initial sound in an orally presented word, a component of phonemic awareness. The examiner presents four pictures to the child, names each picture, and then asks the child to identify (i.e., point to or say) the picture that begins with the sound produced orally by the examiner. For example, the examiner says, "This is sink, cat, gloves, and hat. Which picture begins with /s/?" and the student points to the correct picture. The child is also asked to orally produce the beginning sound for an orally presented word that matches one of the given pictures. The examiner calculates the amount of time taken to identify/produce the correct sound and converts the score into the number of initial sounds correct in a minute.

**PALS-PreK – Alphabet Knowledge**

A description of this task was detailed above. Only the indices that were derived from this measure to represent Phonological and Phonemic Awareness discussed below.

1. **Identification of Any Letter Sounds.** The percentage of children in each viewing group who were eligible to take the Letter Sounds task was recorded.
2. **Number of Sounds Correctly Identified.** The number of letter sounds a child was able to identify correctly was recorded.
3. **Letter Sounds Fluency.** Children’s performance on the Letter Sounds subscale was timed. Then, the number of items accurately identified was divided by the number of seconds it took the child to complete each task. This produced a sound identification per second rate. Only those children eligible to attempt the Letter Sounds task were included in those analyses.

Indicators of Print Conventions

**Print and Story Concepts Tasks**

This assessment was adapted from the Head Start FACES Survey (information available online: [http://www.acf.hhs.gov/programs/opre/hs/faces/instruments/child_instru02/language_story.html](http://www.acf.hhs.gov/programs/opre/hs/faces/instruments/child_instru02/language_story.html)) to examine children’s understanding of basic story concepts including book knowledge, print knowledge, and reading comprehension. Book knowledge examined children’s familiarity with storybooks and print conventions such as where the front of the book is, where to begin reading, and differentiating print from pictures. Print knowledge examined children’s knowledge of the mechanics of reading including reading from left to right, top to bottom, and word-by-word pointing. Reading comprehension measured children’s knowledge of a story plot and required them to answer questions based on presented story content (e.g., what is said goodnight to in *Goodnight Moon*) and well as to generate inferences (e.g., how does a character feel) and to make predictions (e.g., what do you think happens next in this story). Different books were used at each testing point: *Goodnight Moon* by Margaret Wise Brown was used at Pre-Test and *Where’s My Teddy?* by Jez Alborough was used at Post-Test. While most questions were based on a scoring system of (0) incorrect and (1) correct, some of the comprehension questions were worth up to 3 points. Each print and story construct was summed to form three scores for analysis: book knowledge, print knowledge, and reading comprehension.
Combined Early Literacy Skills

Get Ready to Read! Screener

This screener, consisting of 20 items, assessed print knowledge (i.e., knowledge of the letters of the alphabet); book knowledge (recognition of how books work including the difference between words and images); phonological awareness (i.e., understanding that spoken words are composed of individual sounds); phonics (i.e., recognition of the sounds letters make); and writing (i.e., understanding how text should look: letters grouped together into words). Each item required the child to select a response from a group of four pictures (or four letters, words, etc.). Example: “These are pictures of a book. Find the one that shows the back of the book.” Example: “Find the letter that makes a tuh sound.” Example: “Some children wrote their name. Find the one that is written the best.” Children were given a score of a (1) for every correct answer provided and a (0) for every incorrect answer provided, with a maximum score of 20 points. Scores greater than 11 are predictive of reading success by 2nd grade.

Analytical Approach

Repeated-measures Analysis of Covariance (ANCOVA) is a procedure that can be used to statistically control for initial group differences when evaluating intervention effects on outcome measures. In these models, both Intervention Group and Child’s Age were included as factors. Three covariates were constructed to extract the variance associated with variables that were found to relate to the outcomes of interest or that significantly varied by group. The classroom literacy environment, the teacher’s education, and a child’s pre-test performance were used as covariates in the analyses. When multiple tests were conducted for each set of outcomes, Bonferroni adjustments of the alpha level were made to reduce Type 1 error rates (i.e., finding a significant difference when one does not exist). For these analyses, only significant effects associated with Group are reported in the text (i.e., Group; Wave by Group, Age by Group). Along with the statistical significance tests, effect sizes are also reported.

Factor Details

☆ Group: This Between-Subjects factor tested for mean differences among the three possible intervention groups. There were 3 levels associated with this factor.
  • The TREATMENT group is composed of children, teachers, and classrooms who participated in the BTL-LI intervention for the first time. There were 95 children in this group.
  • The MAINTENANCE group is composed of children, teachers, and classrooms who also participated in the BTL-LI intervention; however, the teachers had previously received intervention training and mentoring between 1 and 3 years prior. There were 90 children in this group.
  • The CONTROL group is composed of children, teachers, and classrooms who did not participate in any BTL-LI intervention training. There were 113 children in this group.

☆ Age: This Between-Subjects factor tests for mean differences among different age groups of children. There are 2 levels associated with this factor.
• The **FOURS** group of children is composed of 156 children who are between 46 months and 59 months. On average, children in this group were 54.1 months (i.e., 4.5 years; SD = 3.7).

• The **FIVES** group of children is composed of 140 children who are between 60 months and 74 months. On average, children in this group were 65.6 months (i.e., 5.0 years; SD = 3.8).

**Wave:** This Within-Subjects factor tests for mean differences associated with gains (or losses) from pretest to post-test. There are 2 levels associated with this factor.

• The **PRETEST** was administered prior to participation in any of the intervention materials.

• The **POST-TEST** was administered at the end of participation in any of the intervention materials.

**Interpreting the Results**

**Main Effects:** Main effects represent the simple effect of a particular factor (i.e., Group, Wave, Age) on a dependent variable (i.e., outcome). This ‘main effect’ is the effect of the factor alone averaged across the levels of other factors.

**How Do I Interpret Main Effects?** In this study, there were 3 factors whose main effects were examined for statistical significance: Group, Age, and Wave. When a factor’s main effect was significant, the means for each level of that factor holding constant all other factor effects were reported. This becomes a significant issue when interpreting what the Group and Age main effects signify. Remember that a main effect represents the effect of a factor averaged across the levels of all other factors. When interpreting a main effect of Group, the means are averaged across both levels of Wave. The same is true for the main effect of Age: means for each level of Age represent the average performance for each level when Group and Wave are held constant. Main effects can become confusing when there is a time factor included in the statistical model (i.e., Wave with 2 levels: pre-test and post-test). It is important to remember that a Group main effect can only be interpreted such that the average performance of each level of Group throughout the entire intervention period differed significantly. Therefore, when graphs or tables include a Group or Age main effect, the values for each level of these two factors (i.e., treatment, control, maintenance; 4-year-olds, 5-year-olds) represent performance across both pre-test and post-test.

**Interaction Effects:** An interaction effect represents the variation among the differences between means for different levels of one factor over different levels of the other factor. For instance, in the Wave by Group interaction, the outcomes are evaluated for the presence of differences by each group across each wave. When an interaction effect is statistically significant, it is usually referred to as ‘moderating’ the effect of the other variable. The effects associated with the manipulated factor of interest (i.e., the factor that researchers manipulated; non-manipulated factors are those that could never be manipulated by a researcher – gender, age) must be examined for potential moderators. When a moderator effect is present, it signifies that the pattern of results varies either in strength or direction across each level of the factor that is moderated (in this example, the factor that is moderated is Group).
Results

ELLCO Classroom Environment 2007 to 2008

Four subscales from the ELLCO were evaluated for significant differences across groups.

General Classroom Environment

All classrooms improved from pre-test to post-test (Figure 1), with treatment classes showing the most improvement. Maintenance classrooms started higher and remained at that level. Control classrooms demonstrated slight growth from pretest to post-test. There was also a significant difference by Group.

\[ F(2, 30) = 4.73, \ p < .05, \]
\[ F(2, 30) = 14.17, \ p < 0.001 \]

Figure 1. General Classroom Environment Subscale Differences by Group Across Wave
Literacy Environment

All classrooms improved from pre-test to post-test (Figure 2), with treatment classes showing the most improvement. Maintenance classrooms started higher and remained so at the post-test. Control classrooms demonstrated slight growth from pretest to post-test. There was also a significant difference by Group.

\[ F(2, 30) = 4.73, \ p < 0.05 \]

\[ F(2, 30) = 14.17, \ p < 0.001 \]

\[ F(2, 30) = 6.64, \ p < 0.01 \]

\[ F(2, 30) = 24.33, \ p < 0.001 \]

Figure 2. Literacy Environment Subscale Differences Across Group by Wave

Language, Literacy, and Curriculum

All classrooms improved from pre-test to post-test (Figure 3), with treatment classes improving most. Maintenance classrooms started higher and remained so at the post-test. Control classrooms demonstrated little growth from pretest to post-test. There was also a significant difference by Group.

\[ F(2, 30) = 6.64, \ p < 0.01 \]

\[ F(2, 30) = 24.33, \ p < 0.001 \]

Figure 3. Language, Literacy, and Curriculum Subscale Differences by Group Across Wave
**Literacy Activities**

All classrooms improved from pre-test to post-test (Figure 4), with treatment classes improving most. Maintenance classrooms started higher and remained so at the post-test\(^8\) (i.e., these group by wave differences were marginally significant). Control classrooms demonstrated little growth from pretest to post-test. There was also a significant difference by Group\(^9\).

![Literacy Activities Graph](image)

**Figure 4. Literacy Activity Subscale Differences by Group Across Wave**

---

**ELLCO Classroom Environment 2006 to 2008**

Figures 5 and 6 are based on a subset of 7 teachers who were trained in either 2005 (n = 3) or 2006 (n = 4) to implement the BTL curriculum.

![Literacy Environment Graph](image)

**Figure 5. Literacy Environment Over Time for 7 Trained Teachers**

---

\(^8\) F(2, 30) = 3.13, p < 0.06

\(^9\) F(2, 30) = 11.07, p < 0.001
Once trained to use the BTL Classroom Intervention materials, teachers maintain classrooms that are characterized by high-quality literacy materials and, perhaps more importantly, language interactions that facilitate literacy development.

**ELLCO Classroom Environment Conclusion**

Teachers who were new to the BTL Classroom Intervention were able to create classrooms that provided preschool children with “optimal support for their language and literacy development” (p. 1; Smith, Dickinson, Sangeorge, & Anastasopoulos, 2002). These classrooms were characterized by more and higher quality reading and writing materials and activities as well as multiple teacher-child interactions that are known to facilitate oral language development and early literacy skill acquisition. Teachers in Main Maintenance classrooms were able to maintain the structural components indicative of a strong Literacy Environment from previous years to this project year. The BTL Classroom Intervention has always been successful in helping teachers create literacy environments that would be capable of supporting children’s developing literacy abilities. More importantly, the mentoring provided during the 2007 and 2008 project year indicated that teachers were able to move beyond changing the structural features of the environment into systematically and consistently altering both the general classroom environment as well as the literacy-enriching interactions that are crucial to supporting children’s optimal language and literacy-skill development.

First, the General Classroom Environment subscale indicated that trained teachers introduced greater intentionality in the physical organization of the classroom; provided children with multiple opportunities for choice and for taking initiative; and used more positive management strategies. This subscale more generally evaluated the classroom environment including whether this environment could successfully support general child development and positive classroom experiences for both teachers and children.

Next, specific environmental features that supported a high quality literacy environment were scored using the Literacy Environment subscale. This subscale provides a quick inventory of the types of literacy-related items or supports that were found in each classroom. Both treatment and maintenance
teachers created literacy environments that incorporated designated book areas; provided a number of books that featured varied topics and that were in good physical condition; made these books easily accessible to children; and provided a variety of writing tools that were also easily accessible.

The remaining subscales assessed the various interactions that occurred in classrooms centered around language and literacy topics including interactions between teachers and children and between children and the materials and activities available to them in their classrooms. First, the Language, Literacy, and Curriculum subscale measured a teachers’ ability to facilitate oral language; to adopt positive and intentional approaches to book reading, writing, curriculum integration, and assessment; and to facilitate home support for literacy. These types of interactions help to create a language- and literacy-rich environment that, in other research, has been linked to positive child development outcomes. Teachers in both the treatment and the maintenance groups consistently scored between proficient and exemplary on these key teacher-child interactional items.

Second, the number of positive Literacy Activities in both treatment and maintenance classrooms also increased from pretest to posttest and were substantially higher than control teacher classrooms. Teachers implementing the BTL Classroom Intervention engaged in more and longer full-group and one-to-one book-reading sessions; modeled writing; provided writing assistance and opportunities for writing; and set aside time for children to look at books alone or with a classmate.

Taken together, these findings suggest that teachers who have been trained to deliver the BTL Classroom Intervention are primed and able to make changes to the structural features of their classrooms including general classroom management abilities, specific environmental features that are necessary for literacy support, and increased opportunities for literacy activities. In addition, these structural changes also resulted in process quality changes as indexed by higher-quality language- and literacy-rich interactions and experiences. Finally, maintenance teachers who had been trained in previous years were able to sustain the kind of high-quality environment needed to fundamentally shift young children’s early language and literacy trajectories.

---

**Child Literacy Outcomes**

Children were assessed at two time points: fall 2007, prior to any intervention participation, and again in March or April of 2008, toward the end of the academic year. Where available, benchmarks and expected developmental ranges have been indicated. Tables are provided at the conclusion of each Indicator section. These tables listed all indicators associated with a larger construct (i.e., larger constructs included Vocabulary Knowledge, Letter Knowledge, Phonemic Awareness, Print Conventions, Combined Early Literacy Task). Each table contained a column of assessments that measured aspects of the overall indicator. For instance, the Letter Knowledge Indicator was comprised of Upper Case Total letters named, Upper Case letter naming rate; Any Lower Case letters named?; Lower Case Total letters named; and Lower Case letter naming rate. For each assessment listed in a table, there were columns for each of the three intervention groups (i.e., Treatment, Maintenance, Control). Within each intervention group, there were also columns that indicated whether scores were the pretest average or the post-test average. Tables were constructed using the following information:

All means at both the pretest and the post-test controlled for pre-intervention knowledge, teacher
characteristics associated with an individual child’s classroom, and the classroom literacy environment. Indicator tables were created for each of the 5 sets of Indicators:

a. **Indicator of Language Development:** Language Development consisted of one indicator that measured children’s expressive vocabulary knowledge.

b. **Indicators of Letter Knowledge:** There were five difference indices of letter knowledge. Not all indicators were administered to every child. Administration depended on whether children were able to complete enough items on the previous task. Specifically:
   - All children were administered the Upper Case letter naming task. This task involved presenting all Upper Case alphabet letters to a child. The child was asked to name as many as he/she could. The data collector scored each of the 26 letters as correct or incorrect. The data collector also measured how long (in seconds) it took a child to complete this task.
   - If the child accurately identified 16 Upper Case letters, he/she was eligible to attempt the Lower Case letter knowledge task. This task involved presenting the child with all Lower Case letters in random order. The child was asked to name as many of these letters as she/he could. In addition to scoring whether or not the child was able to accurately identify each lower case letter, the data collector measured (in seconds) how long the child took.
   - If the child was able to accurately identify 9 or more Lower Case letters, he/she was eligible for the Letter Sounds task. This task involved presenting the child with 23 alphabet letters and 3 digraphs in random order. The child was asked to produce the sound associated with each. In addition to scoring whether or not the child was able to accurately identify the letter sounds, the data collector also measured (in seconds) how long it took for the child to complete the task.

c. **Indicators of Phonemic Awareness:** There were three difference indices of letter knowledge. The first index, the IGDI Initial Sounds Fluency task, was administered to all children. The PALS preK Letter Sounds task was only administered to those children who were able to identify 16 Upper Case and 9 Lower Case letters accurately. Not all indicators were administered to every child.

d. **Indicators of Print Conventions:** These indicators included book knowledge, print knowledge, and story comprehension that were administered while reading a book together.

e. **Combined Early Literacy Task:** This measure was administered to all children and was designed to measure their knowledge of book conventions, print conventions, letter knowledge, phonological and phonemic awareness, and early writing skills.

Table 2. Percentage of children who were able to identify at least one item correctly on each task.

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th>Maintenance</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
</tr>
<tr>
<td><strong>Upper Case</strong></td>
<td>94.7%</td>
<td>97.9%</td>
<td>82.2%</td>
</tr>
<tr>
<td><strong>Lower Case</strong></td>
<td>54.7%</td>
<td>84.2%</td>
<td>26.7%</td>
</tr>
<tr>
<td><strong>Letter Sounds</strong></td>
<td>51.6%</td>
<td>76.8%</td>
<td>21.1%</td>
</tr>
</tbody>
</table>

10 The criteria for proceeding to the next subtest on this overall measure was the same at both the pre-test and the post-test (e.g., to be administered the Lower Case subtest, children at both the pre-test and post-test needed to accurately identify 16 Upper Case letters).

11 $\chi^2 = 7.33, p < 0.05$

12 $\chi^2 = 6.82, p < 0.05$

13 $\chi^2 = 6.03, p < 0.05$

14 $\chi^2 = 19.68, p < 0.001$

15 $\chi^2 = 23.79, p < 0.001$

16 $\chi^2 = 18.19, p < 0.001$
Indicators of Language Development

IGDI Picture Naming

This task evaluated young children’s expressive vocabulary knowledge. Children were asked to name or label as many picture cards as they could in one minute. Overall, performance was highest for children in the control group (i.e., 15.34) followed by children in the treatment group (i.e., 14.71) and then children in the maintenance group (i.e., 13.48)\(^\text{17}\); however, these results were moderated by child’s Age\(^\text{18}\). All children who were categorized into the FOURS age group scored similarly. The performance of children who were categorized into the FIVES group differed by group: control children outscored both treatment and maintenance children. Follow-up tests indicated that both the treatment and maintenance groups scored significantly lower than the control group while the difference between treatment and maintenance FIVES did not differ significantly. See Figure 7.

\[ F(2, 285) = 4.07, \ p < 0.05 \]
\[ F(4, 285) = 2.61, \ p < 0.05 \]

Figure 7. IGDI Picture Naming Scores Overall and Split by Group and Age
Table 3. Pre-Test and Post-Test Means for the Language Development Indicator by Age and Group

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th></th>
<th>Maintenance</th>
<th></th>
<th>Control</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
<td>Post</td>
</tr>
<tr>
<td>FOURS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIVES</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGDI Picture Naming</td>
<td>14.26</td>
<td>16.74</td>
<td>13.61</td>
<td>13.59</td>
<td>16.34</td>
<td>17.86</td>
</tr>
<tr>
<td>OVERALL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGDI Picture Naming</td>
<td>13.85</td>
<td>15.57</td>
<td>13.59</td>
<td>13.36</td>
<td>14.64</td>
<td>16.03</td>
</tr>
</tbody>
</table>
**Indicators of Letter Knowledge**

**Upper Case Letter Knowledge**

**Upper Case Total Scores.**

There was a significant 2-way interaction between group and child’s age\(^{19}\) (Figure 8). For children classified as FOUR, the treatment group outperformed both the control and the maintenance groups. For children classified as FIVE, all groups scored similarly; that is, there were no significant differences in their knowledge of Upper Case letter names.

\[ F(2, 288) = 3.68, \ p < 0.05 \]

**Figure 8. Upper Case Letters Split by Group and Age**
Upper Case Letter Naming Speed\textsuperscript{20}.

There was a significant group by wave interaction\textsuperscript{21}. All three groups significantly improved from pretest to post-test. Children in the treatment group named Upper Case letters the fastest followed by control group children and then maintenance group children. See Figure 9.

\textsuperscript{20} Note that the letter naming speed represents the rate at which children were identifying the letters and is not necessarily a reflection of how many letters were named; that is, it extrapolates the rate based on the total time spent on the task and the number of letters named correctly. Incorrect responses were factored into their overall naming rates so that a rate of 22 letters per minute does not necessarily mean that they correctly identified 22 letters.

\textsuperscript{21} F(2, 212) = 3.89, p < 0.05

Figure 9. Upper Case Letter Naming Speed by Group Across Wave
Lower Case Letter Knowledge

Any Knowledge of Lower Case Letters?

A larger majority of Treatment children were able to identify any Lower Case Letters (i.e., 84.2%) compared with Maintenance (i.e., 55.6%) and Control (i.e., 61.3%) children at the post-test. When conducting this analysis for both age groups, more younger children (i.e., FOURS) in the treatment group correctly identified at least one Lower Case letter (i.e., 80.0%) when compared with their peers in both the control (i.e., 50.9%) and maintenance groups (i.e., 50.7%). The differences for older children were marginally significant but evidenced the same patterns; that is, 86.2% of the FIVES in the treatment group were able to accurately identify any Lower Case compared with 76.5% of the FIVES in the maintenance group and 70.7% of FIVES in the control group.

Lower Case Total Scores.

There was a significant 2-way interaction between group and age (Figure 10). Children in the treatment group who were FOUR were able to identify more Lower Case letters when compared with their peers in both the control and maintenance groups. In contrast, children in the maintenance group who were classified as FIVES outperformed their peers in both the treatment and control groups.

Figure 10. Lower Case Letters Split by Age and Group

22 $X^2 = 19.68, p < 0.001$
23 $X^2 = 8.41, p < 0.05$
24 $X^2 = 4.40, p < 0.11$
25 $F(2, 288) = 3.05, p < 0.05$
Lower Case Letter Naming Speed$^{26}$. 

There was a significant 2-way interaction between group and age$^{27}$ (Figure 11). Children in the FOURS treatment group identified lower case letters more quickly than children in the FOURS control and maintenance groups. For both YOUNG and OLD 5s, control group children outperformed their treatment and maintenance group peers.

![Figure 11. Lower Case Letter Naming Speed by Group Across Wave](image)

$^{26}$ Note that the letter naming speed represents the rate at which children were identifying the letters and is not necessarily a reflection of how many letters were named; that is, it extrapolates the rate based on the total time spent on the task and the number of letters named correctly. Incorrect responses were factored into their overall naming rates so that a rate of 22 letters per minute does not necessarily mean that they correctly identified 22 letters.

$^{27}$ $F(2, 71) = 3.20, p < 0.05$
## Table 4. Pre-Test and Post-Test Means for All Letter Knowledge Indices by Age and Group

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th></th>
<th>Maintenance</th>
<th></th>
<th>Control</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
</tr>
<tr>
<td><strong>FOURS</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper Case Total</td>
<td>12.56</td>
<td>17.61</td>
<td>10.88</td>
<td>12.54</td>
<td>11.51</td>
<td>13.91</td>
</tr>
<tr>
<td>Upper Case Rate</td>
<td>12.14</td>
<td>23.63</td>
<td>12.07</td>
<td>15.71</td>
<td>11.81</td>
<td>19.18</td>
</tr>
<tr>
<td>Any Lower Case?</td>
<td>33.3%</td>
<td>80.0%</td>
<td>16.4%</td>
<td>50.7%</td>
<td>24.5%</td>
<td>50.9%</td>
</tr>
<tr>
<td>Lower Case Rate</td>
<td>23.91</td>
<td>32.20</td>
<td>17.15</td>
<td>25.39</td>
<td>16.41</td>
<td>28.95</td>
</tr>
<tr>
<td><strong>FIVES</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper Case Total</td>
<td>10.54</td>
<td>16.49</td>
<td>12.04</td>
<td>16.47</td>
<td>10.86</td>
<td>17.55</td>
</tr>
<tr>
<td>Upper Case Rate</td>
<td>10.59</td>
<td>22.08</td>
<td>12.80</td>
<td>18.41</td>
<td>11.70</td>
<td>22.58</td>
</tr>
<tr>
<td>Any Lower Case?</td>
<td>64.6%</td>
<td>86.2%</td>
<td>70.6%</td>
<td>76.5%</td>
<td>36.2%</td>
<td>70.7%</td>
</tr>
<tr>
<td>Lower Case Total</td>
<td>5.46</td>
<td>12.06</td>
<td>8.03</td>
<td>12.22</td>
<td>5.52</td>
<td>12.75</td>
</tr>
<tr>
<td>Lower Case Rate</td>
<td>18.28</td>
<td>29.40</td>
<td>22.76</td>
<td>24.46</td>
<td>24.37</td>
<td>33.52</td>
</tr>
<tr>
<td><strong>OVERALL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper Case Total</td>
<td>11.55</td>
<td>17.05</td>
<td>11.46</td>
<td>14.50</td>
<td>11.19</td>
<td>15.73</td>
</tr>
<tr>
<td>Upper Case Rate</td>
<td>11.36</td>
<td>22.86</td>
<td>12.43</td>
<td>17.06</td>
<td>11.76</td>
<td>20.88</td>
</tr>
<tr>
<td>Any Lower Case?</td>
<td>54.7%</td>
<td>84.2%</td>
<td>26.7%</td>
<td>55.6%</td>
<td>30.6%</td>
<td>61.3%</td>
</tr>
<tr>
<td>Lower Case Total</td>
<td>6.95</td>
<td>12.83</td>
<td>7.75</td>
<td>10.71</td>
<td>9.29</td>
<td>9.13</td>
</tr>
<tr>
<td>Lower Case Rate</td>
<td>21.10</td>
<td>30.80</td>
<td>19.96</td>
<td>24.92</td>
<td>20.39</td>
<td>31.24</td>
</tr>
</tbody>
</table>
Indicators of Phonemic Awareness

IGDI Initial Sounds Fluency

There was a significant 3-way interaction among group, age, and wave\(^28\) (Figure 12). Both maintenance and control group children who were classified as FOUR outperformed their peers in the treatment group. Children in the maintenance group who were classified as FIVE obtained higher scores followed by FIVEs in the treatment group who, in turn, scored higher than FIVEs in the control group.

\[ F(2, 288) = 3.51, \ p < .01 \]

Figure 12. IGDI Initial Sounds Fluency Split by Age and Group
Letter Sounds Knowledge

Any Knowledge of Letter Sounds?

A larger majority of Treatment children were able to identify any Letter Sounds (i.e., 76.8%) compared with Maintenance (i.e., 51.1%) and Control (i.e., 50.5%) children at the post-test\(^29\). When conducting this analysis for both age groups, more younger children (i.e., FOURs) in the treatment group correctly identified at least one Letter Sound (i.e., 76.7%) when compared with their peers in both the control (i.e., 43.4%) and maintenance groups (i.e., 46.6%)\(^30\). The differences for older children were marginally significant\(^31\) but evidenced the same patterns; that is, 76.9% of the FIVEs in the treatment group were able to accurately identify any Letter Sounds compared with 70.6% of the FIVEs in the maintenance group and 56.9% of FIVEs in the control group.

Letter Sounds Total Score.

Performance was highest for children in the maintenance group (i.e., 5.73) followed by children in the treatment group (i.e., 5.33) and then children in the control group (i.e., 4.08)\(^32\). There was also a significant 2-way interaction between group and age\(^33\) (Figure 13). Children in the treatment group who were categorized as FOURs outperformed their peers in the control and maintenance groups. Children in the maintenance group who were categorized as FIVEs outperformed their peers in the treatment group who, in turn, outperformed those children in the control group.

\(^{29}\) \(\chi^2 = 18.19, p < 0.001\)
\(^{30}\) \(\chi^2 = 9.70, p < 0.01\)
\(^{31}\) \(\chi^2 = 5.70, p < 0.06\)
\(^{32}\) \(F(2, 288) = 4.06, p < 0.05\)
\(^{33}\) \(F(2, 288) = 7.39, p < 0.001\)
There was also a significant 2-way interaction between Group and Wave\textsuperscript{34} (Figure 14). At the post-test, both treatment viewers and maintenance viewers outperformed their control group peers

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure14}
\caption{Letter Sounds by Group Across Wave}
\end{figure}

\textit{Letter Sounds Naming Speed.}

There were no significant differences by group or age for children’s letter sound naming speed.

\textsuperscript{34} F(2, 288) = 11.08, p < 0.05
Table 5. Pre-Test and Post-Test Means for All Phonemic Awareness Indices by Age and Group

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th>Maintenance</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
</tr>
<tr>
<td>FOURS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGDI Initial Sounds Fluency</td>
<td>6.12</td>
<td>8.72</td>
<td>9.26</td>
</tr>
<tr>
<td>Any Letter Sounds?</td>
<td>33.3%</td>
<td>76.7%</td>
<td>9.6%</td>
</tr>
<tr>
<td>Letter Sounds Total</td>
<td>3.52</td>
<td>7.86</td>
<td>3.73</td>
</tr>
<tr>
<td>Letter Sounds Rate</td>
<td>7.40</td>
<td>10.36</td>
<td>6.07</td>
</tr>
<tr>
<td>FIVES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGDI Initial Sounds Fluency</td>
<td>8.98</td>
<td>12.69</td>
<td>5.90</td>
</tr>
<tr>
<td>Any Letter Sounds?</td>
<td>60.0%</td>
<td>76.9%</td>
<td>70.6%</td>
</tr>
<tr>
<td>Letter Sounds Total</td>
<td>2.84</td>
<td>7.08</td>
<td>5.04</td>
</tr>
<tr>
<td>Letter Sounds Rate</td>
<td>5.95</td>
<td>10.58</td>
<td>9.41</td>
</tr>
<tr>
<td>OVERALL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGDI Initial Sounds Fluency</td>
<td>7.55</td>
<td>10.70</td>
<td>7.58</td>
</tr>
<tr>
<td>Any Letter Sounds?</td>
<td>51.6%</td>
<td>76.8%</td>
<td>21.1%</td>
</tr>
<tr>
<td>Letter Sounds Total</td>
<td>3.18</td>
<td>7.47</td>
<td>4.39</td>
</tr>
<tr>
<td>Letter Sounds Rate</td>
<td>6.67</td>
<td>10.47</td>
<td>7.74</td>
</tr>
</tbody>
</table>
Indicators of Print Conventions

Book Knowledge

There was a significant main effect of Group\(^{35}\) (Figure 15). Children in the treatment group outperformed their peers in the maintenance and control groups.

Print Knowledge

There was a significant 2-way interaction between Group and Age\(^{36}\) (Figure 16). The FOURs treatment and maintenance groups obtained higher scores on print knowledge when compared with control group FOURs. The FIVEs treatment and control groups outperformed their peers in the maintenance group.

\(^{35}\) F(2, 288) = 5.29, \(p < 0.01\)

\(^{36}\) F(2, 288) = 3.79, \(p < 0.05\)
Comprehension

There was a significant 2-way interaction between group and wave\(^3\) (Figure 17). Children in both the treatment and control groups outperformed their peers in the maintenance group at the post-test.

\[ F(2, 288) = 15.26, \ p < 0.01 \]

*Figure 17. Comprehension Scores Split by Group*

---

\(^3\) \(F(2, 288) = 15.26, \ p < 0.01\)
Table 6. Pre-Test and Post-Test Means for All Print Conventions Indices by Age and Group

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th></th>
<th></th>
<th>Maintenance</th>
<th></th>
<th></th>
<th>Control</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
<td>Post</td>
<td></td>
</tr>
<tr>
<td>FOURs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Book Knowledge</td>
<td>2.21</td>
<td>2.39</td>
<td>1.85</td>
<td>2.11</td>
<td>1.99</td>
<td>1.58</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Print Knowledge</td>
<td>0.67</td>
<td>1.33</td>
<td>0.58</td>
<td>1.25</td>
<td>0.59</td>
<td>0.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension</td>
<td>4.46</td>
<td>5.58</td>
<td>5.22</td>
<td>4.54</td>
<td>4.25</td>
<td>4.91</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>FIVEs</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Book Knowledge</td>
<td>2.45</td>
<td>2.66</td>
<td>2.46</td>
<td>2.37</td>
<td>2.29</td>
<td>2.14</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Print Knowledge</td>
<td>0.81</td>
<td>1.30</td>
<td>0.18</td>
<td>0.98</td>
<td>0.84</td>
<td>1.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension</td>
<td>5.03</td>
<td>5.63</td>
<td>5.22</td>
<td>4.54</td>
<td>5.08</td>
<td>5.72</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OVERALL</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Book Knowledge</td>
<td>2.33</td>
<td>2.53</td>
<td>2.16</td>
<td>2.24</td>
<td>2.14</td>
<td>1.86</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Print Knowledge</td>
<td>0.74</td>
<td>1.32</td>
<td>0.38</td>
<td>1.12</td>
<td>0.72</td>
<td>1.00</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comprehension</td>
<td>4.75</td>
<td>5.61</td>
<td>5.05</td>
<td>4.41</td>
<td>4.67</td>
<td>5.32</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Combined Early Literacy Skills

Get Ready to Read! Screener

There was a significant main effect of group\textsuperscript{38}. Children in the treatment group outperformed their peers in the maintenance group who, in turn, outperformed their peers in the control group. This effect was moderated by wave of assessment\textsuperscript{39} (Figure 18). Treatment and maintenance children scored higher than their control group peers at the post-test.

\textsuperscript{38} F(2, 288) = 3.74, p < 0.05
\textsuperscript{39} F(4, 285) = 5.42, p < 0.001

\textbf{Figure 18. Get Ready to Read Split by Age and Group}
Table 7. Pre-Test and Post-Test Means for the Combined Early Literacy Indicator by Age and Group

<table>
<thead>
<tr>
<th></th>
<th>Treatment</th>
<th>Maintenance</th>
<th>Control</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Pre</td>
<td>Post</td>
<td>Pre</td>
</tr>
<tr>
<td>FOURS</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Get Ready to Read</td>
<td>9.84</td>
<td>11.78</td>
<td>9.01</td>
</tr>
<tr>
<td>FIVES</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Get Ready to Read</td>
<td>9.18</td>
<td>11.80</td>
<td>8.77</td>
</tr>
<tr>
<td>OVERALL</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Get Ready to Read</td>
<td>9.51</td>
<td>11.79</td>
<td>8.89</td>
</tr>
</tbody>
</table>
Results Summary Table

On the next two pages, two tables contain effect size estimates for each outcome. An effect size is an objective, standardized, and metric-free index of the practical significance of a result. It reflects the magnitude, or size, of group differences (Hedges, 2008). This type of information can help researchers and policy makers determine whether a particular difference between two groups is big and meaningful or whether the difference is actually an artifact of a large sample size. Standardized effect sizes reflect the number of standard deviation units that separate two groups. A standard deviation reflects the dispersion of children’s scores around a group mean using an index of the expected variation around that mean. A small standard deviation indicates that children’s scores are closely clustered around the mean value while a large standard deviation indicates that the spread of their scores is relatively wide. About 68% of children’s scores will fall between one standard deviation above and one standard deviation below the mean while 95% of children’s scores will fall between two standard deviations above and two standard deviations below the mean.

In this study, standardized effect sizes are reported as a way to contextualize the magnitude of differences in an equivalent fashion across measures or participants. Cohen’s $d$ (Cohen, 1988) was selected because it is one of the most widely used effect size indices in the literature. When making comparisons involving two groups of children who participated in two different educational interventions (i.e., Intervention A or Intervention B), obtaining an effect size of 1.0 (with Intervention A children outperforming Intervention B children) indicates that Intervention A children scored, on average, a standard deviation higher than Intervention B children.

Tables 8 and 9 below contain effect sizes associated with two different comparisons: 1) overall by group and 2) within each age level by group. Positive values indicate that either the treatment group or the maintenance group scored higher than the control group. Negative values indicate that the control group scored higher than the treatment group or the maintenance group. For example:

1. The effect size describing the difference between treatment and control group means for Upper Case Total is 0.47. This effect size indicates that children in the treatment group scored just under a ½ standard deviation higher than their control group peers.
2. The effect size describing the difference between maintenance and control group means for Upper Case Total is -0.17. This effect size indicates that children in the control group scored .17 standard deviation units higher than their maintenance group peers.

Cohen (1988) suggested benchmarks for interpreting whether an effect size is meaningful or trivial and, if meaningful, how big an effect size was (i.e., small, moderate, or large). On the next page, there is a legend with these benchmarks as well as a color key indicating, at-a-glance, which effects were small, moderate, or large.
### Table 8. Effect Size Estimates for Indicators of Language Development and Letter Knowledge

<table>
<thead>
<tr>
<th></th>
<th>FOURS (46 – 59 months)</th>
<th>FIVES (60 – 74 months)</th>
<th>OVERALL</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Language Development</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>IGDI Picture Naming</td>
<td>d = 0.10</td>
<td>d = -0.44</td>
<td>d = -0.34</td>
</tr>
<tr>
<td></td>
<td>d = -0.05</td>
<td>d = -0.94</td>
<td>d = -0.90</td>
</tr>
<tr>
<td><strong>Letter Knowledge</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Upper Case Total</td>
<td>d = 0.47</td>
<td>d = 0.15</td>
<td>d = 0.19</td>
</tr>
<tr>
<td></td>
<td>d = -0.17</td>
<td>d = 0.30</td>
<td>d = -0.16</td>
</tr>
<tr>
<td>Upper Case Naming Speed</td>
<td>d = 0.92</td>
<td>d = -0.47</td>
<td>d = 0.15</td>
</tr>
<tr>
<td></td>
<td>d = -0.016</td>
<td>d = -0.60</td>
<td>d = -0.27</td>
</tr>
<tr>
<td>Lower Case Total</td>
<td>d = 0.35</td>
<td>d = -0.07</td>
<td>d = 0.17</td>
</tr>
<tr>
<td></td>
<td>d = -0.17</td>
<td>d = 0.19</td>
<td>d = -0.11</td>
</tr>
<tr>
<td>Lower Case Naming Speed</td>
<td>d = 0.33</td>
<td>d = -0.39</td>
<td>d = -0.02</td>
</tr>
<tr>
<td></td>
<td>d = -0.07</td>
<td>d = -0.42</td>
<td>d = -0.27</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Effect Size</th>
<th>Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.20 to 0.49</td>
<td>Small</td>
</tr>
<tr>
<td>0.50 to 0.79</td>
<td>Moderate</td>
</tr>
<tr>
<td>0.80 and Above</td>
<td>Large</td>
</tr>
<tr>
<td>&lt; -0.10</td>
<td>Control group scored higher</td>
</tr>
<tr>
<td>Between -0.20 and 0.20</td>
<td>Trivial</td>
</tr>
</tbody>
</table>

- **Effect Size Interpretation**:<ref>![Table 8. Effect Size Estimates for Indicators of Language Development and Letter Knowledge](https://example.com/table8)</ref>
Table 9. Effect Size Estimates for Indicators of Phonemic Awareness, Print Conventions, and Get Ready to Read

<table>
<thead>
<tr>
<th>Effect Size</th>
<th>Interpretation</th>
<th>FOURS (46 – 59 months)</th>
<th>FIVES (60 – 74 months)</th>
<th>OVERALL</th>
</tr>
</thead>
<tbody>
<tr>
<td>&lt; -0.20</td>
<td>Control group scored higher</td>
<td>Trivial</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Between -0.20 and 0.20</td>
<td>Trivial</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.20 to 0.49</td>
<td>Small</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.50 to 0.79</td>
<td>Moderate</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0.80 and Above</td>
<td>Large</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Phonemic Awareness**

- **IGDI Initial Sounds Fluency**
  - Treatment vs. Control: $d = -0.23$  
  - Maintenance vs. Control: $d = 0.02$  
  - **OVERALL**: $d = 0.00$

- **Letter Sounds Total**
  - Treatment vs. Control: $d = 0.08$  
  - Maintenance vs. Control: $d = -0.08$  
  - **OVERALL**: $d = 0.45$

- **Letter Sounds Naming Speed**
  - Treatment vs. Control: $d = 0.04$  
  - Maintenance vs. Control: $d = -0.10$  
  - **OVERALL**: $d = 0.08$

**Print Conventions**

- **Book Knowledge**
  - Treatment vs. Control: $d = 0.56$  
  - Maintenance vs. Control: $d = 0.19$  
  - **OVERALL**: $d = 0.46$

- **Print Knowledge**
  - Treatment vs. Control: $d = 0.41$  
  - Maintenance vs. Control: $d = 0.28$  
  - **OVERALL**: $d = 0.14$

- **Comprehension**
  - Treatment vs. Control: $d = 0.31$  
  - Maintenance vs. Control: $d = 0.00$  
  - **OVERALL**: $d = 0.13$

**Combined Early Literacy**

- **Get Ready to Read**
  - Treatment vs. Control: $d = 0.64$  
  - Maintenance vs. Control: $d = 0.00$  
  - **OVERALL**: $d = 0.42$
Bridging the Gap

Contextualizing the relations between the intervention and key early literacy outcomes of interest can be accomplished by comparing child performance with benchmarks and developmental ranges, when available. Another way to contextualize the findings involves comparing the magnitude, or size, of group differences (Hedges, 2008; Hill, Bloom, Black, & Lipsey, 2008). Effect sizes help researchers and policy-makers determine whether a particular group difference is practical or meaningful or whether the difference is actually an artifact of a large sample size.

Policy-Relevant Performance Gaps

Establishing criteria that allow researchers and policy-makers to judge the effectiveness of an educational program or intervention is critical. Typically, minority children and children from low income families evidence substantial deficits in their early language development and literacy skills when compared with their more advantaged peers. For instance, Hart and Risley (1995) found that the amount and quality of talk that parents engaged in with their infants and toddlers from 6 months to three years predicted the size of these children’s vocabularies at age 3. Specifically, children of professional parents had vocabularies that averaged 1,116 words; children from working-class families had vocabularies that averaged 749 words; and children from low-income (or “welfare”) families had vocabularies half the size of children from professional families (i.e., 525 words). This disadvantage continued to follow these children through their transition to formal schooling (Walker, Greenwood, Hart, & Carta, 1994); it grew larger over the years (i.e., Matthew effects; Stanovich, 1986); and was nearly impossible to overcome even with substantial intervention efforts.

National estimates of preschoolers’ early literacy abilities were taken from published reports using data from the Early Childhood Longitudinal Survey – Birth Cohort (i.e., ECLS-B; NCES, 2006). The underlying early literacy skills that were measured in both that study and the present study included indicators of letter knowledge, print conventions, and a combined early literacy skills task. Because the actual measures used in each study differed, it was necessary to standardized all outcome data. Once standardized, comparisons between both studies using a common framework became possible. Recall that the standardized effect size represents the number of standard deviation units that separate the means of two groups.

1. Letter Knowledge: In the ECLS-B, letter knowledge was examined by having children identify both letter names and letter sounds. In the present study, three subtest scores associated with the PALS preK Alphabet Knowledge task were standardized and aggregated to form one indicator of Letter Knowledge (i.e., Upper Case, Lower Case, and Letter Sound Naming).

2. Print Conventions: In the ECLS-B, print conventions were measured through a series of questions targeting young children’s understanding of what print represents and how it works (e.g., how to orient the book, discriminating print from pictures, reading left to right). In the present study, two subtests derived from the Print and Story Concepts tasks were standardized and aggregated to form one indicator of Print Conventions (i.e., book knowledge and print knowledge).

3. General Early Literacy Ability: In the ECLS-B, early literacy skills were evaluated using a combination of letter recognition, receptive and expressive language, letter sounds, and early reading. In the present study, the Get Ready to Read scores were standardized and combined...
with the indicators of Letter Knowledge and Print Conventions to match the procedure used by in the ECLS-B methods.

### Table 10. Effect Size Estimates Contrasting National Estimates with Study Estimates

<table>
<thead>
<tr>
<th>Identified Gaps Between Different Groups of Children</th>
<th>Size of the Gap&lt;sup&gt;a&lt;/sup&gt;</th>
<th>BTL: T vs. C</th>
<th>BTL: M vs. C</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Letter Knowledge</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA vs. AA achievement gap</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA vs non-EA achievement gap</td>
<td>0.39</td>
<td>0.62</td>
<td>0.18</td>
</tr>
<tr>
<td>Low SES vs. Middle SES achievement gap</td>
<td>0.39</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low SES vs. High SES achievement gap</td>
<td>0.98</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Print Conventions</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA vs. AA achievement gap</td>
<td>0.21</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA vs non-EA achievement gap</td>
<td>0.35</td>
<td>0.69</td>
<td>0.24</td>
</tr>
<tr>
<td>Low SES vs. Middle SES achievement gap</td>
<td>0.37</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low SES vs. High SES achievement gap</td>
<td>0.96</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Combined Early Literacy Task</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA vs. AA achievement gap</td>
<td>0.25</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EA vs non-EA achievement gap</td>
<td>0.39</td>
<td>1.51</td>
<td>0.66</td>
</tr>
<tr>
<td>Low SES vs. Middle SES achievement gap</td>
<td>0.40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low SES vs. High SES achievement gap</td>
<td>1.00</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<sup>a</sup>All metrics in this table represent standardized effect sizes (i.e., Cohen’s d)

**Effect Sizes and the Achievement Gap**

As indicated in the tables above, the BTL Classroom Intervention provided a powerful, effective, and engaging intervention that helped economically disadvantaged preschoolers bridge the achievement gap between their performance on each of the three indicators and the scores of their more advantaged peers. The BTL Classroom Intervention was designed to help reduce the achievement gaps typically found between minority and majority subgroups of children and between children from lower SES and middle SES families. Recently, researchers have proposed that the effect sizes associated with an intervention or educational reform should be compared with the size of known achievement gaps in order to judge whether an intervention is meaningful or worth implementing (Hill et al., 2008; Konstantopoulos & Hedges, 2008). Across the three indicators of key early literacy skills, the estimated impact of the BTL Classroom Intervention was remarkably consistent and universally larger than the documented gaps between European American and African American preschoolers or between preschoolers from low SES families and from middle SES families.
Putting the Last Pieces of the Puzzle Together

Establishing that an effect exists and is similar in magnitude to the size of the difference between two populations of interest (e.g., low SES vs. middle SES children) is important. Comparing these two effects (i.e., comparing the magnitude of the differences in favor of the intervention with the magnitude of the differences between low SES and middle SES) helps to establish that the intervention can normatively result in an effect that is similar to or bigger than the observed achievement gap. The next step in the puzzle is determining whether the observed gains meet the criterion level associated with a particular skill. Did participating in the intervention help preschoolers obtain scores that were either at or above the established benchmarks or inside the expected developmental ranges associated with a particular skill. For example, Get Ready to Read scores in preschool that are at or above 11 points have been found to be predictive of later conventional reading success. At the start of the intervention, all children on average scored 9.21. At the post-test, those who participated in the intervention scored above the 11-point benchmark (i.e., treatment = 11.79; maintenance = 11.29). Across 20 comparisons, preschoolers in the treatment group reached established benchmarks 83.3% of the time; preschoolers in the maintenance group reached the established benchmarks 77.8% of the time; and children in the control group reached the established benchmarks 66.7% of the time. See Figure 19.

![Figure 19. Benchmark Comparisons Across Groups](image-url)
Discussion

*Did the BTL Classroom Intervention Impact Teachers’ Literacy-Related Behaviors and the Organization of Their Classrooms?*

Teachers using the BTL Classroom Materials and receiving the mentoring support evidenced significant changes across all four subscales of the ELLCO. The Literacy Environment subscale involved an assessment of the classroom’s layout and contents including availability, content, and diversity of reading, writing, and listening materials. The General Classroom Environment subscale measured the organization of the classroom, a child’s opportunities for choice and self-initiative, appropriate classroom management strategies, and an overall positive climate. The Literacy Activities subscale measured the number and length of full-group and one-to-one book reading sessions as well as whether children were writing on their own or with assistance as well as whether teachers were modeling positive writing behaviors. The Language, Literacy, and Curriculum subscale measured oral language facilitation, presence of and approaches to reading, writing, and curriculum integration, and recognizing diversity and bridging the home-school environments. Maintenance classrooms were able to sustain high-quality literacy environments from previous years to this project year and treatment classrooms, with extended mentoring and support, were able to achieve high-quality literacy environments. Changes were found from the most superficial structural components to increased quantity and higher quality enriching interactions that have been previously linked to children’s optimal growth in literacy and other cognitive and social domains. To determine whether these positive changes translated into gains on key early literacy skills, children’s abilities across a wide array of early literacy measures were measured.

*Did the BTL Classroom Intervention Impact Children’s Early Literacy Skills?*

Children between 46 months and 59 months whose teachers were new to the BTL Classroom Intervention obtained higher scores when compared with their maintenance and control group peers on nearly all outcomes across Indicators of Language Development, Letter Knowledge, Phonemic Awareness, Print Conventions, and a Combined Early Literacy Skills Screener. Specifically, 4 year old children identified more Picture Names, Upper and Lower Case letters, and Letter Sounds. Their ability to name Upper and Lower Case Letters and Letter Sounds was faster than their peers in the maintenance and control groups. They were better able to demonstrate the mechanics of reading (i.e., Print Knowledge) and to identify key actual and inferential story content. Finally, they scored higher than their peers on a combined early literacy skills screener called the Get Ready to Read achieving scores that are predictive of future reading success.

Older children whose teachers had previously participated in the BTL Classroom Intervention (i.e., Maintenance Group) obtained higher scores when compared with their treatment and control group peers on Indicators of Letter Knowledge and the Combined Early Literacy Skills Screener. Specifically, children who were older than 5 ¼ years were able to identify more Upper and Lower Case letters and Letter Sounds, to name Upper Case letters and Letter Sounds more quickly, and to accurately answer items on the Get Ready to Read screener. While there were several instances when older children in the control group outperformed their older peers in the treatment group, the maintenance group, or both,
there were no consistent patterns, suggesting that these differences were more idiosyncratic than systematic.

Changing teacher behavior is a challenging undertaking. In this intervention, teachers were provided with extensive mentoring and support, high quality books and materials, and an easy-to-use supplemental literacy curriculum. As noted above, children who were in either BTL Classroom intervention group evidenced higher scores on a variety of tasks. Understanding the patterns of findings was enhanced when a child’s age was considered. Younger children obtained higher scores when they were in classrooms where teachers were using the intervention for the first time. It is possible that these children are attending a high-quality child care environment for the first time. As such, because these children started the project with scores indicating a serious risk of later reading failure, it is likely that any intervention (particularly one that is known to be both effective and engaging) is going to have powerful effects on these young children’s skills. These children may be experiencing a literacy- and language-enriching environment on a regular basis for the first time. In fact, effect sizes (i.e., an index of the magnitude or size of an effect) for the younger children averaged between 0.34 and 0.44, indicating that participation in the intervention produced changes that were roughly 1/3rd to 2/5th of a standard deviation higher than their control group peers. Meaningful change can also be thought of as the amount of variance in child outcomes that was attributable to the intervention. This intervention accounted for between 10.9% and 19.4% of the variance in child outcomes.

Unlike younger children, older children did better when they were in classrooms with teachers who had previously participated in the project. In these classrooms, teachers were able to maintain a high degree of quality from one year to the next. This quality and stability may have provided an environment that supported these children’s skills particularly because they have already spent at least a year at their current child care center and have experienced a year without any extra language- or literacy-enriching experiences. The magnitude or size of the effect for older children averaged 0.54, a relatively sizable effect accounting for 29.4% of the variance in child outcomes. It is likely that the intervention effects are especially pronounced because teachers who had prior experience with the intervention had developed a certain facility with the materials and were better able to help develop their older learners’ skills. Further, older learners who have not had the benefit of any other literacy instruction are, compared with their younger peers, at a marked disadvantage. It may be that teachers new to the BTL intervention were struggling not only with changing the general literacy environment in the classroom but also trying to incorporate the special needs of older preschoolers who have an extra year of disadvantage behind them. The average magnitude of the effect found for older children is particularly exciting as it ranges between 10% and 20% higher than the effects found for younger children. This means that experienced teachers were successful at helping these older children catch up to their younger peers as well as to standards or benchmarks that children this age need to obtain to ensure conventional reading success.

This study contributes additional evidence regarding the effectiveness of using BTL alone (i.e., pure television exposure; Linebarger et al., 2004; Uchikoshi, 2006) and in combination (i.e., exposure and supplemental classroom materials; Linebarger, 2006; Prince et al., 2001) to support young children’s burgeoning early literacy skills. More importantly, BTL has consistently helped young children who are at substantial risk for later reading failure acquire the key early literacy skills needed for school and later life success. The power of this intervention is a function of its engaging characters, stories, and sketches in combination with early literacy content that is carefully interwoven throughout the televised content as well as infused into teacher materials and classroom manipulatives. Children from low-income and minority backgrounds spend more time watching television and report that the experience is more relaxing and of more value in comparison to children from middle income and majority backgrounds. As
such, BTL represents a powerful tool for these children when they are acquiring the key early literacy skills necessary for learning to read. For instance, both maintenance and treatment group children at the post-test obtained Get Ready to Read scores that were predictive of later conventional reading success (i.e., scores greater than 11). The gains from pre-test to post-test averaged 24% for treatment children, 27% for maintenance children, and 14% for control group children. The standard developmental gain from fall to spring (without any intervention) is approximately 15% in the general population.

**Other Thoughts Regarding the BTL Classroom Intervention**

In years past, effects associated with viewing BTL or participating in a classroom intervention were moderated by a child’s initial risk status (e.g., Linebarger et al., 2004; Prince et al., 2001). To explain the moderated effects, analyses associated with the child’s risk for later reading failure were computed. In the current sample, 32.6% of preschoolers were considered at significant risk for later reading failure; 30.7% of preschoolers were considered at marginal risk; and 36.7% of preschoolers were considered not to be at risk. It is not clear whether Mississippi preschoolers’ literacy skills are generally improving or whether the sample recruited for the 2007-2008 project was less at risk than average. Future evaluations should continue to monitor this emerging trend. Regardless of the reason, it is encouraging that the percentage of children at risk for later reading failure has dropped from nearly all but three children at-risk (Prince et al., 2001) to just about one-third of children at-risk.

**Recommendations for Future Research**

There are some research issues that should be addressed when conducting additional intervention studies. Selection and sampling strategies are important issues to any research design, particularly one that uses a quasi-experimental framework. Because it is highly unlikely that children can be randomly assigned to groups (i.e., children are in classrooms and classrooms are assigned to condition), it is imperative that classrooms involved in a study are randomly assigned to a treatment or control condition. While preparing the available data for analyses, it was determined that, despite best efforts to match control and treatment classrooms on a variety of demographic variables (i.e., poverty rates, location, target age, teacher education), there were still some variables that differed significantly across the groups (e.g., child’s age, pre-test ELLCO scores). Although these variables were statistically controlled, it is important to interpret the results presented here with caution. It is possible that differences are due to other variables that were measured here (e.g., older children in the treatment group may naturally have stronger literacy skills; classrooms with a better literacy environment may provide children in those classrooms with an advantage) or other variables that were not (e.g., parental education). To strengthen the research design and validate the findings of previous studies, this study, and future studies, it is important to make a concerted effort to keep the process as rigorous as possible.

In addition to changes in the way the research is conducted, additional changes may be necessary in the ways in which the intervention is delivered, particularly for teachers who have previously participated in the intervention. It may be that first year teachers received extensive mentoring and support while second year teachers receive less support. Further, the largest changes on ELLCO scores are associated with the general literacy environment (e.g., displays, books). These environmental variables were quickly and substantively changed during the first year of intervention. The more challenging behaviors to change are those related to the language, literacy, and curricular environment including the quantity and quality of language- and literacy-promoting strategies. It is possible that encouraging these types of
interactions will take more mentoring support than is currently possible or available. One way to examine this possibility is to code mentor field notes associated with both treatment and maintenance classrooms. Currently, UPenn staff are working on this coding scheme.

Finally, children whose teachers were supplementing their regular instruction with the BTL classroom intervention demonstrated gains across most early literacy measures with the exception of vocabulary knowledge. In its on-air program, BTL focuses heavily on code-related skill acquisition. Characters spend time sounding words out and reading aloud, small segments include words that morph into other words in the same word families or with the same vowel or consonant sounds. The greater emphasis on phonological and phonemic awareness skills makes these word properties more salient and may leave little time left over or, alternatively, little cognitive capacity left over to encode new words for both their code-related properties and their vocabulary or oral language properties. It would be expected that over time, as children are repetitively exposed to the code-related properties of words, they would be able to devote less time to these code properties and more time to the conceptual understanding of the words.

To Sum It Up

It has long been known that high quality early childhood education programs help young children experiencing significant and chronic poverty and disadvantage bridge the gap between their readiness for school and their more advantaged peers’ readiness for school (e.g., Perry Preschool Project, Abecedarian Project). These programs are typically composed of services designed to comprehensively surround children and their families with, among other services (e.g., health and social services), cognitively stimulating toys and materials as well as positive and sustained language-, literacy-, and prosocial-promoting experiences and interactions. The specific early literacy achievement gaps presented in Table 10 indicate that African American preschoolers are underperforming their European American peers by approximately 1/4th of a standard deviation across the 3 indicators of early literacy achievement while children from low SES homes are underperforming their peers living in middle SES families by 2/5th of a standard deviation. While these gaps may seem relatively small in preschool, there is substantial evidence that they are persistent, resistant to intervention, and widening as children progress through school. Walker et al (1994) found that early language deficits identified at age 3 and linked to family SES predicted language development, verbal ability, and academic achievement throughout the early elementary school years. Stanovich (1986) labeled this phenomenon as the Matthew effect, proposing that children who had more and positive early literacy experiences are more frequently and intensively rewarded for these early accomplishments while children who lack these cumulative experiences and successes find reading less enjoyable, struggle to make sense of what they are reading, and are often unable to benefit from and even utilize new educational experiences effectively. Essentially, children whose early literacy achievements are slowed or delayed progressively decline, children whose early literacy achievements come quickly and frequently progressively improve, resulting in ever-widening differences between their reading, school, and life trajectories.

The BTL Classroom Intervention described and evaluated in this study has the potential to profoundly alter young economically disadvantaged children’s early literacy achievements, bridging the gap between these achievements and their faster peers who are either European American or living in middle and upper SES homes while, at the same time, placing them on a more positive literacy

---

40 The Matthew effect refers to a passage in the Bible attributed to Jesus where he was presenting a parable associated with talents: “For unto every one that hath shall be given, and he shall have abundance: but from him that hath not shall be taken away even that which he hath.” Matthew 25:29
trajectory. While there are no current long-term studies of whether these child achievements are maintained, extrapolating from other successful early literacy interventions with similar short-term effects suggests that these gains will be maintained and will help these at-risk preschoolers to switch from the “poor get poorer” trajectory to the ‘rich get richer’ trajectory. The evidence is quite clear that this intervention (i.e., materials accompanied by sustained and intensive mentoring support) substantially alters teachers’ behaviors. These behaviors, in turn, create daily environments for children that include more and higher quality language- and literacy-promoting interactions. These changes have been maintained up to three years later (e.g., Figure 6). Favorable changes in classroom environments and teacher behaviors are closely linked in this study, as well as in previous studies, to positive changes and accelerating growth of at-risk preschoolers’ early literacy skills (Linebarger, 2006, 2007; Linebarger et al., 2004; Prince et al., 2001). Based on these substantial, pervasive, and consistent classroom and teacher effects, it is highly likely that children exposed to this intervention as preschoolers will continue to benefit not only by engaging in literacy experiences and interactions provided by their teachers but also by actively choosing to “select, shape, and evoke their own environments” (p. 381; Stanovich, 1986). Specifically, exposure to and active participation in the BTL Classroom Intervention has the potential to shift these at-risk children’s trajectories to mirror more closely the trajectories of children who are academically successful by providing developmentally appropriate and highly engaging content that, through a series of self-reinforcing experiences and events, supports their burgeoning early literacy skills and, perhaps even more importantly, increases their desires and motivations to continuously and actively solicit new literacy-specific as well as educationally-general experiences (Stanovich, 1986).
References


Konstantopoulos, S., & Hedges, L. V. (2008). How large an effect can we expect from school reforms? *Teachers College Record*, 110, 1613–1640


The contents of this document were developed under a cooperative agreement between the US Department of Education, the Corporation for Public Broadcasting, and the Public Broadcasting System for the Ready to Learn Initiative, PR# U295A050003. However, these contents do not necessarily represent the policy of the Department of Education and you should not assume endorsement by the Federal Government.