MAKING a DIFFERENCE 2003/2004: EXECUTIVE SUMMARY

Prepared by:
Stephanie Wehry, Ph.D.
Assistant Director for Research
Florida Institute of Education
at the
University of North Florida
# Table of Contents

## Introduction
- ELLM at a Glance ........................................... 2
- Measurement .................................................. 3
- Phonological Awareness Population of ELLM Children 4
- Guidelines for Statistical Procedures .................. 5

## ELLM Children .................................................. 6

## Results
- TERA-3 .................................................................. 7
  - General Summary of TERA-3 Results ................... 12
- ALRI ..................................................................... 13
  - Children Selected for Targeted Instruction in Phonological Awareness: TERA-3 Results 14

## Cross-Cohort Comparison
- Learning Hierarchy in Emergent Literacy Abilities 17
- Declining Kindergarten TERA-3 Meaning Subtest Scores 18
- Three Years of ELLM Data ................................. 18
- Implications of the Cross-Cohort Comparison ....... 20

## Summary of Cohort-Specific Results
- Preschool ......................................................... 22
- Kindergarten .................................................... 23
- First Grade ...................................................... 24
Making a Difference 2003/2004: Executive Summary

Introduction

Improving the achievement of all children, particularly those considered at high risk of academic failure, continues to be a local, state, and national priority. We know when children come to school healthy, eager, and ready to learn, they are much more likely to succeed in school and become responsible, productive citizens. Being ready for school begins long before children enter kindergarten. Readiness includes the development of skills and knowledge in many areas. Specifically, developing the language and literacy skills needed to become proficient readers is crucial to success in school. Any child who does not learn to read early and read well will not easily master other important skills and knowledge. Yet, as many as one third of children entering kindergarten are underprepared for the challenges they will face and are at risk of school failure.

The Early Literacy and Learning Model (ELLM) is designed to improve the language and literacy skills of 3-year-old, 4-year-old, kindergarten, and first-grade children who live in low-income communities and who are often at risk of academic failure. This report provides detailed information about literacy-related outcomes of ELLM children and provides a summary of the successes of the program during the 2003/2004 school year.
ELLM at a Glance

ELLM has been used with urban and rural children in a wide array of childcare centers (subsidized, faith-based, Head Start, and Early Intervention Pre-Kindergarten), as well as with kindergarten and first-grade children in high-needs urban elementary schools.

For program evaluation purposes, ELLM children are organized into three cohorts: 4-year-old preschool, kindergarten, and first-grade children. While ELLM serves both 3- and 4-year-old preschoolers, only 4-year-old children eligible for Florida public kindergarten the following school year are assessed. Hereafter, preschool refers only to children assessed as part of the evaluation.

The Test of Early Reading Ability-3, Form A (TERA-3) was administered in a pretest/posttest design to measure children’s reading readiness and to evaluate the effectiveness of ELLM. Pretest data were collected in early September and posttest data were collected from mid-April to early May.

Additionally, the locally developed Alphabet Letter Recognition Inventory (ALRI) was administered to 4-year-old preschool and kindergarten children in a pretest/posttest design. The pretest results were used to inform instructional practice and the posttest results to evaluate the effectiveness of ELLM.

Furthermore, children whose TERA-3 pretest scores ranked in the lowest quartile were selected for targeted instruction in phonological awareness.
Measurement

Two instruments were used to measure children’s reading readiness, the Test of Early Reading Ability-3 (TERA-3) and the Alphabet Letter Recognition Inventory (ALRI). Trained assessors using scannable forms administered the TERA-3 to all children in a one-on-one setting. Tests were electronically scored. Classroom teachers using scannable forms administered the ALRI tests to 4-year-old preschool and kindergarten children, and the tests were electronically scored.

TERA-3

The TERA-3 is a norm-referenced test that assesses components of early developing reading skills, including familiarity with the letters of the alphabet and numerals, discovery of the arbitrary conventions used in reading and writing English, and recognizing that print conveys information, ideas, and thought. The test is composed of three subtests: the Alphabet subtest, the Conventions of Print subtest, and the Meaning subtest, each measuring one of the three components. A composite score, the Reading Quotient, is the unweighted sum of the three subtest scores.

Table 1 provides categories delimiting the lower, middle two, and upper quartiles of the TERA-3 Reading Quotient and TERA-3 subtest scores of the normative population. These categories are used in this report to indicate the range of early reading ability of the ELLM children.

Table 1
Scale of TERA-3 Reading Quotient and TERA-3 Subtests by Ability Categories

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>Very Poor</th>
<th>Poor</th>
<th>Below Average</th>
<th>Average</th>
<th>Above Average</th>
<th>Superior</th>
<th>Very Superior</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of Scores</td>
<td>2.3</td>
<td>6.9</td>
<td>16.1</td>
<td>49.5</td>
<td>16.1</td>
<td>6.9</td>
<td>2.3</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Percentiles</th>
<th>2nd or lower</th>
<th>2nd to 9th</th>
<th>9th to 25th</th>
<th>25th to 75th</th>
<th>75th to 91st</th>
<th>91st to 98th</th>
<th>98th or higher</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percent of Scores</td>
<td>2.3</td>
<td>6.9</td>
<td>16.1</td>
<td>49.5</td>
<td>16.1</td>
<td>6.9</td>
<td>2.3</td>
</tr>
</tbody>
</table>
**ALRI**

The ALRI is a locally developed test measuring children’s ability to recognize the upper- and lowercase letters of the alphabet when presented in non-alphabetic order. Children’s teachers administered the test one-on-one to children using alphabet letter flashcards. The uppercase letters were presented first in a fixed, non-alphabetic order, followed by the lowercase letters.

ALRI scores are reported in four recognition categories, 0 to 13 letters, 14 to 26 letters, 27 to 39 letters, and 40 to 52 letters. The *Early Childhood Longitudinal Study-Kindergarten* (ECLS-K) provides national benchmarks for alphabet letter recognition. Using a random sample of upper- and lowercase letters, scores are categorized as proficient if children recognized at least 75% of the sampled letters.\(^1\) Reports from the ECLS-K data also state 66% of the children (39% of children living in low-income families) entering kindergarten for the first time were proficient.\(^2\) In this report, ALRI posttest scores are compared to these national benchmarks.

**Phonological Awareness Population of ELLM Children**

Researchers discovered over the history of ELLM that children with very low scores on the TERA-3 pretests showed little, if any, improvement on the posttest. Therefore, beginning in 1999/2000 and continuing on through 2003/2004, children with very low TERA-3 pretest scores were selected for targeted instruction in phonological awareness.

---


Guidelines for Statistical Procedures

All statistically significant mean differences are relative to a Type I error rate (alpha) of .05. However, because of the large sample sizes, small - even trivial - differences can be statistically different. To determine the importance of differences, effect sizes are reported. Effect size is the mean difference relative to the standard deviation of the test. Cohen classified effect sizes of 20-to-49 percent of a standard deviation as small, effect sizes between 50 and 79 percent of a standard deviation as medium, and effect sizes of 80 percent of a standard deviation or higher as large. Differences resulting in small, medium, and large effect sizes are meaningful.

Moreover, Whitehurst and Massetti, in a critique of Head Start, noted when evaluation designs lack control or comparison groups, small effect sizes of 20-to-25 percent of a standard deviation should not be attributed to programs. Small effects could easily be associated with regression toward the mean, increased familiarity with tests and assessment procedures in general, or ordinary maturation and experiences. Because this evaluation of ELLM does not involve a comparison or control group, only effects larger than one third of a standard deviation are attributed to the children’s participation in ELLM.

ELLM CHILDREN

Table 2 presents the demographic information of the ELLM children by cohort.

Table 2
*Demographic Information of the ELLM Children by Cohort*

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Number</th>
<th>Gender % Boys</th>
<th>Age % Older</th>
<th>Ethnicity % Black</th>
<th>Attrition Rate</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschool</td>
<td>129</td>
<td>52</td>
<td>-</td>
<td>92</td>
<td>13%</td>
</tr>
<tr>
<td>Kindergarten</td>
<td>188</td>
<td>47</td>
<td>16</td>
<td>100</td>
<td>19%</td>
</tr>
<tr>
<td>First Grade</td>
<td>184</td>
<td>47</td>
<td>18</td>
<td>100</td>
<td>18%</td>
</tr>
</tbody>
</table>

There were more children in kindergarten and first grade than in preschool. When combined across cohorts, 49% of the ELLM children were boys and 51% were girls. However, the preschool cohort had a larger percentage of boys than girls. The preschool cohort was also the only ethnically diverse cohort. The percentage of children who may have been retained in grade was very similar across kindergarten and the first grade. The rate of attrition was larger in kindergarten and first grade than in preschool. Additionally, there was evidence that kindergarten children whose TERA-3 Reading Quotient, Alphabet, and Meaning subtests pretest scores were low were more likely to be unavailable for testing in the spring. Therefore, results from the kindergarten cohort should be interpreted cautiously. There was also evidence that the first-grade children whose TERA-3 Alphabet subtest pretest scores were low were more likely to be unavailable for testing in the spring. However, scores of the other TERA-3 subtests were not affected by the attrition of the first-grade children.
RESULTS

TERA-3

Table 3 presents summary statistics and ANOVA results for the TERA-3 Reading Quotient and subtest scores by cohorts.

Table 3
Summary Statistics and ANOVA Results for the TERA-3 Reading Quotient and Subtest Scores

<table>
<thead>
<tr>
<th>Cohort</th>
<th>TERA-3 Test</th>
<th>Pretest Mean Percentile</th>
<th>Posttest Mean Percentile</th>
<th>Probability Means Differ</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschool</td>
<td>Reading Quotient</td>
<td>26th</td>
<td>37th</td>
<td>&lt;.0001*</td>
<td>0.306φ</td>
</tr>
<tr>
<td></td>
<td>Alphabet</td>
<td>46th</td>
<td>67th</td>
<td>&lt;.0001*</td>
<td>0.532φφ</td>
</tr>
<tr>
<td></td>
<td>Conventions</td>
<td>22nd</td>
<td>26th</td>
<td>.0507*</td>
<td>0.150</td>
</tr>
<tr>
<td></td>
<td>Meaning</td>
<td>27th</td>
<td>29th</td>
<td>.5921</td>
<td>-</td>
</tr>
<tr>
<td>Kindergarten</td>
<td>Reading Quotient</td>
<td>13th</td>
<td>21st</td>
<td>&lt;.0001*</td>
<td>0.340φφ</td>
</tr>
<tr>
<td></td>
<td>Alphabet</td>
<td>32nd</td>
<td>60th</td>
<td>&lt;.0001*</td>
<td>0.660φφ</td>
</tr>
<tr>
<td></td>
<td>Conventions</td>
<td>13th</td>
<td>21st</td>
<td>.0961</td>
<td>0.350φφ</td>
</tr>
<tr>
<td></td>
<td>Meaning</td>
<td>14th</td>
<td>10th</td>
<td>.0052*</td>
<td>-0.230</td>
</tr>
<tr>
<td>First-Grade</td>
<td>Reading Quotient</td>
<td>18th</td>
<td>28th</td>
<td>&lt;.0001*</td>
<td>0.360φφ</td>
</tr>
<tr>
<td></td>
<td>Alphabet</td>
<td>51st</td>
<td>46th</td>
<td>.0701</td>
<td>0.118</td>
</tr>
<tr>
<td></td>
<td>Conventions</td>
<td>26th</td>
<td>39th</td>
<td>&lt;.0001*</td>
<td>0.353φφ</td>
</tr>
<tr>
<td></td>
<td>Meaning</td>
<td>6th</td>
<td>16th</td>
<td>.0618*</td>
<td>0.596φφ</td>
</tr>
</tbody>
</table>

* Indicates significant difference at the $\alpha = .05$ level.  
φ Denotes a small effect.  
φφ Denotes a medium effect.  
φφφ Denotes a large effect.  
+ Denotes an effect that may be attributed to children’s participation in ELLM.  
 Denotes a percentile ranking higher than the 25th percentile.  
 Denotes a percentile ranking higher than the 50th percentile, the national average.

There are seven effects large enough to be attributed to the children’s participation in ELLM classrooms: one in the preschool cohort, three in the kindergarten cohort, and three in the first-grade cohort. Additionally, the TERA-3 Alphabet subtest posttest mean scores of the preschool and kindergarten children rank at least 10 percentiles above the national average, and the posttest mean score of the first grade children nearly matches the national average. Table 4 on the next page more concisely summarizes the effects attributable to the children’s participation in ELLM.
Table 4  
*Summary of the Improvements in Emergent Literacy That Can Be Attributed to Children’s Participation in ELLM*

<table>
<thead>
<tr>
<th></th>
<th>Reading Quotient</th>
<th>Alphabet</th>
<th>Conventions</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschool</td>
<td></td>
<td>0.532⁺</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kindergarten</td>
<td>0.340⁺</td>
<td>0.660⁺</td>
<td>0.350⁺</td>
<td>-</td>
</tr>
<tr>
<td>First-Grade</td>
<td>0.360⁺</td>
<td>-</td>
<td>0.353⁺</td>
<td>0.596⁺</td>
</tr>
</tbody>
</table>

The numbers in the cells are the effect sizes.
- Indicates an effect attributable to the children’s participation in ELLM classrooms.
- Indicates there were no more than 30% of the posttest scores ranked below the 25th percentile.
- Indicates there were at least 20% fewer scores ranked below the 25th percentile in the spring.

The preschool summary indicates the only preschool meaningful improvement attributable to the children’s participation in ELLM was in the emergent literacy abilities measured by the TERA-3 Alphabet subtest. The kindergarten children made meaningful improvement attributable to their participation in ELLM in two areas: the Alphabet and the Conventions of Print subtests. The first-grade children also made meaningful improvement attributable their participation in ELLM in two areas: the TERA-3 Conventions of Print and Meaning subtests. Therefore, the preschool and kindergarten children’s achievement overlaps in the abilities of the Alphabet subtest, and the kindergarten and first-grade children’s achievement overlaps in the abilities of the Conventions of Print subtest.

To better understand where the improvement in reading readiness occurred, TERA-3 scores are displayed in seven ability categories: three categories representing the lowest 25 percentiles, one category representing the middle 50 percentiles, and three categories representing the highest 25 percentiles (page 3, see Table 1). Figure 1 displays the TERA-3 Alphabet subtest posttest scores of the preschool and kindergarten children, Figure 2 displays the TERA-3 Convention of Print subtest posttest scores of the kindergarten and first-grade children, and Figure 3 displays the TERA-3 Meaning subtest posttest scores of the first-grade children.
Both preschool and kindergarten children have fewer posttest scores in the bottom quartile than in the national normative population (light blue bars) and the percentage of kindergarten scores (dark blue bars) is actually one half the national percentage. Moreover, the percentage of kindergarten scores in the top quartile matches the national population. Over 40% of the preschool scores (striped bars) are in the top quartile, and 24% of their scores rank above the 90th percentile. Both cohorts match the national normative population, indicating the ELLM children closed the achievement gap in the emergent literacy abilities of alphabet letter knowledge.
Figure 2.
The percentage of TERA-3 Conventions of Print subtest posttest scores of the kindergarten (n=188), first-grade (n=184), and national normative population in the seven categories of the ability scale.

The percentage of kindergarten scores (dark blue bars) in the bottom quartile is more than twice the percentage in the national normative population (light blue bars). The distribution of the kindergarten scores remains centered in the bottom quartile as indicated by the 21st percentile ranking of the posttest mean score (see page 7, Table 3). The percentage of first-grade scores (violet bars) almost matches the national normative population percentage; however, there are fewer scores in the top quartile and more scores in the Average category than in the national population. The distribution of the first-grade posttest scores is centered in the Average category as indicated by the 39th percentile ranking of the posttest mean score (see page 7, Table 3).
The percentage of first-grade scores (violet bars) in the bottom quartile is more than twice the percentage in the national normative population (light blue bars); however, the percentage is almost 30% less than at the beginning of the school year. In the spring, more scores have moved into the *Average* category, and 12% more scores are in the top quartile. However, the distribution of first-grade scores remains centered in the bottom quartile as indicated by the 16\textsuperscript{th} percentile ranking of the posttest mean score (see page 7, Table 3).

The information presented in Figures 1-3 does not indicate the changes made because the distributions of the initial status scores are not presented. Therefore, Figure 4 on the next page displays the changes in percentile ranking of the pretest and posttest mean scores across TERA-3 tests and cohorts.
Figure 4 shows the changes in relative ranking for the TERA-3 subtest mean scores across the cohorts. Clearly, the largest increases in ranking relative to the national normative population are the TERA-3 Alphabet subtest improvements of the preschool and kindergarten children. However, with the exception of the Alphabet subtest, the most consistent improvements across the measured emergent literacy abilities occur in the first grade. The most uneven improvements occur in kindergarten, and the preschool children show fairly flat improvements in emergent literacy abilities other than alphabet letter knowledge.

**General Summary of TERA-3 Results**

By far, the preschool and kindergarten children’s improvements on the TERA-3 Alphabet subtest were the largest improvements across TERA-3 subtests and cohorts. The kindergarten and first-grade children’s improvements on the TERA-3 Conventions of Print subtest were much more modest, in part, because of the lower initial status of the scores. The first-grade improvement of 60% of a standard deviation on the TERA-3 Meaning subtest resulted in the posttest mean score ranking at the 16th percentile; however, more than one third of the posttest scores ranked at or below the 2nd percentile.
ALRI

The children in preschool and kindergarten also have ALRI posttest scores. To determine the range of alphabet letter recognition ability of ELLM preschool and kindergarten children, ALRI scores are displayed in Figure 5 using four recognition categories: 0-13 letters, 14-26 letters, 27-39 letters, and 40-52 letters recognized.

Figure 5.
Alphabet Letter Recognition Inventory posttest scores of 163 ELLM kindergarten and 142 preschool children.

ALRI benchmarks from the Early Childhood Longitudinal Study – Kindergarten:
✓ Recognition of at least 75% of the sampled letters is defined as proficient.
✓ 66% of children in the national ECLS-K sample were proficient upon entering kindergarten.
✓ 39% of children living in low-income families in the national ECLS-K sample were proficient upon entering kindergarten.

At the end of the preschool year:

- The typical ELLM preschool child recognized 82% of the letters; therefore was proficient.
- 74% of the ELLM preschool children were proficient.
- 39% of the ELLM preschool children recognized all upper- and lowercase letters.
- 5% of the ELLM preschool children recognized fewer than 10 letters.
At the end of the kindergarten year:

- The typical ELLM kindergarten child recognized 94% of the letters; therefore was proficient.
- 93% of the ELLM kindergarten children were proficient.
- 69% of the ELLM kindergarten children recognized at least 50 letters.

**Children Selected for Targeted Instruction in Phonological Awareness: TERA-3 Results**

Children whose TERA-3 pretest scores ranked in the lowest quartile were selected for targeted instruction in phonological awareness. Table 5 presents the demographic information of the selected ELLM children.

<table>
<thead>
<tr>
<th>Table 5</th>
<th>Demographic Information of the ELLM Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cohort</td>
<td>Number</td>
</tr>
<tr>
<td>Preschool</td>
<td>24</td>
</tr>
<tr>
<td>Kindergarten</td>
<td>68</td>
</tr>
<tr>
<td>First Grade</td>
<td>36</td>
</tr>
</tbody>
</table>

The percentage of kindergarten children selected for targeted instruction in phonological awareness is almost twice the percentage in the other two cohorts, indicating the TERA-3 Reading Quotient and Alphabet subtest pretest scores of the kindergarten children were lower than the pretest scores of the other two cohorts. In all cohorts, a larger percentage of boys were selected for targeted instruction in phonological awareness than in the general ELLM core population, and all of the selected children were Black. Additionally, over 50% of the first graders and 25% of the kindergarteners were older than the traditional age of their cohort peers. Because the TERA-3 scores are normed by the age of the children rather than their grade level, these children could actually be performing adequately if measured against children in their cohort peers rather than against their same-aged peers.

Table 6 summarizes the TERA-3 results of the children selected for targeted instruction in phonological awareness by subtest and cohort.
Table 6
*Summary of the Improvements in Emergent Literacy That Can Be Attributed to Children’s Participation in ELLM and Targeted Instruction in Phonological Awareness*

<table>
<thead>
<tr>
<th></th>
<th>Reading Quotient</th>
<th>Alphabet</th>
<th>Conventions</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschool</td>
<td>0.764 +</td>
<td>1.125 +</td>
<td>-</td>
<td>0.417 +</td>
</tr>
<tr>
<td>Kindergarten</td>
<td>0.512 +</td>
<td>1.201 +</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>First-Grade</td>
<td>0.613 +</td>
<td>-</td>
<td>0.935 +</td>
<td>-</td>
</tr>
</tbody>
</table>

The numbers in the cells are the effect sizes.

+ Indicates an effect attributable to the children’s participation in ELLM classrooms.

- Indicates there were no more than 35% of the posttest scores ranked below the 25th percentile.

- Indicates there were at least 20% fewer scores ranked below the 25th percentile in the spring.

Targeted instruction in phonological awareness primarily addresses the emergent literacy abilities measured by the TERA-3 Alphabet subtest. Improvements across other TERA-3 subtests, especially in the kindergarten and first grade could result from participation in ELLM and targeted instruction in phonological awareness plus additional intensive literacy instruction provided in these high-needs elementary schools. However, because targeted instruction in phonological awareness primarily addresses alphabet letter knowledge, the only distributions of the preschool and kindergarten children’s Alphabet subtest posttest scores are displayed in Figure 6 on the following page using the seven ability categories of the TERA-3 (see page 3, Table 1).
In the fall of the school year, almost 100% of the TERA-3 Alphabet subtest scores were in the bottom quartile because of the selection criterion for participation in targeted instruction in phonological awareness. These pretest scores from both cohorts are seen as measures of the children’s opportunity to learn. The posttest scores indicate just one third of the scores remain in the bottom quartile at posttest with the preschool scores even moving into the top quartile. Most of the kindergarten change is from the bottom quartile to the Average category which spans the inner two quartiles.
SUMMARY OF COHORT-SPECIFIC SIGNIFICANT TERA-3 RESULTS OF CHILDREN SELECTED FOR TARGETED INSTRUCTION IN PHONOLOGICAL AWARENESS

- The TERA-3 Reading Quotient, Alphabet, and Meaning subtest mean scores of the preschool cohort indicated significant and meaningful improvements of 76%, 112%, and 42% of a standard deviation, respectively. All improvements can be attributed to the children’s participation in ELLM and targeted instruction in phonological awareness.

- The TERA-3 Reading Quotient and Alphabet subtest mean scores of the kindergarten cohort indicated significant and meaningful improvements of 51% and 120% of a standard deviation, respectively. These improvements can be attributed to the children’s participation in ELLM and targeted instruction in phonological awareness.

- The TERA-3 Reading Quotient and Conventions of Print subtests mean scores of the first-grade cohort indicated significant and meaningful improvements of 61% and 94% of a standard deviation, respectively. These improvements can be attributed to the children’s participation in ELLM and targeted instruction in phonological awareness.

CROSS-COHORT COMPARISON

Learning Hierarchy in Emergent Literacy Abilities

Table 7 summarizes the improvements in the TERA-3 subtest scores attributable to participation in ELLM across the cohorts.

Table 7  
TERA-3 Improvements in Emergent Literacy Attributable to ELLM Across Cohorts

<table>
<thead>
<tr>
<th>TERA-3 Subtest</th>
<th>Cohort</th>
<th>Posttest Ranking</th>
<th>Change in Ranking</th>
<th>Posttest Scores in Bottom Quartile</th>
<th>Effect Size</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alphabet</td>
<td>Preschool</td>
<td>67th</td>
<td>21</td>
<td>19%</td>
<td>.53</td>
<td>Distribution matches national distribution</td>
</tr>
<tr>
<td></td>
<td>Kindergarten</td>
<td>60th</td>
<td>26</td>
<td>12%</td>
<td>.66</td>
<td>Distribution matches national distribution</td>
</tr>
<tr>
<td>Conventions</td>
<td>Kindergarten</td>
<td>21st</td>
<td>8</td>
<td>59%</td>
<td>.35</td>
<td>Distribution is mostly Average and below</td>
</tr>
<tr>
<td></td>
<td>First Grade</td>
<td>39th</td>
<td>13</td>
<td>26%</td>
<td>.35</td>
<td>Distribution is mostly Average and below</td>
</tr>
<tr>
<td>Meaning</td>
<td>First Grade</td>
<td>16th</td>
<td>10</td>
<td>53%</td>
<td>.60</td>
<td>Distribution is mostly Average and below</td>
</tr>
</tbody>
</table>

The summary presented on the previous page supports the suggestion of the existence of a hierarchy in the teaching and learning of the skills measured by the three...
TERA-3 subtests when using standardized scores, once an achievement gap has been established. The first-grade cohort sustains the alphabet letter knowledge gains of the preschool and kindergarten children. Moreover, when the mean scores exhibit low initial status such as with the Conventions of Print subtest, only more modest gains are achieved over the school year. However, the kindergarten Conventions of Print gains appear to be sustained and built upon by the first grade. If significant improvement is not made until the first grade, as is the case in the TERA-3 Meaning subtest scores, the meaningful improvement still leaves most of the children far behind their national peers.

**Declining Kindergarten TERA-3 Meaning Subtest Scores**

A second, equally important issue is the widening of the achievement gap (compared to the national normative population) that occurs over the kindergarten year in the emergent literacy abilities measured by the TERA-3 Meaning subtest. Among several possible explanations for the puzzle of the TERA-3 Meaning subtest scores of the kindergarten children, two are easily investigated using extant ELLM core data. One explanation simply suggests the TERA-3 Meaning subtest scores of the ELLM preschool children are not sustained over the summer before the children start their kindergarten year. The second, equally simple explanation, suggests the core ELLM preschool children do not attend the core elementary schools in large enough numbers to affect the initial emergent literacy status of the kindergarten children.

**Three Years of ELLM Data**

ELLM TERA-3 scores from core sites are available for three consecutive years, allowing a qualitative look. The three-year data indicate the percentile rankings of the TERA-3 mean scores are fairly consistent within the cohorts across the three years. However, the percentile rankings of the TERA-3 mean scores of the kindergarten children are much more stable across the years than the rankings of the preschool scores.

*Stability of Preschool Scores*

The differences of the percentile rankings of the preschool mean posttest scores and the kindergarten mean pretest scores the following school year are at least 13 percentiles across the three TERA-3 subtests. This indicates, if most of the ELLM preschool children attend the core ELLM kindergarten classes, the preschool scores are not very stable, indicating that the skills acquired in preschool are not sustained without frequent reinforcement.
Stability of Kindergarten Scores

These same data covering three years of ELLM implementation from the three cohorts allow a quasi-longitudinal look at scores from preschool through kindergarten to the first grade. However, it is more likely the kindergarten and first-grade children are the same children than the preschool and kindergarten children are the same because the kindergarten and first-grade classes meet in the same elementary schools.

If the kindergarten and first-grade children are mostly the same children, then, across the TERA-3 subtests, the mean scores indicate the average loss in relative ranking over the summer from kindergarten to first grade is four percentiles, and that the loss is fairly consistent across TERA-3 subtests. Therefore, skills acquired in kindergarten are more stable than skills acquired in preschool.

The Teaching and Learning of Kindergarten Children

However, the three-year data could just as easily indicate ELLM kindergarten classes in the core do not enroll enough ELLM preschool children from the previous school year to benefit from their school readiness. In this case, the TERA-3 Meaning subtest puzzle is mostly a kindergarten problem. Thus, the kindergarten teachers have the dual responsibility to close a gap in emergency literacy achievement that arose long before the children’s kindergarten year and to not allow the gap to widen further in any emergent literacy abilities over the kindergarten year. Thus, acquiring emergent literacy skills sooner than the kindergarten year is precursor to elevating standardized scores in all emergent literacy abilities.

ELLM kindergarten teachers, over the three years of extant TERA-3 data, have prepared the children in the emergent literacy abilities of alphabet letter knowledge, and the ELLM kindergarten children’s scores rank above their national peers at the end of the school year. However, ELLM kindergarten children may not be ready for the higher-level task of application, that is, of using the recently acquired alphabet letter knowledge to correctly respond to items measuring emergent comprehension and vocabulary. It is therefore important to note that alphabet letter knowledge is only a necessary skill but it is not sufficient to the acquisition of reading.
Implications of the Cross-Cohort Comparison

- A longitudinal study using the *Making a Difference: 2002/2003* and *Making a Difference: 2003/2004* data should provide answers to some of the posited questions. At the very least, a longitudinal study could provide an indication of the percentage of ELLM kindergarten children who were in ELLM preschool classes the previous school year.

- It is not sufficient to find out what percentage of the core kindergarten children were in ELLM classes in the previous year, but we also need to determine where the children were if not in ELLM classes – whether in other preschool classes or not in preschool at all. A longitudinal study following ELLM preschool children to kindergarten classes could inform us about the stability of the children’s preschool emergent literacy scores and the impact of the children’s level of school readiness in emergent literacy on their kindergarten literacy achievement.

- A possible solution to the kindergarten puzzle is to implement ELLM in more sites, especially sites in feeder patterns to the high-needs elementary schools in the core sample.

- ELLM developers need to enhance the ELLM literacy curriculum in ways that explicitly address the emergent literacy abilities measured by the TERA-3 Meaning subtest which is seen as a measure of emergent comprehension and vocabulary.

    Providing kindergarten teachers with high initial status kindergarteners allows teachers to spend more time helping children acquire all of the emergent literacy skills expected of all kindergarten children rather than just trying to close the gap in children’s skills compared to their more advantaged peers.
SUMMARY OF SPECIFIC CROSS-COHORT COMPARISON

Emergent Literacy Achievement

- Children in the preschool and kindergarten cohorts made significant and meaningful improvement attributable to their participation in ELLM in the emergent literacy abilities measured by the TERA-3 Alphabet subtest.

- Children in the kindergarten and first-grade cohorts made significant and meaningful improvement attributable to their participation in ELLM in the emergent literacy abilities measured by the TERA-3 Conventions of Print subtest.

- Only the children in first-grade cohort made significant and meaningful improvement attributable to their participation in ELLM in the emergent literacy abilities measured by the TERA-3 Meaning subtest.

- The children’s improvement in emergent literacy skills indicates a possible teaching and learning hierarchy (ranging from the content of the Alphabet to the Conventions of Print to the Meaning subtests) exists for children experiencing gaps in achievement compared to national population, which include their more advantaged peers.

Possible Explanations of the Achievement Differences across the Cohorts

- The use of standardized scores plays a role in establishing the hierarchy in the teaching and learning of emergent literacy skills for children in classrooms where most children experience gaps in achievement compared to national normative population, which include their more advantaged peers.

- The core ELLM preschool sample does not sufficiently feed the core ELLM kindergarten classes to significantly elevate the mean initial emergent literacy status of the core ELLM kindergarten children.

- Core ELLM kindergarten teachers close the gap in emergent literacy skills in the content of alphabet letter knowledge. However, they do not, at the same time, close similar gaps that have accumulated over years in other emergent literacy abilities.

Future ELLM Research and Development

- Conduct longitudinal studies using extant data to investigate the core ELLM preschool, kindergarten, and first-grade cohort progression.

- Longitudinally follow ELLM preschool children to kindergarten to investigate the stability of emergent literacy skills acquired during preschool.

- Enhance the ELLM literacy curriculum to more explicitly address emergent comprehension and vocabulary development.

- Expand ELLM to include more preschool sites that feed high-needs public elementary schools in Duval County.
SUMMARY OF COHORT-SPECIFIC RESULTS

Making a Difference ELLM evaluations over the last three years document ELLM successes in improving the early literacy achievement of children from low-income families. In particular, the reports indicate ELLM preschool and kindergarten children excel in alphabet letter knowledge as exemplified by their end-of-the-year scores on the TERA-3 Alphabet subtest and ALRI. In fact, the percentage of ELLM preschool and kindergarten children’s scores ranked below the 25th percentile is less than in the national normative population and the percentage of ELLM preschool and kindergarten children’s scores ranked above the 75th percentile is more than in the national normative population. Additionally over the last two years, at the end of preschool more than 70% of ELLM children were proficient at alphabet letter recognition. This exceeds the 66% of children in a national sample who were proficient upon entering kindergarten.

The following list provides specific TERA-3 and ALRI achievements of the ELLM preschool, kindergarten, and first-grade children.

Preschool Cohort

- The TERA-3 Reading Quotient mean scores indicated a significant and meaningful improvement of 30% of a standard deviation.

- The TERA-3 Alphabet subtest mean scores indicated a significant and meaningful improvement of 53% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  
  At the end of the school year:
  - Fewer ELLM than national normative population Alphabet subtest scores ranked below the 25th percentile.
  - 5% fewer ELLM than national normative population Alphabet subtest scores ranked at or below the 9th percentile.
  - Almost 20% more ELLM than national normative population Alphabet subtest scores ranked above the 75th percentile.
  - 24% of the ELLM Alphabet subtest scores ranked at or above the 90th percentile.

- The TERA-3 Conventions of Print subtest mean scores indicated a significant but not a meaningful improvement from pretest to posttest.

- The TERA-3 Meaning subtest mean scores did not indicate significant improvement from pretest to posttest.
At the end of the preschool year:

- The typical ELLM preschool child recognized 82% of the letters; therefore was proficient.
- 74% of the ELLM preschool children were proficient.
- 39% of the ELLM preschool children recognized all upper- and lowercase letters.
- 5% of the ELLM preschool children recognized fewer than 10 letters.

**Kindergarten Cohort**

- The TERA-3 Reading Quotient mean scores showed a significant and meaningful improvement of 34% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  - 12% fewer ELLM Reading Quotient posttest than pretest scores ranked below the 25th percentile.

- The TERA-3 Alphabet subtest mean scores showed a significant and meaningful improvement of 66% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  - At the end of the school year:
    - Fewer ELLM than national normative population Alphabet subtest scores ranked below the 25th percentile.
    - 2% fewer ELLM than national normative population Alphabet subtest scores ranked at or below the 9th percentile.
    - The percentage of ELLM Alphabet subtest scores ranked above the 75th percentile exceeded the percentage in the national normative population.

- The TERA-3 Conventions of Print subtest mean scores showed a significant and meaningful improvement of 35% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  - At the end of the school year:
    - 12% fewer ELLM Conventions of Print subtest scores ranked below the 25th percentile than in the fall.

- The TERA-3 Meaning subtest mean scores indicated significant and meaningful decline in ranking relative to the national normative population from pretest to posttest.
  - At the end of the school year:
    - 75% of the ELLM Meaning subtest scores ranked below the 25th percentile. This is 18% more than at the beginning of the school year.
    - 41% of the ELLM Meaning subtest scores ranked at or below the 9th percentile. This is 18% more than at the beginning of the school year.

- At the end of the kindergarten year:

  - The typical ELLM kindergarten child recognized 94% of the letters; therefore was proficient.
  - 93% of the ELLM kindergarten children were proficient.
  - 69% of the ELLM kindergarten children recognized at least 50 letters.
First-Grade Cohort

- The TERA-3 Reading Quotient mean scores showed a significant and meaningful improvement of 36% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  - 11% fewer ELLM Reading Quotient posttest than pretest scores ranked below the 25th percentile.

- The TERA-3 Conventions of Print subtest mean scores showed a significant and meaningful improvement of 35% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  At the end of the school year:
  - 26% fewer ELLM Conventions of Print subtest scores ranked below the 25th percentile.
  - The percentage of ELLM Conventions of Print subtest scores ranked below the 25th percentile nearly matched the percentage in the national normative population.

- The TERA-3 Meaning subtest mean scores showed a significant and meaningful improvement of 60% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  At the end of the school year:
  - 28% fewer Meaning subtest scores ranked below the 25th percentile.
MAKING A DIFFERENCE 2003/2004: TECHNICAL REPORT

Prepared by:
Stephanie Wehry, Ph.D.
Assistant Director for Research
Florida Institute of Education
at the
University of North Florida

## Table of Contents

### Introduction

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ELLM at a Glance</td>
<td>2</td>
</tr>
<tr>
<td>Duval County Core</td>
<td>2</td>
</tr>
<tr>
<td>Measurement</td>
<td>3</td>
</tr>
<tr>
<td>Phonological Awareness Population of ELLM Children</td>
<td>4</td>
</tr>
<tr>
<td>Evaluation Questions</td>
<td>6</td>
</tr>
<tr>
<td>Guidelines for Statistical Procedures</td>
<td>6</td>
</tr>
<tr>
<td>Organization of the Report</td>
<td>7</td>
</tr>
</tbody>
</table>

### The Evaluation of ELLM 2003/2004: Preschool

<table>
<thead>
<tr>
<th>Topic</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Attrition of Children</td>
<td>8</td>
</tr>
<tr>
<td><strong>Question 1</strong>: Who were the 2003/2004 ELLM preschool children?</td>
<td>9</td>
</tr>
<tr>
<td><strong>Question 2</strong>: Was ELLM effective in improving the reading readiness of ELLM preschool children based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?</td>
<td>10</td>
</tr>
<tr>
<td><strong>Question 3</strong>: How did ELLM preschool children compare to existing national benchmarks in the recognition of the upper- and lowercase letters of the alphabet at the end of the school year?</td>
<td>18</td>
</tr>
<tr>
<td><strong>Question 4</strong>: Who were the ELLM preschool children selected for targeted instruction in phonological awareness, and was ELLM effective in improving their reading readiness based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?</td>
<td>20</td>
</tr>
<tr>
<td>Preschool Summary</td>
<td>26</td>
</tr>
<tr>
<td>Specific Preschool Results</td>
<td>28</td>
</tr>
</tbody>
</table>
# Table of Contents

**The Evaluation of ELLM 2003/2004: Kindergarten**

- Attrition of Children .................................................. 30

  **Question 1:** Who were the 2003/2004 ELLM kindergarten children? 31

- **Question 2:** Was ELLM effective in improving the reading readiness of ELLM kindergarten children based on improved TERA-3 Reading Quotient and TERA-3 subtest scores? 33

- **Question 3:** How did ELLM kindergarten children compare to existing national benchmarks in the recognition of the upper- and lowercase letters of the alphabet at the end of the school year? 45

- **Question 4:** Who were the ELLM kindergarten children selected for targeted instruction in phonological awareness, and was ELLM effective in improving their reading readiness based on improved TERA-3 Reading Quotient and TERA-3 subtest scores? 47

  - Kindergarten Summary .................................................. 52
  - Specific Kindergarten Results ....................................... 54

**The Evaluation of ELLM 2003/2004: First Grade**

- Attrition of Children .................................................. 57

- **Question 1:** Who were the 2003/2004 ELLM first-grade children? 58

- **Question 2:** Was ELLM effective in improving the reading readiness of ELLM first-grade children based on improved TERA-3 Reading Quotient and TERA-3 subtest scores? 59

- **Question 3:** How did ELLM kindergarten children compare to existing national benchmarks in the recognition of the upper- and lowercase letters of the alphabet at the end of the school year? (Not applicable for the first grade.)

- **Question 4:** Who were the ELLM first-grade children selected for targeted instruction in phonological awareness, and was ELLM effective in improving their reading readiness based on improved TERA-3 Reading Quotient and TERA-3 subtest scores? 59

  - First-Grade Summary .................................................... 72
  - Specific First-Grade Results ......................................... 74
## TABLE OF CONTENTS

### CROSS-COHORT COMPARISON

**Question 5**: Were there similarities and differences in the demographic characteristics and emergent literacy achievement of the preschool, kindergarten, and first-grade ELLM children?  
- 76

Specific Cross-Cohort Results  
- 83
SECTION II: TECHNICAL REPORT

INTRODUCTION

Improving the achievement of all children, particularly those considered at high risk of academic failure, continues to be a local, state, and national priority. We know when children come to school healthy, eager, and ready to learn, they are much more likely to succeed in school and become responsible, productive citizens. Being ready for school begins long before children enter kindergarten. Readiness includes the development of skills and knowledge in many areas. Specifically, developing the language and literacy skills needed to become proficient readers is crucial to success in school. Any child who does not learn to read early and read well will not easily master other important skills and knowledge. Yet, as many as one third of children entering kindergarten are underprepared for the challenges they will face and are at risk of school failure.

The Early Literacy and Learning Model (ELLM) is designed to improve the language and literacy skills of 3-year-old, 4-year-old, kindergarten, and first-grade children who live in low-income communities and who are often at risk of academic failure. This report provides detailed information about literacy-related outcomes of ELLM children and provides a summary of the success of the program during the 2003/2004 school year.
**ELLM at a Glance**

ELLM has been used with urban and rural children in a wide array of childcare centers (subsidized, faith-based, Head Start, and Early Intervention Pre-Kindergarten), as well as with kindergarten and first-grade children in high-needs urban elementary schools.

- For program evaluation purposes, ELLM children are organized into three cohorts: 4-year-old preschool, kindergarten, and first-grade children.

- The *Test of Early Reading Ability-3, Form A* (TERA-3) was administered in a pretest/posttest design to measure children’s reading readiness and to evaluate the effectiveness of ELLM. Pretest data were collected in early September and posttest data were collected from mid-April to early May.

- Additionally, the locally developed *Alphabet Letter Recognition Inventory (ALRI)* was administered to 4-year-old preschool and kindergarten children in a pretest/posttest design. The pretest results were used to inform instructional practice and the posttest results to evaluate the effectiveness of ELLM.

- Furthermore, children whose TERA-3 pretest scores ranked in the lowest quartile were selected for targeted instruction in phonological awareness.

**Duval County Core ELLM**

In the 2000/2001 school year, ELLM served a core group of Duval County sites, including a microcosm of the possible childcare centers and kindergarten and first-grade classrooms in both faith-based and public elementary schools. During the 2003/2004 school year, ELLM continued to serve this core group, and ELLM was implemented at eight sites in 32 classes serving approximately 600 children.
Measurement

Two instruments were used to measure children’s reading readiness, the *Test of Early Reading Ability-3* (TERA-3) and the Alphabet Letter Recognition Inventory (ALRI). Trained assessors using scannable forms administered the TERA-3 to all children in a one-on-one setting. Tests were electronically scored. Classroom teachers using scannable forms administered the ALRI tests to 4-year-old preschool and kindergarten children, and the tests were electronically scored.

**TERA-3**

The TERA-3 is a norm-referenced test that assesses components of early developing reading skills, including familiarity with the letters of the alphabet and numerals, discovery of the arbitrary conventions used in reading and writing English, and recognizing that print conveys information, ideas, and thought. The test is composed of three subtests: the Alphabet subtest, the Conventions of Print subtest, and the Meaning subtest, each measuring one of the three components. A composite score, the Reading Quotient, is the unweighted sum of the three subtest scores.

- **The Alphabet Subtest** measures whether children can differentiate alphabet letters from numbers or designs; recognize names of letters; isolate beginning, middle, or ending sounds; discern the number of sounds in a word; discern the number of syllables in a word; and point to or say words on demand.

- **The Conventions of Print Subtest** measures whether children can differentiate upper- and lowercase letters, understand book orientation and parts of books, know the function of punctuation, know the function of upper- and lowercase letters, identify appropriate and inappropriate punctuation, and select appropriately spelled words to complete sentences.

- **The Meaning Subtest** measures whether children can use labels, figures, or logos as early or proto-reading activity, identify correct use of relational vocabulary, understand isolated word meanings, use sentences to correctly identify pictures, identify correct topics of stories, use text to predict what comes next, understand the different uses of text, select appropriate words to complete sentences, read and answer questions based on the reading, read a story and paraphrase and/or retell it, read two sentences and construct a sentence combining both, and use knowledge of syntax and semantics to correctly identify a well-constructed sentence.
TERA-3 can be administered to children as young as three years, six months and as old as eight years, six months. Because children undergo rapid development over the age span of the TERA-3, there are 14 different norming populations that cover the age range of the test. Depending upon the time between pretest and posttest and the age of the child at pretest, a posttest score may be ranked relative to a norming population that is between one and four age increments older than the pretest norming population. This process adjusts the standardized scores for the maturation of the children between the pretest and posttest and allows gains resulting from normal maturation to be separated from gains resulting from program effectiveness.

TERA-3 Reading Quotient scores are reported as norm-referenced, standardized scores with a mean of 100 and a standard deviation of 15. The TERA-3 Subtest scores are reported as norm-referenced, standardized scores with a mean of 10 and a standard deviation of three. Because the scores represent the ranking of scores relative to a national norming population, a change in scores represents a change in ranking relative to a norming population. It does not represent an absolute gain in knowledge.

Table 1 provides categories delimiting the lower, middle two, and upper quartiles of the TERA-3 Reading Quotient and TERA-3 subtest scores of the norming populations. These categories are used in this report to indicate the range of early reading ability of the ELLM children.

Table 1
*Scale of TERA-3 Reading Quotient and TERA-3 Subtests by Ability Categories*

<table>
<thead>
<tr>
<th>Categories</th>
<th>Very Poor</th>
<th>Poor</th>
<th>Below Average</th>
<th>Average</th>
<th>Above Average</th>
<th>Superior</th>
<th>Very Superior</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reading Quotient</strong></td>
<td>Below 70</td>
<td>70-79</td>
<td>80-89</td>
<td>90-110</td>
<td>111-120</td>
<td>121-130</td>
<td>Above 130</td>
</tr>
<tr>
<td><strong>Subtest Score Intervals</strong></td>
<td>Below 4</td>
<td>4-5</td>
<td>6-7</td>
<td>8-12</td>
<td>13-14</td>
<td>15-16</td>
<td>17-20</td>
</tr>
<tr>
<td><strong>Percentiles</strong></td>
<td>2\textsuperscript{nd} or lower</td>
<td>2\textsuperscript{nd} to 9\textsuperscript{th}</td>
<td>9\textsuperscript{th} to 25\textsuperscript{th}</td>
<td>25\textsuperscript{th} to 75\textsuperscript{th}</td>
<td>75\textsuperscript{th} to 91\textsuperscript{st}</td>
<td>91\textsuperscript{st} to 98\textsuperscript{th}</td>
<td>98\textsuperscript{th} or higher</td>
</tr>
<tr>
<td><strong>Percent of Scores</strong></td>
<td>2.3</td>
<td>6.9</td>
<td>16.1</td>
<td>49.5</td>
<td>16.1</td>
<td>6.9</td>
<td>2.3</td>
</tr>
</tbody>
</table>
ALRI

The ALRI is a locally developed test measuring children’s ability to recognize the upper- and lowercase letters of the alphabet when presented in non-alphabetic order. Children’s teachers administered the test one-on-one to children using alphabet letter flashcards. The uppercase letters were presented first in a fixed, non-alphabetic order, followed by the lowercase letters.

ALRI scores are reported in four recognition categories, 0 to 13 letters, 14 to 26 letters, 27 to 39 letters, and 40 to 52 letters. The Early Childhood Longitudinal Study-Kindergarten (ECLS-K) provided national benchmarks for alphabet letter recognition. Using a random sample of upper- and lowercase letters, scores were categorized as proficient if children recognized at least 75% of the sampled letters.\(^1\) Reports from the ECLS-K data also stated 66% of the children (39% of children living in poor families) entering kindergarten for the first time were proficient.\(^2\) In this report, ALRI posttest scores are compared to these national benchmarks.

Phonological Awareness Population of ELLM Children

Researchers discovered over the history of ELLM that children with very low scores on the TERA-3 pretests showed little, if any, improvement on the posttest. Therefore, beginning in 1999/2000 and continuing through 2003/2004, children with very low TERA-3 pretest scores were selected for targeted instruction in phonological awareness.


Evaluation Questions

The TECHNICAL REPORT of the evaluation of ELLM is organized by a set of evaluation questions. There are five broad questions, which are:

Question 1: Who were the 2003/2004 ELLM children?

Question 2: Was ELLM effective in improving the reading readiness of ELLM children based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?

Question 3: How did ELLM children compare to existing national benchmarks in the recognition of the upper- and lowercase letters of the alphabet at the end of the school year?

Question 4: Who were the ELLM children selected for targeted instruction in phonological awareness, and was ELLM effective in improving their reading readiness based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?

Question 5: Were there similarities and differences in the demographic characteristics and emergent literacy achievement of the preschool, kindergarten, and first-grade ELLM children?

Guidelines for Statistical Procedures

All descriptive statistics are reported to one decimal place; however, calculations were carried out to machine precision. In some instances, calculations using reported data might differ from reported results because of rounding.

All statistically significant mean differences are relative to a Type I error rate (alpha) of .05. However, because of the large sample sizes, small - even trivial - differences might be statistically different. To determine the importance of differences, effect sizes are reported. (Effect size is the mean difference relative to the standard deviation of the test.) Cohen classified effect sizes of 20 to 49 percent of a standard deviation as small, effect sizes between 50 and 79 percent as medium, and effect sizes of 80 percent or higher as large. Differences resulting in small, medium, and large effect sizes are meaningful.

Moreover, Whitehurst and Massetti, in a critique of Head Start, noted when evaluation designs lack control or comparison groups, small effect sizes of 20 to 25 percent should not be attributed to programs. Small effects could easily be associated with regression toward the mean, increased familiarity with tests and assessment procedures in general, or ordinary maturation and experiences. Because the evaluation of ELLM does not involve a comparison or control group, only effects larger than one third of a standard deviation are attributed to the children’s participation in ELLM.

**Organization of the Report**

Because previous *Making a Difference* reports found that the effectiveness of ELLM varied across cohorts and TERA-3 subtests, *Making a Difference 2003/2004* reports data that are not aggregated across cohorts. Instead, this *TECHNICAL REPORT* has a stand-alone section that answers the first four evaluation questions for each cohort. A final section that provides an overall summary and discussion of the cohort findings and addresses the last evaluation question follows these sections.

---


Table 1p shows the contribution of ELLM preschool sites and classes to the evaluation of ELLM in the 2003/2004 school year. While both 3- and 4-year-old preschool children are served by ELLM, only 4-year-old children, eligible for Florida public kindergarten in 2004/2005, were assessed as part of program evaluation. Hereafter, preschool refers only to the 4-year-old children assessed. There were TERA-3 pretest scores for 146 ELLM preschool children and posttest scores for 129 ELLM children.

Table 1p
The Number of ELLM Preschool Sites, Teachers, and Assessed Children

<table>
<thead>
<tr>
<th>Attrition</th>
<th>Number of Sites</th>
<th>Number of Classes</th>
<th>Number of Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete Scores</td>
<td>6</td>
<td>9</td>
<td>129</td>
</tr>
<tr>
<td>Incomplete Scores</td>
<td>5</td>
<td>7</td>
<td>17</td>
</tr>
</tbody>
</table>

Attrition of Children

The discrepancy between the number of pretest and posttest scores represents the attrition of 12% of the preschool children. To determine if the attrition occurred at random, the TERA-3 pretest scores of the ELLM children who were not administered the TERA-3 posttest were compared to pretest scores of ELLM children who were administered the TERA-3 posttest. Table 2p reports the results of the comparison.

Table 2p
Summary Statistics and ANOVA Results: ELLM Preschool TERA-3 Reading Quotient and TERA-3 Subtest Pretest Scores by Attrition Category

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean Scores Complete (129)</th>
<th>Mean Scores Incomplete (17)</th>
<th>Probability Means Differ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>90.4</td>
<td>86.5</td>
<td>0.2700</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>9.7</td>
<td>8.3</td>
<td>0.1242</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>7.7</td>
<td>7.4</td>
<td>0.6714</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>8.2</td>
<td>8.1</td>
<td>0.8137</td>
</tr>
</tbody>
</table>

Numbers in parentheses indicate the number of scores.

There was no evidence to suggest an underlying mechanism contributed to the attrition of the preschool children.
Question 1: Who were the 2003/2004 ELLM preschool children?

The answer to this evaluation question involves a description of the children by age, gender, and ethnicity.

**Age**

The categorization of preschool children as four year olds does not indicate the age of the children. Because of the September 1 birthday cut-off for children attending Florida public kindergarten, the typical 4-year-old preschool child is between 48 and 60 months old on September 1 of the school year. Figure 1p shows the ages of the 4-year-old ELLM children who had both TERA-3 pretest and posttest scores.

![Figure 1p](image-url)

*Figure 1p.* Ages of 4-year-old children in months on September 1, 2003 (n = 129 preschool children).

Almost 55% of the children were older than the 54-month median age of the typical age span for Florida 4-year-old preschool children.
**Gender**

Of the 129 preschool children with TERA-3 pretest and posttest scores, 48% were girls and 52% were boys.

**Ethnicity**

The ethnicity of the ELLM preschool children was categorized as Black, White, Hispanic, Asian, Native American, and Other. Ninety-two percent of the ELLM preschool children were Black and 8% were White.

**Question 2: Was ELLM effective in improving the reading readiness of ELLM preschool children based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?**

Table 3 provides summary statistics for the TERA-3 pretest and posttest scores of the ELLM preschool children with both scores. Data were analyzed as a repeated-measures ANOVA design to determine if there were statistically significant differences in the TERA-3 Reading Quotient and TERA-3 subtest pretest and posttest mean scores. Effect sizes are provided for mean scores that represent differences.

**Table 3**

| Summary Statistics and ANOVA Results: ELLM Preschool TERA-3 Reading Quotient and TERA-3 Subtests |
|----------------------------------|-----------------|-----------------|-----------------|-----------------|
|                                   | Pretest Means   | Posttest Means  | Probability Means Differ | Effect Size     |
| Reading Quotient                  | 90.4            | 95.0            | <.0001*                  | 0.3066          |
| Alphabet Subtest                 | 9.7             | 11.3            | <.0001*                  | 0.5326+++       |
| Conventions Subtest              | 7.7             | 8.1             | .0507*                   | 0.1500          |
| Meaning Subtest                  | 8.2             | 8.3             | .5921                    | -               |

* Denotes there was a difference in the pretest and posttest means α = .05.
† Denotes a small effect.
‡ Denotes a medium effect.
§ Denotes an effect that may be attributed to children’s participation in ELLM.
□ Denotes a significant and meaningful difference.

The pretest mean scores of the TERA-3 Reading Quotient, Alphabet, Conventions of Print, and Meaning subtests rank at the 26th, 46th, 22nd, and 27th percentiles, respectively. The posttest mean scores of the Reading Quotient, Alphabet, Conventions of Print, and Meaning subtests rank at the 37th, 67th, 26th, and 29th percentiles, respectively. In the fall, the TERA-3 Reading Quotient, Alphabet, and Meaning subtest mean scores rank in the broad Average category, spanning the two inner quartiles. In the
spring, all scores are in the *Average* category, and the Alphabet subtest mean score ranks 17 percentiles higher than the 50th percentile ranking of the national norming population’s mean score.

The differences in the pretest and posttest mean scores of the TERA-3 Reading Quotient, Alphabet, and Conventions of Print subtests were statistically significant, but only the differences in the mean scores of the TERA-3 Reading Quotient and Alphabet subtest were large enough to be meaningful.

To better understand where the improvement in reading readiness occurred, TERA-3 Reading Quotient and Alphabet subtest scores are displayed in seven ability categories: three categories representing the lowest 25 percentiles, one category representing the middle 50 percentiles, and three categories representing the highest 25 percentiles. (See Table 1, page 4.) Figures 2p-4p display the TERA-3 Reading Quotient score, and Figures 5p-7p display the TERA-3 Alphabet subtest scores.
Preschool: TERA-3 Reading Quotient Scores in the Bottom Quartile

Figure 2p displays the percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the three categories of the bottom quartile. While the percentage of posttest scores in these categories (dark blue bars) remains higher than in the national norming population (light blue bars), there are 19% fewer scores in these categories in the spring than in the fall (striped bars), and the percentage of scores in the bottom two categories (the Very Poor and Poor categories), representing scores ranked at or below the 9th percentile, is about 4% more than in the national norming population.

Figure 2p.
The percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the three categories of the bottom quartile (n=129 preschool scores).
Preschool: TERA-3 Reading Quotient Scores in the Top Quartile

Figure 3p displays the percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the three categories of the top quartile. The percentage of posttest scores in these categories (dark blue bars) remains lower than in national norming population (light blue bars), and there is very little change in the percentage of scores in the top quartile from pretest (striped bars) to posttest.

![Bar chart showing the percentage of TERA-3 Reading Quotient scores in the top quartile for pretest, posttest, and national norming population.]

**Figure 3p.**
The percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the three categories of the top quartile (n=129 preschool scores).
Preschool: The Distribution of TERA-3 Reading Quotient Scores

Figure 4p shows the percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the categories of the ability scale. While the distribution of the posttest scores (dark blue bars) remains shifted toward the low side of the ability scale, it has shifted closer to the national norming population (light blue bars). Most of the changes in the percentage of scores occur in the bottom four categories. There are fewer posttest scores in the bottom quartile and more in the broad *Average* category. This indicates the ELLM preschool children (especially those with scores in the lowest five categories) were closing the gap in achievement measured by the TERA-3 Reading Quotient.

*Figure 4p.* The percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the categories of the ability scale (*n*=129 preschool scores).
Preschool: TERA-3 Alphabet Subtest Scores in the Bottom Quartile

Figure 5p shows the percentage of TERA-3 Alphabet subtest pretest, posttest, and national norming population scores in the three categories of the bottom quartiles. There are about 6% fewer posttest scores (dark blue bars) in the bottom quartile than in the national norming population (light blue bars). Additionally, the percentage of posttest scores in the bottom two categories (the Very Poor and Poor categories), representing scores ranked at or below the 9th percentile, is almost 5% less than in the national norming population.

![Graph showing the percentage of TERA-3 Alphabet pretest, posttest, and national norming population scores in the three categories of the bottom quartile.](image)

38.8% of the Pretest Scores
19.4% of the Posttest Scores
25.3% of the National Scores

**Figure 5p.**
The percentage of TERA-3 Alphabet pretest, posttest, and national norming population scores in the three categories of the bottom quartile (n=129 preschool scores).
Preschool: TERA-3 Alphabet Subtest Scores in the Top Quartile

Figure 6p shows the percentage of TERA-3 Alphabet subtest pretest, posttest, and national norming population scores in the three categories of the top quartile. The percentage of posttest scores (dark blue bars) in the top quartile is 19% more than in the national population (light blue bars), and 24% of the posttest scores are in the top two categories (the *Superior* and *Very Superior* categories), representing scores ranked at or above the 90th percentile.

![Bar chart showing TERA-3 Alphabet subtest scores in the top quartile for pretest, posttest, and national norming population.](chart.png)

**Figure 6p.**
The percentage of TERA-3 Alphabet pretest, posttest, and national norming population scores in the three categories of the top quartile (n=129 preschool scores).
Preschool: The Distribution of TERA-3 Alphabet Subtest Scores

Figure 7p shows the percentage TERA-3 Alphabet subtest pretest, posttest, and national norming population scores in the categories of the ability scale. The percentage of the posttest scores (dark blue bars) in all categories is better than in the national norming population (light blue bars). The entire distribution of posttest scores has shifted to the high side of the ability scale, which indicates the ELLM preschool children closed the gap in achievement measured by the TERA-3 Alphabet subtest.

Figure 7p.
The percentage of TERA-3 Alphabet pretest, posttest, and national norming population scores in the categories of the ability scale (n=129 preschool scores).
Question 3: How did ELLM preschool children compare to existing national benchmarks in the recognition of the upper- and lowercase letters of the alphabet at the end of the school year?

In 2003/2004, 115 ELLM preschool children had Alphabet Letter Recognition Inventory (ALRI) pretest and posttest scores. The ALRI pretest mean score indicates the typical ELLM preschool child recognized about 48% of the letters. The ALRI posttest mean score indicates the typical preschool child recognized 82% of the letters. The Early Childhood Longitudinal Study-Kindergarten (ECLS-K) defined proficient as recognizing 75% of the sampled letters; therefore, at the end of the school year, the typical ELLM preschool child was proficient. ECLS-K researchers also reported that 39% of children living in poor families were proficient when entering kindergarten for the first time. Inspection of Figure 8p indicates 36% of the ELLM preschool children were proficient upon entering their 4-year-old preschool year, which nearly matches the percentage of proficient children living in poor families entering kindergarten for the first time in the ECLS-K national sample.

To determine the range of the alphabet letter recognition ability of ELLM preschool children, ALRI scores are displayed in Figure 8p on the next page using four recognition categories: 0-13 letters, 14-26 letters, 27-39 letters, and 40-52 letters. Inspection of Figure 8p indicates that at the end of the school year 74% of the scores are in the 40-to-52 letters recognized category (recognizing at least 75% of the letters; therefore, proficient). Additionally, 39% of the ELLM preschool children recognized all of the letters, and only 5% recognized fewer than 10 letters.
Figure 8p.
Alphabet Letter Recognition Inventory pretest and posttest scores (n=115 preschool scores).
Question 4: Who were the ELLM preschool children selected for targeted instruction in phonological awareness, and was ELLM effective in improving their reading readiness based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?

The answer to the first part of this evaluation question involves a description of the 24 preschool children selected for targeted instruction in phonological awareness by age and gender.

Age

As previously mentioned, the categorization of the preschool children as four year olds does not indicate the age of the children. Because of the September 1 birthday cut-off for children attending Florida public kindergarten, the typical 4-year-old preschool child is between 48 and 60 months old on September 1 of the school year. Figure 9p shows the ages of the 24 ELLM preschool children selected for targeted instruction in phonological awareness.

Figure 9p.
Ages of children in months on September 1, 2003 (n = 24 preschool children selected for targeted instruction in phonological awareness).
Almost 51% of the preschool children selected for targeted instruction in phonological awareness were older than the 54-month median age of the typical Florida 4-year-old preschooler.

**Gender**

All of the 24 preschool children selected for targeted instruction in phonological awareness were *Black*, 42% were girls, and 58% were boys. Forty-nine percent of the complete ELLM preschool sample were boys; therefore, the scores of boys are disproportionately represented in the bottom quartile of ELLM scores.

**TERA-3 Results: Phonological Awareness Population**

The answer to the second part of the evaluation question comes from the TERA-3 scores of these 24 preschool children. Table 4p provides summary statistics for the TERA-3 pretest and posttest scores of the ELLM preschool children selected for targeted instruction in phonological awareness. Data were analyzed as a repeated-measures ANOVA design to determine if there were statistically significant differences in the TERA-3 pretest and posttest mean scores. Effect sizes are provided for mean scores that represent differences.

**Table 4p**

*Summary Statistics and ANOVA Results: ELLM Preschool TERA-3 Reading Quotient and TERA-3 Subtests (Phonological Awareness Population)*

<table>
<thead>
<tr>
<th></th>
<th>Pretest Means</th>
<th>Posttest Means</th>
<th>Probability Means Differ</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>72.7</td>
<td>84.2</td>
<td>&lt;.0001*</td>
<td>0.764φφ+</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>5.6</td>
<td>9.0</td>
<td>&lt;.0001*</td>
<td>1.125φφφ+</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>5.9</td>
<td>6.5</td>
<td>.0961</td>
<td>0.208†</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>5.8</td>
<td>7.0</td>
<td>.0052*</td>
<td>0.417†</td>
</tr>
</tbody>
</table>

* Indicates significant difference at the α = .05 level.
† Denotes a small effect.
φφ Denotes a medium effect.
φφφφ Denotes a large effect.
†† Denotes an effect that may be attributed to children’s participation in ELLM and targeted instruction in phonological awareness.
□ Indicates a statistically significant and meaningful effect.

Targeted instruction in phonological awareness primarily addresses the content measured by the TERA-3 Alphabet subtest. In the fall, the Alphabet subtest mean score of the preschool children selected for targeted instruction in phonological awareness ranked at the 7th percentile in the *Poor* category of the bottom quartile of the ability scale. In the
spring, the Alphabet subtest mean score ranked below the mean ranking of the national norming population, but still in the broad *Average* category of the ability continuum. Additionally, the ELLM preschool children’s scores improved on the TERA-3 Reading Quotient and Meaning subtest.

Statistical analyses indicated all three of these increases were significant and meaningful. The TERA-3 Alphabet subtest improvement of 113% of a standard deviation greatly contributed to the improvement of 76% of a standard deviation in the Reading Quotient mean scores. Because targeted instruction in phonological awareness primarily addresses the content measured by the TERA-3 Alphabet subtest, only the distribution of those scores will be further described. To better understand where the improvement occurred, the TERA-3 Alphabet subtest scores are shown in Figure 10p using the categories described in Table 1, page 4.
Preschool Children Selected for Targeted Instruction in Phonological Awareness: TERA-3 Alphabet Subtest Scores

Figure 10p shows the percentage of TERA-3 Alphabet subtest pretest, posttest, and national norming population scores in the categories of the ability scale. There are 67% fewer posttest scores (dark blue bars) in the bottom quartile than pretest scores (striped bars). In the fall, 38% of the scores (striped bars) are in the bottom two categories (the Very Poor and Poor categories), representing scores ranked at or below the 9th percentile, but no scores remain in these categories in the spring (dark blue bars). Additionally, in the spring, the percentage of scores in the top quartile is half of the percentage in the national norming population.

![Figure 10p](image_url)

The percentage of TERA-3 Alphabet pretest, posttest, and national norming population scores in the categories of the ability scale (n=24 scores of preschool children selected for targeted instruction in phonological awareness).
**ALRI Results: Phonological Awareness Population**

In 2003/2004, 20 ELLM preschool children selected for targeted instruction in phonological awareness had Alphabet Letter Recognition Inventory (ALRI) pretest and posttest scores. The ALRI pretest mean score indicates the typical ELLM preschool child in this group of children recognized about 14% of the letters. The ALRI posttest mean score indicates the typical preschool child recognized 63% of the letters. The *Early Childhood Longitudinal Study-Kindergarten* (ECLS-K) defined proficient as recognizing 75% of the sampled letters; therefore, by the end of the school year, the typical ELLM preschool child selected for targeted instruction in phonological awareness was well on the way toward proficiency.

To determine the range of alphabet letter recognition ability of these ELLM preschool children, ALRI scores are displayed in Figure 11p on the next page using four recognition categories: 0-13 letters, 14-26 letters, 27-39 letters, and 40-52 letters. Inspection of Figure 11p indicates 35% of the ELLM preschool children selected for targeted instruction in phonological awareness were proficient at the end of the year. The end of the 4-year-old school year is somewhat similar to the beginning of the kindergarten year. In the spring, the percentage of ELLM preschool children selected for targeted instruction in phonological awareness who were proficient at letter recognition nearly matches the 39% of proficient children living in poor families in the ECLS-K national sample of children entering kindergarten for the first time.
Figure 11p.
Alphabet Letter Recognition Inventory pretest and posttest scores of 20 preschool children selected for targeted instruction in phonological awareness.
Preschool Summary

Table 5p summarizes the changes in the percentage of TERA-3 scores in the bottom quartile from pretest to posttest.

Table 5p
Summary of ELLM Preschool Children’s TERA-3 Scores in the Top Three Quartiles

<table>
<thead>
<tr>
<th>TERA-3 Test</th>
<th>Percent of Pretest Scores Ranked Average or Above</th>
<th>Percent of Posttest Scores Ranked Average or Above</th>
<th>Percent of Change</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>41.9</td>
<td>61.2</td>
<td>19.4</td>
<td>0.306⁹</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>61.2</td>
<td>80.6</td>
<td>19.4</td>
<td>0.532++</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>43.4</td>
<td>52.7</td>
<td>9.3</td>
<td>0.150</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>65.9</td>
<td>73.6</td>
<td>7.8</td>
<td></td>
</tr>
</tbody>
</table>

* Indicates a small effect.
++ Indicates a medium effect.
+ Indicates an improvement that can be attributed to children’s participation in ELLM.
 Denotes the percentage of ELLM children’s scores matches the percentage of scores in the national population.

The large effect size of the TERA-3 Alphabet subtest clearly was the major contributing factor in the improvement of the TERA-3 Reading Quotient scores. The Alphabet subtest pretest mean score ranked just lower than the 50th percentile of the national norming population. The Alphabet subtest posttest mean score ranked at the 67th percentile. This TERA-3 achievement coupled with the 74% of the preschool children who were proficient at the end of the school year indicates the ELLM 4-year-old preschoolers’ ability in alphabet letter knowledge ranks above the national sample of children.
Table 6p presents the improvements made by ELLM preschool children selected for targeted instruction in phonological awareness and summarizes the changes in the percentage of TERA-3 scores of the in the bottom quartile (ranked above the 9th percentile) from pretest to posttest.

Table 6p
Summary of TERA-3 Scores of the ELLM Preschool Children Selected for Targeted Instruction in Phonological Awareness

<table>
<thead>
<tr>
<th>TERA-3 Test</th>
<th>Percent of Pretest Scores Average or Above and Above the 9th Percentile in Parentheses</th>
<th>Percent of Posttest Scores Average or Above and Above the 9th Percentile in Parentheses</th>
<th>Percent of Change</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>0.0 (4.2)</td>
<td>25.0 (66.7)</td>
<td>25.0 (62.6)</td>
<td>0.764 +</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>0.0 (62.5)</td>
<td>66.7 (100.0)</td>
<td>66.7 (37.5)</td>
<td>1.125 +++</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>4.2 (66.7)</td>
<td>29.2 (62.5)</td>
<td>25.0 (4.2)</td>
<td>0.208 b</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>8.3 (54.2)</td>
<td>54.2 (75.0)</td>
<td>45.8 (20.8)</td>
<td>0.417 +++</td>
</tr>
</tbody>
</table>

The numbers in parenthesis represent percentage of scores ranked above the 9th percentile.

++ Indicates a small effect.

+++ Indicates a medium effect.

+++] Indicates a large effect.

* Indicates an improvement that can be attributed to children’s participation in ELLM and targeted instruction in phonological awareness.

[+] Denotes there were at least 25 percent fewer scores in the bottom quartile at posttest.

The scores of the children selected for targeted instruction in phonological awareness showed significant and meaningful improvement in the TERA-3 Reading Quotient and all TERA-3 subtests. The improvement in the TERA-3 Reading Quotient, Alphabet, and Meaning subtests can be attributed to the children’s participation in ELLM and the targeted instruction in phonological awareness.

In the fall, all TERA-3 Alphabet subtest scores ranked in the bottom quartile and 63% ranked at or below the 9th percentile. At the end of the school year, 67% of the TERA-3 Alphabet subtest scores ranked higher than the bottom quartile, and no scores ranked at or below the 9th percentile.
SPECIFIC PRESCHOOL RESULTS

Question 1: Who were the 2003/2004 ELLM preschool children?

- The ages of the ELLM preschool children were evenly spread across the typical range of Florida 4-year-old preschool children.
- 52% of the children were boys and 48% were girls.
- 92% of the children were Black and 8% were White.
- There were complete TERA-3 scores for 129 preschool children (12% attrition).

Question 2: Was ELLM effective in improving the reading readiness of ELLM preschool children based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?

- All TERA-3 posttest mean scores ranked in the broad Average ability category that spans the inner two quartiles of the ability scale.
- The TERA-3 Reading Quotient mean scores indicated a significant and meaningful improvement of 30% of a standard deviation.
  - There were 19% fewer ELLM TERA-3 Reading Quotient scores in the bottom quartile in the spring than in the fall.
- The TERA-3 Alphabet subtest mean scores indicated a significant and meaningful improvement of 53% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  - At the end of the school year:
    - Fewer ELLM Alphabet subtest scores were in the bottom quartile than in the national norming population.
    - 5% fewer ELLM Alphabet subtest scores ranked at or below the 9th percentile than in the national norming population.
    - Almost 20% more ELLM Alphabet subtest scores were in the top quartile than in the national norming population.
    - 24% of the ELLM Alphabet subtest scores ranked at or above the 90th percentile.
    - The ELLM Alphabet subtest mean score ranked 17 percentiles higher than the national norming population average.
- The TERA-3 Conventions of Print subtest mean scores indicated a significant but not meaningful improvement from pretest to posttest.
- The TERA-3 Meaning subtest mean scores did not indicate significant improvement from pretest to posttest.
Question 3: How did ELLM preschool children compare to existing national benchmarks in the recognition of the upper- and lowercase letters of the alphabet at the end of the school year?

At the end of the school year:
- The typical ELLM preschool child recognized 82% of the letters. ECLS-K researchers defined proficient as the recognition of at least 75% of sampled letters; therefore, the typical preschool ELLM child was proficient.
- 74% of the ELLM preschool children were proficient compared to 66% of all children (39% of children living in poor families) entering kindergarten for the first time in the national ECLS-K sample.
- 39% of the ELLM preschool children recognized all letters.
- 5% of the ELLM preschool children recognized fewer than 10 letters.

Question 4: Who were the ELLM preschool children selected for targeted instruction in phonological awareness, and was ELLM effective in improving their reading readiness based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?

- The 24 ELLM preschool children selected for targeted instruction in phonological awareness (18.6% of the ELLM preschool children) were Black, slightly older, and included a larger percentage of boys than the complete ELLM preschool sample.
- The TERA-3 Reading Quotient, Alphabet, and Meaning subtest mean scores indicated significant and meaningful improvements of 76%, 112%, and 42% of a standard deviation, respectively. All improvements can be attributed to the children’s participation in ELLM and targeted instruction in phonological awareness.
- There were 67% fewer Alphabet subtest scores in the bottom quartile in the spring than in the fall. Additionally, at the end of the school year, there were no Alphabet subtest scores ranked at or below the 9th percentile.
- At the end of the school year, the typical ELLM preschool child selected for targeted instruction in phonological awareness recognized 63% of the letters and was therefore moving toward proficiency.
- 35% of the ELLM children selected for targeted instruction in phonological awareness were proficient at the end of the year.

Table 1k shows the contribution of ELLM kindergarten sites and classes to the evaluation of ELLM kindergarten in the 2003/2004 school year. There were TERA-3 pretest scores for 233 ELLM kindergarten children and posttest scores for 188 ELLM children.

Table 1k
The Number ELLM Kindergarten Sites, Teachers, and Assessed Children

<table>
<thead>
<tr>
<th>Attrition</th>
<th>Number of Sites</th>
<th>Number of Classes</th>
<th>Number of Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete Scores</td>
<td>4</td>
<td>11</td>
<td>188</td>
</tr>
<tr>
<td>Incomplete Scores</td>
<td>3</td>
<td>10</td>
<td>45</td>
</tr>
</tbody>
</table>

Attrition of Children

The discrepancy between the number of pretest and posttest scores represents the attrition of 19% of the kindergarten children. To determine if the attrition occurred at random, the TERA-3 pretest scores of the ELLM children who were not administered the TERA-3 posttest were compared to pretest scores of ELLM children who were administered the TERA-3 posttest. Table 2k reports the results of the comparison.

Table 2k
Summary Statistics and ANOVA Results: ELLM Kindergarten TERA-3 Reading Quotient and TERA-3 Subtest Pretest Scores by Attrition Category

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean Scores Complete (188)</th>
<th>Mean Scores Incomplete (45)</th>
<th>Probability Means Differ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>82.8</td>
<td>77.6</td>
<td>.0320*</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>8.6</td>
<td>7.5</td>
<td>.0596</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>6.6</td>
<td>5.8</td>
<td>.0574</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>6.8</td>
<td>6.3</td>
<td>.1093</td>
</tr>
</tbody>
</table>

* Indicates significant difference at the $\alpha = .05$ level.
Numbers in parentheses indicate the number of scores.

There was evidence suggesting that an underlying mechanism contributed to the attrition of the children. Children with incomplete scores had significantly lower TERA-3 Reading Quotient scores than children with complete scores ($\alpha = .05$). Furthermore, children with incomplete scores had significantly lower TERA-3 Alphabet and Conventions of Print subtest scores than children with complete score ($\alpha = .10$).
Therefore, results from this cohort should be interpreted cautiously. However, results from the scores of ELLM children selected for targeted instruction in phonological awareness indicate if the lower scoring children were not lost to attrition their results would be in line with the children with lower initial status who had complete scores and made significant and meaningful improvement.

**Question 1: Who were the 2003/2004 ELLM kindergarten children?**

The answer to this evaluation question involves a description of the kindergarten children by age, gender, and ethnicity.

**Age**

The categorization of the children as kindergarteners does not indicate the age of the children. Because of the September 1 birthday cut-off for children attending Florida public kindergarten, the typical kindergarten child is between 60 and 72 months old on September 1 of the school year. Figure 1k shows the ages of the ELLM kindergarten children who had both TERA-3 pretest and posttest scores.

![Ages of children in months on September 1, 2003](image)

*Figure 1k.* Ages of children in months on September 1, 2003 (n = 188 kindergarten children).
Even though more than 16% of the children were older than the typical range of Florida kindergarten children, 57% of the children in the typical 60- to-72 month span were also older than the 66-month median age of the typical Florida age span for kindergarten children.

**Gender**

Of the 188 ELLM kindergarten children with TERA-3 pretest and posttest scores, 51% were girls and 49% were boys.

**Ethnicity**

The ethnicity of the ELLM children was categorized as Black, White, Hispanic, Asian, Native American, and Other. All of the ELLM kindergarten children were Black.

**Question 2: Was ELLM effective in improving the reading readiness of ELLM kindergarten children based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?**

Table 3k provides summary statistics for the TERA-3 pretest and posttest scores of the ELLM kindergarten children with both scores. Data were analyzed as a repeated-measures ANOVA design to determine if there were statistically significant differences in the TERA-3 Reading Quotient and TERA-3 subtest pretest and posttest mean scores. Effect sizes are provided for mean scores that represent differences.

Table 3k

**Summary Statistics and ANOVA Results: ELLM Kindergarten TERA-3 Reading Quotient and TERA-3 Subtest Scores**

<table>
<thead>
<tr>
<th></th>
<th>Pretest Means</th>
<th>Posttest Means</th>
<th>Probability Means Differ</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>82.8</td>
<td>87.8</td>
<td>&lt;.0001</td>
<td>0.340*</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>8.6</td>
<td>10.6</td>
<td>&lt;.0001</td>
<td>0.660**</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>6.6</td>
<td>7.6</td>
<td>&lt;.0001</td>
<td>0.350**</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>6.8</td>
<td>6.1</td>
<td>&lt;.0001</td>
<td>-0.230</td>
</tr>
</tbody>
</table>

* Denotes there was a difference in the pretest and posttest mean scores α=.05. 
† Denotes a small effect. 
‡ Denotes a medium effect. 
+++ Denotes an effect that may be attributed to children’s participation in ELLM. 
★★ Denotes a significant and meaningful effect.

The pretest mean scores of the TERA-3 Reading Quotient, Alphabet, Conventions of Print, and Meaning subtests rank at the 13th, 32nd, 13th, and 14th percentiles, respectively. The posttest mean scores of the Reading Quotient, Alphabet, Conventions
of Print, Conventions of Print, and Meaning subtests rank at the 21st, 60th, 21st, and 10th percentiles, respectively. The differences in the pretest and posttest TERA-3 mean scores of all tests are statistically different, and all differences except the Meaning subtest difference are positive. All effect sizes are meaningful.

To better understand where the improvement in reading readiness occurred, TERA-3 Reading Quotient and subtest scores are displayed in seven ability categories: three categories representing the lowest 25 percentiles, one category representing the middle 50 percentiles, and three categories representing the highest 25 percentiles. (See Table 1, page 4.) Figures 2k-4k display the TERA-3 Reading Quotient scores, Figures 5k-7k display the Alphabet subtest scores, Figures 8k-10k display the Conventions of Print subtest scores, and Figures 11k-12k display the Meaning subtest scores.
Kindergarten: TERA-3 Reading Quotient Scores in the Bottom Quartile

Figure 2k displays the percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the three categories of the bottom quartile. While the percentage of posttest scores in these categories (dark blue bars) remains higher than in the national norming population (light blue bars), there are 12% fewer scores in these categories in the spring than in the fall (striped bars), and the percentage of scores in the bottom two categories (the Very Poor and Poor categories), representing scores ranked at or below the 9th percentile, is about 13% fewer than in the fall.

![Bar Chart]

Figure 2k.
The percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the three categories of the bottom quartile (n=188 kindergarten scores).
Kindergarten: TERA-3 Reading Quotient Scores in the Top Quartile

Figure 3k displays the percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the three categories of the top quartile. While percentage of posttest scores in these categories (dark blue bars) remains lower than in national norming population (light blue bars), there are 3% more scores in these categories in the spring than in the fall (striped bars).

Figure 3k.
The percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the three categories of the top quartile (n=188 kindergarten scores).
Kindergarten: The Distribution of TERA-3 Reading Quotient Scores

Figure 4k shows the percentage of the TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the categories of the ability scale. While the distribution of the posttest scores (dark blue bars) remains shifted toward the low side of the ability scale, it has shifted closer to the national norming population (light blue bars). Most of the changes in the percentage of scores occur in the bottom four categories. In the spring, there are fewer scores in the bottom quartile and more in the broad Average category. This indicates the ELLM kindergarten children (especially those with scores in the lowest five categories) were closing the gap in achievement measured by the TERA-3 Reading Quotient.

Figure 4k.
The percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the categories of the ability scale (n=188 kindergarten scores).
Kindergarten: TERA-3 Alphabet Subtest Scores in the Bottom Quartile

Figure 5k shows the percentage of TERA-3 Alphabet subtest pretest, posttest, and national norming population scores in the three categories of the bottom quartile. The percentage of posttest scores (dark blue bars) in the bottom quartile is less than half of the percentage in the national norming population (light blue bars). Additionally, the percentage of posttest scores in the bottom two categories (the Very Poor and Poor categories), representing scores ranked at or below the 9th percentile, is 2% less than in the national norming population.

![Figure 5k](chart.png)

*Figure 5k.* The percentage of TERA-3 Alphabet pretest, posttest, and national norming population scores in the three categories of the bottom quartile (n=188 kindergarten scores).
Kindergarten: TERA-3 Alphabet Subtest Scores in the Top Quartile

Figure 6k shows the percentage of TERA-3 Alphabet subtest pretest, posttest, and national norming population scores in the three categories of the top quartile. The percentage of posttest scores (dark blue bars) matches the national population (light blue bars); however there are fewer scores in the top two categories (the Superior and Very Superior categories), representing scores ranked at or above the 90th percentile, than in the national population.

Figure 6k.
The percentage of TERA-3 Alphabet pretest, posttest, and national norming population scores in the three categories of the top quartile (n=188 kindergarten scores).
Kindergarten: The Distribution of TERA-3 Alphabet Subtest Scores

Figure 7k shows the percentage of the Alphabet subtest pretest, posttest, and national norming population scores in the categories of the ability scale. The percentage of the posttest scores (dark blue bars) is better than in the national norming population (light blue bars) in all categories except the Superior and Very Superior categories. The distribution of posttest scores almost matches the national population, which indicates the kindergarten ELLM children closed the gap in achievement measured by the TERA-3 Alphabet subtest.

![Figure 7k](image)

*Figure 7k.*
The percentage of TERA-3 Alphabet pretest, posttest, and national norming population scores in the categories of the ability scale (n=188 kindergarten scores).
Kindergarten: TERA-3 Conventions of Print Subtest Scores in the Bottom Quartile

Figure 8k shows the percentage of TERA-3 Conventions of Print subtest pretest, posttest, and national norming population scores in the three categories of the bottom quartile. The percentage of posttest scores (dark blue bars) in the bottom quartile is 12% less than in the fall (striped bars). However, the percentage of scores in the bottom two categories (the *Very Poor* and *Poor* categories), representing scores ranked at or below the 9th percentile, shows very little change, with 37% in the fall and 35% in the spring.

![Figure 8k](image-url)

*Figure 8k.*
The percentage of TERA-3 Conventions of Print subtest pretest, posttest, and national norming population scores in the three categories of the bottom quartile (n=188 kindergarten scores).
Kindergarten: TERA-3 Conventions of Print Subtest Scores in the Top Quartile

Figure 9k displays the percentage of TERA-3 Conventions of Print subtest pretest, posttest, and national norming population scores in the three categories of the top quartile. The percentage of posttest scores in these categories (dark blue bars) remains lower than in the national norming population (light blue bars) but is 7% more than in the fall (striped bars).

![Bar chart showing percentages of TERA-3 Conventions of Print subtest scores in the top quartile.](chart)

**Figure 9k.**
The percentage of TERA-3 Conventions of Print subtest pretest, posttest, and national norming population scores in the three categories of the top quartile (n=188 kindergarten scores).
Kindergarten: The Distribution of TERA-3 Conventions of Print Subtest Scores

Figure 10k shows the percentage of the TERA-3 Conventions of Print subtest pretest, posttest, and national norming population scores in the categories of the ability scale. While the distribution of the posttest scores (dark blue bars) remains shifted toward the low side of the ability score, it has shifted closer to the national norming population (light blue bars). The biggest change in the percentage of scores is the 10% fewer scores in the Below Average category. These scores are now spread across the Average, Above Average, and Superior categories. This indicates the ELLM children with scores in the bottom quartile were closing the gap in achievement measured by the TERA-3 Conventions of Print subtest.

Figure 10k.
The percentage of TERA-3 Conventions of Print subtest pretest, posttest, and national norming population scores in the categories of the ability scale (n=188 kindergarten scores).
Kindergarten: TERA-3 Meaning Subtest Scores in the Bottom Quartile

Figure 11k shows the percentage of TERA-3 Meaning subtest pretest, posttest, and national norming population scores in the three categories of the bottom quartile. In the spring, over 75% of the scores are in these categories (dark blue bars), which is 18% more than in the fall (striped bars). Additionally, the percentage of posttest scores in the bottom two categories (the Very Poor and Poor categories), representing scores ranked at or below the 9th percentile, increased 16% from the fall (striped bars).

Figure 11k.
The percentage of TERA-3 Meaning subtest pretest, posttest, and national norming population scores in the three categories of the bottom quartile (n=188 kindergarten scores).
Kindergarten: The Distribution of TERA-3 Meaning Scores

Figure 12k shows the percentage of the TERA-3 Meaning subtest pretest, posttest, and national norming population scores in the categories of the ability scale. There are 2% more posttest scores (dark blue bars) in the top quartile than pretest scores (striped bars). However, the entire distribution has shifted to the lower end of the ability scale, indicating the gap in achievement measured by the TERA-3 Meaning subtest widened over the kindergarten year for the ELLM children.

Figure 12k.
The percentage of TERA-3 Meaning subtest pretest, posttest, and national norming population scores in the categories of the ability scale (n=188 kindergarten scores).
Question 3: How did ELLM kindergarten children compare to existing national benchmarks in the recognition of the upper- and lowercase letters of the alphabet at the end of the school year?

In 2003/2004, 163 ELLM kindergarten children had Alphabet Letter Recognition Inventory (ALRI) pretest and posttest scores. The ALRI pretest mean score indicates the typical ELLM kindergarten child recognized about 55% of the letters. The ALRI posttest mean score indicates the typical kindergarten child recognized 94% of the letters. The Early Childhood Longitudinal Study-Kindergarten (ECLS-K) defined proficient as recognizing 75% of the sampled letters; therefore, at the end of the school year, the typical ELLM kindergarten child was proficient. ECLS-K researchers also reported that 39% of children living in poor families were proficient when entering kindergarten for the first time. Inspection of Figure 13k indicates 34% of the ELLM kindergarten children were proficient at the beginning of the year, which is slightly less than the percentage of proficient kindergarten children living in poor families in the national ECLS-K sample. However, the 34% of ELLM children includes all ELLM kindergarten children, regardless of when they entered kindergarten.

To determine the range of alphabet letter recognition ability of ELLM kindergarten children, ALRI scores are displayed in Figure 13k on the next page using four recognition categories: 0-13 letters, 14-26 letters, 27-39 letters, and 40-52 letters. Inspection of Figure 13k indicates 93% of the posttest scores are in the 40-to-52 letters recognized category (recognizing at least 75% of the letters; therefore, proficient). Additionally, 69% of the ELLM kindergarten children recognized at least 50 letters.
Figure 13k.
Alphabet Letter Recognition Inventory pretest and posttest scores of 163 ELLM kindergarten children.
Question 4: Who were the ELLM kindergarten children selected for targeted instruction in phonological awareness, and was ELLM effective in improving their reading readiness based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?

The answer to the first part of this evaluation question involves a description of the 68 kindergarten children selected for targeted instruction in phonological awareness by age and gender.

**Age**

As previously mentioned, the categorization of the children as kindergarteners does not indicate the age of the children. Because of the September 1 birthday cut-off for children attending Florida public kindergarten, the typical kindergarten child is between 60 and 72 months old on September 1 of the school year. Figure 14k shows the ages of the 68 kindergarten ELLM children selected for targeted instruction in phonological awareness.

![Figure 14k](image)

*Ages of children in months on September 1, 2003 (n = 68 ELLM kindergarten children selected for targeted instruction in phonological awareness).*
Over 26% of the kindergarten children selected for targeted instruction in phonological awareness were older than the typical range for kindergarten children indicating some of these children have been retained in grade. Because the TERA-3 is normed by the age of the child rather than the grade level, these children could be performing adequately when measured against children in kindergarten rather than against their same-age peers.

**Gender**

Of the 68 kindergarten children selected for targeted instruction in phonological awareness, 40% were girls, and 60% were boys. In the complete ELLM kindergarten sample, 49% were boys. Therefore, the scores of boys are disproportionately represented in the bottom quartile of ELLM scores.

**TERA-3 Results: Phonological Awareness Population**

The answer to the second part of the evaluation question comes from the TERA-3 scores of these 68 children. Table 4k provides summary statistics for the TERA-3 pretest and posttest scores of the ELLM kindergarten children selected for targeted instruction in phonological awareness. Data were analyzed as a repeated-measures ANOVA design to determine if there were statistically significant differences in the TERA-3 pretest and posttest mean scores. Effect sizes are provided for mean scores that represent differences.

Table 4k

<table>
<thead>
<tr>
<th>Summary Statistics and ANOVA Results: ELLM Kindergarten TERA-3 Reading Quotient and TERA-3 Subtests (Phonological Awareness Population)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Reading Quotient</strong></td>
</tr>
<tr>
<td>67.7</td>
</tr>
<tr>
<td>4.8</td>
</tr>
<tr>
<td>5.4</td>
</tr>
</tbody>
</table>

* Indicates significant difference at the \( \alpha = .05 \) level.
+ Indicates an improvement that can be attributed to children’s participation in ELLM and targeted instruction in phonological awareness.
φ Denotes a small effect.
φφ Denotes a medium effect.
φφφ Denotes a large effect.
Indicates a statistically significant and meaningful effect.
Targeted instruction in phonological awareness primarily addresses the content measured by the TERA-3 Alphabet subtest, but the kindergarten children also improved on the Reading Quotient test. Statistical analyses indicated both of these increases were significant and meaningful. Mirroring the findings of the general ELLM kindergarten population, the scores of the ELLM kindergarten children selected for targeted instruction in phonological awareness lost significant and meaningful ranking relative to the national norming population in the content measured by the TERA-3 Meaning subtest.

The TERA-3 Alphabet subtest improvement of 120% of a standard deviation clearly drove the improvement of 51% of a standard deviation in the Reading Quotient mean scores, which occurred in spite of the negative effect of the TERA-3 Meaning subtest scores. To better understand where the improvement occurred, the TERA-3 Alphabet subtest scores are shown in Figure 15k using the categories described in Table 1, page 4.
Kindergarten Children Selected for Targeted Instruction in Phonological Awareness: TERA-3 Alphabet Subtest Scores

Figure 15k shows the percentages of the TERA-3 Alphabet subtest pretest, posttest, and national norming population scores in the four ability categories with 100% of the scores. There are 60% fewer posttest scores (dark blue bars) in the bottom quartile than pretest scores (striped bars), and the percentage of posttest scores in the bottom quartile is only 5% greater than in the national population (light blue bars). In the fall, 74% of the scores (striped bars) are in the bottom two categories (the Very Poor and Poor categories), representing scores ranked at or below the 9th percentile, but only 18% of the scores remain in these categories in the spring (dark blue bars).

Figure 15k.
The percentage of TERA-3 Alphabet subtest pretest, posttest, and national norming population scores in the bottom four categories of the ability scale (n=68 scores of kindergarten children selected for targeted instruction in phonological awareness).
**ALRI Results: Phonological Awareness Population**

In 2003/2004, 60 ELLM kindergarten children selected for targeted instruction in phonological awareness had Alphabet Letter Recognition Inventory (ALRI) pretest and posttest scores. The ALRI pretest mean score indicates the typical ELLM kindergarten child in this group recognized about 34% of the letters. The ALRI posttest mean score indicates the typical kindergarten child recognized 88% of the letters. The *Early Childhood Longitudinal Study-Kindergarten* (ECLS-K) defined *proficient* as recognizing 75% of the sampled letters; therefore, the typical ELLM kindergarten child selected for targeted instruction in phonological awareness was *proficient* in letter recognition at the end of the school year.

To determine the range of alphabet letter recognition ability of ELLM kindergarten children, ALRI scores are displayed in Figure 16k using four recognition categories: 0-13 letters, 14-26 letters, 27-39 letters, and 40-52 letters. Inspection of Figure 16k indicates almost 83% of the posttest scores are in the 40-to-52 letters recognized category (recognizing at least 75% of the letters; therefore, *proficient*).

![Figure 16k](image-url)

*Figure 16k.* Alphabet Letter Recognition Inventory pretest and posttest scores of the 60 ELLM kindergarten children selected for targeted instruction in phonological awareness.
Kindergarten Summary

Table 5k summarizes the changes in the percentage of TERA-3 scores in the bottom quartile from pretest to posttest.

Table 5k
Summary of ELLM Kindergarten Children’s TERA-3 Scores in the Top Three Quartiles

<table>
<thead>
<tr>
<th>TERA-3 Test</th>
<th>Percent of Pretest Scores Average or Above</th>
<th>Percent of Posttest Scores Average or Above</th>
<th>Percent of Change</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>30.3</td>
<td>42.6</td>
<td>12.2</td>
<td>0.340</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>57.5</td>
<td>87.8</td>
<td>30.3</td>
<td>0.660</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>29.3</td>
<td>41.0</td>
<td>11.7</td>
<td>0.350</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>39.4</td>
<td>21.8</td>
<td>-17.6</td>
<td>-0.230</td>
</tr>
</tbody>
</table>

+ Denotes an effect that can be attributed to the children’s participation in ELLM classes.
‰ Denotes a medium effect.
Denotes at least 25 percent fewer scores in the bottom quartile at posttest.
Denotes the percentage of ELLM children matches the percentage in the national population

The largest change in the percentage of scores in the bottom quartile occurred in the TERA-3 Alphabet subtest. At the end of the school year, less than 25% of the ELLM kindergarten children’s scores were in the bottom quartile. In fact, the distribution of the Alphabet subtest scores matched the distribution of national norming population scores.

The second largest change in the percentage of scores in the bottom quartile occurred in the TERA-3 Meaning subtest. At the end of the school year, almost 80% of the ELLM kindergarten children’s scores ranked at or below the 25th percentile, which is 17% more than at the beginning of the school year.
Table 6k presents the improvements made by the children selected for targeted instruction in phonological awareness and summarizes the changes in the percentage of TERA-3 scores in the bottom quartile (ranked above the 9th percentile) from pretest to posttest.

Table 6k
Summary of TERA-3 Scores of the ELLM Kindergarten Children Selected for Targeted Instruction in Phonological Awareness

<table>
<thead>
<tr>
<th>TERA-3 Test</th>
<th>Percent of Pretest Scores Average or Above and Above the 9th Percentile in Parentheses</th>
<th>Percent of Posttest Scores Average or Above and Above the 9th Percentile in Parentheses</th>
<th>Percent of Change</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>0.0 (0.0)</td>
<td>8.8 (27.9)</td>
<td>8.8 (27.9)</td>
<td>0.512φ⁺</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>8.8 (26.5)</td>
<td>69.1 (82.3)</td>
<td>60.3 (55.9)</td>
<td>1.201φφφ⁻</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>4.4 (29.4)</td>
<td>10.3 (30.8)</td>
<td>5.9 (1.5)</td>
<td>0.191φ⁴⁻</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>7.4 (50.0)</td>
<td>2.9 (32.4)</td>
<td>-17.7 (-4.4)</td>
<td>-0.211φ⁻</td>
</tr>
</tbody>
</table>

The numbers in parenthesis represent percentage of scores ranked above the 9th percentile.

⁺ Denotes an effect that can be attributed to the children’s participation in ELLM and targeted instruction in phonological awareness.

φ Denotes a small effect.

φφ Denotes a medium effect.

φφφ Denotes a large effect.

□ Denotes there were at least 15 percent fewer scores in the bottom quartile at posttest.

The scores of children selected for targeted instruction in phonological awareness showed significant and meaningful improvement in the TERA-3 Alphabet subtest. The largest change in the percentage of scores in the bottom quartile matches the largest effect size – both occurred in the scores of the TERA-3 Alphabet subtest. At the end of the school year, there were 60% fewer scores ranked at or below the 50th percentile than at the beginning of the school year. Additionally, there were only 5% more of these children’s posttest scores in the bottom quartile than in the national norming population.

Again, for the children selected for targeted instruction in phonological awareness, the second largest change in the percentage of scores in the bottom quartile occurred in the scores of the TERA-3 Meaning subtest. At the end of the school year, 97% of the TERA-3 Meaning subtest scores ranked at or below the 25th percentile.
SPECIFIC KINDERGARTEN RESULTS

Question 1: Who were the 2003/2004 ELLM kindergarten children?

- 16% of the ELLM kindergarten children were older than the typical age range of Florida kindergarten children.
- 49% of the ELLM kindergarten children were boys and 51% were girls.
- All of the ELLM kindergarten children were Black.
- There were complete TERA-3 scores for 188 kindergarten children (19% attrition). Additionally, there was evidence suggesting an underlying mechanism for the attrition. Children with incomplete scores had lower initial status.

Question 2: Was ELLM effective in improving the reading readiness of ELLM kindergarten children based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?

- The differences between the pretest and posttest mean scores of all TERA-3 tests were significant and meaningful.
- The TERA-3 Reading Quotient mean scores showed a significant and meaningful improvement of 34% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  - 12% fewer ELLM Reading Quotient posttest scores were in the bottom quartile than pretest scores.
- The TERA-3 Alphabet subtest mean scores showed a significant and meaningful improvement of 66% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  - Fewer ELLM Alphabet subtest scores were in the bottom quartile than in the national norming population.
  - 2% fewer ELLM Alphabet subtest scores ranked at or below the 9th percentile than in the national norming population.
  - The percentage of ELLM Alphabet subtest scores in the top quartile exceeded the percentage in the national norming population.
  - The mean ELLM Alphabet subtest score ranked 10 percentiles higher than the mean ranking of the national norming population.
- The TERA-3 Conventions of Print subtest mean scores showed a significant and meaningful improvement of 35% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  - At the end of the school year:
    - 12% fewer ELLM Conventions of Print subtest scores were in the bottom quartile than in the fall.
Over the school year, the TERA-3 Meaning subtest mean scores indicated significant and meaningful decline in ranking relative to the national norming population. At the end of the school year:
- 75% of the ELLM Meaning subtest scores were in the bottom quartile. This is 18% more than at the beginning of the school year.
- 41% of the ELLM Meaning subtest scores ranked at or below the 9th percentile. This is 18% more than at the beginning of the school year.

**Question 3: How did ELLM kindergarten children compare to existing national benchmarks in the recognition of the upper- and lowercase letters of the alphabet at the end of the school year?**

At the end of the school year:
- The typical ELLM kindergarten child recognized 94% of the letters. ECLS-K researchers defined *proficient* as recognizing at least 75% of sampled letters; therefore, the typical kindergarten ELLM child was *proficient*.
- 93% of the ELLM kindergarten children were *proficient*.
- 69% of the ELLM kindergarten children at least recognized 50 letters.

**Question 4: Who were the ELLM kindergarten children selected for targeted instruction in phonological awareness, and was ELLM effective in improving their reading readiness based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?**

The 68 ELLM kindergarten children selected for targeted instruction in phonological awareness (36.2% of the kindergarten ELLM children) included a larger percentage of boys and older children than the complete ELLM kindergarten sample.

The TERA-3 Reading Quotient and Alphabet subtest mean scores indicated significant and meaningful improvements of 51% and 120% of a standard deviation, respectively. These improvements can be attributed to the children’s participation in ELLM and targeted instruction in phonological awareness. At the end of the school year:
- 60% fewer of the Alphabet subtest scores were in the bottom quartile than in the fall, and the percentage is only 5% greater than in the national norming population.
- 56% fewer of the ELLM Alphabet subtest scores ranked at or below the 9th percentile than in the fall.
Mirroring the complete ELLM kindergarten sample, the TERA-3 Meaning subtest mean scores indicated a significant and meaningful decline in ranking relative to the national norming population.

At the end of the school year, the typical ELLM kindergarten child selected for targeted instruction in phonological awareness recognized 88% of the letters; therefore, was proficient.

83% of the ELLM kindergarten children selected for targeted instruction in phonological awareness were proficient at the end of the year.

Table 1f shows the contribution of ELLM first-grade sites and classes to the evaluation of ELLM in the 2003/2004 school year. There were TERA-3 pretest scores from 225 ELLM first-grade children and 184 posttest scores.

Table 1f
The Number of ELLM First-Grade Sites, Teachers, and Assessed Children

<table>
<thead>
<tr>
<th>Attrition</th>
<th>Number of Sites</th>
<th>Number of Classes</th>
<th>Number of Children</th>
</tr>
</thead>
<tbody>
<tr>
<td>Complete Scores</td>
<td>3</td>
<td>12</td>
<td>184</td>
</tr>
<tr>
<td>Incomplete Scores</td>
<td>3</td>
<td>10</td>
<td>41</td>
</tr>
</tbody>
</table>

Attrition of Children

The discrepancy between the number of pretest and posttest scores represents the attrition of 18% of the first-grade children. To determine if the attrition occurred at random, the TERA-3 pretest scores of the ELLM children who were not administered the TERA-3 posttest were compared to pretest scores of ELLM children who were administered the TERA-3 posttest. Table 2f reports the results of the comparison.

Table 2f
Summary Statistics and ANOVA Results: ELLM TERA-3 Reading Quotient and TERA-3 Subtest Pretest Scores by Attrition Category

<table>
<thead>
<tr>
<th>Test</th>
<th>Mean Scores Complete (184)</th>
<th>Mean Scores Incomplete (41)</th>
<th>Probability Means Differ</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>86.0</td>
<td>81.6</td>
<td>.0881</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>10.1</td>
<td>9.3</td>
<td>.0306*</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>8.1</td>
<td>7.5</td>
<td>.2970</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>5.3</td>
<td>4.6</td>
<td>.1610</td>
</tr>
</tbody>
</table>

* Indicates significant difference at the $\alpha = .05$ level.
Numbers in parentheses indicate the number of scores.

The Alphabet subtest mean score of the children not available in the spring was lower than the children who were available. However, since there were no significant differences in the pretest mean scores of the remaining TERA-3 subtests, the suggestion of an underlying mechanism contributing to the attrition may not be founded.
Question 1: Who were the 2003/2004 ELLM first-grade children?

The answer to this evaluation question involves a description of the children by age, gender, and ethnicity.

*Age*

The categorization of the children as first graders does not indicate the age of the children. Because of the September 1 birthday cut-off for children attending Florida public kindergarten, the typical first-grade child is between 72 and 84 months old on September 1 of the school year. Figure 1f displays the ages of the ELLM first-grade children who had both TERA-3 pretest and posttest scores.

![Figure 1f: Ages of the first-grade children in months on September 1, 2003 (n = 184).](image)

Even though more than 18% of the children were older than the typical age range of Florida first-grade children, 55% of the children in the typical 72-to-84 month age span were older than the 78-month median age in the typical span for first-grade children.
**Gender**

Of the 184 first-grade children with TERA-3 pretest and posttest scores, 53% were girls and 47% were boys.

**Ethnicity**

The ethnicity of the ELLM children was categorized as *Black, White, Hispanic, Native American, Asian,* and *Other.* All of the ELLM first-grade children were *Black.*

**Question 2: Was ELLM effective in improving the reading readiness of ELLM first-grade children based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?**

Table 3f provides summary statistics for the TERA-3 pretest and posttest scores of the ELLM first-grade children with both scores. Data were analyzed as a repeated-measures ANOVA design to determine if there were statistically significant differences in the TERA-3 Reading Quotient and TERA-3 subtest pretest and posttest mean scores. Effect sizes are provided for mean scores that represent differences.

**Table 3f**

*Summary Statistics and ANOVA Results: ELLM First-Grade TERA-3 Reading Quotient and TERA-3 Subtest*

<table>
<thead>
<tr>
<th>TERA-3 Test</th>
<th>Pretest Mean</th>
<th>Posttest Mean</th>
<th>Probability</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>86.0</td>
<td>91.4</td>
<td>&lt;.0001⁺</td>
<td>0.360⁺⁺</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>10.1</td>
<td>9.7</td>
<td>.0048⁻</td>
<td>-0.118⁻⁻</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>8.1</td>
<td>9.2</td>
<td>&lt;.0001⁺</td>
<td>0.353⁺⁺</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>5.3</td>
<td>7.0</td>
<td>&lt;.0001⁺</td>
<td>0.596⁺⁺</td>
</tr>
</tbody>
</table>

* Denotes there was a difference in the pretest and posttest means, α=.05.
⁺ Denotes a small effect.
⁺⁺ Denotes a medium effect.
⁺⁺⁺ Denotes an effect that may be attributed to children’s participation in ELLM.
Denotes a significant and meaningful difference

The pretest mean scores of the TERA-3 Reading Quotient, Alphabet, Conventions of Print, and Meaning subtests rank at the 18th, 51st, 26th, and 6th percentiles, respectively. The posttest mean scores of the Reading Quotient, Alphabet, Conventions of Print, and Meaning subtests rank at the 28th, 46th, 39th, and 16th percentiles, respectively. The differences in the pretest and posttest TERA-3 mean scores of all subtests are statistically significant and all differences except the Alphabet subtest difference are positive. The
negative difference in the Alphabet subtest do not represent large enough effect to be meaningful. Meaningful effect sizes range from 0.360 to 0.596, and all can be attributed to the children’s participation in ELLM.

To better understand where the improvement in reading readiness occurred, TERA-3 Reading Quotient and subtest scores are displayed in seven ability categories: three categories representing the lowest 25 percentiles, one category representing the middle 50 percentiles, and three categories representing the highest 25 percentiles. (See Table 1, page 4). Figures 2f-4f display the Reading Quotient scores, Figures 5f-7f display the Convention of Print subtest scores, and Figures 8f-10f display the Meaning subtest scores.

**First Grade: TERA-3 Reading Quotient Scores in the Bottom Quartile**

Figure 2f displays the percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the three categories of the bottom quartile.

![Figure 2f](data:image/png;base64,iVBORw0KGgoAAAANSUhEUgAAAiwAAAAHgCAYAAABPFQPjAAAAGXRFWHRTb2Z0d2FyZQBBZG9iZSBJbWFnZVJlYWR5ccllPAAAAyElEQVR42mOy8/8w8DP1w/BWwAEGQkgEEV6qEIDQhCVAECdYcAhIwAABkADgADgAAAAAElFTkSuQmCC)

*Figure 2f.*
The percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the three categories of the bottom quartile (n=184 first-grade scores).
While the percentage of posttest scores in these categories (dark blue bars) remains higher than in the national norming population (light blue bars), there are 11% fewer scores in these categories in the spring than in the fall (striped bars), and the percentage of scores in the bottom two categories (the Very Poor and Poor categories), representing scores ranked at or below the 9th percentile, is about 11% more than in the national norming population.

**First Grade: TERA-3 Reading Quotient Scores in the Top Quartile**

Figure 3f displays the percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the three categories of the top quartile. While the percentage of posttest scores in these categories (dark blue bars) remains lower than in national norming population (light blue bars), there are 4% more scores in these categories in the spring than in the fall (striped bars).

![Bar chart showing the percentage of TERA-3 Reading Quotient scores in the top quartile for Pretest, Posttest, and National norms.

Figure 3f.
The percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the three categories of the top quartile (n=184 first-grade scores).
First Grade: The Distribution of TERA-3 Reading Quotient Scores

Figure 4f shows the percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the categories of the ability scale. While the distribution of the posttest scores (dark blue bars) remains shifted toward the low side of the ability scale, it has shifted closer to the national norming population (light blue bars). This indicates the ELLM first-grade children (especially those with scores in the lowest five categories) were closing the gap in achievement measured by the TERA-3 Reading Quotient.

![Bar chart showing the distribution of TERA-3 Reading Quotient scores](chart)

**Figure 4f.**
The percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the categories of the ability scale (n=184 first-grade scores).
First Grade: TERA-3 Conventions of Print Subtest Scores in the Bottom Quartile

Figure 5f shows the percentage of TERA-3 Conventions of Print subtest pretest, posttest, and national norming population scores in the three categories of the bottom quartile. There are 26% fewer scores in these categories in the spring (dark blue bars) than in the fall (striped bars). The percentage of posttest scores in the bottom quartile matches the national norming population (light blue bars). Additionally, the percentage of scores in the bottom two categories (the Very Poor and Poor categories), representing scores ranked at or below the 9th percentile, is 2% less than in the national norming population.

Figure 5f.
The percentage of TERA-3 Conventions of Print subtest pretest, posttest, and national norming population scores in the three categories of the bottom quartile (n=184 first-grade scores).
First Grade: TERA-3 Conventions of Print Subtest Scores in the Top Quartile

Figure 6f displays the percentage of TERA-3 Conventions of Print pretest, posttest, and national norming population scores in the three categories of the top quartile. The percentage of posttest scores in these categories (dark blue bars) remains lower than in the national norming population (light blue bars) and nearly the same as pretest scores (striped bars). However, there are 3% more scores ranked at or above the 90th percentile than in the spring.

Figure 6f.
The percentage of TERA-3 Conventions of Print subtest pretest, posttest, and national norming population scores in the three categories of the top quartile (n=184 first-grade scores).
First Grade: The Distribution of TERA-3 Conventions of Print Subtest Scores

Figure 7f shows the percentage of TERA-3 Conventions of Print subtest pretest, posttest, and national norming population scores in the categories of the ability scale. While the distribution of the posttest scores (dark blue bars) remains shifted toward the low side of the ability scale, it has shifted closer to the national norming population (light blue bars) than the fall scores (striped bars). This is especially true of the categories in the bottom quartile and the Average category. This indicates the ELLM children were closing the gap in achievement measured by the TERA-3 Conventions of Print subtest.

Figure 7f.
The percentage of TERA-3 Conventions of Print subtest pretest, posttest, and national norming population scores in the categories of the ability scale (n=184 first-grade scores).
First Grade: TERA-3 Meaning Subtest Scores in the Bottom Quartile

Figure 8f shows the percentage of TERA-3 Meaning subtest pretest, posttest, and national norming population scores in the three categories of the bottom quartile. There are 28% fewer scores in these categories in the spring (dark blue bars) than in the fall (striped bars). Additionally, the percentage of scores in the bottom two categories (the Very Poor and Poor categories), representing scores ranked at or below the 9th percentile, is 16% less than in the fall (striped bars).

![Bar chart showing TERA-3 Meaning subtest scores in the bottom quartile.](chart)

Figure 8f.
The percentage of TERA-3 Meaning subtest pretest, posttest, and national norming population scores in the three categories of the bottom quartile (n=184 first-grade scores).
First Grade: TERA-3 Meaning Subtest Scores in the Top Quartile

Figure 9f displays the percentage of TERA-3 Meaning subtest pretest, posttest, and national norming population scores in the three categories of the top quartile. While the percentage of posttest scores in these categories (dark blue bars) remains lower than in the national norming population (light blue bars), there is a 12% increase in scores in the top quartile from fall testing (striped bars).

**Figure 9f:**
The percentage of TERA-3 Meaning subtest pretest, posttest, and national norming population scores in the three categories of the top quartile (n=184 first-grade scores).
First Grade: The Distribution of TERA-3 Meaning Subtest Scores

Figure 10f shows the percentage of the TERA-3 Meaning subtest pretest, posttest, and national norming population scores in the categories of the ability scale. While the distribution of the posttest scores (dark blue bars) remains shifted toward the low side of the ability scale, it has shifted closer to the national norming population (light blue bars) than the fall scores (striped bars) indicating first-grade ELLM children (especially those with scores in the bottom quartile and Average categories) were closing the gap in achievement in the content measured by the TERA-3 Meaning subtest.

![Diagram showing the distribution of TERA-3 Meaning subtest scores in different ability categories](image)

**Figure 10f.**
The percentage of TERA-3 Meaning subtest pretest, posttest, and national norming population scores in the categories of the ability scale (n=184 first-grade scores).
Question 4: Who were the ELLM first-grade children selected for targeted instruction in phonological awareness, and was ELLM effective in improving their reading readiness based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?

The answer to the first part of this evaluation question involves a description of the 36 first-grade children selected for targeted instruction in phonological awareness by age and gender.

*Age*

As previously mentioned, the categorization of the children as first graders does not indicate the age of the children. Because of the September 1 birthday cut-off for children attending Florida public kindergarten, the typical first-grade child is between 72 and 84 months old on September 1 of the school year. Figure 11f displays the ages of the first-grade ELLM children selected for targeted instruction in phonological awareness.

*Figure 11f.*
Ages of the children in months on September 1, 2003 (n = 36 first-grade children selected for targeted instruction in phonological awareness).
Over 50% of the first-grade children selected for targeted instruction in phonological awareness were older than the typical range for first-grade children, indicating some of these children have been retained in grade. Because the TERA-3 is normed by the age of the child rather than grade level, these children could actually be performing adequately when measured against children in the first-grade rather than against their same-age peers.

**Gender**

Of the 36 first-grade children selected for targeted instruction in phonological awareness, 44% were girls and 56% were boys. Forty-seven percent of the complete ELLM first-grade population were boys; therefore, the scores of boys are disproportionately represented in the bottom quartile of ELLM scores.

**TERA-3 Results: Phonological Awareness Population**

The answer to the second part of the evaluation question comes from the TERA-3 scores from these 36 children. Table 4f provides summary statistics for the TERA-3 pretest and posttest scores of the ELLM first-grade children selected for targeted instruction in phonological awareness. Data were analyzed as a repeated-measures ANOVA design to determine if there were statistically significant differences in the TERA-3 pretest and posttest mean scores. Effect sizes are provided for mean scores that represent differences.

Table 4f

*Summary Statistics and ANOVA Results: ELLM First-Grade TERA-3 Reading Quotient and TERA-3 Subtests (Phonological Awareness Population)*

<table>
<thead>
<tr>
<th>Subtest</th>
<th>Pretest Means</th>
<th>Posttest Means</th>
<th>Probability Means Differ</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>66.9</td>
<td>76.1</td>
<td>&lt;.0001*</td>
<td>0.613**</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>7.3</td>
<td>7.9</td>
<td>.0701</td>
<td>0.213</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>4.3</td>
<td>7.1</td>
<td>&lt;.0001*</td>
<td>0.935***</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>2.9</td>
<td>3.8</td>
<td>.0618*</td>
<td>0.278</td>
</tr>
</tbody>
</table>

* Indicates significant difference at the $\alpha = .05$ level.
+ Denotes a small effect.
+++ Denotes a large effect.
+ Denotes an effect that may be attributed to children’s participation in ELLM.
 Denotes a significant and meaningful difference.

Targeted instruction in phonological awareness primarily addresses the content measured by the TERA-3 Alphabet subtest, but these first-grade children improved in all
TERA-3 subtest content areas. Statistical analyses indicated the improvement in TERA-3 Reading Quotient and Conventions of Print subtest scores was significant, meaningful, and attributable to the children’s participation in ELLM. Therefore, to better understand where the improvement occurred, the TERA-3 Reading Quotient scores are shown in Figure 12f using the categories described in Table 1, page 4.

**First-Grade Children Selected for Targeted Instruction in Phonological Awareness: TERA-3 Reading Quotient Scores**

Figure 12f shows the percentage of pretest, posttest, and national norming population scores in the four ability categories, with 100% of the TERA-3 Reading Quotient scores. There are 8% fewer posttest scores (dark blue bars) in the bottom quartile than pretest scores (striped bars). However, there are 39% fewer posttest scores in the bottom two categories (the *Very Poor* and *Poor* categories), representing scores at or below the 9th percentile.

Figure 12f.
The percentage of TERA-3 Reading Quotient pretest, posttest, and national norming population scores in the bottom four categories of the ability scale (n=36 scores of first-grade children selected for targeted instruction in phonological awareness).
First-Grade Summary

Table 5f summarizes the changes in the percentage of TERA-3 Reading Quotient and subtests scores in the bottom quartile from pretest to posttest.

Table 5f
Summary of ELLM First-Grade Children’s TERA-3 Scores in the Top Three Quartiles

<table>
<thead>
<tr>
<th>TERA-3 Test</th>
<th>Percent of Pretest Scores Above Average</th>
<th>Percent of Posttest Scores Above Average</th>
<th>Percent of Change</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>38.0</td>
<td>49.0</td>
<td>11.0</td>
<td>0.360*</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>88.6</td>
<td>91.8</td>
<td>3.3</td>
<td>-0.118</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>48.4</td>
<td>73.9</td>
<td>25.5</td>
<td>0.353**</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>18.5</td>
<td>46.7</td>
<td>28.3</td>
<td>0.596***</td>
</tr>
</tbody>
</table>

* Denotes a small effect.
** Denotes a medium effect.
+ Indicates improvement that can be attributed to children’s participation in ELLM.
Denotes there were at least 25 percent fewer scores in the bottom quartile at posttest.
Denotes the percentage of ELLM children matches the percentage in the national population.

Even though the TERA-3 Alphabet subtest posttest mean score was not as high as the pretest mean score, there were 3% fewer scores in the bottom quartile at the end of the year. Additionally, there were 16% fewer ELLM first-grade children’s scores in the bottom quartile than in the national norming population.

At the end of the year, there were 25% fewer TERA-3 Conventions of Print subtest scores in the bottom quartile. Additionally, the percentage of ELLM first-grade children’s scores in the bottom quartile nearly matched the national norming population.

The largest effect size and the largest change in percentage of scores in the bottom quartile occurred with the TERA-3 Meaning subtest scores. However, at the end of the year, the typical first-grade ELLM child’s score ranked at the 16th percentile nationally, which ranks in the Below Average category of the bottom quartile.
Table presents the improvements made by ELLM first-grade children selected for targeted instruction in phonological awareness and summarizes the changes in the percentage of TERA-3 scores of the children selected for targeted instruction in phonological awareness in the bottom quartile (ranked above the 9th percentile) from pretest to posttest.

Table 6f
*Summary of TERA-3 Scores of the First-Grade Children Selected for Targeted Instruction in Phonological Awareness*

<table>
<thead>
<tr>
<th>TERA-3 Test</th>
<th>Percent of Pretest Scores <em>Average</em> or Above (Above the 9th Percentile in Parentheses)</th>
<th>Percent of Posttest Scores <em>Average</em> or Above (Above the 9th Percentile in Parentheses)</th>
<th>Percent of Change</th>
<th>Effect Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Reading Quotient</td>
<td>0.0 (2.8)</td>
<td>8.3 (41.7)</td>
<td>8.3 (38.9)</td>
<td>0.61³⁶⁵</td>
</tr>
<tr>
<td>Alphabet Subtest</td>
<td>50.0 (91.7)</td>
<td>63.9 (97.2)</td>
<td>13.9 (5.6)</td>
<td>0.21³⁶⁵</td>
</tr>
<tr>
<td>Conventions Subtest</td>
<td>0.0 (19.4)</td>
<td>36.1 (72.2)</td>
<td>36.1 (52.8)</td>
<td>0.93³⁶⁵</td>
</tr>
<tr>
<td>Meaning Subtest</td>
<td>0.0 (2.8)</td>
<td>16.7 (22.2)</td>
<td>16.7 (19.4)</td>
<td>0.27³⁶⁵</td>
</tr>
</tbody>
</table>

The numbers in parenthesis represent percentage of scores ranked above the 9th percentile.

³ Denotes a small effect.
⁶ Denotes a medium effect.
⁹ Denotes a large effect.
+ Indicates improvement that can be attributed to children’s participation in ELLM.
☐ Denotes there were at least 15 percent fewer scores in the bottom quartile at posttest.

At the end of the year, there were 14% more TERA-3 Alphabet subtest scores in the top three quartiles, and there were fewer scores ranked at or below the 9th percentile than in the national norming population.

The largest change in percentage of scores and the largest effect size are associated with the TERA-3 Conventions of Print subtest. There were over 50% fewer posttest scores ranked at or below the 9th percentile than pretest scores. The Conventions of Print posttest mean score of the first-grade children selected for targeted instruction in phonological awareness ranked at the 17th percentile, which is 15 percentiles above the ranking of the mean pretest score.
Question 1: Who were the 2003/2004 ELLM first-grade children?

- 18% of the ELLM first-grade children were older than the typical age range of Florida first-grade children.
- 47% of the ELLM first-grade children were boys and 53% were girls.
- All of the ELLM first-grade children were Black.
- There were complete TERA-3 scores for 184 first-grade children (19% attrition).

Question 2: Was ELLM effective in improving the reading readiness of ELLM first-grade children based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?

- The differences between the pretest and posttest mean scores of TERA-3 Reading Quotient, Conventions of Print, and Meaning subtests were significant and meaningful.
- The TERA-3 Reading Quotient mean scores showed a significant and meaningful improvement of 36% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  - 11% fewer ELLM Reading Quotient posttest scores were in the bottom quartile than pretest scores.
- The TERA-3 Conventions of Print subtest mean scores showed a significant and meaningful improvement of 35% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  - At the end of the school year:
    - 26% fewer ELLM Conventions of Print subtest scores were in the bottom quartile.
    - 2% fewer ELLM Conventions of Print subtest scores ranked at or below the 9th percentile than in the national norming population.
    - The percentage of ELLM Conventions of Print subtest scores in the bottom quartile nearly matched the percentage in the national norming population.
- The TERA-3 Meaning subtest mean scores showed a significant and meaningful improvement of 60% of a standard deviation. This improvement can be attributed to the children’s participation in ELLM classes.
  - At the end of the school year:
    - 28% fewer Meaning subtest scores were in the bottom quartile.
- The TERA-3 Alphabet subtest mean scores indicated significant but not meaningful decline in ranking relative to the national norming population.
Question 4: Who were the ELLM first-grade children selected for targeted instruction in phonological awareness, and was ELLM effective in improving their reading readiness based on improved TERA-3 Reading Quotient and TERA-3 subtest scores?

- The 36 ELLM first-grade children selected for targeted instruction in phonological awareness (19.6% of the first-grade ELLM children) included a larger percentage of boys and older children than the complete ELLM first-grade sample.
  - 50% of the children selected for targeted instruction in phonological awareness were older than the typical range for Florida first-grade children indicating these children have been retained in grade.

- The TERA-3 Reading Quotient and Conventions of Print subtests mean scores indicated significant and meaningful improvements of 61% and 94% of a standard deviation, respectively. These improvements can be attributed to the children’s participation in ELLM and targeted instruction in phonological awareness.
  At the end of the school year:
  - 36% fewer ELLM TERA-3 Conventions of Print subtest scores were in the bottom quartile.
  - 56% fewer of the ELLM Conventions of Print subtest scores ranked at or below the 9th percentile.
SIMILARITIES AND DIFFERENCES ACROSS COHORTS

Question 5: Were there similarities and differences in the demographic characteristics and emergent literacy achievement of the preschool, kindergarten, and first-grade ELLM children?

Table 2 presents the demographic information of the ELLM children by cohort, allowing for a comparison of the cohorts.

Table 2
Demographic Information of the ELLM Children by Cohort

<table>
<thead>
<tr>
<th>Cohort</th>
<th>Number</th>
<th>Gender % Boys</th>
<th>Age Older %</th>
<th>Ethnicity % Black</th>
<th>Attrition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschool</td>
<td>129</td>
<td>52</td>
<td>92</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>Kindergarten</td>
<td>188</td>
<td>47</td>
<td>16</td>
<td>100</td>
<td>19</td>
</tr>
<tr>
<td>First Grade</td>
<td>184</td>
<td>47</td>
<td>18</td>
<td>100</td>
<td>18</td>
</tr>
</tbody>
</table>

There were more children in the kindergarten and first-grade cohorts than in the preschool cohort. When combined across cohorts, 49% of the ELLM children were boys and 51% were girls. However, only the preschool cohort had a larger percentage of boys than girls. The preschool cohort was the only ethnically diverse cohort. The percentage of children who may have been retained in grade was very similar across the kindergarten and first-grade cohorts. The rate of attrition was larger in the kindergarten and first-grade cohorts than in the preschool cohort. Additionally, there was evidence that kindergarten children whose TERA-3 Reading Quotient, Alphabet, and Meaning subtests pretest scores were low were more likely to be unavailable in the spring. Therefore, results from this cohort should be interpreted cautiously. There was also evidence that first-grade children whose TERA-3 Alphabet subtest pretest scores were low were more likely to be unavailable in the spring. However, scores from the other TERA-3 subtests were not affected by the attrition of the first-grade children.
Table 3 presents a comparison of the emergent literacy achievement of the ELLM children across cohorts and TERA-3 tests.

Table 3
Summary of TERA-3 Emergent Literacy Achievement of the ELLM Children by Cohort

<table>
<thead>
<tr>
<th></th>
<th>Reading Quotient</th>
<th>Alphabet</th>
<th>Conventions</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschool</td>
<td>0.306</td>
<td>0.532*</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>Kindergarten</td>
<td>0.340*</td>
<td>0.660*</td>
<td>0.350*</td>
<td>-0.230</td>
</tr>
<tr>
<td>First-Grade</td>
<td>0.360*</td>
<td>-0.118</td>
<td>0.353*</td>
<td>0.596*</td>
</tr>
</tbody>
</table>

The numbers in the cells are the effect sizes.

+ Indicates an effect attributable to the children’s participation in ELLM classrooms.
- Indicates there were no more than 30% of the posttest scores in the bottom quartile.
- Indicates there were at least 20% fewer scores in the bottom quartile in the spring.

The preschool summary indicates the only meaningful improvement attributable to the children’s participation in ELLM was in the emergent literacy content measured by the TERA-3 Alphabet subtest. The kindergarten cohort made meaningful improvement attributable to the children’s participation in ELLM in two content areas: the TERA-3 Alphabet and the Conventions of Print subtests. The first-grade cohort also made meaningful improvement attributable to the children’s participation in ELLM in two content areas: the TERA-3 Conventions of Print and Meaning subtests. Therefore, the preschool and kindergarten cohorts’ achievement overlaps in the Alphabet subtest content, and the kindergarten and first-grade cohorts’ achievement overlaps in the Conventions of Print subtest content. The summary suggests a hierarchy exists in the TERA-3 subtests, but the publishers specifically indicate there is not a hierarchy. Furthermore, the process used to standardize test scores assures any developmental hierarchy in the skills measured is eliminated because the scores are ranked against a norming population of same-aged children rather than by content mastery. However, there may be a hierarchy in the emergent literacy continuum measured using TERA-3 standardized scores of the academically at-risk population of young children served by ELLM.

To better understand the achievement differences across cohorts and TERA-3 tests, a more detailed comparison that includes the initial status of the children is presented in a four-panel chart in Figure 1 on the next page, which is a foldout page.
The preschool children’s growth parallels the kindergarten children’s growth in the abilities measured by the TERA-3 Alphabet subtest. Both cohorts had posttest mean scores ranked higher than the national average.

Likewise the kindergarten children’s growth parallels the first-grade children’s growth in the abilities measured by the Conventions of Print subtest. However, the first-grade scores show much higher initial status and higher mean posttest ranking than the kindergarten scores.
Figure 2 shows the changes in relative ranking for each TERA-3 subtest mean score across the cohorts. Clearly, the largest increases in ranking relative to the national norming population are the TERA-3 Alphabet subtest improvements of the preschool and kindergarten cohorts. However, with the exception of the Alphabet subtest, the most consistent improvements across the measured emergent literacy content occurred in the first-grade cohort. The most uneven improvement occurred in the kindergarten cohort, and the preschool cohort shows fairly flat improvement in content areas other than alphabet letter knowledge.

The information summarized in Table 3 and in Figures 1-2 supports the suggestion of a hierarchy existing in the teaching and learning of the skills measured by the three TERA-3 subtests when using standardized rather than raw scores. However, the information does not explain the widening of the achievement gap (compared to the
national norming population) that occurs over the kindergarten year in the emergent literacy content measured by the TERA-3 Meaning subtest.

Among several possible explanations for the puzzle of the TERA-3 Meaning subtest scores of the kindergarten cohort, two are easily investigated using extant ELLM core data. One explanation simply suggests the TERA-3 Meaning subtest scores of the ELLM preschool children do not sustain their ranking relative to a national norming population over the summer before the children enter their kindergarten year. The second equally simple explanation suggests the core ELLM preschool children do not attend the core elementary schools in large enough numbers to affect the initial emergent literacy status of the core kindergarten children. ELLM TERA-3 scores from core sites are available for three consecutive years. Table 4 presents the percentile rankings of the pretest and posttest mean scores of the preschool and kindergarten cohorts for the 2001/2002, 2002/2003, and 2004/2004 school years.

Table 4
Percentile Rankings of the TERA-3 Subtest Pretest and Posttest Mean Scores for Three School Years

<table>
<thead>
<tr>
<th>Cohort</th>
<th>School Year</th>
<th>Alphabet Pretest Percentile</th>
<th>Alphabet Posttest Percentile</th>
<th>Conventions of Print Pretest Percentile</th>
<th>Conventions of Print Posttest Percentile</th>
<th>Meaning Pretest Percentile</th>
<th>Meaning Posttest Percentile</th>
</tr>
</thead>
<tbody>
<tr>
<td>Preschool</td>
<td>2001/2002</td>
<td>29</td>
<td>54</td>
<td>21</td>
<td>27</td>
<td>25</td>
<td>32</td>
</tr>
<tr>
<td></td>
<td>2002/2003</td>
<td>38</td>
<td>73</td>
<td>26</td>
<td>34</td>
<td>34</td>
<td>37</td>
</tr>
<tr>
<td></td>
<td>2003/2003</td>
<td>46</td>
<td>67</td>
<td>22</td>
<td>26</td>
<td>27</td>
<td>29</td>
</tr>
<tr>
<td>Kindergarten</td>
<td>2001/2002</td>
<td>22</td>
<td>53</td>
<td>13</td>
<td>27</td>
<td>17</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>2002/2003</td>
<td>37</td>
<td>59</td>
<td>14</td>
<td>29</td>
<td>18</td>
<td>9</td>
</tr>
<tr>
<td></td>
<td>2003/2003</td>
<td>32</td>
<td>58</td>
<td>13</td>
<td>21</td>
<td>14</td>
<td>10</td>
</tr>
</tbody>
</table>

The range of percentile rankings of the TERA-3 mean scores are fairly consistent within the cohorts across the three years. However, the percentile rankings of the TERA-3 mean scores of the kindergarten cohort are much more stable across the years than the preschool score rankings. The differences of the preschool mean posttest rankings and the kindergarten mean pretest rankings the following school year are at least 13 percentiles across the three TERA-3 subtests. For example, the 2001/2002 preschool Alphabet subtest posttest mean score ranked at the 54th percentile. If these children entered core kindergarten classes the following year, they would be part of the 2002/2003 core kindergarten class whose mean Alphabet pretest score ranked at the 37th percentile. This
represents a drop of 17 percentiles in ranking. If most of the ELLM preschool children attended the core ELLM kindergarten classes, the mean scores of preschool children lose, on average, at least 13 percentiles in relative ranking over the summer vacation. This indicates preschool scores are not very stable and that the skills acquired are not sustained without frequent reinforcement. Conversely, the stability of the kindergarten scores across the three school years is surprising.

These same data covering three years of ELLM implementation from the three core cohorts allow a quasi-longitudinal look at scores from preschool through kindergarten to first grade. Table 5 presents percentile rankings of TERA-3 scores of the 2001/2002 preschool, 2002/2003 kindergarten, and 2003/2004 first-grade cohorts. However, it is more likely the kindergarten and first-grade children are the same children than the preschool and kindergarten children are the same because the kindergarten and first-grade classes meet in the same elementary schools.

Table 5
Changes in the Percentile Ranking of the TERA-3 Mean Scores of Preschool, Kindergarten, and First-Grade Children over Consecutive School Years

<table>
<thead>
<tr>
<th>School Year</th>
<th>Cohort</th>
<th>Alphabet</th>
<th>Conventions</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>2003/2004</td>
<td>First Grade</td>
<td>51-46</td>
<td>26-39</td>
<td>6-16</td>
</tr>
</tbody>
</table>

Bold numbers indicate successive years from kindergarten to first grade. Indicates successive years from preschool to kindergarten.

Table 5 shows that if the kindergarten and first-grade children are mostly the same children, across the TERA-3 subtests, the mean scores indicate average loss in relative ranking over the summer from kindergarten to first grade is four percentiles and that the loss is fairly consistent across TERA-3 subtests. However, if the preschool and kindergarten children are mostly the same children, mean scores indicate the average loss in relative ranking over the summer from preschool to kindergarten is 14 percentiles. This indicates skills acquired in kindergarten are more stable than skills acquired in preschool. However, the information in Table 5 could just as easily indicate that large numbers of core ELLM preschool children do not become core ELLM kindergarteners.

It is reasonable to assume the more modest loss in relative ranking of three to eight percentiles over the summer months, and that the ELLM kindergarten classes in the core
do not enroll enough ELLM preschool children from the previous school year to benefit from their school readiness. This implies the TERA-3 Meaning subtest puzzle is mostly a kindergarten problem. Thus, the kindergarten teachers have the dual responsibility to close a gap in emergency literacy achievement that arose long before the children’s kindergarten year and not to allow the gap to widen further in any emergent literacy content area over the kindergarten year.

Moreover, the TERA-3 subtests measure independent constructs that are interrelated. The items on the subtests progressively increase in difficulty, and as they do, the abilities in emergent literacy that are not overtly assessed become increasingly necessary for item understanding and response. ELLM kindergarten teachers, over the three years of extant TERA-3 data, have prepared the children in the emergent literacy area of alphabet letter knowledge, and the ELLM kindergarten children’s scores rank above their national peers at the end of each school year. However, ELLM kindergarten children may not be prepared for the higher-level task of application, that is, of using the recently acquired alphabet letter knowledge to correctly respond to items measuring emergent comprehension and vocabulary. It is important to note that alphabet letter knowledge is only a necessary skill; it is not sufficient to the acquisition of reading. Thus, acquiring emergent literacy skills sooner is precursor to elevating standardized scores in all emergent literacy content areas. Therefore, ELLM developers need to enhance the literacy curriculum in ways that address the constructs measured by the TERA-3 Meaning subtest, which is seen as a measure of emergent comprehension and vocabulary.

A true longitudinal study using the 2002/2003 and 2003/2004 Making a Difference data should provide answers to some of the posited questions. At the very least, a simple longitudinal study should provide a snapshot answer indicating the percentage of ELLM kindergarten children who were in ELLM preschool classes the previous year. In the meantime, it is also reasonable to assume there is a hierarchy associated with skills measured by the three TERA-3 subtests when using standardized scores.

Therefore, another challenge is to implement ELLM in more sites, especially sites in feeder patterns to the high-needs elementary schools in the core sample. It is not enough to find out what percentage of the core kindergarten children were in ELLM preschool classes the previous year, but also to determine where the children were if not in ELLM classes – whether in other preschool classes or not in preschool at all. Providing
kindergarten teachers with high initial status kindergarteners allows teachers to spend more time helping children acquire all of the emergent literacy skills expected of all kindergarten children rather than just trying to close the gap in children’s skills compared to their more advantaged peers.

**SPECIFIC CROSS-COHORT RESULTS**

Demographics of ELLM Children

- There were fewer children in the preschool cohort than in either the kindergarten or first-grade cohorts.
- When combined across cohorts, 49% of the ELLM children were boys.
- The preschool cohort was the only racially diverse cohort.
- The percentage of children who were possibly retained in grade was similar across kindergarten and first-grade cohorts.

Emergent Literacy Achievement

- Children in the preschool and kindergarten cohorts made significant and meaningful improvement attributable to their participation in ELLM in the content measured by the TERA-3 Alphabet subtest.
- Children in the kindergarten and first-grade cohorts made significant and meaningful improvement attributable to their participation in ELLM in the content measured by the TERA-3 Conventions of Print subtest.
- Only the children in the first-grade cohort made significant and meaningful improvement attributable to their participation in ELLM in the content measured by the TERA-3 Meaning subtest.
- The children’s improvement in emergent literacy skills indicates a possible teaching and learning hierarchy (ranging from the content of the Alphabet to the Conventions of Print to the Meaning subtests) exists for children living in poverty situations that make them at risk for future academic failure.
Explaining Achievement Differences across the Cohorts

- The use of standardized scores creates the hierarchy in the teaching and learning of emergent literacy skills for children in classrooms where most children live in poverty situations that make them at risk for future academic failure.

- The core ELLM preschool sample does not sufficiently feed the core ELLM kindergarten classes to significantly elevate the mean initial emergent literacy status of the core ELLM kindergarten children.

- Core ELLM kindergarten teachers close the gap in emergent literacy skills in the content area of alphabet letter knowledge. However, they may not be able, at the same time, to close similar gaps that have accumulated over years in other emergent literacy abilities.

Future ELLM Research

- Conduct longitudinal studies using extant data to investigate core ELLM preschool, kindergarten, and first-grade cohort progression.

- Longitudinally follow ELLM preschool children to kindergarten to investigate the stability of emergent literacy skills acquired during preschool.

- Enhance the ELLM literacy curriculum to address emergent comprehension and vocabulary development.

- Expand ELLM to include more preschool sites that feed high-needs public elementary schools in Duval County.