

The Early Literacy and Learning Model (ELLM) Initiative

*Making a Difference
1999/2000 and 2000/2001*

Center for Early Literacy and Learning
Florida Institute of Education at the University of North Florida
UNF College of Education and Human Services

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

The *Early Literacy and Learning Model (ELLM)* Initiative
Making a Difference

Contents

SECTION I: EXECUTIVE SUMMARY

Introduction	1
<i>ELLM</i> at a Glance.....	1
Overview of the Evaluation of the <i>Early Learning and Literacy Model (ELLM)</i>	2
1999/2000	3
<i>TERA-2</i> Results	3
Alphabet Letter Recognition Inventory Results.....	3
2000/2001	4
<i>TERA-2</i> Results	4
Phonemic Awareness Instruction Population.....	4
Validation Study.....	5
Alphabet Letter Recognition Inventory Results.....	5
Technology: <i>The Waterford Early Reading Program (WERP)</i>	6
Evaluation Research Questions.....	6
Conclusions	6
Comparison Study of the 1999/2000 and 2000/2001 School Year Data	7
Evaluation Research Question	7
Conclusion.....	7

Overall Summary and Conclusion	8
Appendix	
Table 1—Summary Statistics: <i>TERA-2</i> and Alphabet Letter Recognition Inventory.....	A1
Figure 1—Percentile Rankings of <i>ELLM</i> Cohorts in 1999/2000 and 2000/2001 School Years	A2

SECTION II: TECHNICAL REPORT

Overview of the Evaluation of the <i>ELLM</i> Initiative.....	1
1999/2000	3
<i>TERA-2</i> Results.....	3
Alphabet Letter Recognition Inventory Results	8
2000/2001	12
<i>TERA-2</i> Results	12
Phonemic Awareness Instruction Population.....	17
Validation Study.....	23
Alphabet Letter Recognition Inventory Results.....	29
Technology: <i>The Waterford Early Reading Program (WERP)</i>	33
Evaluation.....	33
Conclusions	39
Comparison Study of the 1999/2000 and 2000/2001 School Year Data	40
Evaluation.....	40
Conclusion.....	44
Overall Summary and Conclusions.....	45

The Early Literacy and Learning Model (ELLM) Initiative

Section I: Executive Summary

Center for Early Literacy and Learning
Florida Institute of Education at the University of North Florida
UNF College of Education and Human Services

SECTION I: EXECUTIVE SUMMARY

Introduction

Improving the achievement of all children, particularly those considered to be at high risk of academic failure, continues to be a local, state, and national priority. We know that 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 different areas. However, 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 unprepared for the challenges they will face and are at-risk of school failure.

The *Early Literacy and Learning Model (ELLM)* initiative is designed to improve the language and literacy skills of 3-, 4-, 5-, and 6-year-old 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 1999/2000 and 2000/2001 school years.

ELLM at a Glance

- *ELLM* children were organized into three cohorts: pre-kindergarten (representing 4-year-old children), kindergarten, and first grade.
- The *Test of Early Reading Ability (TERA-2)* and the Alphabet Letter Recognition Inventory were administered in a pretest/posttest design in order to measure children’s reading readiness achievement and to assess the impact of the *ELLM* initiative.
- Children whose *TERA-2* pretest scores ranked in the lowest quartile were targeted for specific phonemic awareness instruction.

The *ELLM* program has been used with urban 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 in Duval and Orange counties. During the 2001/2002 school year, the *ELLM* program will be used in approximately 200 classrooms serving about 3,000 at-risk children in northeast Florida.

Overview of the Evaluation of the *Early Literacy and Learning Model (ELLM)*

During the 1999/2000 and 2000/2001 school years, the approximately 1,100 children enrolled in *ELLM* classrooms were organized into three cohorts: pre-kindergarten (representing 4-year-old children), kindergarten, and first grade. Several instruments were used to measure children's achievement and to assess the impact of the *ELLM* initiative. In 1999/2000, the *Test of Early Reading Ability (TERA-2)*, a standardized, norm-referenced test, and the Alphabet Letter Recognition Inventory were administered to the *ELLM* children in a pretest/posttest design. The pretests were given early in the fall and the posttests were given the following spring. The *TERA-2* was administered to all *ELLM* children, and the Alphabet Letter Recognition Inventory was administered to all *ELLM* pre-kindergarten and kindergarten children.

During the 2000/2001 school year, random samples of *ELLM* children were also given the *Expressive One-Word Picture Vocabulary Test*, the *Woodcock Reading Mastery Tests: Visual-Auditory Learning*, and the *TERA-3*, a revised version of the *TERA-2*. The *Expressive One-Word Picture Vocabulary Test* and *Woodcock Reading Mastery Tests: Visual-Auditory Learning* were administered in a pretest/posttest design. These additional tests were administered in order to validate the continued use of *TERA-2* and *TERA-3* to evaluate the *ELLM* initiative.

Researchers discovered over the history of the *ELLM* initiative that children with low scores on the *TERA-2* pretest showed little, if any, improvement on the *TERA-2* posttest. Therefore, beginning in 1999/2000 and continuing in 2000/2001, children with low *TERA-2* pretest scores were selected for targeted instruction in phonemic awareness. This additional instruction led to an improvement in reading readiness scores for this population of children.

The *ELLM* initiative also included the *Waterford Early Reading Program*, a research-based early intervention reading curriculum that uses computer technology, in nearly half of the *ELLM* classrooms. The program focuses on phonemic awareness and other issues related to early literacy including knowledge of the letters of the alphabet, oral language, and print awareness. Classrooms were selected to receive the *Waterford Early Reading Program* based on grant criteria, donor identification, eligibility, and the extent to which at-risk children are served.

The combined 1999/2000 and 2000/2001 school year data provided the opportunity to answer the research question: Was the *ELLM* initiative equally effective in both school years?

1999/2000

***TERA-2* Results**

The 1999/2000 school year data included 488 children with both *TERA-2* pretest and posttest scores: 158 pre-kindergarten children, 256 kindergarten children, and 74 first grade children. The reading readiness score of the typical *ELLM* child, when compared to the national norms, ranked in the lower range of “average” ability (near the 37th percentile) on the *TERA-2* pretest and in the upper range of the “average” ability (near the 63rd percentile) on the *TERA-2* posttest. Even though the score of the typical *ELLM* child ranked as “average” in ability on both the *TERA-2* pretest and posttest, statistical analysis of the means indicated *ELLM* children achieved significant gains in reading readiness. (See Appendix -Table 1 for summary statistics and Appendix - Figure 1 for percentiles for each cohort.)

TERA-2 results were further sorted into seven major categories of the national norms: three categories representing the lowest 25 percentiles or “poor” rating, one category representing the middle 50 percentiles or “average” rating, and three categories representing the highest 25 percentiles or “superior” rating. Inspection of the pretest and posttest scores by these categories indicated *ELLM* children across the entire ability continuum improved in reading readiness.

Alphabet Letter Recognition Inventory Results

The 1999/2000 school year data included pretest and posttest Alphabet Letter Recognition Inventory scores for 326 pre-kindergarten and kindergarten children. *ELLM* children in both cohorts achieved significant gains in alphabet letter recognition on the posttest. In 1999/2000, the typical *ELLM* child recognized 85% of the upper and lower case letters at posttest. (See Appendix - Table 1 for summary statistics.)

Children demonstrating a competency level of at least 90% letter recognition (47 letters) were considered proficient. Fifty-nine percent of the *ELLM* pre-kindergarten children and 81% of the *ELLM* kindergarten children demonstrated this level of mastery at posttest.

2000/2001

***TERA-2* Results**

The 2000/2001 school year data included 603 children with both *TERA-2* pretest and posttest scores: 257 pre-kindergarten children, 214 kindergarten children, and 132 first grade children. The reading readiness score of the typical *ELLM* child, when compared to the national norms, ranked in the lower range of “average” ability (near the 31st percentile) on the *TERA-2* pretest and in the middle range of “average” ability (near the 50th percentile) on the *TERA-2* posttest. Even though the score of the typical *ELLM* child ranked as “average” in reading readiness ability on both the *TERA-2* pretest and posttest, statistical analysis of the means indicated these *ELLM* children achieved significant gains in reading readiness. (See Appendix -Table 1 for summary statistics and Appendix - Figure 1 for percentiles for each cohort.)

TERA-2 results were further sorted into seven major categories of the national norms: three categories representing the lowest 25 percentiles or “poor” rating, one category representing the middle 50 percentiles or “average” rating, and three categories representing the highest 25 percentiles or “superior” rating. Inspection of the pretest and posttest scores by these categories indicated children across the entire ability continuum improved in reading readiness.

Phonemic Awareness Instruction Population

In the 2000/2001 school year data, there were 103 *ELLM* children for whom both pretest and posttest *TERA-2* scores were available and who were selected for targeted phonemic awareness instruction: 50 from the pre-kindergarten cohort, 38 from the kindergarten cohort, and 15 from the first grade cohort. The typical *ELLM* child in this subset, when compared to the national norms, had a pretest *TERA-2* score ranked as “poor” in reading readiness ability (2nd to 7th percentile) and a posttest score ranked in the lower range of the “average” ability (near the 25th percentile). *ELLM* children in all three cohorts achieved significant gains in reading readiness. Results of further analyses, however, indicated targeted phonemic awareness instruction was more effective (as measured by the *TERA-2* posttest scores) with pre-kindergarten children who had low *TERA-2* pretest scores than with first grade children with similar pretest scores.

Validation Study

In the 2000/2001 school year, random samples of *ELLM* children were selected for a validation study of the use of *TERA* for evaluation of the *ELLM* initiative. Two additional tests, the *Expressive One-Word Picture Vocabulary Test* and the *Woodcock Reading Mastery Tests: Visual-Auditory Learning* were administered to random samples of *ELLM* children in a pretest/posttest design. Additionally, *TERA-2* was revised as *TERA-3*. A correlation study of these three tests with *TERA-2* and of the four tests with the Alphabet Letter Recognition Inventory posttest was conducted using data from the random samples of the *ELLM* population.

All three tests, *TERA-3*, *Expressive One-Word Picture Vocabulary Test*, and *Woodcock Reading Mastery Tests: Visual-Auditory Learning* were positively correlated with the *TERA-2*; therefore, the abilities measured by the tests overlap to some degree. Not surprisingly, the strongest positive correlation was between the *TERA-2* and the *TERA-3*. Because alphabet letter recognition is a basic tenet of reading readiness and of the *ELLM* initiative, it is also not surprising that among the four tests, the *TERA-2* and *TERA-3* had the first and second strongest positive correlations, respectively, with the Alphabet Letter Recognition Inventory posttest scores. Based on these correlations, either *TERA-2* or *TERA-3* is considered to be the best instrument to evaluate the *ELLM* initiative.

Additionally, *ELLM* children with pretest and posttest scores on the *Woodcock Reading Mastery Tests: Visual-Auditory Learning* and the *Expressive One-Word Picture Vocabulary Test* achieved significant gains in their ability to form associations between visual stimuli and oral responses and their English speaking vocabulary, respectively. These gains further confirm the overlapping relationship between reading readiness and ability to form associations between visual stimuli and oral responses and between reading readiness and English speaking vocabulary.

Alphabet Letter Recognition Inventory Results

The 2000/2001 school year data included pretest and posttest Alphabet Letter Recognition Inventory scores for 618 pre-kindergarten and kindergarten children: 326 children in the pre-kindergarten cohort and 292 children in the kindergarten cohort. Children in both cohorts achieved significant gains in alphabet letter recognition. In 2000/2001, the typical *ELLM* child recognized 81% of the upper and lower case letters at posttest. (See Appendix - Table 1 for summary statistics.) Children demonstrating a competency level of at least 90% letter recognition (47 letters) were considered proficient. Forty-seven percent of the *ELLM* pre-kindergarten children and 77% of the *ELLM* kindergarten children demonstrated this level of proficiency at posttest.

Technology: *The Waterford Early Reading Program (WERP)*

The evaluation of *WERP* was based on *TERA-2* posttest scores relative to participation in *WERP* and the amount of time children used the program. The developers of the program recommend children spend 15 minutes daily on a personalized program. Therefore, based on a 180-day school year, children should spend 2,700 total minutes using *WERP* over the school year. However, the typical *ELLM* child used the program 1,445 minutes – far less than the amount recommended by the program’s developers.

Evaluation Research Questions

Evaluation of the use of *WERP* involved answering three questions: 1) Was the mean *TERA-2* posttest score of *ELLM* children in classrooms that used *WERP* higher than that of *ELLM* children in classrooms that did not use *WERP*? 2) Among *ELLM* children who used *WERP* was the *TERA-2* posttest score associated with the amount of time children used the program? 3) Among *ELLM* children in classrooms that used *WERP* was the access to and use of *WERP* equitable?

Conclusions

1. After accounting for initial differences in *TERA-2* pretest scores, the mean *TERA-2* posttest score was higher for *ELLM* children in classrooms that used *WERP* than in classrooms that did not use *WERP*. This result may be misleading, however, because *WERP* was not randomly assigned to classrooms. Additionally, the amount of time children used *WERP* was not associated with the mean *TERA-2* posttest score.
2. Children whose *TERA-2* pretest scores were among the lowest *TERA-2* pretest scores received targeted phonemic awareness instruction; however, these children also had comparatively high use of *WERP*, as measured in the total time *WERP* was used over the school year. Yet, among this subset of *ELLM* children, there was no difference in the mean *TERA-2* posttest score of the children who used *WERP* and those who did not use *WERP*.

Typically, *ELLM* children in classrooms that used *WERP* did not use the program 15 minutes daily as suggested by the program’s developers. In fact, the amount of time children used the program was associated with their *TERA-2* pretest scores – children with low *TERA-2* pretest scores used the program more. Moreover, there was no difference in the amount of time boys and girls used *WERP*. Therefore, results of data analysis indicated access to and use of *WERP* was equitable.

Comparison Study of the 1999/2000 and 2000/2001 School Year Data

Evaluation Research Question

The combined 1999/2000 and 2000/2001 school year data provided an opportunity to compare the effectiveness of the *ELLM* initiative over two school years. The data include *TERA-2* scores from three distinct groups of *ELLM* children: those who were only in *ELLM* classrooms during the 1999/2000 school year, those who were only in *ELLM* classrooms during the 2000/2001 school year, and those who were in *ELLM* classrooms during both the 1999/2000 and the 2000/2001 school years. *TERA-2* scores from children who were in *ELLM* classrooms for the two school years were excluded from the study; therefore, only those children who were in *ELLM* classrooms during the 1999/2000 or during the 2000/2001 school years were included. The comparison study answered the research question: Was the *ELLM* initiative equally effective in both school years?

Conclusion

The *ELLM* initiative was effective with children from all three cohorts in both the 1999/2000 and 2000/2001 school years. A comparison of the effectiveness of the *ELLM* initiative over the 1999/2000 and 2000/2001 school years depended on both the cohort of children and the level of their *TERA-2* pretest scores. In general, the *ELLM* initiative was more effective during the 1999/2000 school year with children who had low *TERA-2* pretest scores. Moreover, the mean *TERA-2* posttest score of the 1999/2000 pre-kindergarten cohort was higher than all three 2000/2001 cohorts and was higher than the 1999/2000 kindergarten cohort. The *ELLM* initiative was equally effective with all three cohorts during the 2000/2001 school year.

Overall Summary and Conclusions

Analyses of the 1999/2000 and 2000/2001 school year data from *ELLM* children in all three cohorts indicated the following:

1. In the 1999/2000 school year, *ELLM* children with *TERA-2* scores across the ability continuum had improved reading readiness scores at the end of the school year.
2. In the spring of the 1999/2000 school year, the typical *ELLM* child recognized 85% of the upper and lower case letters of the alphabet.
3. In the 2000/2001 school year, *ELLM* children with *TERA-2* scores across the ability continuum had improved reading readiness scores at the end of the school year.
4. In the 2000/2001 school year, *ELLM* children selected for targeted phonemic awareness instruction had improved reading readiness scores at the end of the school year. However, the targeted phonemic awareness instruction was more effective with pre-kindergarten children than with first grade children.
5. In the spring of the 2000/2001 school year, the typical *ELLM* child recognized 81% of the upper and lower case letters of the alphabet.
6. Access to and use of the *Waterford Early Reading Program (WERP)* was equitable in *ELLM* classrooms that used the program. The mean reading readiness achievement in classrooms that used *WERP* was higher than in classrooms that did not use *WERP*. These results could be misleading, however, because *WERP* was not randomly assigned to classrooms. Furthermore, the use of *WERP* was not effective when used in addition to targeted phonemic awareness instruction.

A comparison of the effectiveness of the *ELLM* initiative over the two school years depended on the *TERA-2* pretest scores and cohort. In general, the *ELLM* initiative was more effective during the 1999/2000 school year for children with low *TERA-2* pretest scores. The *ELLM* initiative was very effective with the pre-kindergarten cohort during the 1999/2000 school year and was equally effective among the three cohorts during the 2000/2001 school year.

Appendix

Table 1
Summary Statistics: *TERA-2* and Alphabet Letter Recognition Inventory

	Cohort	Sample Size	Pretest	Posttest	Probability Value
	<i>TERA-2</i>				
1999/2000	<i>ELLM</i> Children	488	94.83	105.32	<0.0001*
	Pre-Kindergarten	158	97.85	109.41	<0.0001*
	Kindergarten	256	93.37	102.92	<0.0001*
	First Grade	74	93.45	104.89	<0.0001*
2000/2001	<i>ELLM</i> Children	603	92.00	100.44	<0.0001*
	Pre-Kindergarten	257	90.90	99.82	<0.0001*
	Kindergarten	214	90.05	99.53	<0.0001*
	First Grade	132	97.31	103.13	<0.0001*
	Alphabet Letter Recognition Inventory				
1999/2000	<i>ELLM</i> Children	326	28.51	44.86	<0.0001*
	Pre-Kindergarten	123	26.08	39.03	<0.0001*
	Kindergarten	203	29.98	48.39	<0.0001*
2000/2001	<i>ELLM</i> Children	618	25.24	41.86	<0.0001*
	Pre-Kindergarten	326	23.34	37.44	<0.0001*
	Kindergarten	292	27.37	46.78	<0.0001*

* Denotes more than 99% confidence there was a difference in the pretest/posttest means.

National Percentiles of *ELLM* Cohorts in 1999/2000 and 2000/2001

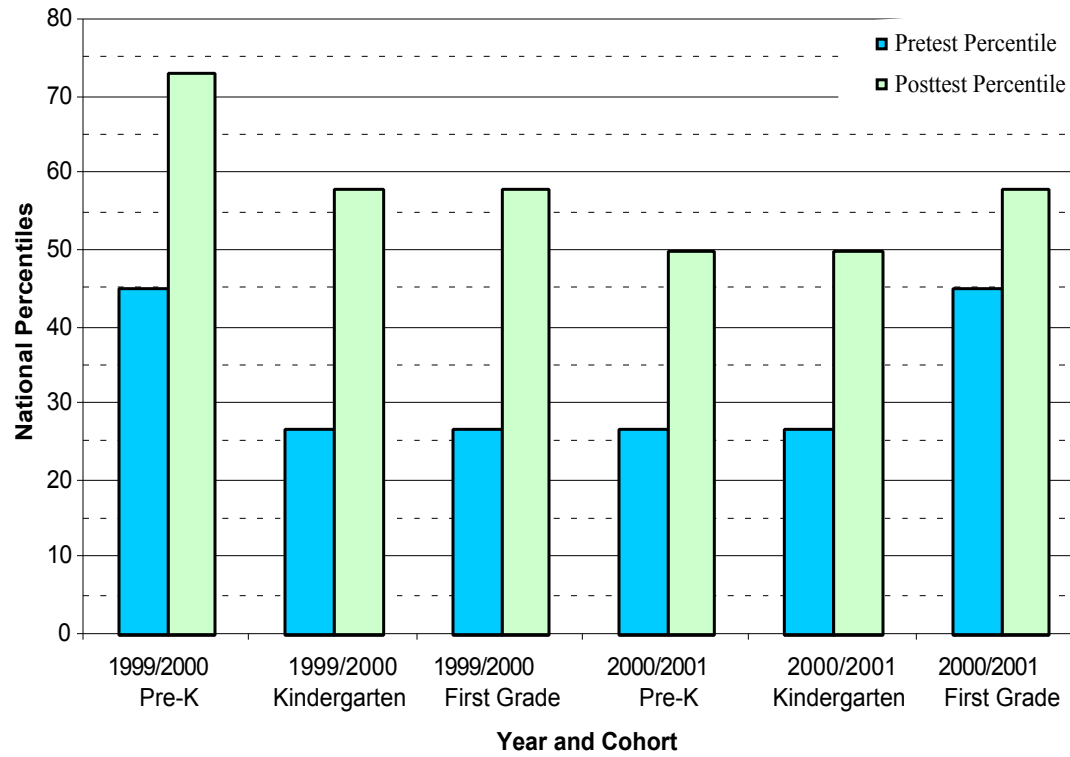


Figure 1

The Early Literacy and Learning Model (ELLM) Initiative

Section II: Technical Report

Center for Early Literacy and Learning
Florida Institute of Education at the University of North Florida
UNF College of Education and Human Services

SECTION II: TECHNICAL REPORT

Overview of the Evaluation of *ELLM* Initiative

The approximately 1,100 children enrolled in *ELLM* classrooms during the 1999/2000 and 2000/2001 school years, were organized into three cohorts: pre-kindergarten (representing 4-year-old children), kindergarten, and first grade. Several instruments were used to measure children's achievement and to assess the impact of the *ELLM* initiative. In 1999/2000, the *Test of Early Reading Ability (TERA-2)*, a standardized, norm-referenced measure of ability to attribute meaning to printed symbols, knowledge of the alphabet and its functions, and understanding of the conventions of print; and the Alphabet Letter Recognition Inventory were administered to the *ELLM* children in a pretest/posttest design. The pretests were given early in the fall and the posttests the following spring. The *TERA-2* was given to all *ELLM* children, and the Alphabet Letter Recognition Inventory was given to all pre-kindergarten and kindergarten children.

In 2000/2001, in addition to the *TERA-2* and the Alphabet Letter Recognition Inventory, random samples of *ELLM* children were administered the *Expressive One-Word Picture Vocabulary Test*, a measure of English speaking vocabulary; the *Woodcock Reading Mastery Tests: Visual-Auditory Learning*, a measure of the ability to form associations between visual stimuli and oral responses; and the *TERA-3*, a revised version of the *TERA-2*. The *Expressive One-Word Picture Vocabulary Test* and *Woodcock Reading Mastery Tests: Visual-Auditory Learning* were also given in a pretest/posttest design. These additional tests were administered in order to validate the continued use of *TERA-2* and *TERA-3* to evaluate the *ELLM* initiative.

The scores from these four tests: *TERA-2*, *TERA-3*, *Expressive One-Word Picture Vocabulary Test*, and *Woodcock Reading Mastery Tests: Visual-Auditory Learning*, all follow a normal distribution with a mean of 100 and a standard deviation of 15. Therefore, all scores can be sorted into the same seven categories across their respective ability continuums. Table 1 displays the seven categories along with the interval of scores and percent of the national population used to establish norms that fall into each of the categories.

Table 1
Categories of the *TERA-2*, *TERA-3*, *Expressive One-Word Picture Vocabulary Test*, and *Woodcock Reading Mastery Tests: Visual-Auditory Learning*

	Categories						
	Very Poor	Poor	Below Average	Average	Above Average	Superior	Very Superior
Score Interval	Below 70	70-79	80-89	90-110	111-120	121-130	Above 130
Percent of Population	2.34	6.87	16.12	49.51	16.12	6.87	2.34

Researchers discovered over the history of the *ELLM* initiative that children with low scores on the *TERA-2* pretest showed little, if any, improvement on the *TERA-2* posttest. Therefore, beginning in 1999/2000 and continuing in 2000/2001, children with low *TERA-2* pretest scores were selected for targeted instruction in phonemic awareness, and improvement in reading readiness scores for this selected population was noted.

The *ELLM* initiative also included the *Waterford Early Reading Program*, a research-based early intervention reading curriculum that uses computer technology, in nearly half of the *ELLM* classrooms. The program focuses on phonemic awareness and other issues related to early literacy including knowledge of the letters of the alphabet, oral language, and print awareness. Classrooms were selected to receive the *Waterford Early Reading Program* based on grant criteria, donor identification, eligibility, and the extent to which at-risk children are served.

The combined 1999/2000 and 2000/2001 school year data provided the opportunity to answer the research question: Was the *ELLM* initiative equally effective in both school years?

1999/2000

***TERA-2* Results**

In the 1999/2000 school year, 488 children from the three cohorts at *ELLM* sites had both pretest and posttest *TERA-2* Reading Quotient scores. Table 2 provides summary statistics for these two measures. The pretest and posttest scores were analyzed as a repeated-measures ANOVA design in order to determine if there was a statistically significant difference in the pretest and posttest means. The *ELLM* children in all three cohorts achieved significant gains in reading readiness during the 1999/2000 school year.

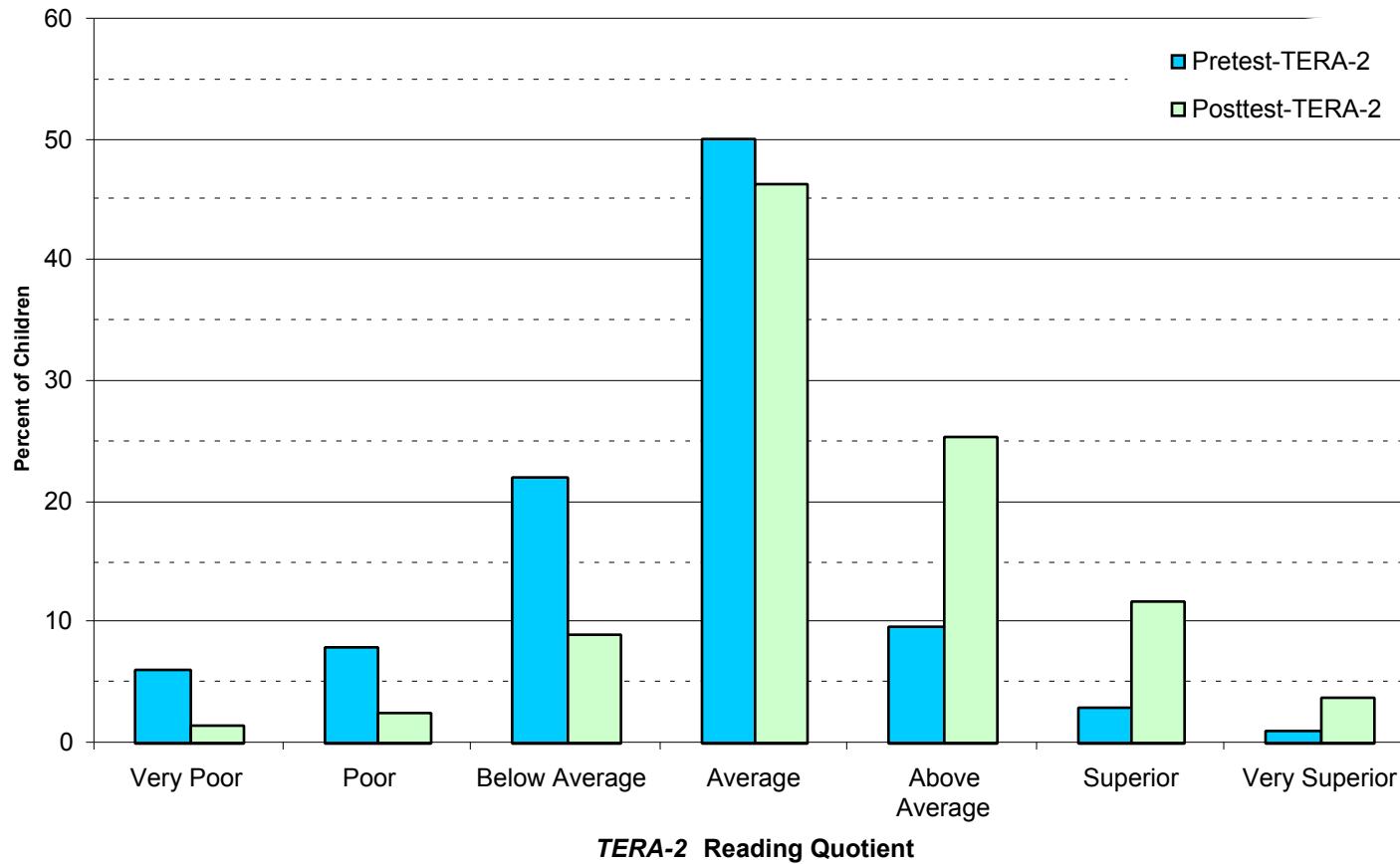
Table 2
Summary Statistics and ANOVA Results: *TERA-2*

	Sample Size	Pretest Mean	Posttest Mean	Probability Value
All <i>ELLM</i> Children	488	94.83	105.32	<0.0001*
Pre-Kindergarten	158	97.85	109.41	<0.0001*
Kindergarten	256	93.37	102.92	<0.0001*
First Grade	74	93.45	104.89	<0.0001*

* Denotes more than 99% confidence there was a difference in the pretest/posttest means.

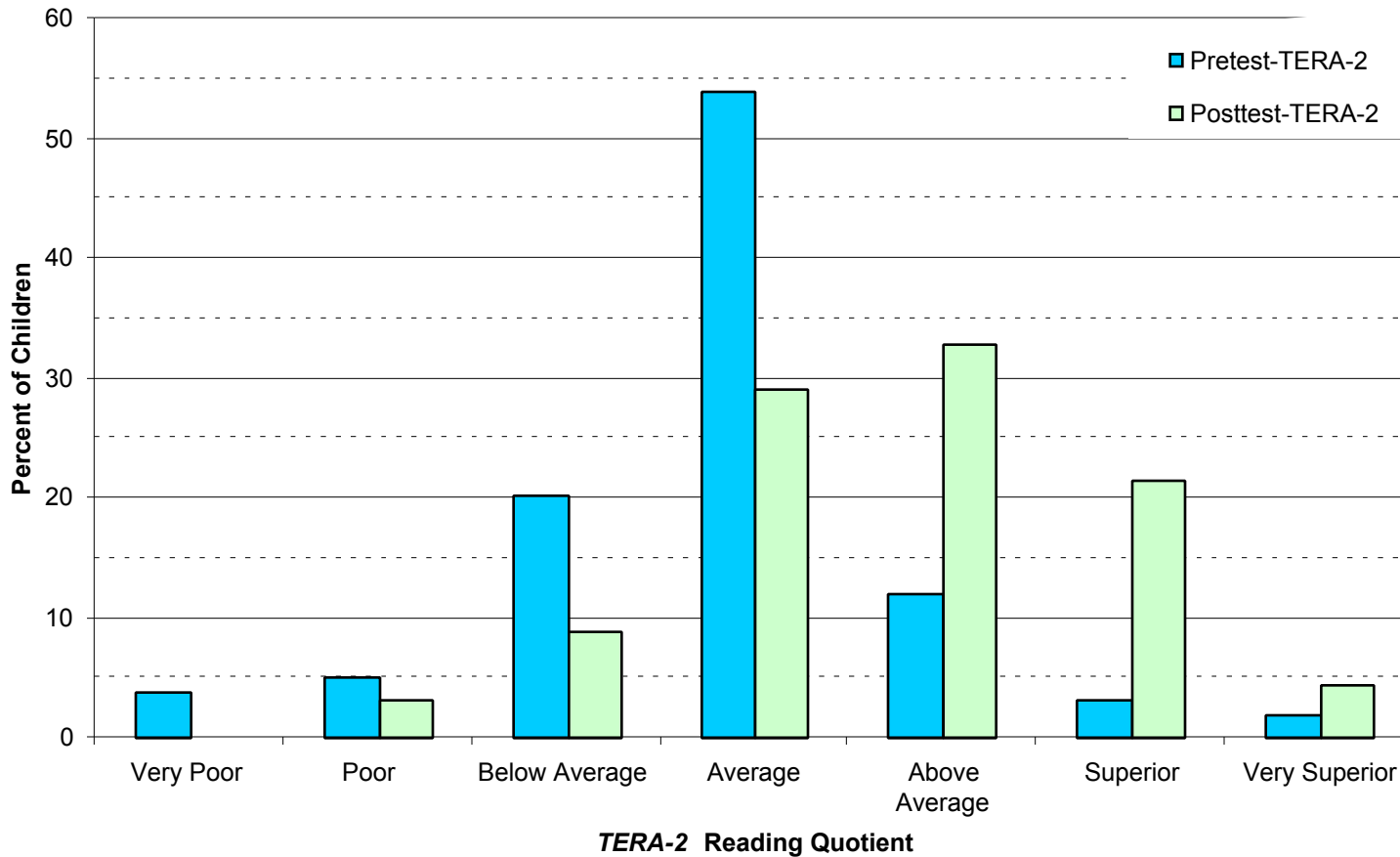
In order to determine if improvement in reading readiness occurred across the ability continuum (measured by *TERA-2*), scores from the *TERA-2* are displayed in seven major categories: three categories representing the lowest 25 percentiles or “poor” rating, one category representing the middle 50 percentiles or “average” rating, and three categories representing the highest 25 percentiles or “superior” rating (see Figures 1-4). Inspection of Figures 1-4 indicates children with pretest scores across the range of abilities had improved reading readiness scores at posttest.

ELLM Pre-Kindergarten, Kindergarten, and First Grade Children 1999/2000



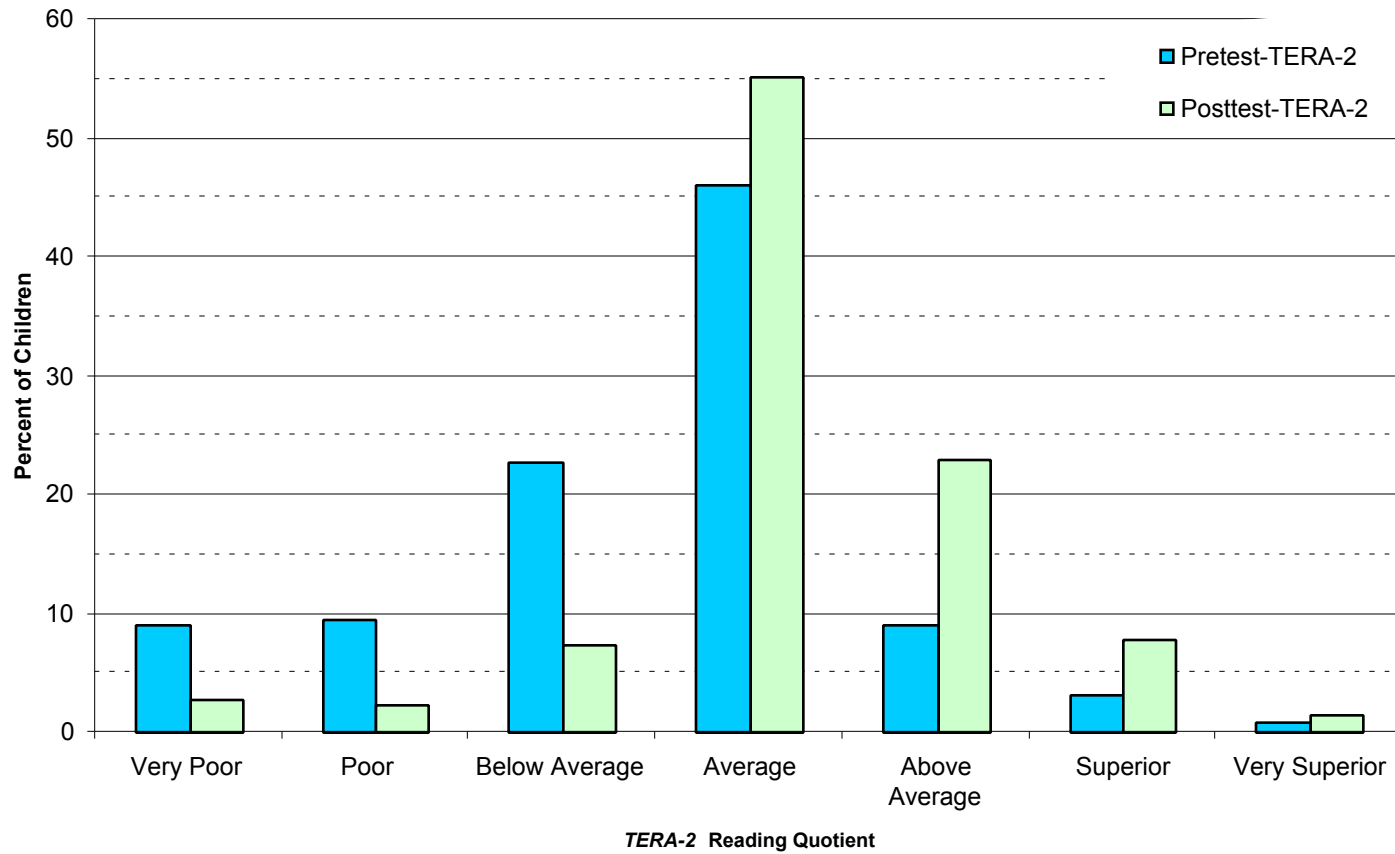
In the 1999/2000 school year data, there were 488 children representing the three cohorts from all ELLM sites for whom both pretest and posttest TERA-2 Reading Quotient scores are available. In the fall of 1999, 36.3% of the TERA-2 pretest scores of the ELLM children were in the lowest quartile (at or below "Below Average"), and 12.9% of the pretest scores were in the highest quartile (at or above "Above Average") of the national norms. In the spring of 2000, 13.5% of the TERA-2 posttest scores of the ELLM children were in the lowest quartile, and 40.8% of the posttest scores were in the highest quartile of the national norms. Therefore, among these 488 scores, 22.8% of the scores in the lowest quartile on the pretest shifted to a higher quartile on the posttest. There was also a 27.9% increase in scores in the highest quartile.

ELLM Pre-Kindergarten Children 1999/2000



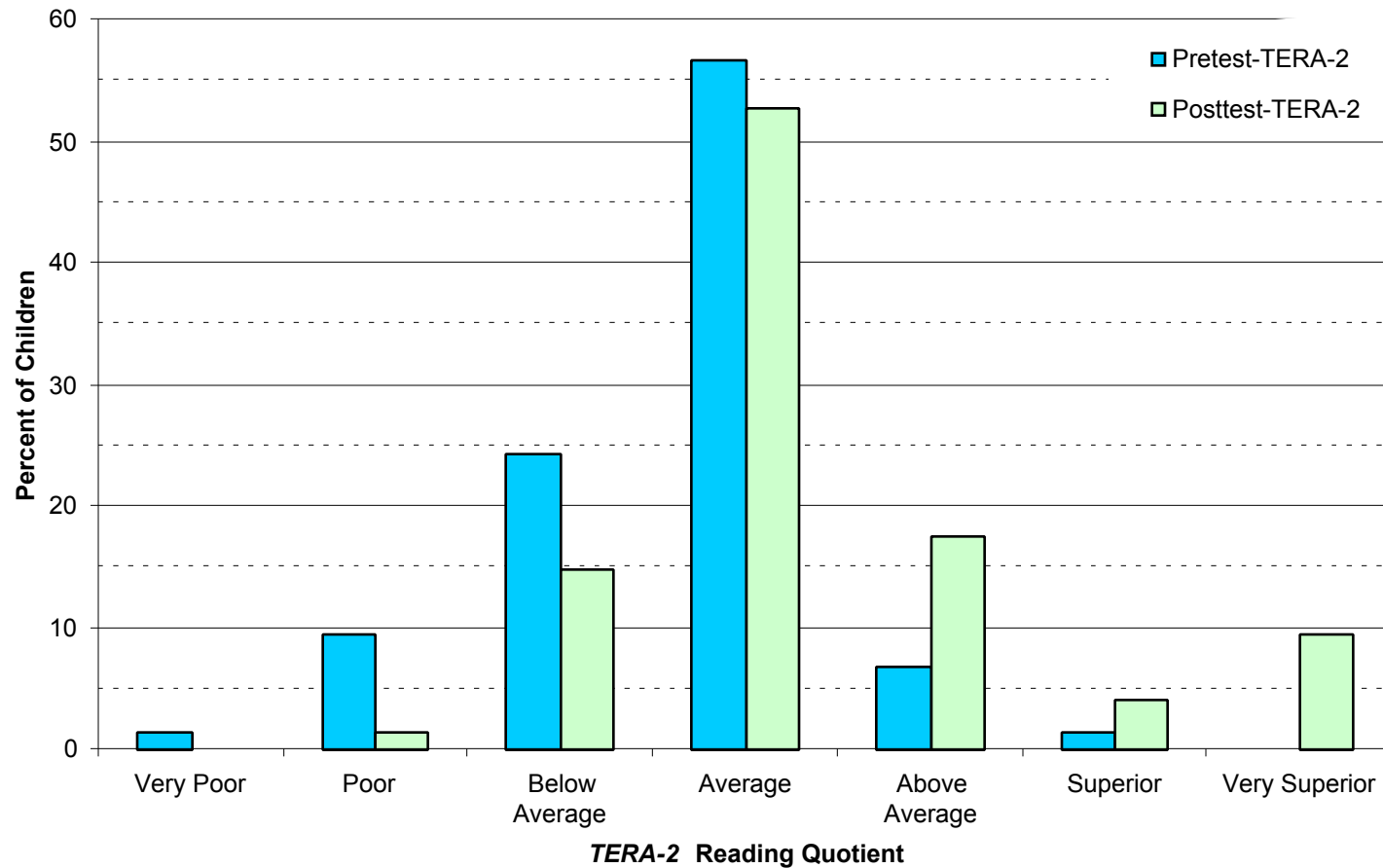
In the 1999/2000 school year data, there were 158 children representing the pre-kindergarten cohorts from all ELLM sites for whom both pretest and posttest TERA-2 Reading Quotient scores are available. In the fall of 1999, 29.1% of the TERA-2 pretest scores of the ELLM pre-kindergarten children were in the lowest quartile (at or below "Below Average"), and 17.1% of the pretest scores were in the highest quartile (at or above "Above Average") of the national norms. In the spring of 2000, 12.0% of the TERA-2 posttest scores of the ELLM pre-kindergarten children were in the lowest quartile, and 58.9% of the posttest scores were in the highest quartile of the national norms. Therefore, among these 158 scores, 17.1% of the scores in the lowest quartile on the pretest shifted to a higher quartile on the posttest. There was also a 41.8% increase in scores in the highest quartile.

ELLM Kindergarten Children 1999/2000



In the 1999/2000 school year data, there were 256 children representing the kindergarten cohorts from all ELLM sites for whom both pretest and posttest TERA-2 Reading Quotient scores are available. In the fall of 1999, 41.0% of the TERA-2 pretest scores of the ELLM kindergarten children were in the lowest quartile (at or below "Below Average"), and 12.9% of the pretest scores were in the highest quartile (at or above "Above Average") of the national norms. In the spring of 2000, 12.5% of the TERA-2 posttest scores of the ELLM kindergarten children were in the lowest quartile, and 32.4% of the posttest scores were in the highest quartile of the national norms. Therefore, among these 256 scores, 28.5% of the scores in the lowest quartile on the pretest shifted to a higher quartile on the posttest. There was a 19.5% increase in scores in the highest quartile.

ELLM First Grade Children 1999/2000



In the 1999/2000 school year data, there were 74 children representing the first grade cohorts from all ELLM sites for whom both pretest and posttest TERA-2 Reading Quotient scores are available. In the fall of 1999, 35.1% of the TERA-2 pretest scores of the ELLM first grade children were in the lowest quartile (at or below "Below Average"), and 8.1% of the pretest scores were in the highest quartile (at or above "Above Average") of the national norms. In the spring of 2000, 16.2% of the TERA-2 posttest scores of the ELLM first grade children were in the lowest quartile, and 31.1% of the posttest scores were in the highest quartile of the national norms. Therefore, among these 74 scores, 18.9% of the scores in the lowest quartile on the pretest shifted to a higher quartile on the posttest. There was also a 23.0% increase in scores in the highest quartile.

Alphabet Letter Recognition Inventory Results

In the 2000/2001 school year, 618 children from the pre-kindergarten and kindergarten cohorts from *ELLM* sites had both pretest and posttest Alphabet Letter Recognition Inventory scores. Table 13 provides summary statistics for these two measures. The pretest and posttest scores were analyzed as a repeated-measures ANOVA design in order to determine if there was a statistically significant difference in the pretest and posttest means. The *ELLM* children in both the pre-kindergarten and kindergarten cohorts achieved significant gains in alphabet recognition during the 2000/2001 school year.

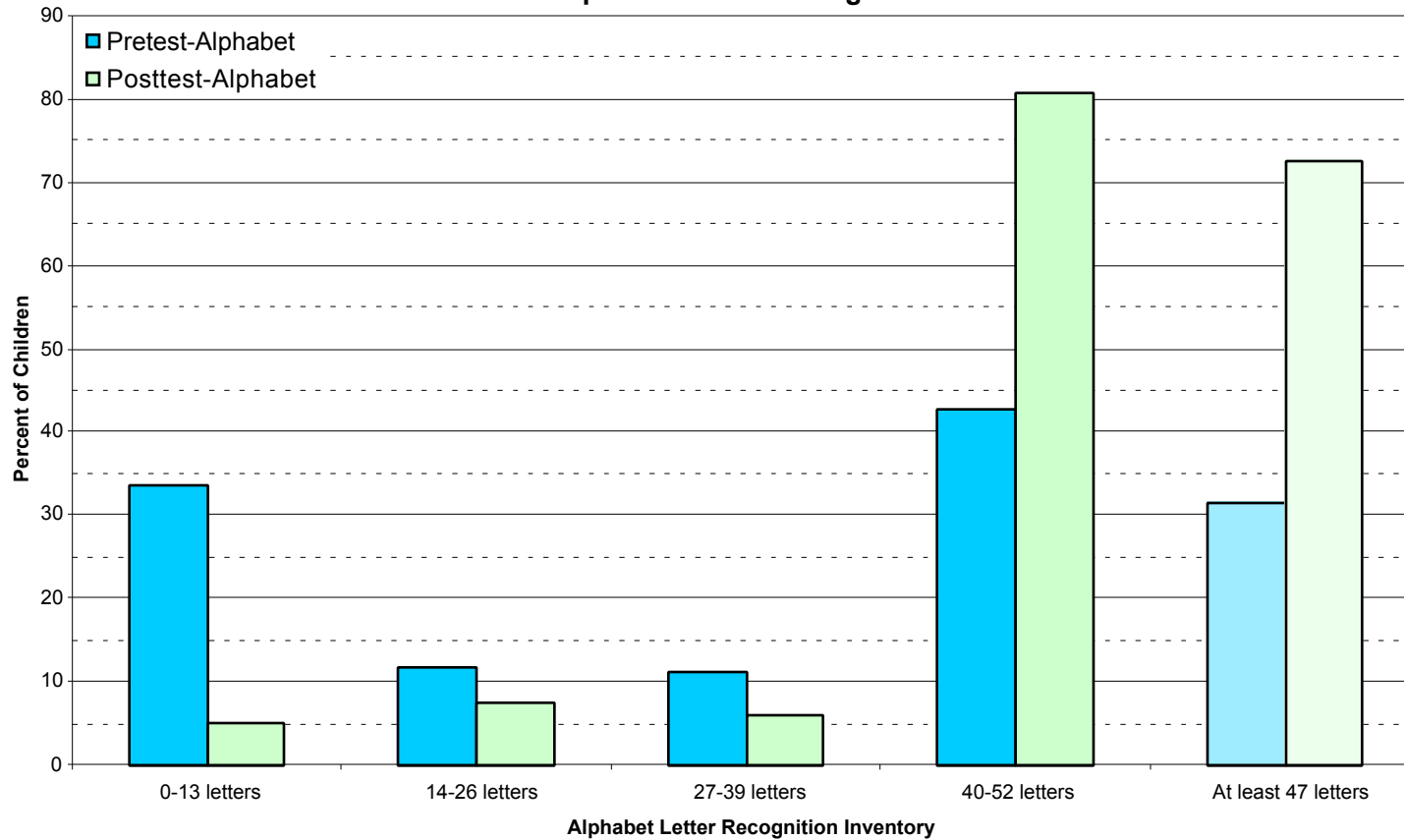
Table 13
Summary Statistics and ANOVA Results: Alphabet Letter Recognition Inventory

	Sample Size	Pretest Mean	Posttest Mean	Probability Value
All <i>ELLM</i> Children	618	25.24	41.86	<0.0001*
Pre-Kindergarten	326	23.34	37.44	<0.0001*
Kindergarten	292	27.37	46.78	<0.0001*

* Denotes more than 99% confidence there was a difference in the pretest/posttest means.

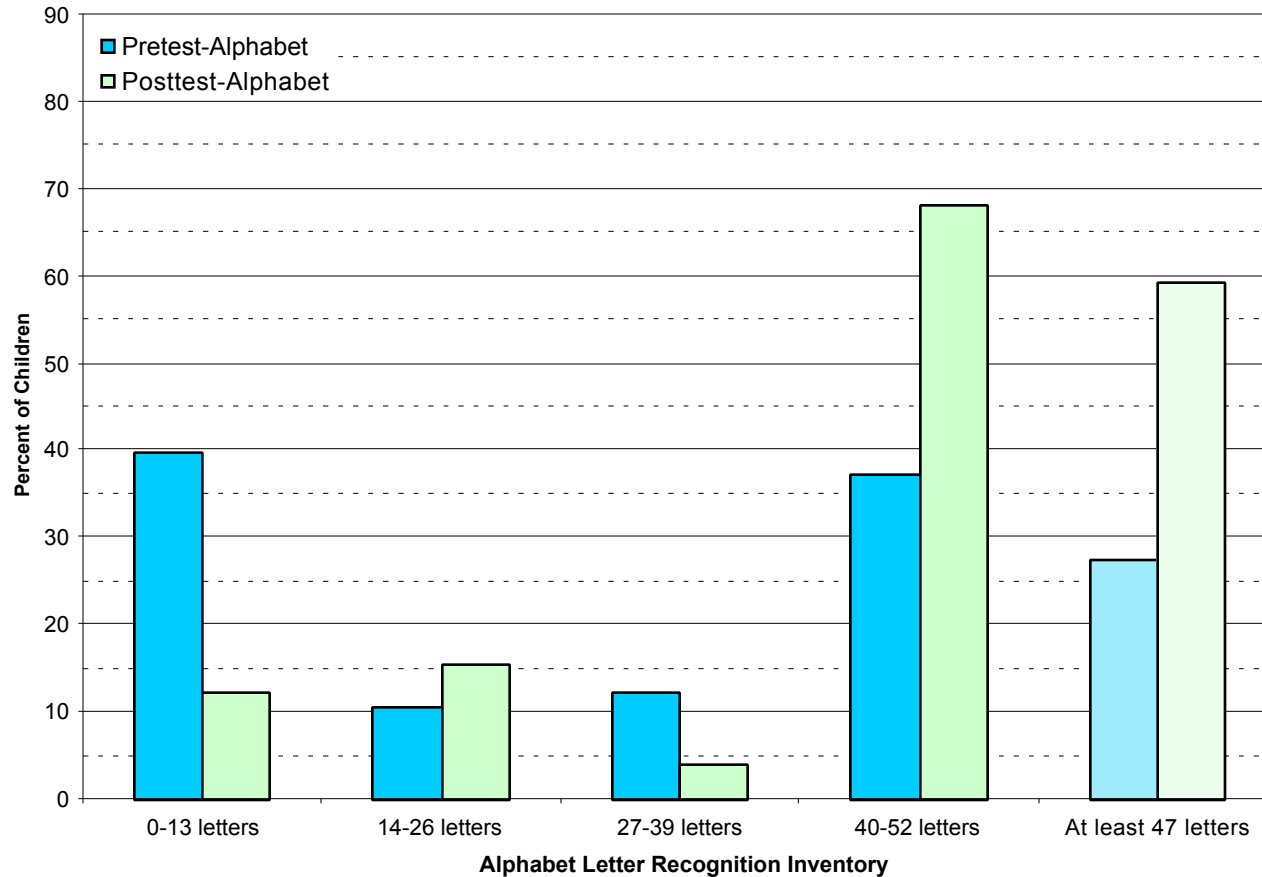
The posttest mean of the Alphabet Letter Recognition Inventory indicated a typical *ELLM* child recognized 81% of the upper and lower case letters. However, in order to determine the range of alphabet letter recognition ability of *ELLM* children, both pre-kindergarten and kindergarten children's scores are displayed using four recognition categories: 0-13 letters, 14-26 letters, 27-39 letters, and 40-52 letters. Because of the importance of this skill, a proficiency category of at least 90% recognition (47 letters) is also displayed (see Figures 18-20). Inspection of Figures 18-20 indicates the largest shift in scores was from the 0-13 letters category to the 40-52 letters category. The two middle categories remained comparatively stable.

**ELLM Pre-Kindergarten and Kindergarten Children 1999/2000
Alphabet Letter Recognition**



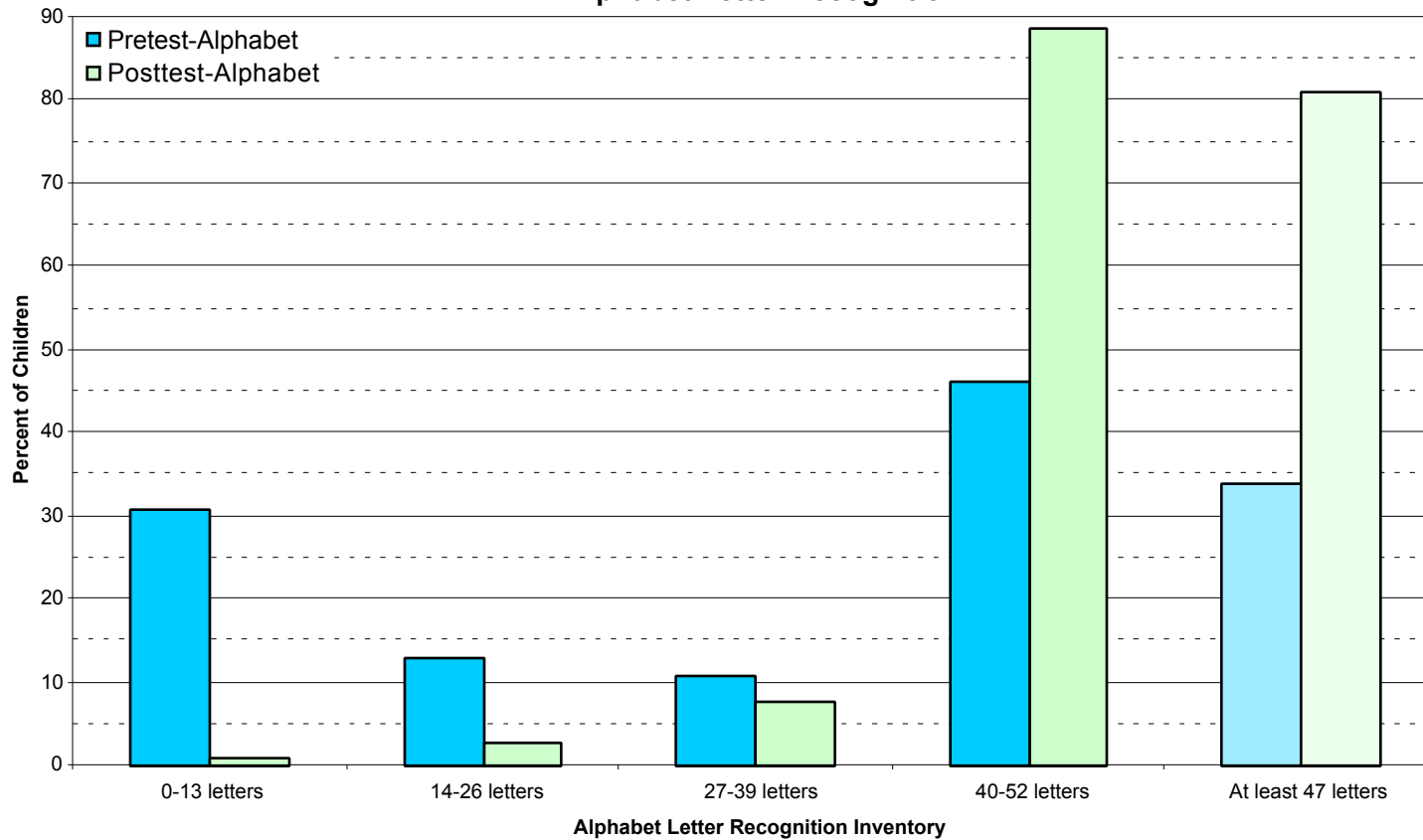
In the 1999/2000 school year data, there were 326 children representing the pre-kindergarten and kindergarten cohorts from all ELLM sites for whom both pretest and posttest Alphabet Letter Recognition Inventory scores are available. Of these scores, 29% in the lowest recognition category on the pretest shifted to a higher category on posttest. There was also an increase of 38% in scores in the highest recognition category. Seventy-three percent (up from 32%) of the combined pre-kindergarten and kindergarten children exhibited at least 90% mastery of the upper and lower case letters at posttest.

**ELLM Pre-Kindergarten Children 1999/2000
Alphabet Letter Recognition**



In the 1999/2000 school year data, there were 123 children representing the pre-kindergarten cohorts from all ELLM sites for whom both pretest and posttest Alphabet Letter Recognition Inventory scores are available. Of these scores, 28% in the lowest recognition category on the pretest shifted to a higher category on the posttest. There was also an increase of 31% in scores in the highest recognition category. Fifty-nine percent (up from 28%) of pre-kindergarten children exhibited at least 90% mastery of the upper and lower case letters at posttest.

**ELLM Kindergarten Children 1999/2000
Alphabet Letter Recognition**



In the 1999/2000 school year data, there were 203 children representing the kindergarten cohorts from all ELLM sites for whom both pretest and posttest Alphabet Letter Recognition Inventory scores are available. Of these scores, 29% in the lowest recognition category on the pretest shifted to a higher category on posttest. There was also an increase of 42% in scores in the highest recognition category. Eighty-one percent (up from 34%) of the kindergarten children exhibited at least 90% mastery of the upper and lower case letters at posttest.

2000/2001

***TERA-2* Results**

In the 2000/2001 school year, 603 children from the cohorts at all *ELLM* sites had both pretest and posttest *TERA-2* Reading Quotient scores. Table 4 provides summary statistics for these two measures. The pretest and posttest scores were analyzed as a repeated-measures ANOVA design in order to determine if there was a statistically significant difference in the pretest and posttest means. The *ELLM* children in all three cohorts achieved significant gains in reading readiness during the 2000/2001 school year.

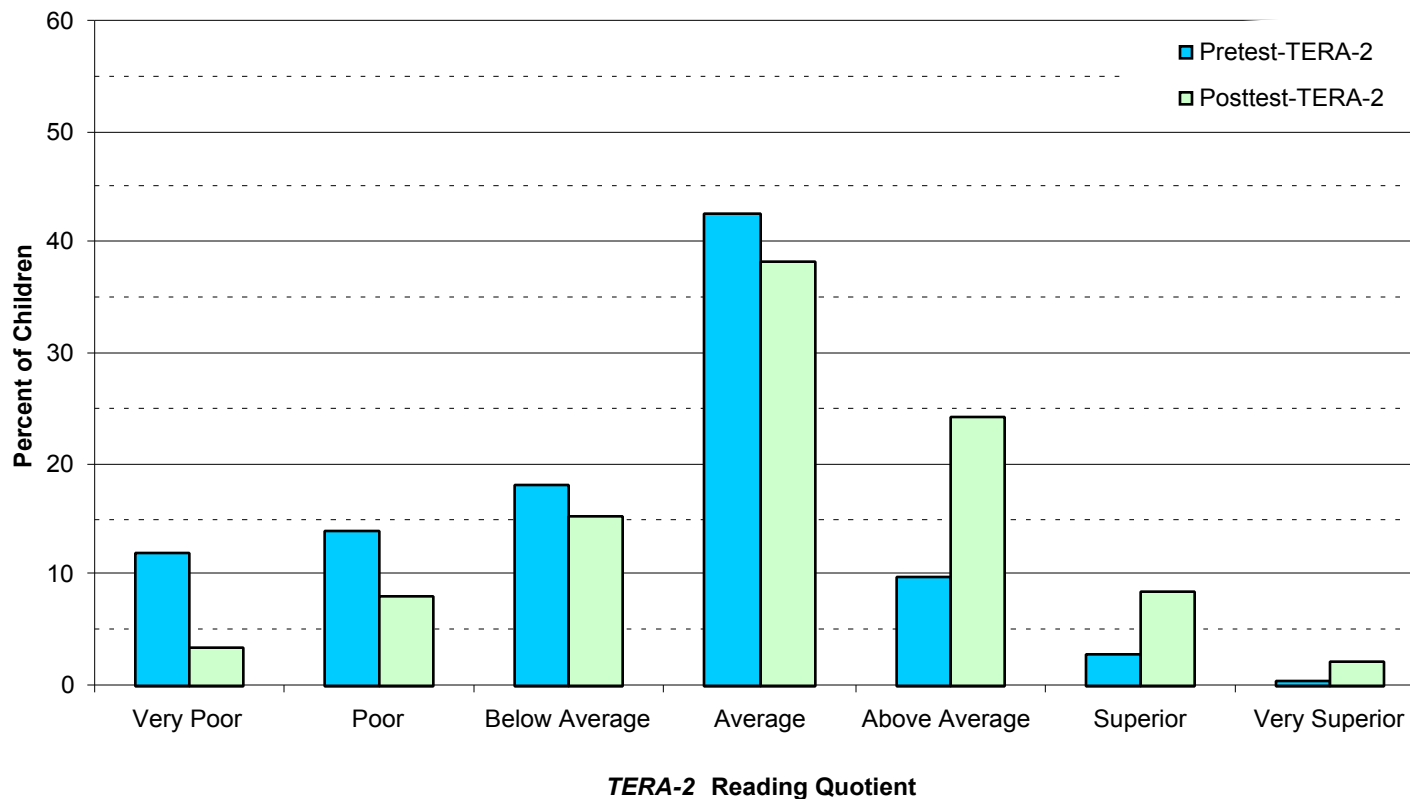
Table 4
Summary Statistics and ANOVA Results: *TERA-2*

	Sample Size	Pretest Mean	Posttest Mean	Probability Value
All <i>ELLM</i> Children	603	92.00	100.44	<0.0001*
Pre-Kindergarten	257	90.90	99.82	<0.0001*
Kindergarten	214	90.05	99.53	<0.0001*
First Grade	132	97.31	103.13	<0.0001*

* Denotes more than 99% confidence there was a difference in the pretest/posttest means.

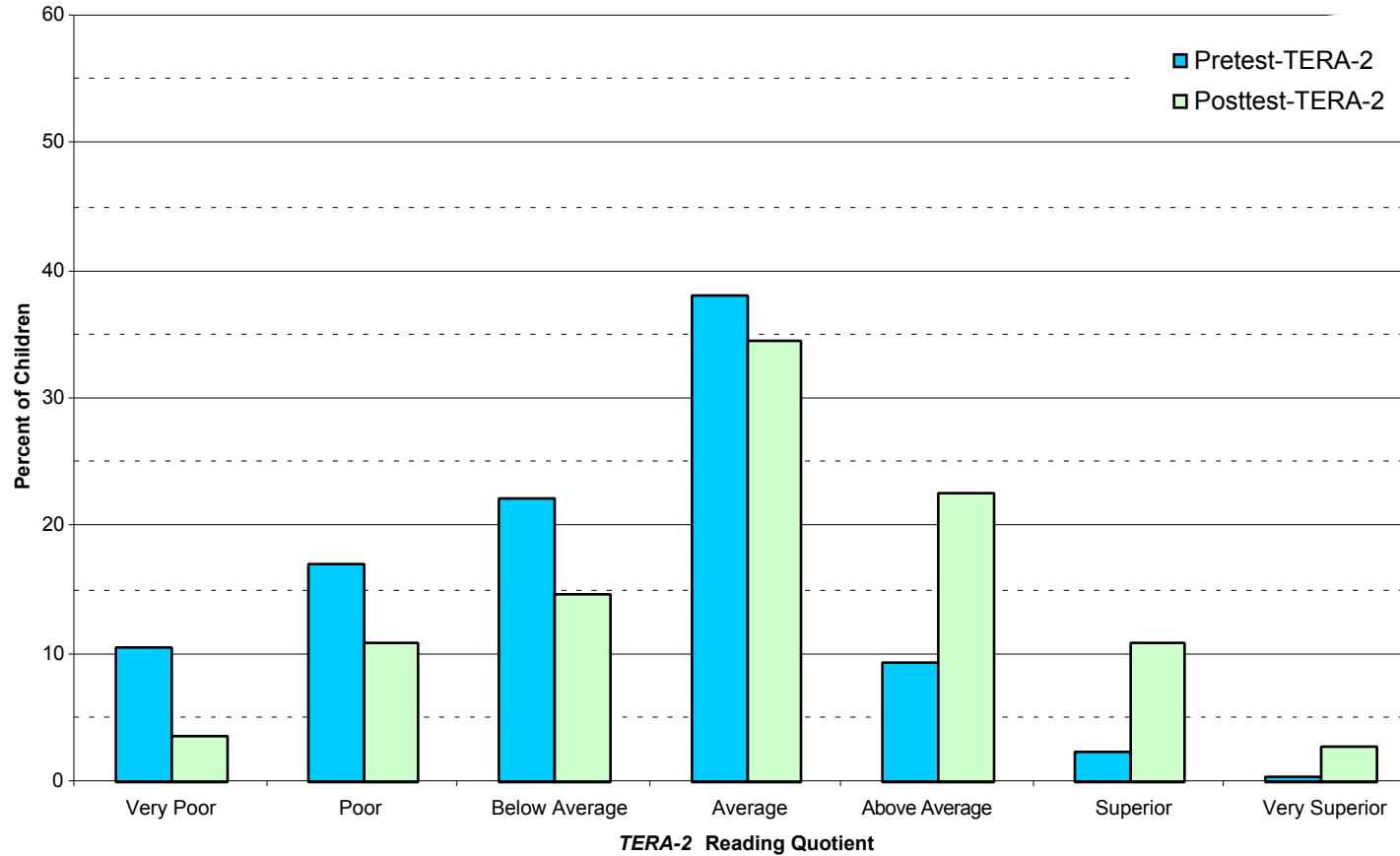
In order to determine if improvement in reading readiness occurred across the ability continuum (measured by *TERA-2*), scores from the *TERA-2* are also displayed in seven major categories: three categories representing the lowest 25 percentiles or “poor” rating, one category representing the middle 50 percentiles or “average” rating, and three categories representing the highest 25 percentiles or “superior” rating (see Figures 8-11). Inspection of Figures 8-11 indicates children with pretest scores across the range of abilities had improved reading readiness scores at posttest.

ELLM Pre-Kindergarten, Kindergarten, and First Grade Children 2000/2001



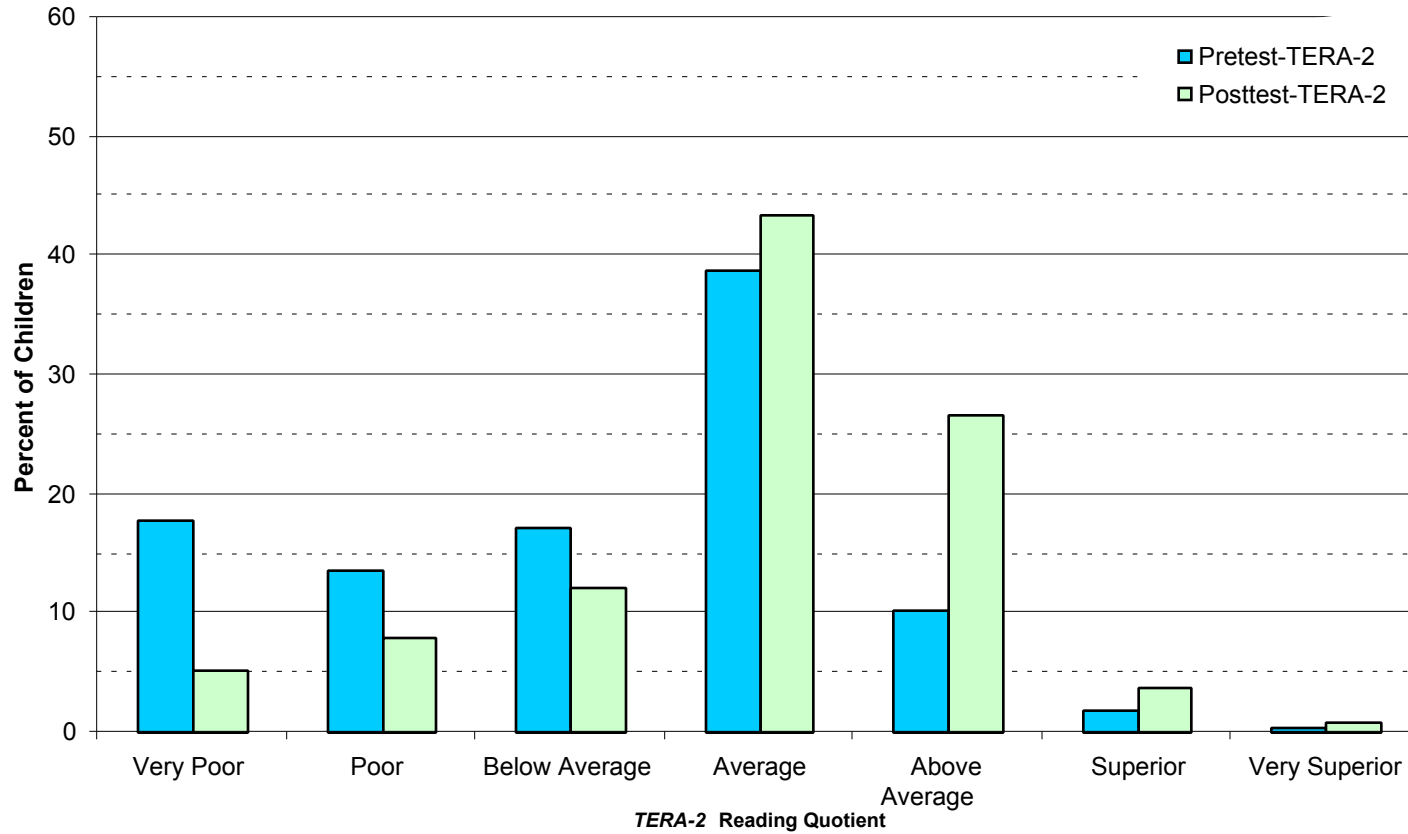
In the 1999/2000 school year data, there were 203 children representing the kindergarten cohorts from all ELLM sites for whom both pretest and posttest Alphabet Letter Recognition Inventory scores are available. Of these scores, 29% in the lowest recognition category on the pretest shifted to a higher category on posttest. There was also an increase of 42% in scores in the highest recognition category. Eighty-one percent (up from 34%) of the kindergarten children exhibited at least 90% mastery of the upper and lower case letters at posttest.

ELLM Pre-Kindergarten Children 2000/2001



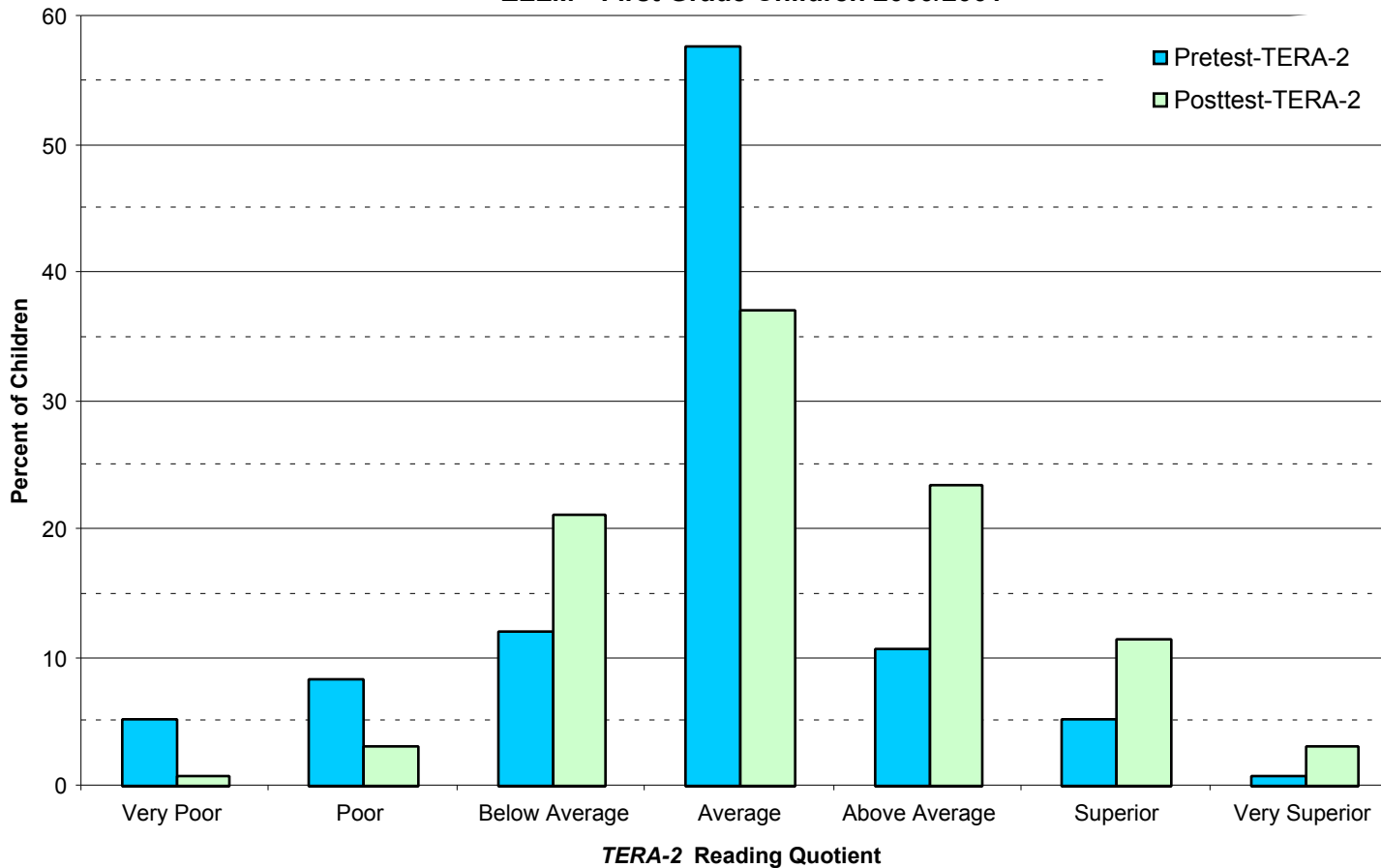
In the 2000/2001 school year data, there were 257 children representing the pre-kindergarten cohorts from all ELLM sites for whom both pretest and posttest TERA-2 Reading Quotient scores are available. In the fall of 2000, 49.8% of the TERA-2 pretest scores of the ELLM pre-kindergarten children were in the lowest quartile (at or below "Below Average"), and 12.1% of the pretest scores were in the highest quartile (at or above "Above Average") of the national norms. In the spring of 2001, 29.2% of the TERA-2 posttest scores of the ELLM pre-kindergarten children were in the lowest quartile, and 36.2% of the posttest scores were in the highest quartile of the national norms. Therefore, among these 257 scores, 20.6% of the scores in the lowest quartile on the pretest shifted to a higher quartile on the posttest. There was also a 24.1% increase in scores in the highest quartile.

ELLM Kindergarten Children 2000/2001



In the 2000/2001 school year data, there were 214 children representing the kindergarten cohorts from all ELLM sites for whom both pretest and posttest TERA-2 Reading Quotient scores are available. In the fall of 2000, 48.6% of the TERA-2 pretest scores of the ELLM kindergarten children were in the lowest quartile (at or below "Below Average"), and 12.6% of pretest scores were in the highest quartile (at or above "Above Average") of the national norms. In the spring of 2001, 25.2% of the TERA-2 posttest scores of the ELLM kindergarten children were in the lowest quartile, and 31.3% of the posttest scores were in the highest quartile of the national norms. Therefore, among these 214 scores, 23.4% of the scores in the lowest quartile on the pretest shifted to a higher quartile on the posttest. There was also an 18.7% increase in scores in the highest quartile.

ELLM First Grade Children 2000/2001



In the 2000/2001 school year data, there were 132 children representing the first grade cohorts from all ELLM sites for whom both pretest and posttest TERA-2 Reading Quotient scores are available. In the fall of 2000, 25.8% of the TERA-2 pretest scores of the ELLM first grade children were in the lowest quartile (at or below "Below Average"), and 16.7% of the pretest scores were in the highest quartile (at or above "Above Average") of the national norms. In the spring of 2001, 25.0% of the TERA-2 posttest scores of the ELLM first grade children were in the lowest quartile, and 37.9% of the posttest scores were in the highest quartile of the national norms. Therefore, among these 132 scores, 0.8% of the scores in the lowest quartile on the pretest shifted to a higher quartile on the posttest. However, there was a 21.2% increase in scores in the highest quartile.

Phonemic Awareness Instruction Population

In the 2000/2001 school year, children who had *TERA-2* pretest scores in the lowest quartile of *ELLM TERA-2* pretest scores were selected for targeted phonemic awareness instruction. Table 5 provides summary statistics for this subset of *TERA-2* pretest and posttest scores. The data were analyzed as a repeated-measures ANOVA design in order to determine if there was a statistically significant difference in the pretest and posttest means. The subset of *ELLM* children who received targeted phonemic awareness instruction achieved significant gains in reading readiness during the 2000/2001 school year.

Table 5
Summary Statistics and ANOVA Results: *TERA-2* (Phonemic Awareness Population)

	Sample Size	Pretest Mean	Posttest Mean	Probability Value
All <i>ELLM</i> Children	103	72.79	88.98	<0.0001*
Pre-Kindergarten	50	73.54	89.20	<0.0001*
Kindergarten	38	70.92	90.47	<0.0001*
First Grade	15	75.00	84.47	0.0002*

* Denotes 99% confidence there was a difference in the pretest/posttest means.

In order to determine if improvement in reading readiness occurred across the ability continuum (measured by *TERA-2*), *TERA-2* scores from the children who received targeted phonemic awareness instruction are displayed in seven major categories: three categories representing the lowest 25 percentiles or “poor” rating, one category representing the middle 50 percentiles or “average” rating, and three categories representing the highest 25 percentiles or “superior rating” (see Figures 12-15). Inspection of Figures 12-15 indicates most scores moved from the “poor” to the “below average” category. Very few scores moved into the upper quartile of the national norms.

Early literacy theory and the *TERA-2* posttest means presented in Table 5 suggest the effectiveness of phonemic awareness instruction may differ among the three *ELLM* cohorts. In particular, the mean achievement of the pre-kindergarten and kindergarten children selected for phonemic awareness instruction may be greater than that of the first grade children. Data from *ELLM* children who received targeted phonemic awareness instruction were analyzed using an ANCOVA design with *TERA-2* pretest scores as the covariate. Tables 6 and 7 summarize these ANCOVA results.

Table 6

ANCOVA Summary: Kindergarten vs. First Grade

Source of Variance	Degrees of Freedom	Mean Square	F-Ratio	Probability
<i>TERA-2</i> Pretest	1	4688.34	25.85	<0.0001
Cohort	1	1004.19	5.54	0.0226*
		Kindergarten	First Grade	
Adjusted Means		91.55	81.73	

* Denotes 95% confidence the *TERA-2* posttest means of the two cohorts were different.

Table 7

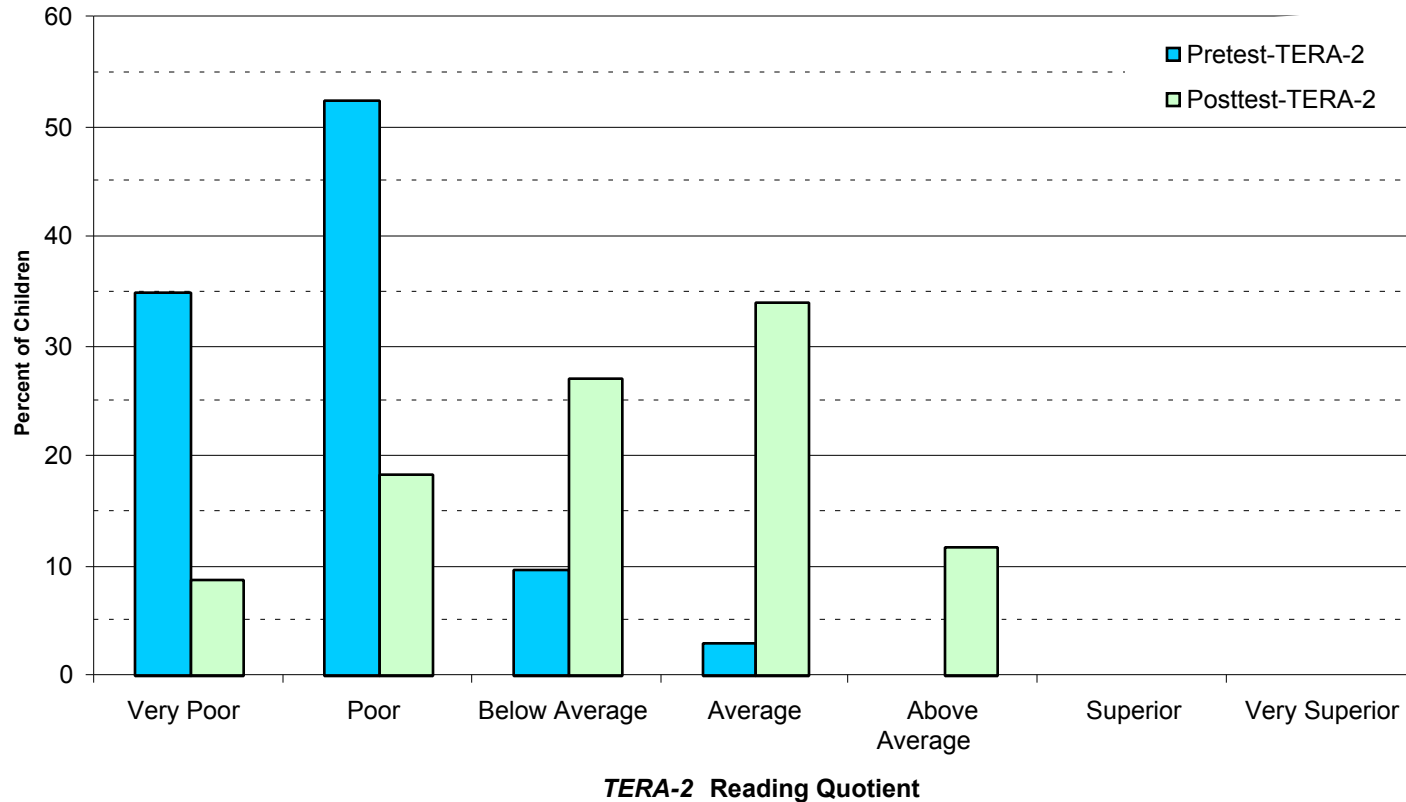
ANCOVA Summary: Pre-Kindergarten vs. First Grade

Source of Variance	Degrees of Freedom	Mean Square	F-Ratio	Probability
<i>TERA-2</i> Pretest	1	814.26	5.00	0.0290
Cohort	1	335.14	2.06	0.1562**
		Pre-Kindergarten	First Grade	
Adjusted Means		89.36	83.94	

** Denotes the *TERA-2* posttest means of the two cohorts were not different with 95% confidence.

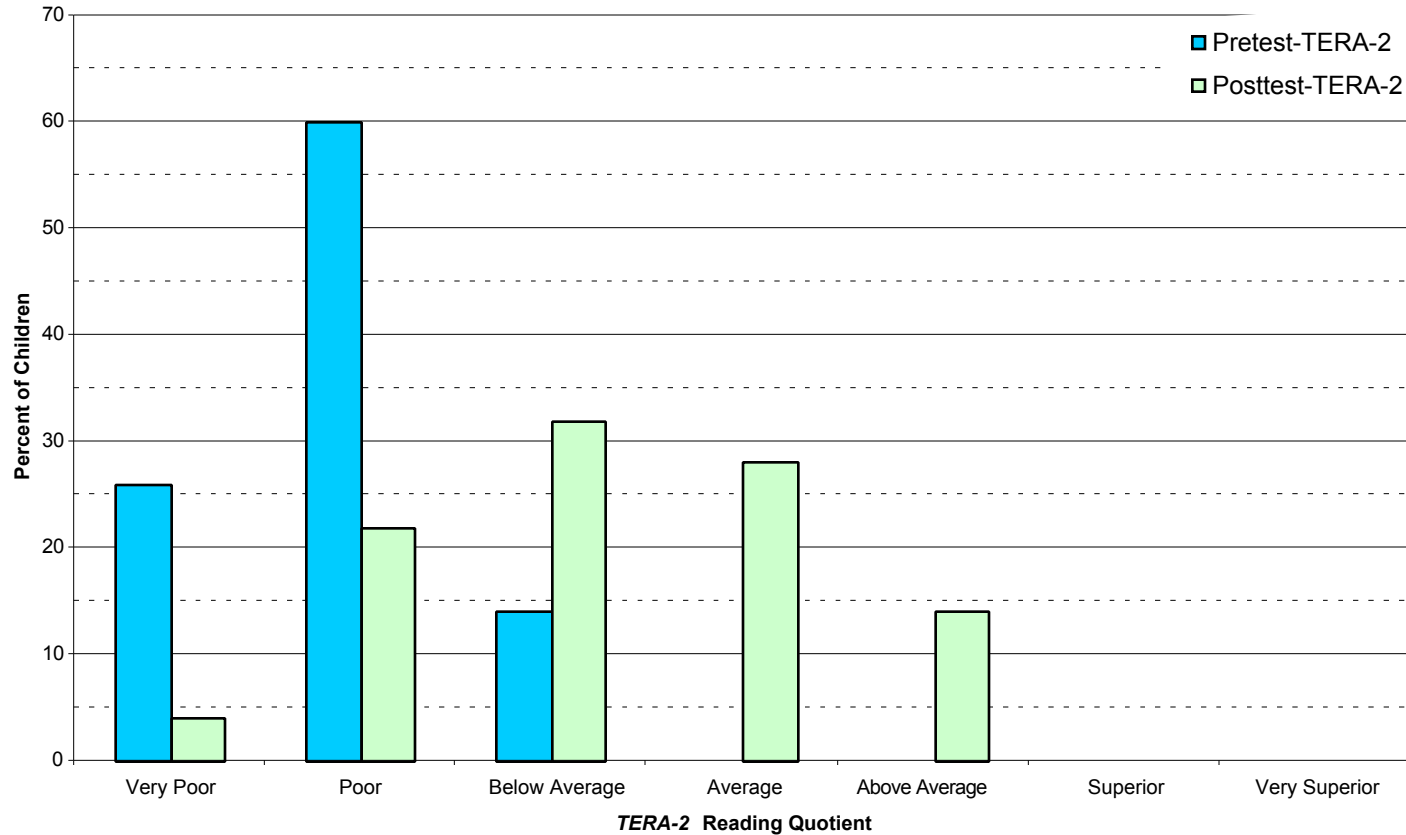
The results indicate phonemic awareness instruction was more effective (as measured by *TERA-2* posttest scores) with pre-kindergarten children who had low *TERA-2* pretest scores than with first grade children having similar pretest scores.

**ELLM Pre-Kindergarten, Kindergarten, and First Grade Children 2000/2001
Phonemic Awareness Instruction**



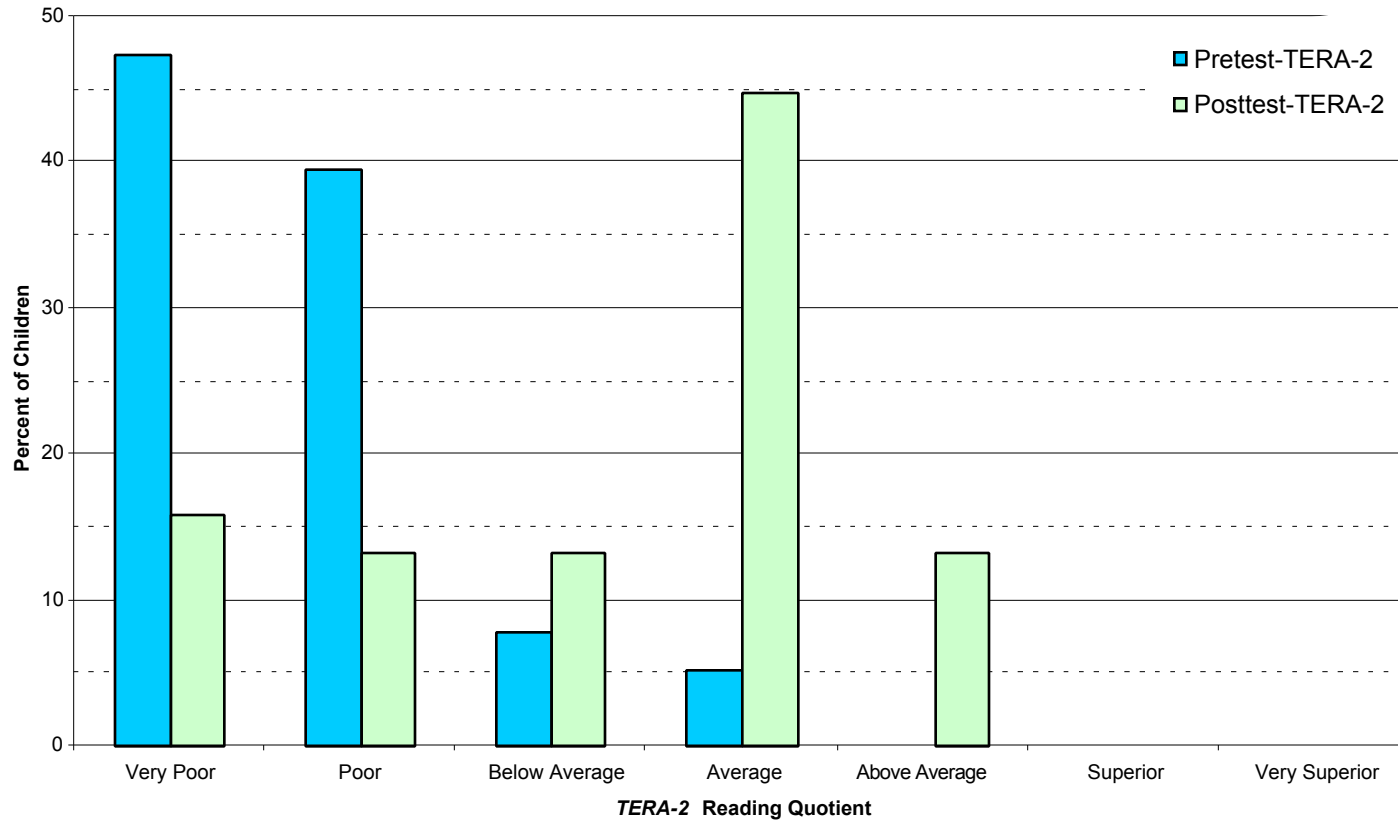
In the 2000/2001 school year data, there were 103 children representing the three cohorts from all ELLM sites who were engaged in targeted phonemic awareness instruction and for whom both pretest and posttest TERA-2 Reading Quotient scores are available. In the fall of 2000, 97.1% of the TERA-2 pretest scores of the selected ELLM children were in the lowest quartile (at or below "Below Average"), and none of the pretest scores were in the highest quartile (at or above "Above Average") of the national norms. In the spring of 2001, 54.4% of the TERA-2 posttest scores of these ELLM children were in the lowest quartile, and 11.7% of the posttest scores were in the highest quartile of the national norms. Therefore, among these 103 scores, 42.7% in the lowest quartile on the pretest shifted to a higher quartile on the posttest. There was also an 11.7% increase in scores in the highest quartile.

**ELLM Pre-Kindergarten Children 2000/2001
Phonemic Awareness Instruction**



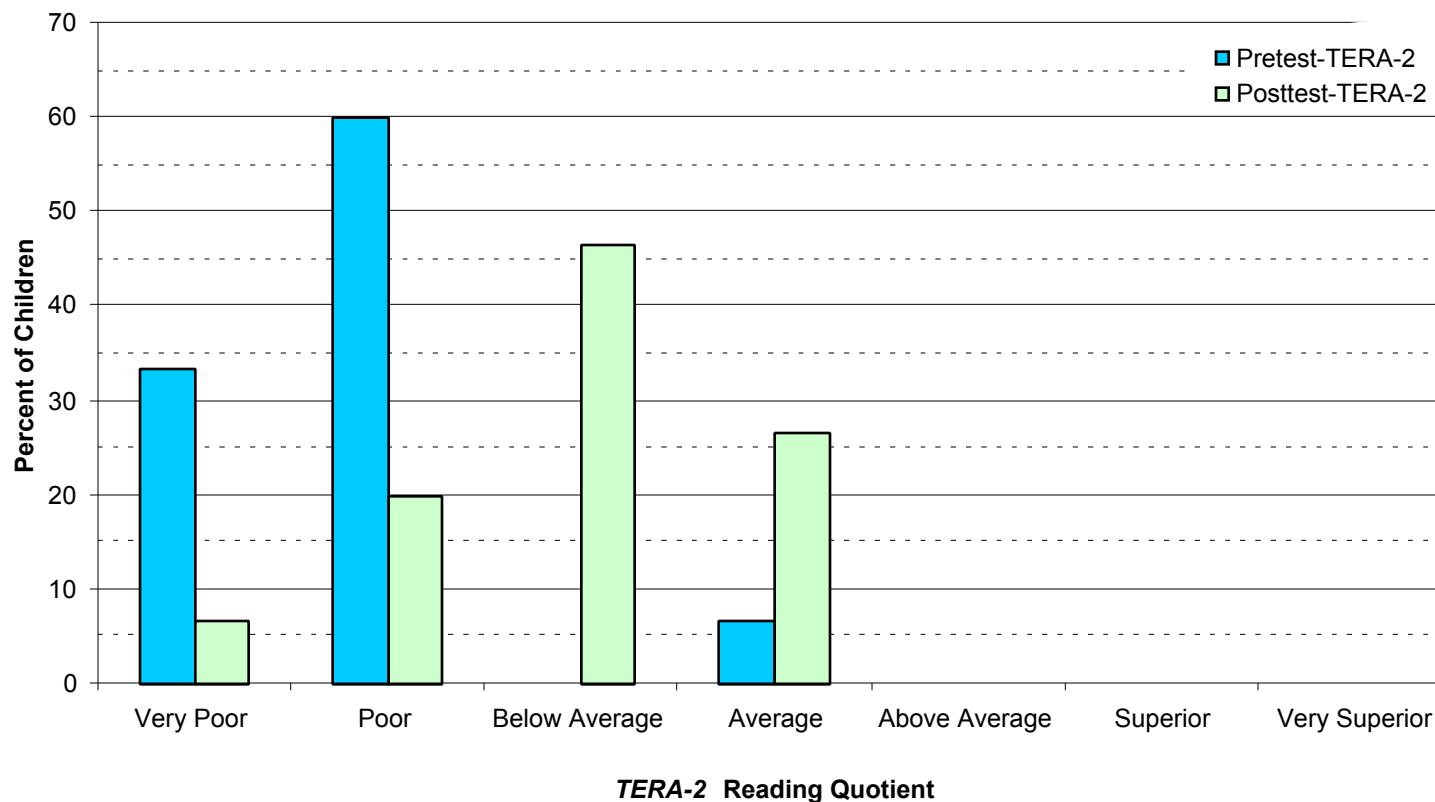
In the 2000/2001 school year data, there were 50 children representing the pre-kindergarten cohorts from all ELLM sites who were engaged in targeted phonemic awareness instruction and for whom both pretest and posttest TERA-2 Reading Quotient scores are available. In the fall of 2000, all of the TERA-2 pretest scores from these selected ELLM pre-kindergarten children were in the lowest quartile (at or below "Below Average"). In the spring of 2001, 58% of the TERA-2 posttest scores from these ELLM pre-kindergarten children were in the lowest quartile, and 14% of the posttest scores were in the highest quartile (at or above "Above Average") of the national norms. Therefore, among these 50 scores, 42% in the lowest quartile on the pretest shifted to a higher quartile on the posttest. There was also a 14% increase in scores in the highest quartile.

**ELLM Kindergarten Children 2000/2001
Phonemic Awareness Instruction**



In the 2000/2001 school year data, there were 38 children representing the kindergarten cohorts from all ELLM sites who were engaged in targeted phonemic awareness instruction and for whom both pretest and posttest TERA-2 Reading Quotient scores are available. In the fall of 2000, 94.7% of the TERA-2 pretest scores from the selected ELLM kindergarten children were in the lowest quartile (at or below "Below Average"), and none of the pretest scores were in the highest quartile (at or above "Above Average") of the national norms. In the spring of 2001, 42.1% of the TERA-2 posttest scores from these ELLM kindergarten children were in the lowest quartile, and 13.2% of the posttest scores were in the highest quartile of the national norms. Therefore, among these 38 scores, 52.6% in the lowest quartile on the pretest shifted to a higher quartile on the posttest. There was also a 13.2% increase in scores in the highest quartile.

**ELLM First Grade Children 2000/2001
Phonemic Awareness Instruction**



In the 2000/2001 school year data, there were 15 children representing the first grade cohorts from all ELLM sites, who were engaged in targeted phonemic awareness instruction and for whom both pretest and posttest TERA-2 Reading Quotient scores are available. In the fall of 2000, 93.3% of the TERA-2 pretest scores from these selected ELLM first grade children were in the lowest quartile (at or below "Below Average"), and none of the pretest scores were in the highest quartile (at or above "Above Average") of the national norms. In the spring of 2001, 73.3% of the TERA-2 posttest scores from the these ELLM first grade children were in the lowest quartile, and none of the posttest scores were in the highest quartile of the national norms. Therefore, among these 15 scores, 20.0% in the lowest quartile on the pretest shifted to a higher quartile on the posttest. However, there was no change in the number of scores in the highest quartile.

Validation Study

In the 2000/2001 school year, random samples of *ELLM* children were selected for a validation study of the use of *TERA-2* for evaluation of the *ELLM* initiative. Additionally, *TERA-2* was revised as *TERA-3*; therefore, a correlation study of the two versions of *TERA* was conducted using the *ELLM* population. The design of the 2000/2001 school year study is presented in Table 8.

Table 8
Validity Study: Evaluation Design

Sample Size	Pretest	Posttest	Additional Posttest
537	<i>TERA-2</i>	<i>TERA-2</i>	None
66	<i>TERA-2</i>	<i>TERA-2</i>	<i>TERA-3</i>
70	<i>Expressive One-Word</i>	<i>Expressive One-Word</i>	<i>TERA-2</i>
64	<i>Visual-Auditory Learning</i>	<i>Visual-Auditory Learning</i>	<i>TERA-2</i>

The correlations of *Expressive One-Word Picture Vocabulary Test*, *Woodcock Reading Mastery Tests: Visual-Auditory Learning*, and *TERA-3* with the *TERA-2* are presented in Table 9. All correlations were positive, with the strongest relationship occurring between *TERA-2* and *TERA-3*.

Table 9
Correlations of the *TERA-3*, *Woodcock Reading Mastery Tests: Visual-Auditory Learning*, and the *Expressive One-Word Picture Vocabulary Test* with the *TERA-2*

	Correlations		
<i>TERA-2 Form B</i>	<i>TERA-3 Form B</i>	<i>Visual-Auditory Learning</i>	<i>Expressive One-Word</i>
	0.671	0.476	0.397
	N=66	N=64	N=70

Research indicates alphabet letter recognition is a cornerstone of early literacy and a fundamental component of reading readiness. Therefore, as a part of the validation process, a correlation study was also conducted between the four measures, *TERA-2*, *TERA-3*, *Woodcock Reading Mastery Tests: Visual-Auditory Learning*, and *Expressive One-Word Picture Vocabulary Test*, with the pre-kindergarten and kindergarten children’s Alphabet Letter Recognition Inventory posttest scores. Table 10 displays these correlations. The *TERA-2* and *TERA-3* scores provided the strongest positive correlations with the Alphabet Letter Recognition Inventory posttest scores. The correlation of the *Expressive One-Word Picture Vocabulary Test* and the Alphabet Letter Recognition Inventory posttest was not statistically different from zero at the 95% confidence level.

Table 10
Correlations of the *TERA-2*, *TERA-3*, *Woodcock Reading Mastery Tests: Visual-Auditory Learning*, and the *Expressive One-Word Picture Vocabulary Test* Scores with the Alphabet Letter Recognition Inventory Posttest Scores

	Correlation			
Alphabet Posttest	<i>TERA-2</i>	<i>TERA-3</i>	<i>Visual-Auditory Learning</i>	<i>Expressive One-Word</i>
	0.532	0.367	0.056	0.275
	N=492	N=38	N=34	N=44
	Probability =0. 0001*	Probability =0. 0233*	Probability =0. 7547	Probability =0. 0713**

* Denotes at least 95% confidence there was a positive relationship between the test scores.

** Denotes at least 90% confidence there was a positive relationship between the test scores.

However, *ELLM* children achieved significant improvement in the abilities measured by the *Expressive One-Word Picture Vocabulary Test* and *Woodcock Reading Mastery Tests: Visual-Auditory Learning*. Tables 11 and 12 provide summary statistics and ANOVA results for the random samples of children assessed using the *Expressive One-Word Picture Vocabulary Test* and *Woodcock Reading Mastery Tests: Visual-Auditory Learning*, respectively.

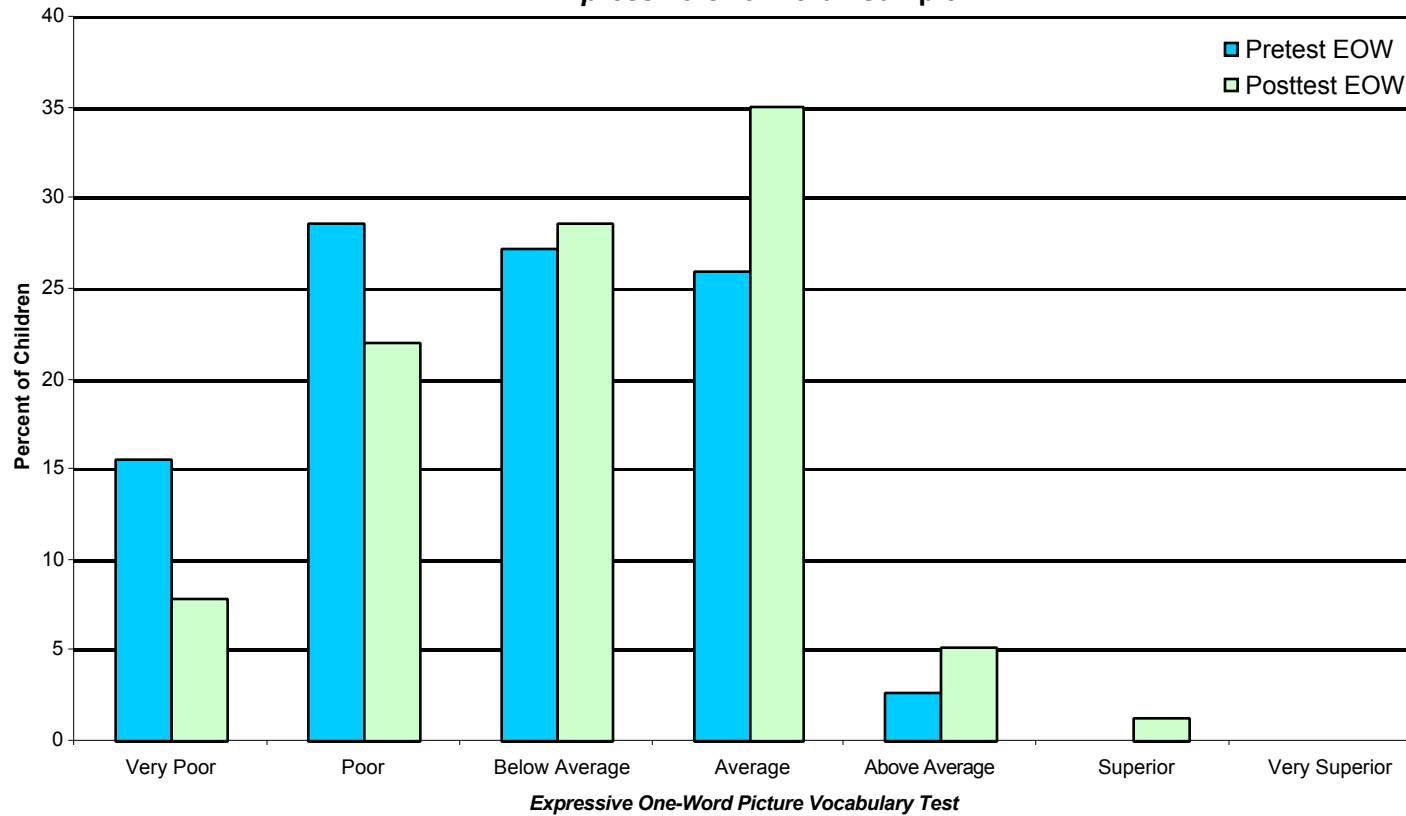
Table 11
Summary Statistics and ANOVA Results: *Expressive One-Word Picture Vocabulary Test*

	Sample Size	Pretest Mean	Posttest Mean	Probability Value
All <i>ELLM</i> Children	77	83.31	87.21	<0.0001*
Pre-Kindergarten	19	86.21	90.37	0.0078*
Kindergarten	28	83.07	87.61	0.0108*
First Grade	30	81.70	84.83	0.0179*

* Denotes at least 98% confidence there was a difference in the pretest/posttest means.

In order to determine if improvement in children’s English speaking vocabulary occurred across the ability continuum (measured by the *Expressive One-Word Picture Vocabulary Test*), scores from the *Expressive One-Word Picture Vocabulary Test* are displayed in seven major categories: three categories representing the lowest 25 percentiles or “poor” rating, one category representing the middle 50 percentiles or “average” rating, and three categories representing the highest 25 percentiles or “superior” rating (see Figure 16). Inspection of Figure 16 indicates there was a systematic shift in the percent of children scoring in lower ability categories to higher ability categories.

ELLM Pre-Kindergarten, Kindergarten, and First Grade Children 2000/2001
Expressive One-Word Sample



In the 2000/2001 school year data, there were 77 children representing a random sample of children from the three cohorts at all ELLM sites for whom both pretest and posttest Expressive One-Word scores are available: 19 representing the pre-kindergarten cohort, 28 representing the kindergarten cohort, and 30 representing the first grade cohort. In the fall of 2000, 71.4% of the Expressive One-Word pretest scores of the ELLM children were in the lowest quartile (at or below "Below Average"), and 2.6% of the pretest scores were in the highest quartile (at or above "Above Average") of the national norms. In the spring of 2001, 58.4% of the Expressive One-Word posttest scores of the ELLM children were in the lowest quartile, and 6.5% of the posttest scores were in the highest quartile of the national norms. Therefore, among these 77 scores, 13.0% of the scores in the lowest quartile on the pretest shifted to a higher quartile on the posttest. There was also a 3.9% increase in scores in the highest quartile.

Table 12

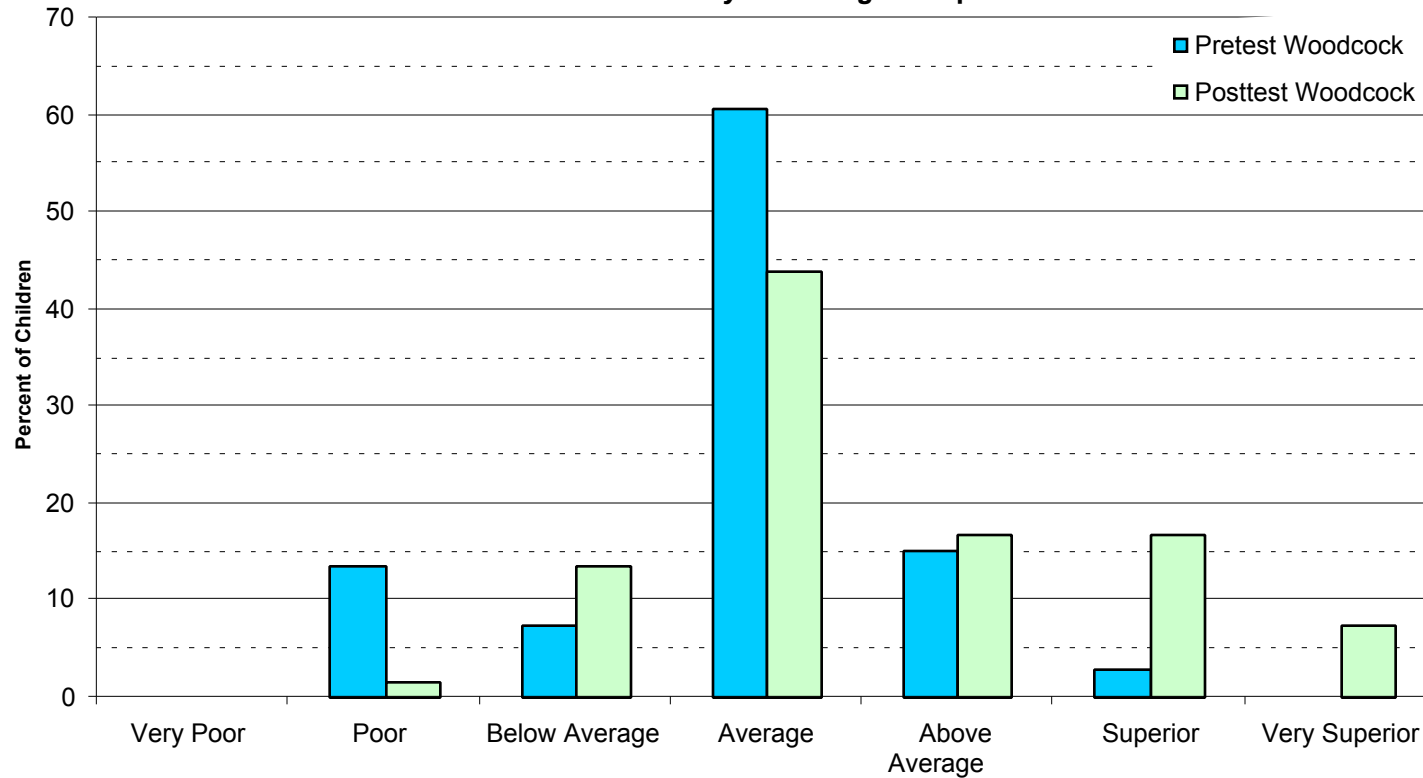
Summary Statistics: *Woodcock Reading Mastery Tests: Visual-Auditory Learning* ANOVA Results

	Sample Size	Pretest Mean	Posttest Mean	Probability Value
All <i>ELLM</i> Children	66	98.41	107.23	<0.0001*
Kindergarten	37	102.11	108.70	0.0012*
First Grade	29	93.69	105.34	0.0006*

* Denotes at least 99% confidence there a difference in the pretest/posttest means.

In order to determine if improvement in children’s ability to form associations between visual stimuli and oral responses occurred across the ability continuum (measured by the *Woodcock Reading Mastery Tests: Visual-Auditory Learning*), scores from the *Woodcock Reading Mastery Tests: Visual-Auditory Learning* are also displayed in seven major categories: three categories representing the lowest 25 percentiles or “poor” rating, one category representing the middle 50 percentiles or “average” rating, and three categories representing the highest 25 percentiles or “superior” rating (see Figure 17). Inspection of Figure 17 indicates there was a systematic shift in the percent of children in lower ability categories to higher ability categories.

**ELLM Kindergarten and First Grade Cohorts 2000/2001
Visual-Auditory Learning Sample**



Woodcock Reading Mastery: Visual-Auditory Learning

In the 2000/2001 school year data, there were 66 children representing a random sample of children from the kindergarten and first grade cohorts at all ELLM sites for whom both pretest and posttest Visual-Auditory Learning scores are available: 37 representing the kindergarten cohort and 29 representing the first grade cohort. In the fall of 2000, 21.2% of the Visual-Auditory Learning pretest scores of the ELLM children were in the lowest quartile (at or below "Below Average"), and 18.2% of the pretest scores were in the highest quartile (at or above "Above Average") of the national norms. In the spring of 2001, 15.2% of the Visual-Auditory Learning posttest scores of the ELLM children were in the lowest quartile, and 40.9% of the posttest scores were in the highest quartile of the national norms. Therefore, among these 66 scores, 6.0% of the scores in the lowest quartile on the pretest shifted to a higher quartile on the posttest. There was also a 22.7% increase in scores in the highest quartile.

Alphabet Letter Recognition Inventory Results

In the 2000/2001 school year, 618 children from the pre-kindergarten and kindergarten cohorts from *ELLM* sites had both pretest and posttest Alphabet Letter Recognition Inventory scores. Table 13 provides summary statistics for these two measures. The pretest and posttest scores were analyzed as a repeated-measures ANOVA design in order to determine if there was a statistically significant difference in the pretest and posttest means. The *ELLM* children in both the pre-kindergarten and kindergarten cohorts achieved significant gains in alphabet recognition during the 2000/2001 school year.

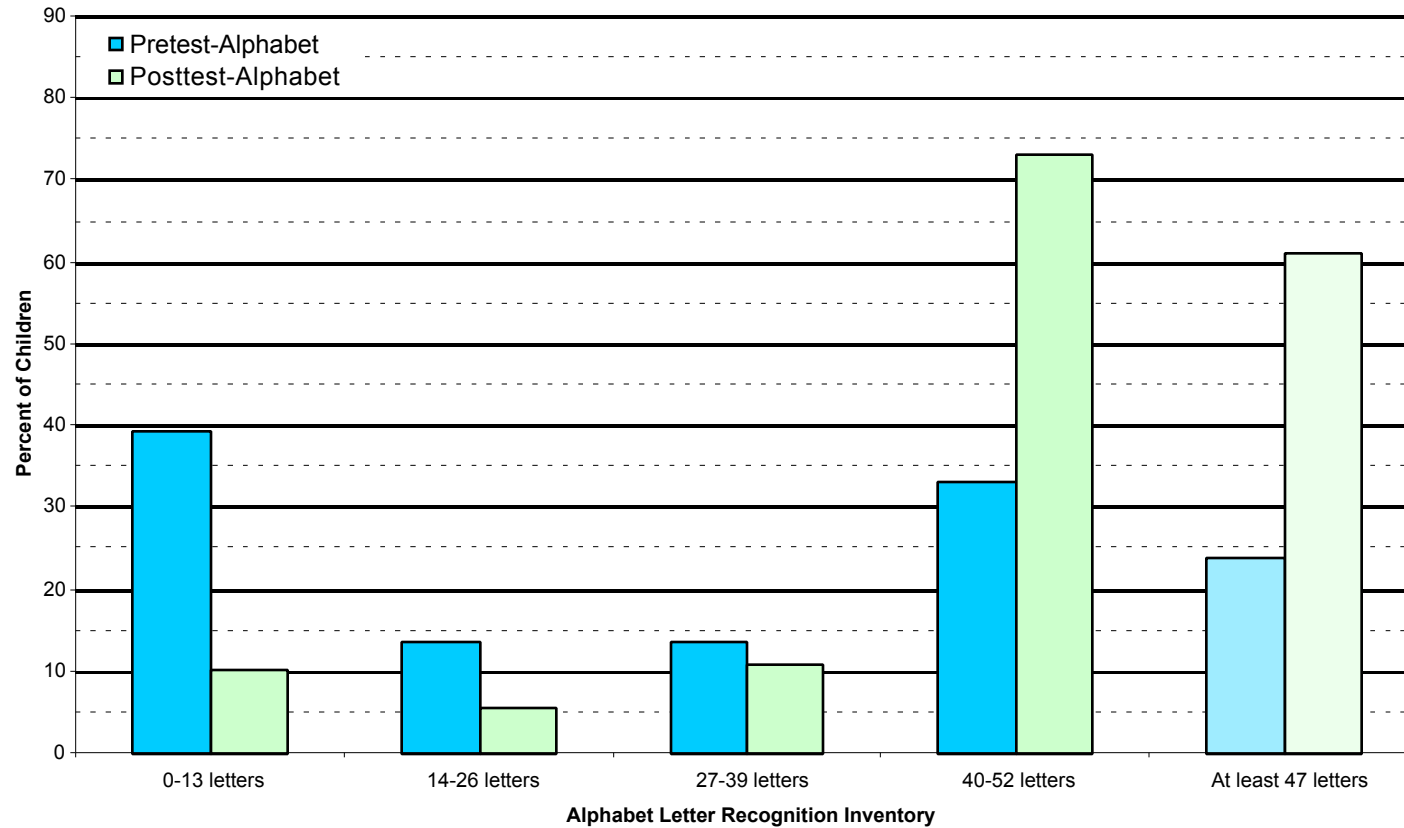
Table 13
Summary Statistics and ANOVA Results: Alphabet Letter Recognition Inventory

	Sample Size	Pretest Mean	Posttest Mean	Probability Value
All <i>ELLM</i> Children	618	25.24	41.86	<0.0001*
Pre-Kindergarten	326	23.34	37.44	<0.0001*
Kindergarten	292	27.37	46.78	<0.0001*

* Denotes more than 99% confidence there was a difference in the pretest/posttest means.

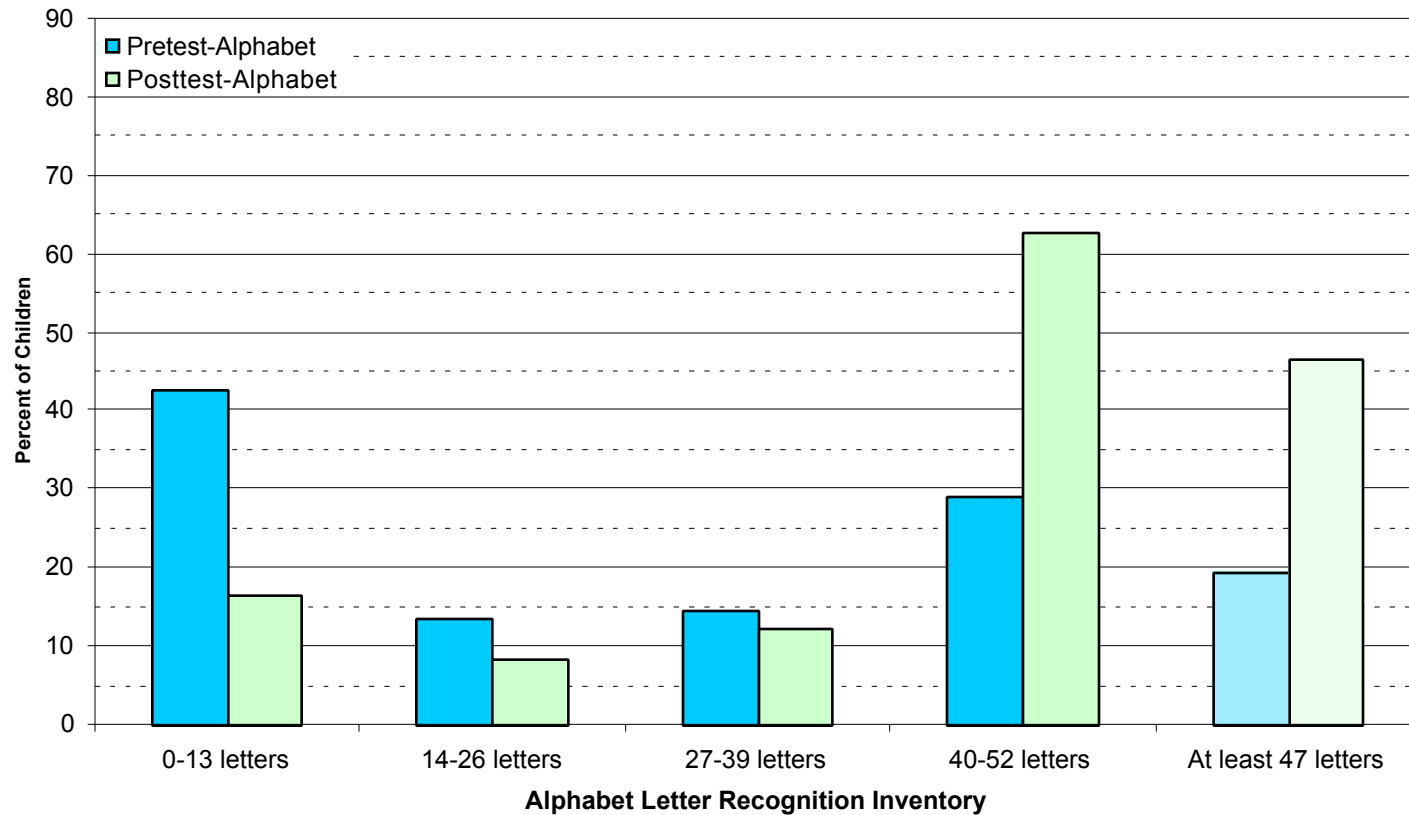
The posttest mean of the Alphabet Letter Recognition Inventory indicated a typical *ELLM* child recognized 81% of the upper and lower case letters. However, in order to determine the range of alphabet letter recognition ability of *ELLM* children, both pre-kindergarten and kindergarten children's scores are displayed using four recognition categories: 0-13 letters, 14-26 letters, 27-39 letters, and 40-52 letters. Because of the importance of this skill, a proficiency category of at least 90% recognition (47 letters) is also displayed (see Figures 18-20). Inspection of Figures 18-20 indicates the largest shift in scores was from the 0-13 letters category to the 40-52 letters category. The two middle categories remained comparatively stable.

**ELLM Pre-Kindergarten and Kindergarten Children 2000/2001
Alphabet Letter Recognition**



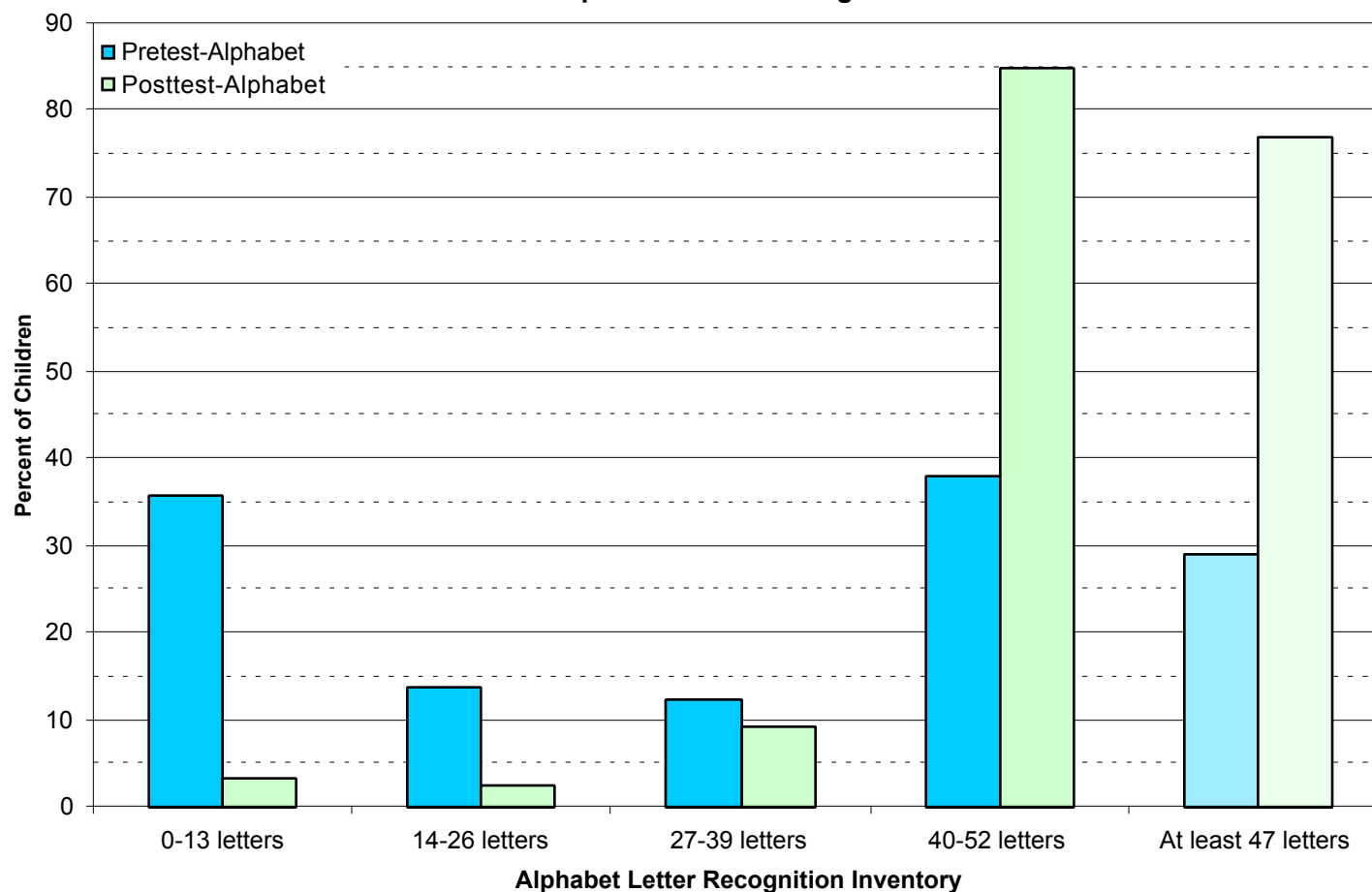
In the 2000/2001 school year data, there were 618 children representing the pre-kindergarten and kindergarten cohorts from all ELLM sites for whom both pretest and posttest Alphabet Letter Recognition Inventory scores are available. Of these scores, 29% in the lowest recognition category on the pretest shifted to a higher category on posttest. There was also an increase of 40% in scores in the highest recognition category. Sixty-one percent (up from 24%) of the combined pre-kindergarten and kindergarten children exhibited at least 90% mastery of the upper and lower case letters at posttest.

**ELLM Pre-Kindergarten Children 2000/2001
Alphabet Letter Recognition**



In the 2000/2001 school year data, there were 326 children representing the pre-kindergarten cohorts from all ELLM sites for whom both pretest and posttest Alphabet Letter Recognition Inventory scores are available. Of these scores, 26% in the lowest recognition category on the pretest shifted to a higher category on posttest. There was also an increase of 34% in scores in the highest recognition category. Forty-seven percent (up from 19%) of the pre-kindergarten children exhibited at least 90% mastery of the upper and lower case letters at posttest.

**ELLM Kindergarten Children 2000/2001
Alphabet Letter Recognition**



In the 2000/2001 school year data, there were 292 children representing the kindergarten cohorts from all ELLM sites for whom both pretest and posttest Alphabet Letter Recognition Inventory scores are available. Of these scores, 33% in the lowest recognition category on the pretest shifted to a higher category on posttest. There was also an increase of 47% in scores in the highest recognition category. Seventy-seven percent (up from 29%) of the kindergarten children exhibited at least 90% mastery of the upper and lower case letters at posttest.

Technology: *The Waterford Early Reading Program (WERP)*

The evaluation of *WERP* was based on *TERA-2* posttest scores relative to participation in *WERP* and the amount of time children used the program. The developers of the program recommend children spend 15 minutes daily on a personalized program. Therefore, based on a 180-day school year, children should spend 2,700 total minutes using *WERP* over the course of the school year. However, the typical *ELLM* child used the program 1,445 minutes – far less than the amount recommended by the program’s developers.

Among the 243 *ELLM* children who had *WERP* in their classrooms and for whom both *TERA-2* pretest/posttest scores are available: 25% of the children used the program 2,159 or more minutes, 25% of the children used the program between 2,158 and 1,038 minutes, 25% of the children used the program between 1,037 and 472 minutes, and 25% of the children used the program less than 472 minutes. Therefore, at least 50% of the children in *ELLM* classrooms that had *WERP* used the program less than 50% of the time recommended by the program’s developers.

Evaluation

Evaluation of the use of *WERP* involved answering three questions: 1) Was the mean *TERA-2* posttest score of *ELLM* children in classrooms that used *WERP* higher than that of *ELLM* children in classrooms that did not use *WERP*? 2) Among *ELLM* children who used *WERP* was the mean *TERA-2* posttest score associated with the amount of time children used the program? 3) Among *ELLM* children in classrooms that used *WERP* was the access to and use of *WERP* equitable?

Question 1: Was the mean *TERA-2* posttest score of *ELLM* children in classrooms that used *WERP* higher than that of *ELLM* children in classrooms that did not use *WERP*?

In order to answer this question, the 603 *ELLM* children with *TERA-2* pretest/posttest scores were categorized as “users” or “non-users” of *WERP*. Table 14 summarizes the mean *TERA-2* pretest and posttest scores of these two groups of *ELLM* children.

Table 14
Summary Statistics: *TERA-2* (*WERP* “Users” and “Non-Users”)

	Sample Size	Pretest Mean	Posttest Mean
<i>WERP</i> “Non-Users”	360	91.60	98.86
<i>WERP</i> “Users”	243	92.59	102.79

The resulting data were analyzed as an ANCOVA design with the *TERA-2* pretest scores as the covariate in order to determine if the mean *TERA-2* posttest scores were different for “users” and “non-users” of *WERP*. Table 15 summarizes the results of the ANCOVA analysis.

Table 15
ANCOVA Summary: *WERP* “Users” and “Non-Users”

Source of Variance	Degrees of Freedom	Mean Square	F-Ratio	Probability Value
<i>TERA-2</i> Pretest	1	79777.63	561.50	<0.0001
“Users/Non-Users”	1	1529.47	10.57	0.0011*
Adjusted Mean	“Non-Users”	“Users”		
	99.13	102.38		

* Denotes 98% confidence the *TERA-2* posttest scores were different.

The ANCOVA results indicated that after controlling for initial differences in *TERA-2* pretest scores, the mean *TERA-2* posttest score of the “users” of *WERP* was higher than that of “non-users.” Because classrooms were not randomly assigned to use *WERP*, however, these results should be interpreted with great care.

Additionally, 103 *ELLM* children with low *TERA-2* pretest scores were selected for targeted instruction in phonemic awareness, and 57 of them were also in classrooms that used *WERP*. In order to determine if the use of *WERP* was an effective instructional strategy when used in addition to targeted phonemic awareness instruction, the 103 children were categorized as “users” or “non-users” of *WERP*. Table 16 presents summary statistics of *TERA-2* scores from these two groups of *ELLM* children.

Table 16
TERA-2 Summary Statistics: *WERP* “Users” and “Non-Users”
(Among Children Targeted for Phonemic Awareness Instruction)

	Sample Size	Pretest Mean	Posttest Mean
“Users”	57	73.33	89.60
“Non-Users”	46	72.11	88.22

The resulting data were analyzed as an ANCOVA design with the *TERA-2* pretest scores as the covariate in order to determine if the mean *TERA-2* posttest scores of the two groups of children were different. Table 17 summarizes the results of the ANCOVA.

Table 17
 ANCOVA Summary: *WERP* “Users” and “Non-Users”
 (Among *ELLM* Children Targeted for Phonemic Awareness Instruction)

Source of Variance	Degrees of Freedom	Mean Square	F-Ratio	Probability Value
<i>TERA-2</i> Pretest	1	3838.85	19.25	<0.0001
“Users/Non-Users”	1	7.12	0.04	0.8506

The results indicated that after accounting for initial differences in *TERA-2* pretest scores, using *WERP* did not significantly impact the *TERA-2* posttest scores of children receiving targeted phonemic awareness instruction.

Therefore, after controlling for initial differences in *TERA-2* pretest scores, children in *ELLM* classrooms that used *WERP* had a higher mean *TERA-2* posttest score than children in *ELLM* classrooms that did not use *WERP*. However, this result can be misleading because *WERP* was not randomly assigned to classrooms. Additionally, among *ELLM* children who were selected for phonemic awareness instruction, after controlling for initial differences in *TERA-2* pretest scores, there was no difference in mean *TERA-2* posttest scores between children who used *WERP* and those who did not use *WERP*.

Question 2: Among *ELLM* children who used *WERP* was the mean *TERA-2* posttest score associated with the amount of time children used the program?

In order to answer this question, *TERA-2* pretest/posttest scores and the amount of time the 243 *ELLM* children used *WERP* were analyzed as a multiple regression. Table 18 summarizes the results of the regression analysis.

Table 18
Multiple Regression Results: The Effect of *TERA-2* Pretest Score and Total Time of *WERP* Use on *TERA-2* Posttest Score

Source of Variance	Degrees of Freedom	Mean Square	F-Ratio	Probability Value
<i>TERA-2</i> Pretest	1	27874.11	210.79	<0.0001
Total Time	1	4.77	0.04	0.8496

The results of the multiple regression indicated that the amount of time the program was used was not related to the *TERA-2* posttest score when the *TERA-2* pretest score was held constant.

Question 3: Among *ELLM* children in classrooms that used *WERP* was the access to and use of *WERP* equitable?

Two variables, *TERA-2* pretest scores and gender, were analyzed to answer this question. The 1999/2000 school year data suggested *ELLM* children with low *TERA-2* pretest scores used *WERP* more than *ELLM* children with higher *TERA-2* pretest scores. Therefore, data from 2000/2001 *ELLM* children were analyzed as a simple linear regression in order to determine if the *TERA-2* pretest scores were associated with the total amount of time children used *WERP*. Table 19 summarizes the results of the regression.

Table 19
The Relationship Between the Amount of Time Children
Used *WERP* and the *TERA-2* Pretest Scores

Parameter	Estimate	<i>t</i>	Probability
<i>TERA-2</i> Pretest	-10.045	-2.04	0.0429*

* Denotes at least 95% confidence the parameter was not zero.

The results of the regression indicated that an increase of one point on the *TERA-2* pretest is associated with a 10-minute decrease in the total time children used *WERP*. Therefore, children with lower *TERA-2* pretest scores (children who could potentially benefit the most from the use of *WERP*) used *WERP* more.

Gender equity provided an additional area of concern with the use of *WERP*. Table 20 presents summary statistics of scores from the 241 *ELLM* boys and girls for whom gender identification was available and who were in classrooms that used *WERP*. Even though *WERP* was not randomly assigned to classrooms, summary statistics presented in Table 20 indicate the number of boys (122) who used *WERP* was nearly the same as the number of girls (119) who used *WERP*.

Table 20
Summary Statistics (*TERA-2* Pretest/Posttest and Total Time):
ELLM Boys and Girls in Classrooms That Used *WERP*

	Sample Size	Pretest Means	Posttest Means	Total Time (<i>WERP</i>) Means (in Minutes)
Boys	122	89.92	101.29	1482.07
Girls	119	95.01	104.24	1429.55

In order to determine if the amount of time boys and girls used *WERP* was statistically different, data were analyzed as an ANCOVA design with the *TERA-2* pretest scores as the covariate. Table 21 summarizes the results of the ANCOVA analysis.

Table 21
ANCOVA Summary: *WERP* Boys and Girls

Source of Variance	Degrees of Freedom	Mean Square	F-Ratio	Probability Value
<i>TERA-2</i> Pretest	1	5214983.05	3.40	0.0665
Gender	1	1581.62	0.00	0.9744*
		Boys		Girls
Adjusted Means		1458.70 minutes		1453.51 minutes

*Denotes more than 97% confidence the mean time boys and girls used *WERP* did not differ.

After adjusting for initial differences in *TERA-2* pretest scores, there was no difference in the mean time that boys and girls used *WERP*. Results confirmed boys and girls had equal access to *WERP* and equitable use of the program.

Conclusions

1. After accounting for initial differences in *TERA-2* pretest scores, the mean *TERA-2* posttest score was higher for *ELLM* children in classrooms that used *WERP* than in classrooms that did not use *WERP*. This result may be misleading, however, because *WERP* was not randomly assigned to classrooms. Additionally, the total amount of time children used *WERP* was not associated with the mean *TERA-2* posttest score.
2. Children whose *TERA-2* pretest scores were among the lowest *TERA-2* pretest scores received targeted phonemic awareness instruction. Furthermore, these children had comparatively high use of *WERP*, as measured in the total time *WERP* was used over the school year. Yet, among this subset of *ELLM* children, there was no difference in the mean *TERA-2* posttest scores of the children who used *WERP* and those who did not use *WERP*.
3. Typically, *ELLM* children in classrooms that used *WERP* did not use the program 15 minutes daily as suggested by the program's developers. In fact, the amount of time children used the program was associated with their *TERA-2* pretest scores – children with low *TERA-2* pretest scores used the program more. Moreover, there was no difference in the amount of time boys and girls used *WERP*. Furthermore, results of data analysis indicated access to and use of *WERP* was equitable.

Comparison Study of the 1999/2000 and 2000/2001 School Year Data

The combined 1999/2000 and 2000/2001 school year data provided an opportunity to compare the effectiveness of the *ELLM* initiative over two school years. The data included *TERA-2* scores from three distinct groups of *ELLM* children: those who were in *ELLM* classrooms only during the 1999/2000 school year, those who were in *ELLM* classrooms only during the 2000/2001 school year, and those who were in *ELLM* classrooms during both the 1999/2000 and the 2000/2001 school years. *TERA-2* scores from children who were in *ELLM* classrooms for the two school years were excluded from the study; therefore, only those children who were in *ELLM* classrooms during the 1999/2000 or during the 2000/2001 school year were included. The comparison study answered the research question: Was the *ELLM* initiative equally effective in both school years?

Evaluation

There were *TERA-2* pretest/posttest scores from 869 children who were in *ELLM* classrooms for only one school year. Table 22 provides summary statistics of these *TERA-2* scores.

Table 22
Summary Statistics: *TERA-2* Pretest/Posttest Scores

School Year	Sample Size	Pretest		Posttest	
		Mean	Standard Deviation	Mean	Standard Deviation
1999/2000	376	94.11	14.21	105.36	14.80
2000/2001	493	90.54	16.48	99.59	16.88

An ANCOVA design using *TERA-2* pretest scores as the covariate and “cohort” and “school year” as categorical variables provided the answer to the research question. The initial ANCOVA model included all main effects, all two-way interaction effects, and the one three-way interaction effect. The ANOVA table for the final model is presented in Table 23.

Table 23
ANCOVA Results Comparing the Effectiveness of the *ELLM* Initiative for the 1999/2000 and 2000/2001 School Years

Source of Variance	Degrees of Freedom	Mean Square	F-Ratio	Probability Value
<i>TERA-2</i> Pretest	1	60933.13	376.79	<0.0001
School Year	1	4124.05	25.50	<0.0001
Cohort	2	514.04	3.18	0.0421
School-Year* <i>TERA-2</i> Pretest	1	3083.74	19.07	<0.0001*
School-Year*Cohort	2	669.35	4.14	0.0163*

* Denotes at least 95% confidence there was an interaction between the two variables.

Because “school year” interacts with both the *TERA-2* pretest and “cohort,” interpreting the main effects of “school year” and “cohort” would be misleading.

The “School Year” and *TERA-2* Pretest Interaction

In order to understand the interaction of the *TERA-2* pretest and “school year,” mean *TERA-2* posttest scores are plotted against high, medium, and low *TERA-2* pretest scores. The mean *TERA-2* posttest scores were calculated using simple linear regression equations for *TERA-2* pretest scores at one standard deviation above the pretest mean (high), at the pretest mean (medium), and at one standard deviation below the pretest mean (low). Figure 21 depicts this plot.

TERA-2 Pretest and "School Year" Interaction

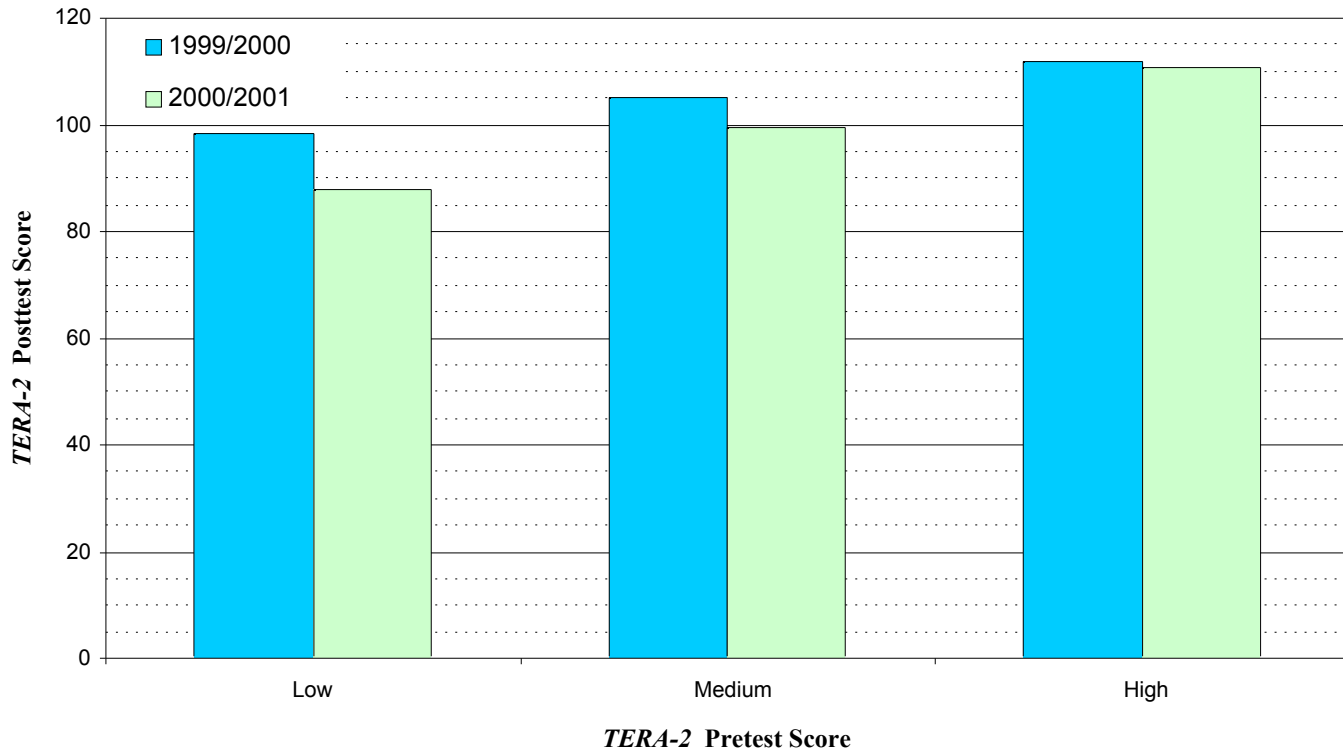


Figure 21

Children with low *TERA-2* pretest scores achieved higher *TERA-2* posttest scores during the 1999/2000 school year; however, there was very little difference between the 1999/2000 school year and 2000/2001 school year *TERA-2* posttest scores of the children with either medium or high *TERA-2* pretest scores. Therefore, as *TERA-2* pretest scores increased, the difference between the mean *TERA-2* posttest scores of the two school years decreased.

The “School Year” and “Cohort” Interaction

In order to understand the “school year” and “cohort” interaction, the mean *TERA-2* pretest scores for each school year and cohort were analyzed as an ANCOVA. Table 24 presents the summary statistics of the *TERA-2* pretest/posttest by school year and cohort.

Table 24
Summary Statistics: *TERA-2* Pretest and Posttest Means for Pre-Kindergarten, Kindergarten, and First Grade Children for the 1999/2000 and 2000/2001 School Years

School Year	Cohort	Sample Size	Pretest Mean	Posttest Mean	Adjusted Posttest Mean
1999/2000	Pre-Kindergarten	125	97.91	109.96	107.38
	Kindergarten	180	91.61	102.23	102.44
	First Grade	71	93.75	105.18	104.45
2000/2001	Pre-Kindergarten	251	90.97	99.96	100.74
	Kindergarten	182	88.49	98.71	101.22
	First Grade	60	94.92	100.73	98.75

The results of the multiple comparisons of the six means are presented in Table 25.

Table 25
ANOVA Results: Multiple Comparison of the School Year by Cohort *TERA-2*
Adjusted Posttest Means

Comparison	Probability Value
Pre-Kindergarten 1999/2000 vs. Pre-Kindergarten 2000/2001	<0.0001*
Pre-Kindergarten 1999/2000 vs. Kindergarten 2000/2001	<0.0001*
Pre-Kindergarten 1999/2000 vs. First Grade 2000/2001	<0.0001*
Pre-Kindergarten 1999/2000 vs. Kindergarten 1999/2000	0.0011*
First Grade 1999/2000 vs. First Grade 2000/2001	0.0110**
First Grade 1999/2000 vs. Pre-Kindergarten 2000/2001	0.0305**

* Denotes the adjusted posttest means were different (with 95% confidence) after adjusting the probability value for multiple testing.

** Denotes the adjusted posttest means were different (with 95% confidence) when the probability value is not adjusted for multiple testing.

The 1999/2000 school year pre-kindergarten cohort achieved a higher adjusted mean *TERA-2* posttest score than the 1999/2000 school year kindergarten cohort and all of the 2000/2001 school year cohorts. Moreover, there was no difference in the adjusted *TERA-2* mean posttest scores among the three 2000/2001 school year cohorts.

Conclusion

Research Question: Was the *ELLM* initiative equally effective in both school years?

The *ELLM* initiative was effective with children from all three cohorts in both the 1999/2000 and 2000/2001 school years. A comparison of the effectiveness of the *ELLM* initiative over the 1999/2000 and 2000/2001 school years depended on both the cohort of children and the level of the *TERA-2* pretest scores. In general, the *ELLM* initiative was more effective during the 1999/2000 school year with children with low *TERA-2* pretest scores. Moreover, the mean adjusted *TERA-2* posttest score of the 1999/2000 pre-kindergarten cohort was higher than all three 2000/2001 cohorts and was higher than the 1999/2000 kindergarten cohort. The *ELLM* initiative was equally effective with all three cohorts during the 2000/2001 school year.

Overall Summary and Conclusions

Analyses of the 1999/2000 and 2000/2001 school year data from children in all three cohorts in *ELLM* classrooms indicated the following:

7. In the 1999/2000 school year, *ELLM* children with *TERA-2* scores across the ability continuum had improved reading readiness scores at the end of the school year.
8. In the spring of the 1999/2000 school year, the typical *ELLM* child recognized 85% of the upper and lower case letters of the alphabet.
9. In the 2000/2001 school year, *ELLM* children with *TERA-2* scores across the ability continuum had improved reading readiness scores at the end of the school year.
10. In the 2000/2001 school year, *ELLM* children selected for targeted phonemic awareness instruction had improved reading readiness scores at the end of the school year. However, the targeted phonemic awareness instruction was more effective with pre-kindergarten children than with first grade children.
11. In the spring of the 2000/2001 school year, the typical *ELLM* child recognized 81% of the upper and lower case letters of the alphabet.
12. Access to and use of the *Waterford Early Reading Program (WERP)* was equitable in *ELLM* classrooms that used the program. The mean reading readiness achievement in classrooms that used *WERP* was higher than in classrooms that did not use *WERP*. These results can be misleading, however, because *WERP* was not randomly assigned to classrooms. Furthermore, the use of *WERP* was not effective when used in addition to targeted phonemic awareness instruction.
13. A comparison of the effectiveness of the *ELLM* initiative over the two school years depended on the *TERA-2* pretest scores and cohort. In general, the *ELLM* initiative was more effective during the 1999/2000 school year for children with low pretest scores. The *ELLM* initiative was very effective with the pre-kindergarten cohort during the 1999/2000 school year and was equally effective among the three cohorts during the 2000/2001 school year.