

Project LEARN

Results from Year 1, 2005 – 2006



Center for Promoting Research to Practice

Lehigh University

Project Evaluation Team

Patricia Manz, Ph.D, Ageliki Nicolopoulou, Ph.D., Edward S. Shapiro, Ph.D.

Co-Project Directors

Karen L. Gischlar, Ed.S., Project Coordinator

December 9, 2006 Revision

ERF Report Year 1: October 2005 to July 2006

1. Who are Early Reading First Children?

The sample for the 2005-2006 school year consisted of 70 students, which was approximately 50% of the population enrolled in Community Services for Children, Inc.'s Early Reading first classrooms. Students were selected for participation in the evaluative aspects of the grant through random assignment. The sample is representative of the current population in student age, gender, and primary language. Specifically, students are represented in the sample as follows:

- Age:
 - 18 three-year-old children (25.7%)
 - 52 four-year-old children (74.3%)
- Gender:
 - 33 male (47.1%) and 37 female (52.9%) students
 - a.) 8 three-year-old males (11.4% of total population)
 - b.) 25 four-year-old males (35.7% of total population)
 - c.) 10 three-year-old females (14.3% of total population)
 - d.) 27 four-year-old females (38.6% of total population)
- Primary language:
 - 26 (37.14%) English Language Learners (ELL), primary language Spanish
 - (Students were identified as ELL through information provided by their parents or primary care givers at the time of enrollment.)
- Absenteeism: Attendance rates are reported by average numbers of days attended by students in the sample over days possible for months January through June 2006.
 - January 2006: 12.69 mean days attended/14 possible days
 - February 2006: 16.44 mean days attended/19 possible days
 - March 2006: 17.87 mean days attended/21 possible days
 - April 2006: 15.10 mean days attended/17 possible days
 - May 2006: 18.56 mean days attended/21 possible days
 - June 2006: 16.44 mean days attended/21 possible days
- Attrition: During the course of the 2005-2006 school year, three students (4.3%) withdrew from the program and, thus, are missing data.

2. How do we know that ERF is working?

Due to the lack of a control group, one cannot directly attribute student progress to the ERF curriculum within the first cohort. Over time, some indication of amount of growth of subsequent cohorts of students in the ERF curriculum will be compared to the initial cohort, which should offer some potential indirect evidence of the impact of the curriculum.

To measure progress, students included in the sample were assessed with the following measures:

- *Peabody Picture Vocabulary Test: Third Edition (PPVT: III)*, AGS Publishing

This test is a widely used measure of receptive vocabulary for standard English and a screening tool of verbal ability. It is an individually administered, norm-referenced instrument that is offered in two parallel forms for reliable testing and retesting. During administration, the student is presented with a four-picture plate as the examiner says a stimulus word that corresponds to one of the pictures in the plate. The student responds by pointing to one of the pictures. The *PPVT: III* requires no oral or written responses and no reading by the examinee. Raw scores are converted to standard scores and a middle-class, typically developing student should receive a score around 100 with a standard deviation of 15. Scores between 85 and 115 are considered within the average range of performance. The *PPVT: III* was administered to all 70 students included in the sample during the winter (January) and spring (June) as a pre- and posttest measure. A standard score and percentile were derived for each student.

- *Test de Vocabulario en Imagenes Peabody (TVIP)*, AGS Publishing

This test is based on the *Peabody Picture Vocabulary Test: Revised* and contains translated items to assess the vocabulary of Spanish-speaking and bilingual students. Administered and scored similarly to the *PPVT: III* described above, the student is shown a four-picture plate in the test easel as the examiner says a corresponding stimulus word. The student responds by pointing to one of the pictures and is not required to answer orally. The *TVIP* was administered to the 25 (35.7%) students who were identified as ELL in school records. The measure, like the *PPVT: III*, was administered in the winter (January) and spring (June) by university trained Spanish speakers and provided a standard score and percentile rank for each student.

- *Phonological Awareness Literacy Screening: PreK (PALS)*, University of Virginia

This test assesses skills that have been demonstrated in the literature as predictive of future reading success. These skills include name writing ability, upper and lower-case alphabet recognition, letter sound and beginning sound production, print and word awareness, rhyme awareness and nursery rhyme awareness. Some of the tasks (e.g., Alphabet Knowledge and Beginning Sound Awareness) require a verbal response from the student, while others (e.g., Rhyme Awareness and Print and Word Awareness) require the student to point to stimuli on a page. *PALS* was administered in the winter (January) and spring (June) as a pre- and posttest to the 4-year-old children whose native language was English ($n=33$; 47.1%). A raw score and descriptive (i.e., below developmental range, within developmental range, or above developmental range) was derived for each student for each of the eight subtests included within this measure.

- *CIRCLE* (Center for Improving the Readiness of Children for Learning and Education) Assessment, University of Texas Medical Health Science Center

CIRCLE is administered via a Personal Digital Assistant (PDA) and picture easel. It assesses a student's progress in letter naming, vocabulary, and phonological awareness and offers the teacher immediate feedback and suggested classroom interventions for addressing individual skill deficits. The student is asked to respond orally or to point to a picture as the administrator records responses on the PDA. The *CIRCLE* measure was designed to provide benchmarks at three points during the school year, fall, winter, and spring. Due to the delayed start of the grant, *CIRCLE* was administered only during the winter (February) and spring (June) of the 2005-2006 school year. Children identified as ELL through school records were administered the Spanish version of *CIRCLE* ($n=25$, 35.7%), while all others took the test in English ($n= 45$; 64.3%). A raw score and descriptive (i.e., full understanding, some understanding, or emerging understanding) was derived for each student for each of the nine subtests (eight subtests for the Spanish version) included within this measure.

- *Individual Growth and Development Indicators (IGDI)*, University of Minnesota

IGDI is an assessment tool that can be used repeatedly over a short period of time to identify children at risk and to inform intervention (i.e., progress monitoring). The student is assessed in picture naming, alliteration, and rhyming, areas which have been indicated in the literature as predictive of future reading success. The measures are timed and require the student to give an oral response (i.e., picture naming) or to point to a picture on the stimulus card (i.e., alliteration and rhyming). Teachers were trained in *IGDI* prior to commencement of the study and they administer it to all children in their classes on a monthly basis. For purposes

of this report, data (i.e., raw score and slope) are reported only for those students included in the sample.

The timeline for assessments for Years 1 and 2 of the grant is provided in the table below:

Fall 2005	Winter 2006 January	Spring 2006 June	Fall 2006 September	Winter 2007 January	Spring 2007 June
	<i>PPVT: 3</i> <i>TVIP</i> <i>PALS</i> <i>CIRCLE</i>	<i>PPVT: 3</i> <i>TVIP</i> <i>PALS</i> <i>CIRCLE</i>	<i>PPVT: 3</i> <i>TVIP</i> <i>PALS</i> <i>CIRCLE</i>	<i>CIRCLE</i>	<i>PPVT: 3</i> <i>TVIP</i> <i>PALS</i> <i>CIRCLE</i>

*Please note that the *IGDI* is administered on a monthly basis by the classroom teachers.

Key terms used throughout the report include:

- *Mean*: The mean is a measure of central tendency and is an average of the scores in the sample. The mean is equal to the sum of the scores divided by the number of scores.
- *Standard Deviation*: The standard deviation (SD) is a measure of statistical dispersion that indicates how “spread out” the values in a data set are. If the data points all are similar to the mean, the SD will be low; conversely, if many of the data points are substantially different from the mean, the SD will be high.
- *Change Score*: The change score is the difference between the pre- and posttest for any given measure. A positive number indicates that students achieved a higher score on the posttest than they did on the pretest, whereas a negative number indicates that students’ scores decreased on the posttest on average.
- *Effect Size*: The effect size indicates the magnitude of the difference between the two mean standard scores (i.e., pre- and posttest) for each measure and a larger effect size indicates a greater difference between mean scores. In fact, a small effect size is about 0.20 to 0.40, a medium effect size is 0.50 to 0.70, and anything above 0.80 can be considered a large effect size. Effect sizes less than 0.20 would be considered not to be different.

3. How have ERF children progressed on each of the outcome measures?

3.1 *Peabody Picture Vocabulary Test – III*

Sample	Pre-test Mean (Standard Deviation)	Posttest Mean (Standard deviation)	Change Score	Effect Size
Total sample	87.20 (14.13)	88.16 (15.15)	0.96	0.07
3-year olds	82.72 (16.62)	83.50 (18.36)	0.78	0.04
4-year olds	88.75 (12.98)	89.84 (13.64)	1.09	0.08
Girls	86.76 (15.43)	89.44 (15.67)	2.68	0.17
Boys	87.70 (12.73)	86.72 (14.64)	-0.98	-0.07
Primary language - English	92.02 (12.30)	90.95 (15.02)	-1.07	-0.08
Primary language - Spanish	78.52 (13.20)	83.04 (14.29)	4.52	0.33

When performance on the *Peabody Picture Vocabulary Test: III (PPVT: III)* is considered for the total sample of 70 students, the mean standard score on pretest (January 2006) was 87.20 (range=55 to 112; SD=14.13) and on posttest (June 2006) was 88.16 (range=49 to 127; SD=15.15). Standard scores ranging from 85 to 115 are considered to be in the low to high average range of performance for the *PPVT: III*. The mean scores for the total sample fell in the low average range for both the pre- and posttest assessment periods, when the sample students were compared to others in their age range in the normative sample. The respective standard deviations for each administration indicated that scores varied at expected levels. (Standard deviation of 15 would be expected). The change score, which is the difference between the pre- and posttest means, for the entire sample of students was 0.96 and the effect size was 0.07. This indicated that there was little difference in pre- and posttest mean scores on the *PPVT: III*, when the performance of all 70 students was considered. In other words, the students' performance on the test did not differ greatly from winter to spring.

When standard scores were considered by student age groups, the pretest mean score for the 3-year-olds was 82.72 (range=55 to 112; SD=16.62) and the posttest mean score was 83.50 (range=60 to 127; SD=18.36). Somewhat greater than expected variability in scores for 3-year-olds was found. The change scores and effect sizes for both groups indicated little difference in mean scores between the winter and spring administrations.

Scores for the *PPVT: III* also were considered by gender. These data showed somewhat larger increases for girls over boys, although effect sizes indicated non-meaningful changes from pre- to posttest periods. The negative numbers for boys indicated that the mean spring standard score was slightly lower than the mean winter score.

Finally, scores were considered by students' primary language. The *PPVT: III* was administered to all 70 students in the sample, regardless of whether primary language was English or Spanish. The students who spoke English as their primary language had mean scores that fell within the average range of performance. The change score (-1.07) and the effect size (-0.08) indicated little difference in mean scores between the English speaking students' performance in the spring, as compared to winter. The students who spoke Spanish as their primary language obtained a mean standard score of 78.52 (range=55 to 99; SD=13.20) on the winter pretest and a mean score of 83.04 (range=54 to 104; SD=14.29) on the spring posttest. Both of these mean scores fell in the moderately low range of performance, but the change score (4.52) and medium effect size (0.33) indicated that the native Spanish speaking students performed better at posttest as a group, than they had on the pretest.

The following table indicates the *percentage* of the current sample that is above a standard score of 85 (as requested by the government):

Sample	Pre-Test	Posttest
Total sample	60.00	57.40
3-year old participants	50.00	33.33
4-year old participants	63.50	66.00
Girls	59.50	61.10
Boys	60.60	53.10
Primary language - English	70.50	65.10
Primary language - Spanish	42.30	44.00

Overall, the percentages of children in the posttest that were above the minimum standard score required decreased when compared to those children who achieved this minimum during the time of pretest. In short, no effects from the curriculum were visible, including effects that may be contributed to children's growth.

3.2 Test de Vocabulario en Imagenes Peabody

Sample	Pre-Test M (SD)	Posttest M (SD)	Change Score	Effect Size
Total sample	84.28 (13.46)	79.65 (15.51)	-4.63	-0.32
3-year old participants	85.83 (8.84)	81.33 (17.39)	-4.50	-0.33
4-year old participants	84.61 (14.89)	79.06 (15.33)	-5.55	-0.37
Girls	79.33 (14.18)	76.55 (14.85)	-2.78	-0.19
Boys	90.50 (10.47)	82.50 (16.20)	-8.00	-0.59

The *Test de Vocabulario en Imagenes Peabody (TVIP)* was administered to students identified as English Language Learners (ELL) by their parents or caregivers at the time of enrollment.. The change score, which is the difference between the pre- and posttest means, for the sample of ELL students was -4.63 and the effect size was -0.32. The medium effect size suggested that there was some change in the mean standard score from pre- to posttest. The negative numbers indicated that the mean score decreased on

the spring administration. Similar outcomes were present for both 3- and 4- year old participants. Comparisons by gender found larger increases for boys over girls.

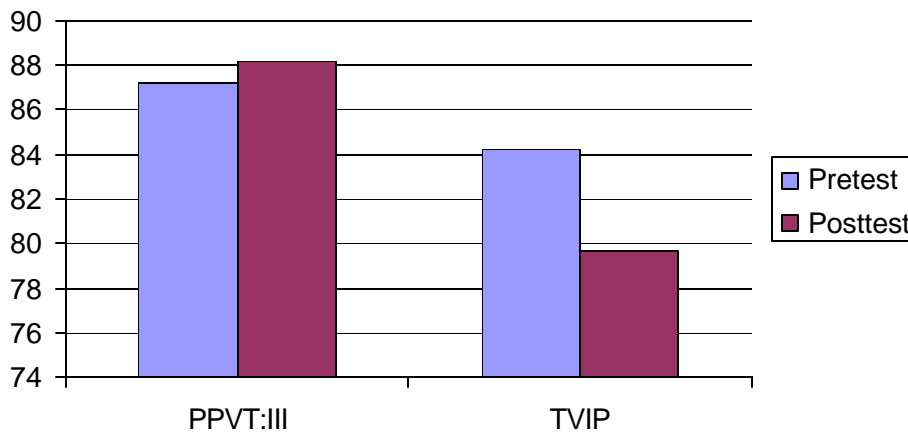
- Relationship to *PPVT-III* performance

Sample	PPVT/TVIP Correlations Pretest	PPVT/TVIP Correlations Posttest
Total sample	.182	.201
3-year old participants	-.418	-.675
4-year old participants	.358	.534*
Girls	-.188	.551
Boys	.349	-.134

* $p < .05$

When the relationship between student performance on the *PPVT: III* and *TVIP* was examined by student age, significant Pearson product-moment correlations were found for the 4-year-olds ($r = .534$) on posttest. This suggested a moderate relationship between standard scores achieved by the 4-year-old students on the tests administered in the spring. When standard scores were correlated by student gender (girls, $n = 11$; boys, $n = 14$) and for the entire sample of students taking both tests ($n = 25$), no significant correlations were found either at pretest or posttest. This indicated that there was not a strong relationship between student performance on the *PPVT: III* and *TVIP*.

PPVT: III & TVIP Pre- and Posttest Mean Standard Score Comparisons



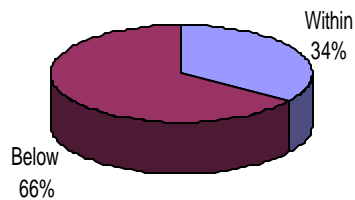
3.3 Phonological Awareness Literacy Screening

The *Phonological Awareness Literacy Screening: PreK (PALS)* was administered only to 4-year-old students who were identified as English speakers by their parents or primary care givers at the time of enrollment. The total number of students assessed with *PALS* in the winter was 32 and in the spring, 30. *Percentages* of students falling within developmental ranges (i.e., above developmental range, within developmental range, and below developmental range) are provided for each domain.

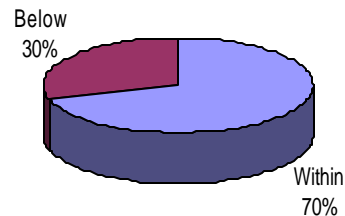
Name Writing

Sample	<u>PRE-TEST</u>		<u>POSTTEST</u>	
	Within DL	Below DL	Within DL	Below DL
4-year-olds	34.38	65.63	70	30
Girls	44.44	55.56	64.71	35.29
Boys	21.43	78.57	76.92	23.08

Name Writing Pretest



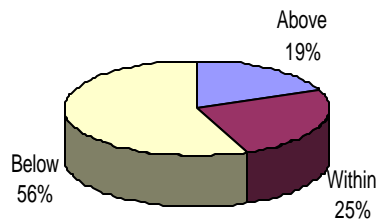
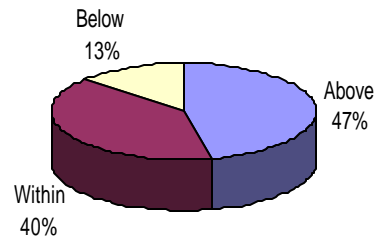
Name Writing Posttest



On the Name Writing measure, students are scored either as within or below level. Unlike the other *PALS* measures, students cannot score above the expected developmental level. From winter pretest to spring posttest, the number of students scoring within level on the Name Writing measure showed increases in the desired direction, with boys' performance on this subtest showing the greatest increase.

*Upper Case Alphabet Recognition***PRE-TEST****POSTTEST**

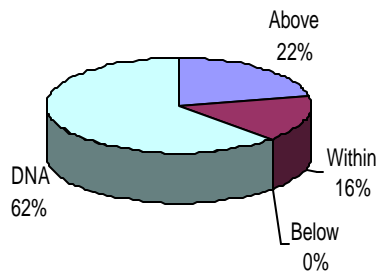
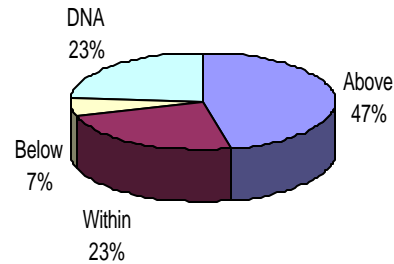
Sample	Above DL	Within DL	Below DL	Above DL	Within DL	Below DL
4-year-olds	18.75	25.00	56.25	46.67	40.00	13.33
Girls	11.11	38.89	50.00	52.94	47.06	0
Boys	28.57	7.14	64.29	38.46	30.77	30.77

Upper Case Pretest**Upper Case Posttest**

The Upper Case Alphabet Recognition measure requires the student to name upper case letters. The letters are presented on a page in random order and the test is untimed. From pretest to posttest, 4-year-old students who were within or above developmental expectations on the Upper Case Alphabet Recognition measure increased from 43.75% to 86.67%. Both boys and girls showed substantial gains on this measure, with girls showing the greatest increase.

*Lower Case Alphabet Recognition***PRE-TEST****POSTTEST**

Sample	Above DL	Within DL	Below DL	Not Administered	Above DL	Within DL	Below DL	Not Administered
4-year-olds	21.88	15.63	0	62.50	46.67	23.34	6.67	23.34
Girls	16.67	22.22	0	61.11	50	33.33	0	16.67
Boys	28.57	7.14	0	64.29	38.46	7.69	15.38	38.46

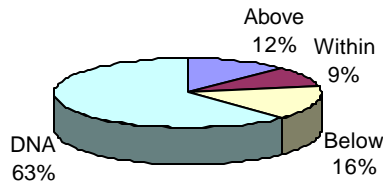
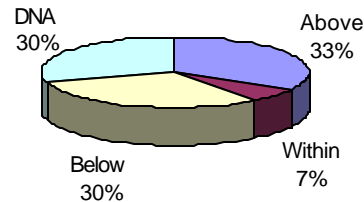
Lower Case Pretest**Lower Case Posttest**

The Lower Case Alphabet Recognition measure requires the student to name lower case letters. The letters are presented on a page in random order and the test is untimed. To be administered the Lower Case Alphabet Recognition measure, the student must have scored 16 or more correct out of 26 possible responses on the Upper Case Alphabet Recognition measure. The percentages recorded in the “not administered” column reflect those students who did not meet criteria for administration.

At winter pretest, more than half (62.50%) of the 4-year-old students did not meet criteria for assessment on the Lower Case Alphabet Recognition measure. In the spring, the percentage of students who did not meet criteria was reduced to about one quarter (23.34%). Overall, substantial increases were found from pretest to posttest, with the largest increases noted among the girls.

*Letter Sounds***PRE-TEST****POSTTEST**

Sample	Above DL	Within DL	Below DL	Not Administered	Above DL	Within DL	Below DL	Not Administered
4-year-olds	12.50	9.38	15.63	62.50	33.33	6.67	30	30
Girls	0	11.11	27.78	61.11	29.40	11.80	47.10	11.80
Boys	28.57	7.14	0	64.29	38.46	0	7.69	53.85

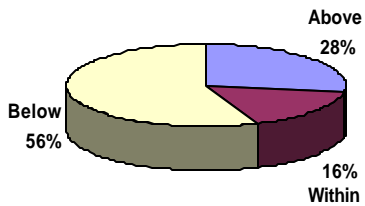
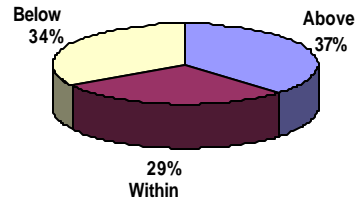
Letter Sounds Pretest**Letter Sounds Posttest**

The Letter Sounds measure requires the student to produce the phoneme associated with a printed grapheme. The letters are presented on a page in random order and the test is untimed. To be administered the Letter Sounds measure, the student must have scored 9 or more correct out of 26 possible responses on the Lower Case Alphabet Recognition measure. The percentages recorded in the “not administered” column reflect those students who did not meet criteria for administration.

Overall improvement in Letter Sounds was noted from pre- to posttest with a total increase from 21.88% to 40% of 4-year-old students within or above the developmental level on the Letter Sounds measure. In addition, substantial decreases of 33% of students not administered the measure from pre- to posttest were found. The greatest improvements were noted for girls. Given the large number of students who did not achieve the criteria for administration, students may benefit from more intensive instruction in identification of lower case letters, as well as letter sounds.

*Beginning Sound Awareness***PRE-TEST****POSTTEST**

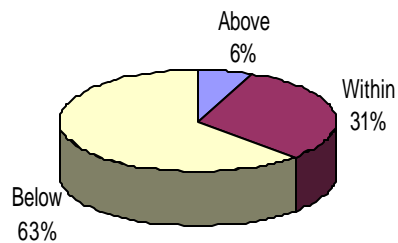
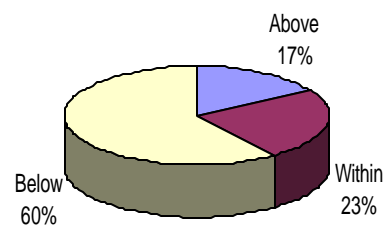
Sample	Above DL	Within DL	Below DL	Above DL	Within DL	Below DL
4-year-olds	28.13	15.63	56.25	36.67	30.00	33.33
Girls	16.67	22.22	61.11	35.29	29.41	35.29
Boys	42.86	7.14	50	38.46	30.77	30.77

Beginning Sounds Pretest**Beginning Sounds Posttest**

The Beginning Sound Awareness task requires the student to match pictures whose names begin with the same initial sound (/m/, /s/, and /b/). At pretest, 43.76% of the 4-year-old students were within or above developmental level and, at posttest, that percentage increased to 66.67%. The boys and girls performed equally well on this measure at posttest, but 33.33% of students remained below developmental level overall, suggesting another area requiring attention.

*Print and Word Awareness***PRE-TEST****POSTTEST**

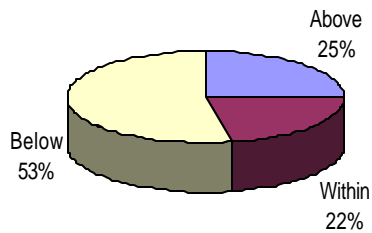
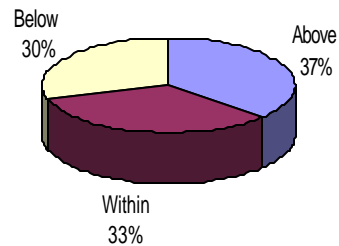
Sample	Above DL	Within DL	Below DL	Above DL	Within DL	Below DL
4-year-olds	6.25	31.25	62.50	16.67	23.33	60
Girls	11.11	27.78	61.11	5.88	35.29	58.82
Boys	0	35.71	64.29	30.77	7.69	61.54

Print Pretest**Print Posttest**

The Print and Word Awareness measure requires the student to locate print on the page of a nursery rhyme book as indicated by the administrator. For instance, the student is asked to find two words in the title of the book that are the same and to point to the “smallest” word. As a group, the 4-year-olds did not improve performance on this measure from winter pretest to spring posttest, and both boys and girls performed similarly. This is an area requiring further intervention within the classrooms.

*Rhyme Awareness***PRE-TEST****POSTTEST**

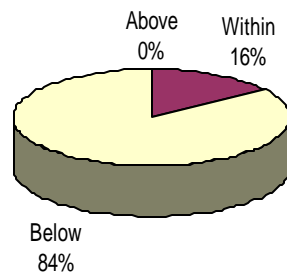
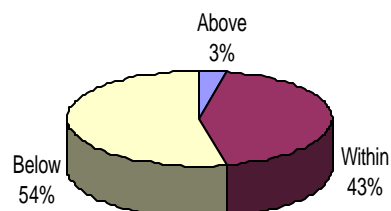
Sample	Above DL	Within DL	Below DL	Above DL	Within DL	Below DL
4-year-olds	25	21.88	53.13	36.67	33.33	30
Girls	22.22	16.67	61.11	29.41	35.29	35.29
Boys	28.57	28.57	42.86	46.15	30.77	23.08

Rhyme Pretest**Rhyme Posttest**

The Rhyme Awareness measure requires the student to choose a rhyming match for a target picture from three possible choices. On the winter pretest, more than half of the 4-year-old students (i.e., 53.13%) were below the developmental level; this percentage decreased to less than one third (i.e., 30%) on the spring posttest. At the time of the posttest administration, the boys outperformed the girls in this area. While 76.92% of the boys scored within or above the expected developmental range, only 64.70% of the girls scored within or above level.

*Nursery Rhyme Awareness***PRE-TEST****POSTTEST**

Sample	Above DL	Within DL	Below DL	Above DL	Within DL	Below DL
4-year-olds	0	15.63	84.38	3.33	43.33	53.33
Girls	0	16.67	83.33	5.88	41.18	52.94
Boys	0	14.29	85.71	0	46.15	53.85

Nursery Rhyme Pretest**Nursery Rhyme Posttest**

The Nursery Rhyme Awareness measure requires the student to supply the missing rhyming word for popular nursery rhymes (e.g., “Jack and Jill” and “Jack Be Nimble”) that are presented orally by the examiner. On pretest, 84.38% of the students scored below developmental level, with equal numbers of boys and girls scoring below the expected range. On posttest, 53.33% of the students remained below level. Once again, the boys and girls performed similarly with almost equal numbers scoring below the developmental level. Students would benefit from classroom instruction in this area.

- Comparison of change (based on raw scores) within each PALS domain

One way of examining the level of improvement over time is to compare changes in raw scores from pre- to posttest. The amount of change is reflected in the effect size, with effect sizes of 0.20 to 0.40 equal to small effects, 0.50 to 0.70, considered moderate, and 0.80 or greater, considered large.

Each of the tables below displays this analysis with a summary chart provided across the subtests.

Name Writing

Sample	Pre-test Mean (Standard Deviation)	Posttest Mean (Standard deviation)	Change Score	Effect Size
Total sample	3.16 (2.76)	5.17 (2.47)	2.01	0.76
Girls	3.94 (2.55)	5.06 (2.19)	1.12	0.47
Boys	2.14 (2.77)	5.31 (2.87)	3.17	1.12

Upper Case Alphabet Recognition

Sample	Pre-test Mean (Standard Deviation)	Posttest Mean (Standard deviation)	Change Score	Effect Size
Total sample	12.63 (8.81)	18.93 (7.83)	6.30	0.76
Girls	13.72 (7.23)	21.35 (4.37)	7.63	1.28
Boys	11.21 (10.63)	15.77 (10.18)	4.56	0.44

Lower Case Alphabet Recognition

Sample	Pre-test Mean (Standard Deviation)	Posttest Mean (Standard deviation)	Change Score	Effect Size
Total sample	18.58 (4.93)	18.57 (5.89)	-0.01	-0.002
Girls	16.00 (4.40)	18.87 (5.06)	2.87	0.61
Boys	22.20 (3.11)	18.00 (7.58)	-4.20	-0.72

Letter Sounds

Sample	Pre-test Mean (Standard Deviation)	Posttest Mean (Standard deviation)	Change Score	Effect Size
Total sample	6.33 (7.35)	8.29 (7.11)	1.96	0.27
Girls	1.57 (2.82)	5.60 (5.08)	4.03	.98
Boys	13.00 (6.44)	15.00 (7.35)	2.00	0.29

Beginning Sound Awareness

Sample	Pre-test Mean (Standard Deviation)	Posttest Mean (Standard deviation)	Change Score	Effect Size
Total sample	4.16 (4.14)	5.97 (3.78)	1.81	0.46
Girls	3.50 (3.81)	6.00 (3.61)	2.50	0.67
Boys	5.00 (4.54)	5.92 (4.15)	0.92	0.21

Print and Word Awareness

Sample	Pre-test Mean (Standard Deviation)	Posttest Mean (Standard deviation)	Change Score	Effect Size
Total sample	5.19 (3.07)	5.93 (2.99)	0.74	0.24
Girls	4.94 (3.33)	5.94 (2.73)	1.00	0.33
Boys	5.50 (2.79)	5.92 (3.43)	0.42	0.13

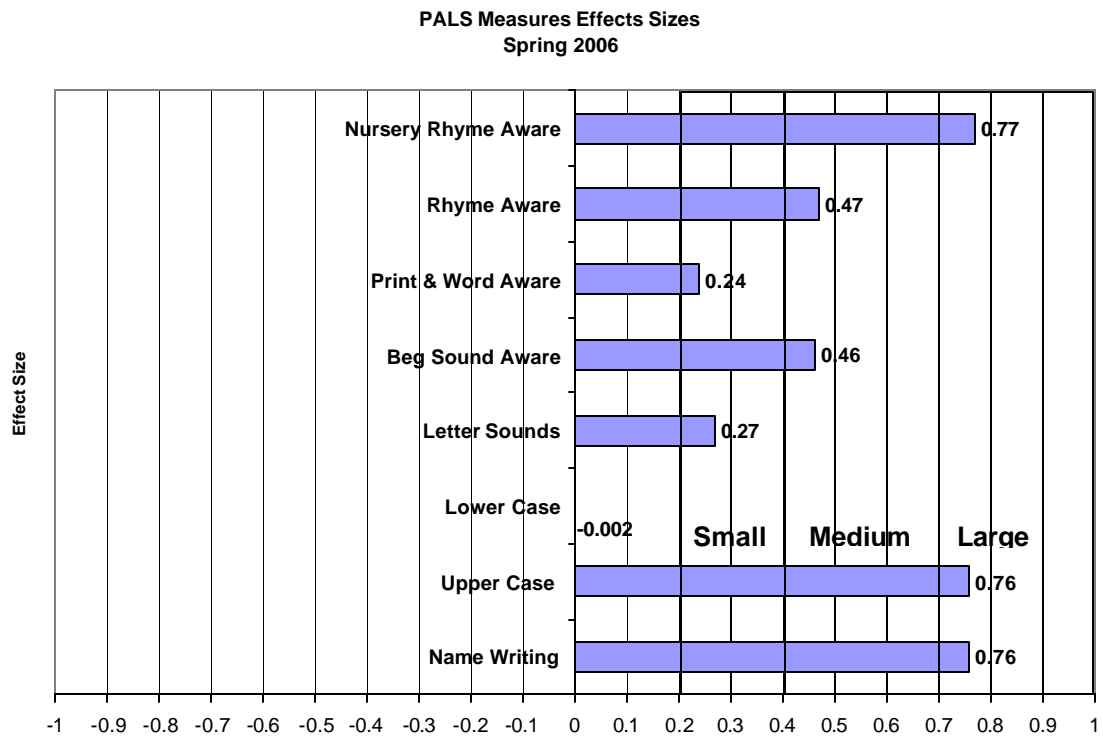
Rhyme Awareness

Sample	Pre-test Mean (Standard Deviation)	Posttest Mean (Standard deviation)	Change Score	Effect Size
Total sample	4.59 (3.56)	6.17 (3.19)	1.58	0.47
Girls	4.06 (3.61)	5.71 (3.37)	1.65	0.47
Boys	5.29 (3.50)	6.77 (2.95)	1.48	0.46

Nursery Rhyme Awareness

Sample	Pre-test Mean (Standard Deviation)	Posttest Mean (Standard deviation)	Change Score	Effect Size
Total sample	3.38 (2.18)	5.10 (2.31)	1.72	0.77
Girls	3.61 (2.38)	5.12 (2.11)	1.51	0.67
Boys	3.07 (1.94)	5.08 (2.63)	2.01	0.87

Summary of Effect Sizes Across Measures



As seen in the chart above, 4-year-olds made moderate to large improvements in Nursery Rhyme, Rhyme Awareness, Beginning Sound Awareness, Upper Case Letters, and Name Writing. Smaller increases were noted in Print & Word Awareness and Letter Sounds. The area still in need of improvement was Lower Case letter recognition.

3.4 Individual Growth and Development Indicators (IGDI)

The following table contains the aggregated slope for each IGDI measure across groups.

Sample	Picture Naming	Alliteration	Rhyming
Total sample	0.27	0.74	1.02
3-year-olds	-0.21	0.25	0.72
4-year-olds	0.44	0.84	1.10
Girls	-0.15	0.51	1.24
Boys	0.74	1.05	0.77
Primary language English	.30	.80	1.01
Primary language Spanish	.20	.61	1.06

The aggregated slope indicates the average trend in score gains for students across *IGDI* measures. Negative numbers indicate areas in which the trend line for a particular group was on a downward slope. As a group, the students made substantial gains on the *IGDI* measures from March through June.

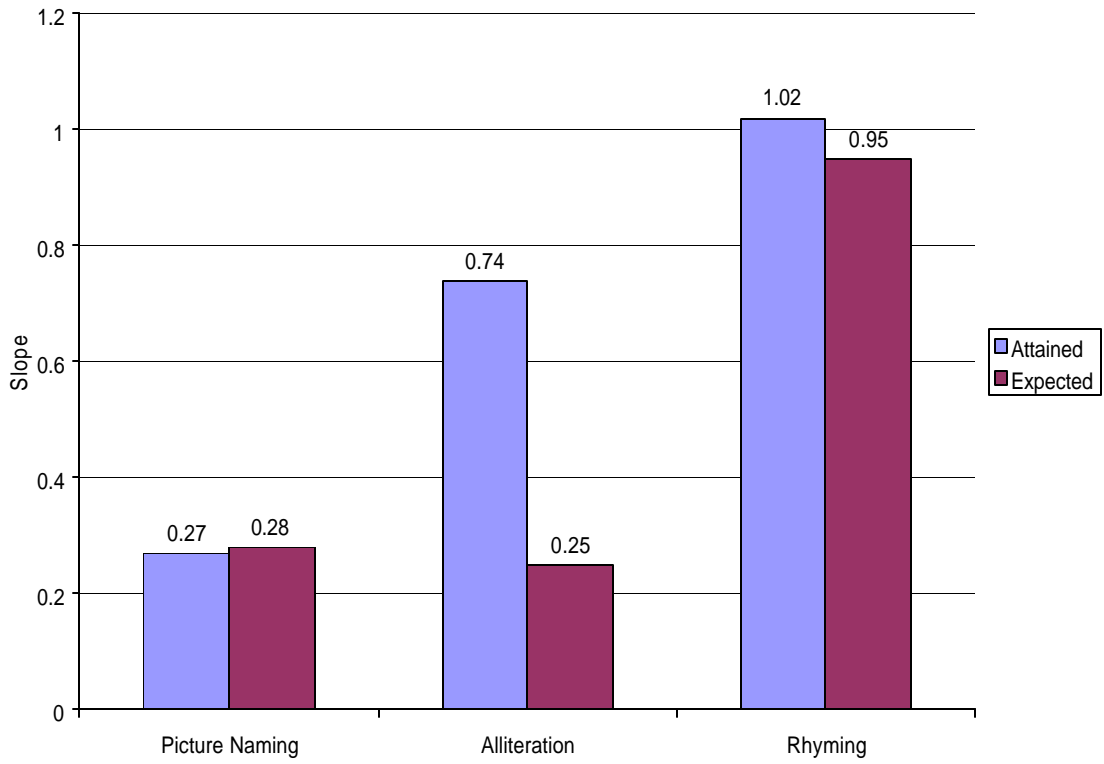
According to the *IGDI* technical report, the average slope for Picture Naming for low income students, at 66 months of age, is 0.28 pictures per month. As a whole, students performed at the expected level on this measure. The 4-year-old students and the boys made notable progress in this area, with aggregated slopes exceeding expectations for Picture Naming.

Average slope for Alliteration, centered at 53 months of age, for low income students is 0.25 alliterations per month. The aggregated slope for all students was 0.74, which greatly exceeded expected slope. The 3-year-old students were on target with a slope of 0.25 alliterations per month, while all other groups exceeded the expected levels.

Finally, average slope for Rhyming, centered at 53 months of age, for low income children, is 0.95 rhymes per month. Once again the total sample exceeded this expected level with an aggregated slope of 1.02.

The chart on the following page illustrates the comparison between expected and attained slopes on all *IGDI* measures.

IGDI Slopes



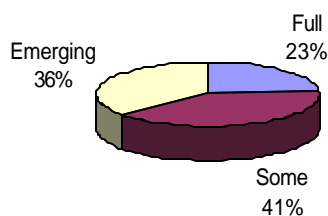
3.5 CIRCLE

The *CIRCLE* assessment was administered to all 70 students in their native language. The following tables contain *percentages* of children performing within each developmental level.

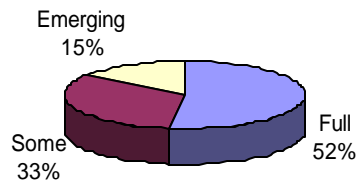
Letter Naming

Sample	Time 1			Time 2			Time 3		
	Full	Some	Emerging	Full	Some	Emerging	Full	Some	Emerging
Total sample				23.1 9	40.58	36.23	52.2 4	32.84	14.93
3-year-olds				22.2 2	33.33	44.44	38.8 9	44.44	16.67
4-year-olds				23.5 3	43.13	33.33	57.1	28.6	14.29
Girls				24.3 2	48.65	27.03	60	37.14	2.86
Boys				21.8 8	31.25	46.88	43.7 5	28.13	28.13
Primary language English				37.2 1	44.19	18.60	61.9 0	27.91	9.30
Primary language Spanish				0	34.62	65.38	36	40	24

Letter Naming Pretest



Letter Naming Posttest



The Letter Naming measure requires the student to name rapidly upper and lower case alphabet letters in one minute's time. The letters are presented in random order on a flip chart and the student's response is recorded as "correct" or "incorrect" by the examiner via a personal digital assistant (PDA). The student's score is the number of correct responses. On the January pretest, 23.19% of the students achieved scores that fell in the full understanding range. On the June posttest, this percentage increased to 52.24%

of students. Likewise, in January 36.23% of students placed in the emerging range and this percentage decreased to 14.93% in June.

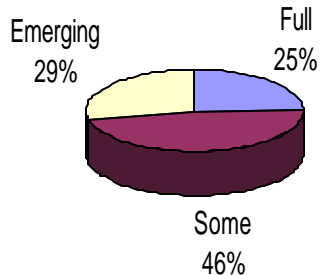
At the time of the January pretest, 55.55% of the 3-year-old students performed in the some to full understanding ranges and this percentage increased to 83.33% on the June posttest. When the 4-year-old students are considered, 66.66% placed in the some to full understanding range of performance on the January pretest and this percentage increased to 85.7% on the June posttest. When students were considered by gender, both boys and girls showed substantial improvements, with girls showing the greatest gains.

Finally, when student performance is considered by primary language group membership, students who were considered native English speakers increased the percentage performing in the full understanding range from 37.21% at January pretest to 61.90% at June posttest. At posttest, 9.30% of the English speaking students performed in the emerging range. The students who were assessed in Spanish made great gains in this area, as well, but it should be noted that on pretest, students were assessed with the Spanish alphabet and on posttest, with the English alphabet. The decision was made to change format due to the fact that most of the children had not been exposed to the Spanish alphabet either at home or in school and were being taught the English alphabet in the classroom. On pretest, 0% of the Spanish speaking students performed in the full understanding range and more than half (i.e., 65.38%) placed in the emerging range. On posttest, the percentage that performed in the emerging range decreased to 24%. In sum, rapid letter naming was an area where students made gains across groups.

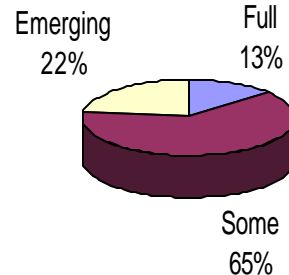
Vocabulary

Sample	Time 1			Time 2			Time 3		
	Full	Some	Emerging	Full	Some	Emerging	Full	Some	Emerging
Total sample				24.64	46.38	28.99	13.43	64.18	22.39
3-year-olds				16.67	55.56	27.78	11.11	66.67	22.22
4-year-olds				27.45	43.14	29.41	14.29	63.27	22.45
Girls				29.73	45.95	24.32	17.14	62.86	20
Boys				18.75	46.88	34.38	9.38	65.63	25
Primary language English				37.21	62.79	0	21.43	69.05	9.52
Primary language Spanish				4	20	76	0	56	44

Vocabulary Pretest



Vocabulary Posttest



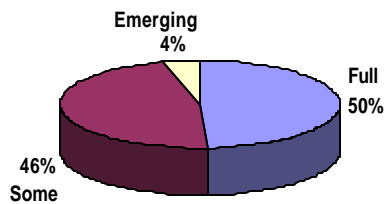
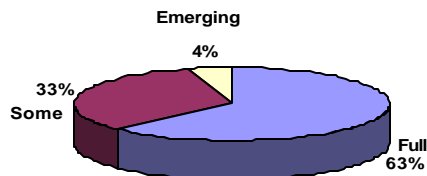
The Vocabulary measure, like Letter Naming, is a timed test. Pictures of objects are presented in a flip chart and the student is asked to name as many pictures as s/he can in one minute's time. The examiner records responses as correct or incorrect on the PDA. The Vocabulary measure also allows the examiner to label an incorrect response as a "sensible error," if appropriate. A sensible error is a response that is incorrect, yet makes sense. For example, on pretest many of the students labeled the picture of the "apron" as a "skirt." Although the answer was incorrect, it was a "sensible" error. Students assessed in Spanish were presented with the same pictures, but were given directions in their native language. The student's score is the number of correct responses, not including those deemed sensible errors.

On January pretest, 24.64% of the students performed in the full understanding range and on posttest, 13.43% of students performed in this range. The majority (i.e., 46.38% on pretest and 64.18% on posttest) achieved scores that fell in the some understanding range of performance, but since nearly one-fourth of students placed in the emerging range on posttest, this suggests an area for intervention.

Although percentages of students in the some understanding range consistently increased across groups (i.e., age, gender, and primary language), percentages of students performing in the full understanding range consistently declined across the groups. When these percentages are considered with student performance on the *Peabody Picture Vocabulary Test: Third Edition (PPVT: III)*, an area of need is indicated.

Listening

Sample	Time 1			Time 2			Time 3		
	Full	Some	Emerging	Full	Some	Emerging	Full	Some	Emerging
Total sample				49.28	46.38	4.35	62.7	32.8	4.48
3-year-olds				33.33	61.11	5.56	50	44.44	5.56
4-year-olds				54.90	41.18	3.92	67.35	28.57	4.08
Girls				56.76	37.84	5.41	65.71	25.71	8.57
Boys				40.63	56.25	3.13	59.38	40.63	0
Primary language English				60.47	37.21	2.33	76.19	21.43	2.38
Primary language Spanish				30.77	61.54	7.69	40	52	8

Listening Pretest**Listening Posttest**

For the Listening measure, the student is presented with two words orally (e.g., book-book or stop-go) and asked to indicate whether the words or the same or not by responding “yes” or “no.” The students who were assessed in Spanish were presented with directions for the assessment in Spanish and with sets of Spanish words. The administrator indicates correct and incorrect responses on the PDA. This measure is untimed.

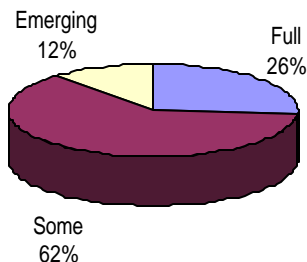
The full sample of children performed quite well on this measure, with 62.7% and 32.8% placing in the full understanding and some understanding ranges respectively at posttest. Some differences in percentages were noted across age groups, with the 4-year-

old students having a greater percentage in the full understanding range, than the 3-year-olds at both pre- and posttest. The greatest difference in performance across groups was evidenced when primary language of the student was considered. The majority of English speaking students placed in the full understanding range on January pre- and June posttest, whereas more than half of the Spanish speaking students placed in the some understanding range on each administration.

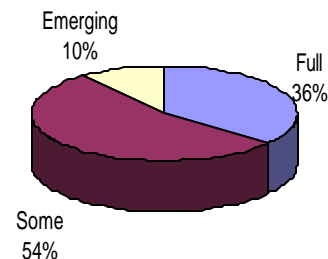
Rhyming 1

Sample	Time 1			Time 2			Time 3		
	Full	Some	Emerging	Full	Some	Emerging	Full	Some	Emerging
Total sample				26.09	62.32	11.59	35.82	53.73	10.45
3-year-olds				16.67	66.67	16.67	16.67	72.22	11.11
4-year-olds				29.41	60.80	9.80	42.86	46.94	10.20
Girls				29.73	64.86	5.41	42.88	45.71	11.43
Boys				21.88	59.38	18.75	28.13	62.50	9.38
Primary language English				23.26	67.44	9.30	42.86	52.38	4.76
Primary language Spanish				30.77	53.85	15.38	24	56	20

Rhyming 1 Pretest



Rhyming 1 Posttest



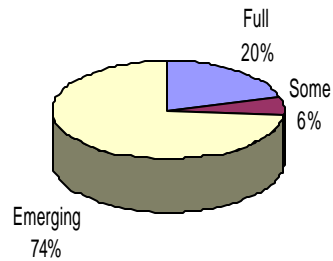
The Rhyming 1 assessment requires the student to decide whether two words presented orally rhyme or not. As with the Listening measure, students assessed in Spanish were presented with directions and sets of words in Spanish. The student is asked to respond “yes” or “no” to the oral prompt and the administrator enters the student’s

response into the PDA. The measure is not timed. When the full sample of 70 students is considered, students increased performance on this measure from January pretest to June posttest. In January, 49.28% of the students and 46.38% of the students placed in the full and some understanding ranges of performance respectively. On posttest in June, the percentage of students placing in the full understanding range increased to 65.67%, while the percentage in the some understanding range decreased to 29.85%, indicating that students made gains in this area. When groups of students were considered across age, gender, and primary language, all groups demonstrated an increase in performance with percentages of students performing in the some understanding range decreasing, while the percentages of students placing in the full understanding range increased.

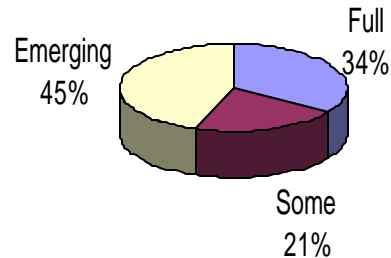
Rhyming 2

Sample	Time 1			Time 2			Time 3		
	Full	Some	Emerging	Full	Some	Emerging	Full	Some	Emerging
Total sample				20.29	5.80	73.91	34.33	20.90	44.78
3-year-olds				0	0	100	16.67	5.56	77.78
4-year-olds				27.45	7.84	64.71	40.82	26.53	32.65
Girls				18.92	8.11	72.97	42.86	14.29	42.86
Boys				3.13	3.13	93.75	25	28.13	46.88
Primary language English				27.91	4.65	67.44	42.86	14.29	42.86
Primary language Spanish				7.69	7.69	84.62	20	32	48

Rhyming 2 Pretest



Rhyming 2 Posttest

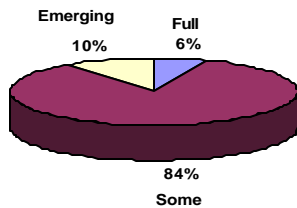
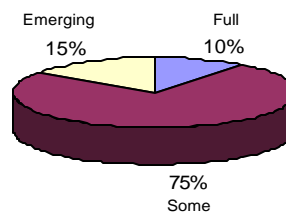


For the Rhyming 2 measure, students are presented with a word orally and are asked to provide a rhyme for that word. The student response may be a nonsense word. As with the other measures, Spanish-speaking students were given directions and words in Spanish. The measure is not timed and the administrator records the student response as correct or incorrect in the PDA.

This task proved to be more difficult for the students than did the Rhyming 1 task. On January pretest, nearly three-fourths (i.e., 73.91%) of the students placed in the emerging range. This percentage decreased to 44.78% on June posttest, but only 34.33% of the 70 students placed in the full understanding range on the Rhyming 2 measure in June. When student performance is considered across age, gender, and primary language groups, all groups made progress on this measure from pretest to posttest. Because the 4-year-olds had greater percentages of students in the full and some understanding ranges on both pre- and posttest than did the 3-year-olds, this may be reflective of developmental differences in the students.

Alliteration

Sample	Time 1			Time 2			Time 3		
	Full	Some	Emerging	Full	Some	Emerging	Full	Some	Emerging
Total sample				5.80	84.06	10.14	10.4	74.6	15.0
3-year-olds				5.56	83.33	11.11	5.56	72.22	22.22
4-year-olds				5.88	84.31	9.80	12.24	75.51	12.24
Girls				8.11	81.08	10.81	11.43	74.29	14.29
Boys				3.13	87.50	9.38	9.38	75	15.63
Primary language English				6.98	79.07	13.95	11.90	73.81	14.29
Primary language Spanish				3.85	92.31	3.85	8	76	16

Alliteration Pretest**Alliteration Posttest**

The Alliteration measure requires the student to decide whether or not two words presented orally begin with the same sound by indicating “yes” or “no.” The student is not required to provide the corresponding letter names. Students assessed with the Spanish measure were provided both directions and sets of words in Spanish. Student responses were recorded on the PDA by the examiner as either correct or incorrect and the student’s score was the number of correct responses.

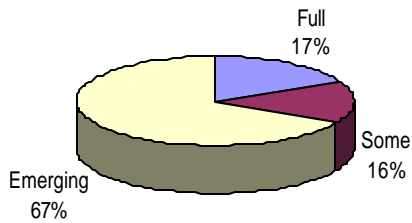
When the sample of 70 students was considered as a whole, the majority performed in the some understanding range on both pre- (84.06%) and on posttest (74.6%). The 4-year-old students performed better than the 3-year-old students during both assessment periods. On posttest, 67.35% of the 4-year-olds performed in the full or some understanding range, compared to 22.23% of the 3-year-olds and this difference may be reflective of developmental patterns. On posttest, a greater percentage of girls (42.86%) achieved scores that fell in the full understanding range than did the boys

(25%), but nearly equal percentages of girls and boys scored within the emerging understanding range (i.e., 42.86% of girls and 46.88% of boys). Further, when students are compared by primary language, on posttest a greater percentage of English speaking students (42.86%) placed in the full understanding range than did the Spanish speaking students (20%). However, similar percentages of students fell within the emerging understanding range on posttest (42.86% of English speaking students and 48% of Spanish speaking students).

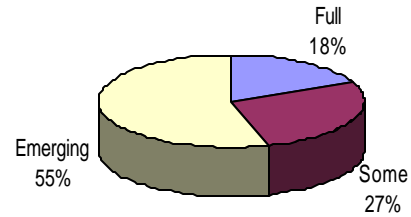
Syllabication

Sample	Time 1			Time 2			Time 3		
	Full	Some	Emerging	Full	Some	Emerging	Full	Some	Emerging
Total sample				17.39	15.94	66.67	17.91	26.87	55.2
3-year-olds				5.56	66.67	27.78	0	61.11	38.89
4-year-olds				9.80	54.90	35.29	26.53	55.10	18.37
Girls				13.51	51.35	35.14	20	51.43	28.57
Boys				3.13	65.63	31.25	18.75	62.50	18.75
Primary language English				11.63	58.14	30.23	16.67	66.67	16.67
Primary language Spanish				3.85	57.69	38.46	24	40	36

Syllabication Pretest



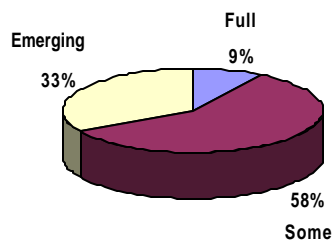
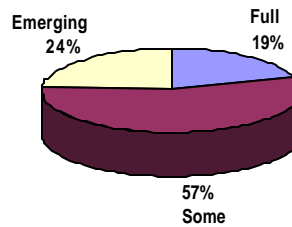
Syllabication Posttest



For the Syllabication measure, the student is presented with a word orally and is asked to repeat that word while clapping the syllables. The administrator first models what is expected by clapping the syllables for the word “cowboy” and then allows the student opportunity to practice. This measure is not timed. The student’s score is the number of correct responses given. When the sample of 70 students is considered, there was little change between performance on the January pretest and June posttest. The majority of the students remained in the emerging range of performance.

Sentences

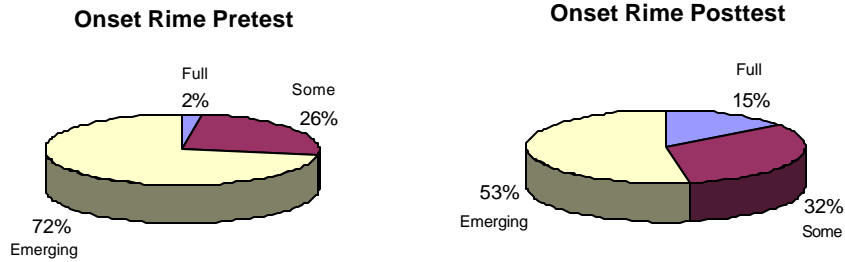
Sample	Time 1			Time 2			Time 3		
	Full	Some	Emerging	Full	Some	Emerging	Full	Some	Emerging
Total sample				8.7	58.0	33.3	19.4	56.7	23.9
3-year-olds				5.56	16.67	77.78	5.56	27.78	66.67
4-year-olds				21.57	15.69	62.75	22.45	26.53	51.02
Girls				18.92	18.92	62.16	20	31.43	48.57
Boys				15.63	12.50	71.88	15.63	21.88	62.50
Primary language English				23.26	13.95	62.79	28.57	30.95	40.48
Primary language Spanish				7.69	19.23	73.08	0	20	80

Sentences Pretest**Sentences Posttest**

The Sentences measure requires the student to listen to a sentence and then repeat it while moving counters to indicate the number of words in the sentence. This task is first modeled by the administrator and then the student is granted time to practice. The task is untimed and student responses are recorded as correct or incorrect. When performance of the 70 students is considered as a whole, there was little change in the percentages of students performing in the three developmental ranges between January pretest and June posttest.

Onset Rime

Sample	Time 1			Time 2			Time 3		
	Full	Some	Emerging	Full	Some	Emerging	Full	Some	Emerging
Total sample				2.3	25.6	72.1	15.0	32.6	53.4
3-year-olds				0	8.3	91.7	0	33.3	66.7
4-year-olds				3.2	32.3	64.5	19.4	32.2	48.4
Girls				0	28.0	72.0	12.5	37.5	50.0
Boys				5.6	22.2	72.2	15.8	26.3	57.9



The Onset Rime measure requires the student to determine the word made from parts spoken by the administrator. The Spanish version of *CIRCLE* does not include the Onset Rime test and, thus, students who were tested in Spanish were not assessed with this measure. When performance of the total sample of students is considered as a whole, there was a slight increase in the percentage of students falling in the full and some understanding ranges from January pretest to June posttest. The percentage of students in the full understanding range increased from 2.3% to 15.0% and the percentage in the some understanding range increased from 25.6% to 32.6%.

- Comparison of change (based on raw scores) within each *CIRCLE* measure

One way of examining the level of improvement over time is to compare changes in raw scores from pre- to posttest. The amount of change is reflected in the effect size, with effect sizes of 0.20 to 0.40 equal to small effects, 0.50 to 0.70 considered moderate, and 0.80 or greater considered large.

Each of the tables below displays this analysis with a summary chart provided across the subtests.

Letter Naming

Sample	Change Score	Effect Size
Total sample	6.94	0.68
3-year old participants	5.50	0.61
4-year old participants	7.49	0.72
Girls	9.54	1.19
Boys	5.75	0.49
Primary language - English	4.96	0.50
Primary language - Spanish	10.18	1.13

Listening

Sample	Change Score	Effect Size
Total sample	0.32	0.24
3-year old participants	0.50	0.35
4-year old participants	0.25	0.19
Girls	0.11	0.08
Boys	0.40	0.30
Primary language - English	0.40	0.31
Primary language - Spanish	0.16	0.12

Vocabulary

Sample	Change Score	Effect Size
Total sample	-0.82	-0.12
3-year old participants	-1.00	-0.16
4-year old participants	-0.74	-0.10
Girls	-1.00	-0.14
Boys	-0.28	-0.04
Primary language - English	-3.09	-0.59
Primary language - Spanish	2.87	0.66

Rhyming 1

Sample	Change Score	Effect Size
Total sample	0.59	0.32
3-year old participants	0.66	0.39
4-year old participants	0.57	0.31
Girls	0.85	0.49
Boys	0.59	0.32
Primary language - English	0.81	0.47
Primary language - Spanish	0.21	0.11

Rhyming 2

Sample	Change Score	Effect Size
Total sample	0.99	0.51
3-year old participants	1.05	0.91
4-year old participants	0.98	0.50
Girls	1.21	0.62
Boys	0.68	0.43
Primary language - English	0.95	0.46
Primary language - Spanish	1.07	0.69

Syllabication

Sample	Change Score	Effect Size
Total sample	0.99	0.55
3-year old participants	0.22	0.16
4-year old participants	1.28	0.68
Girls	1.22	0.55
Boys	1.12	0.74
Primary language - English	1.00	0.54
Primary language - Spanish	0.99	0.57

Alliteration

Sample	Change Score	Effect Size
Total sample	0.20	0.15
3-year old participants	0.17	0.12
4-year old participants	0.22	0.14
Girls	0.33	0.24
Boys	0.09	0.07
Primary language - English	0.32	0.23
Primary language - Spanish	0	0

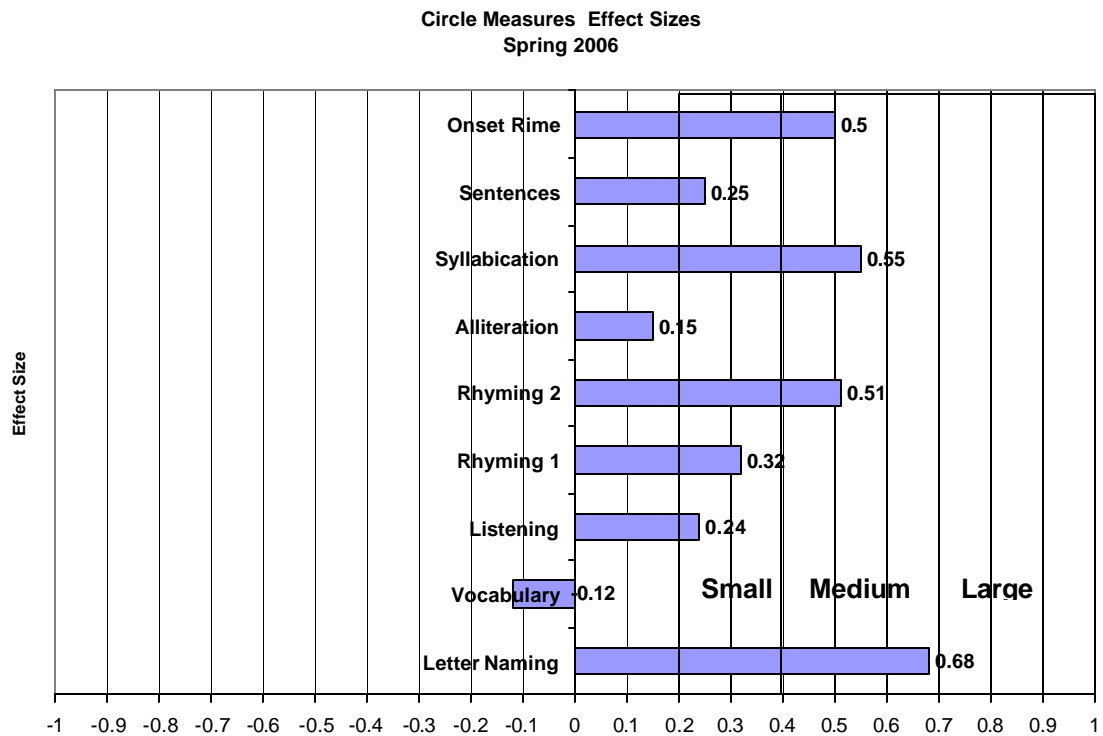
Sentences

Sample	Change Score	Effect Size
Total sample	0.40	0.25
3-year old participants	0.22	0.17
4-year old participants	0.47	0.28
Girls	0.60	0.35
Boys	0.31	0.19
Primary language - English	0.67	0.37
Primary language - Spanish	-0.08	-0.07

Onset Rime

Sample	Change Score	Effect Size
Total sample	0.74	.50
3-year old participants	0.75	.64
4-year old participants	0.74	.48
Girls	0.76	.49
Boys	0.53	0.35

- Summary of Effect Sizes Across Measures



As seen in the chart above, moderate improvements in Letter Naming, Onset Rime, Syllabication, and Rhyming 2 were evident. Smaller increases were noted in Sentences, Rhyming 1, and Listening. The areas still needing improvement are Vocabulary and Alliteration.