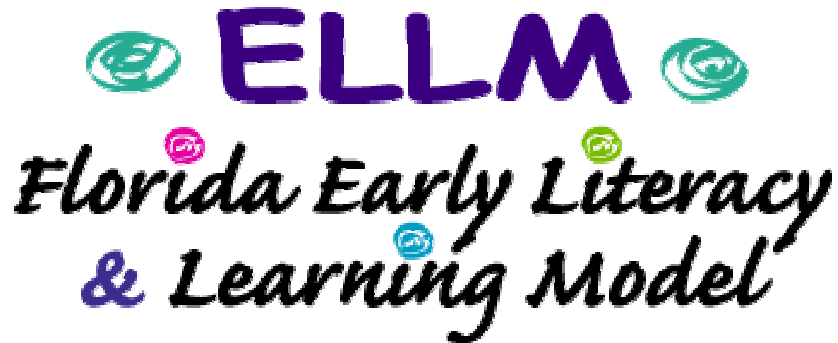


Hillsborough County



Program Evaluation Report (Year 2) 2003-2004

February, 2006

Author:

Christine M. Caffray, Ph.D.

Prepared by:

The Children's Board of Hillsborough County
1002 East Palm Avenue
Tampa, FL 33605

For more information contact:

Christine M. Caffray, Ph.D.
813-229-2884
christine@childrensboard.org

Recommended Citation:

Caffray, C. M., (2006). Hillsborough County Early Literacy and Learning Model (ELLM): Program Evaluation Report, 2003-2004. Tampa: FL. The Children's Board of Hillsborough County.

Permission to copy all or portions of this report is granted as long as this report and The Children's Board of Hillsborough County are acknowledged as the source in any reproduction, quotation or use.



Acknowledgements

The author would like to express appreciation for the efforts of Dr. Janice Wood and the staff at University of North Florida College of Education and Human Services and the Florida Institute of Education in conceptualizing the Early Literacy and Learning Model (ELLM) and for their willingness to share their vision with members of our community. I would also like to acknowledge Luanne Panacek, E.D. and the Board of Directors of the Children's Board of Hillsborough County for pursuing this opportunity and for providing funding and dedicated staff resources to make this program a reality for our community. Special recognition is granted to Amy Haile (Children's Board) for her efforts in program implementation, the literacy coaches (Carole Pandorf, Thelma Seyferth, Mona Posinoff, Kavita Andrews) for their innovation and dedication to the program, members of the Children's Board staff who assisted with data collection, Mary Bailey (Children's Board Research Assistant) for her commendable skills in data management and organization, and the ELLM teachers and directors at the participating child care centers for their commitment to enhancing the quality of literacy instruction. Thanks are granted to the families for their support of their child's participation in the program and evaluation and to the children for their warm smiles, enthusiasm for learning and serving as a constant reminder of the importance and value of this work.

Foreword

This program evaluation report is based on the second year (August 2003- July 2004) of the Hillsborough County Early Learning and Literacy (ELLM) program and it represents the first comprehensive assessment and written manuscript regarding the evaluation findings. Program implementation in the first year (November 2002-July 2003) was of shorter duration due to the timing of start-up activities and although similar data was collected as in the second year, only preliminary analyses were conducted. Throughout the report, where appropriate, second year activities and results are compared to those of the first program year.

While the Childrens' Board of Hillsborough County originally initiated, implemented and managed the Hillsborough County ELLM program for the duration of the first year, in the second year of program operations (i.e., February, 2004), a decision was made to outsource program service delivery and operations and fiscal oversight for the program. Through a competitive selection process, Hillsborough Community College was selected as the grantee for the ELLM program. All ELLM staff members were retained with the program and became employees of the grantee.

Early Literacy and Learning Model (ELLM) of Hillsborough County
Program Evaluation
2003-2004

Background and History

Program Model Description

The Early Learning and Literacy Model (ELLM) is a research-based professional development model for early childhood teachers developed by the University of North Florida (UNF) College of Education and Human Services and the UNF based Florida Institute of Education (UNF-FIE). In partnership with several governmental agencies and private organizations, Head Start childcare centers, public and private childcare centers and two public school districts, UNF implemented the ELLM project in 1996. This collaboration was committed to assisting urban preschool children in acquiring the literacy skills necessary to promote and motivate future reading proficiency through the provision of explicit literacy instruction.

ELLM is a research-based and standards driven professional development support system substantiated by the following research:

- The most important activity for building reading skills is reading aloud to children.
- The single best predictor of reading success is knowledge of letters of the alphabet.
- The second best predictor of reading success is the ability to discriminate phonemes.
- In order to read successfully, children need to understand the concepts of printed word.
- Oral language competence positively affects reading success.
- Family involvement improves student achievement and attitudes about school.

Through the use of a coaching model, the ELLM professional development system provides recurring on-the-job support for teachers working directly with young children to improve children's language and emergent literacy skill development. ELLM provides these teachers with a literacy coach and research-based instructional strategies and materials for use in the classroom. The literacy coaches work in the classrooms on a regular basis and observe classroom instruction and assist teachers in translating research findings into practice through demonstration, observation and feedback, and lesson planning. The literacy program addresses the following five elements of beginning reading including: (1) reading aloud and emergent comprehension, (2) oral language and listening, (3) letter and sound knowledge, (4) phonological awareness, and (5) print concepts and emergent writing.

Since 1996, the UNF-FIE ELLM participation has grown each year and by 2003 included approximately 1,300 children and their families and 72 classroom teachers and assistants in addition to site directors, principals and other stakeholders. In 2002, through

the efforts of the Children’s Board of Hillsborough County, 15 licensed childcare centers located in zip codes 33612-13 that served a high percentage of low-income children in Hillsborough County joined the community of ELLM project participants. The Hillsborough ELLM project specifically targeted childcare teachers of preschool children aged three, four and five years old.

In its inception, Hillsborough ELLM was implemented in 27 classrooms representing approximately 300 children comprised of mostly three and four year olds although two sites had preschool, kindergarten, and first and second graders. The Hillsborough ELLM project provided each of the teachers with weekly coaching sessions that included modeling, observation and feedback, and lesson planning. Each ELLM participating classroom received six books and an accompanying literacy activities packet and teacher and family tips sheets each month. Teachers were provided with an annual multi-session training course and were encouraged to attend the monthly training sessions including “make-n-takes” (demonstration sessions that provided materials and instruction designed to create a literacy activity for their classroom). In addition, each ELLM teacher received an annual financial stipend of up to \$1,000 for their participation and the center received reimbursement for substitute teachers to enable ELLM teachers to attend the training sessions.

Because the Hillsborough County implementation of ELLM was focused on three primary groups of project participants including the teachers, children and families, three separate project goals were established. These goals are as follows:

- **Teacher Goal:** Teachers will demonstrate an understanding of the link between teaching literacy related activities and their student’s mastery of literacy skills and abilities.
- **Child Goal:** To enhance the child’s pre-reading skills (abilities) in order to improve their readiness for school.
- **Family Goal:** Raise awareness of the importance of family involvement in literacy development with their children.

A rigorous evaluation of the program was designed and conducted annually to measure a variety of indicators to help determine the program’s accomplishments and success in achieving these goals.

Methodology

Participants

The ELLM evaluation design was focused on three different sets of participants. Since the model is a professional development program for the teachers, they were the primary intended participants of the program. However, because much of the teacher’s effectiveness rests on their interactions with the children in their classrooms, the children have been included as an additional participant group in the evaluation. Finally, because some emergent

literacy learning may occur at home, the families have also been included as program participants.

ELLM teachers. As indicated previously, ELLM was originally implemented in 27 classrooms within 15 childcare centers. By the end of the second year, enrollment and staffing changes resulted in a total of 26 participating classrooms located in 12 childcare centers. Two of the original centers (i.e., New Life Learning and The Spring) closed after the first year due to low enrollment and a third one (i.e., Children's Center) closed after the pretesting phase had been completed. Within these 26 ELLM classrooms, a total of 38 teachers participated in some or all of the program year. Seven teachers who began the second program year left prior to completion of the program year and another seven teachers entered the program following the beginning of the program year. Of the 24 teachers who participated for the entire Year 2 program period (i.e., June 2003-May 2004), 17 were teachers who had previously participated in the initial year of the ELLM program (i.e., 2002-2003) and seven were new to the program in Year 2.

ELLM children. In Year 1, a total of 239 children were pretested and 204 children were posttested, thereby yielding a matched sample of 182 children for the year. In Year 2, there were a total of 248 children who were pretested. Of these children, 151 were available and retested during the posttesting phase that resulted in a matched sample of 140 children.

ELLM families. A total of 310 families participated in the pretest phase of the first year, however only 64 of these parents completed posttests. This yielded a matched set of 55 families. In the second year, 171 families completed the pretest and 31 of these submitted a posttest (i.e., matched sample).

Measures

Teacher/Classroom measures. Several different observational and self-report measures were employed to assess teachers' performance and adherence to the ELLM program and principles. Two observational instruments were administered by the literacy coaches in order to assess the teachers and their classroom environments. A modified version of the UNF Targeted Instructional Strategies (TIS) measure was used in both the first and second years of the ELLM program, whereas a revised UNF-developed Environmental Materials Checklist, used in the first year, was replaced by the Early Language and Literacy Classroom Observation (ELLCO) in Year 2. To supplement the observational measures, each of the literacy coaches also kept journals of free text case notes for each of the teachers to record qualitative information about their feedback sessions with the teachers. The two self-report teacher measures included a retrospective pretest that assessed their incorporation of the literacy activities into their classroom instruction and an overall program satisfaction survey.

The TIS instrument was modified to measure teachers' frequency of use of transitions (e.g., transitioning children from one activity to another) and the sources of the activities they used with the children in their classroom. The use of transitions was prioritized and measured as a result of classroom management concerns observed by the coaches. In the first year of implementation, this measure was used to only record the frequency of transitions and sources of activities however in the second year a rating scale was employed to assess performance in these areas as well.

The Environmental Checklist was conducted three times during the first program year and was intended to account for changes in the variety of tools used in the classroom that encourage literacy (visual, auditory, verbal and tactile elements). The checklist consisted of over 50 items in 4 categories that facilitate the following: (1) oral language development (e.g., puppets, interactive books), (2) letter sound knowledge (e.g., alphabet, number, word cards), (3) print concept & emergent writing (e.g., lending library, Word Wall) and (4) phonemic awareness (e.g., rhyming cards, interactive song CDs).

The ELLCO, which replaced the Environmental Checklist in Year 2, is a standardized measure that includes three observation instruments that yield total scores comprised of ratings on several subscales. The Literacy Environment Checklist consists of 5 subscales including Book Area, Book Selection, Book Use, Writing Materials and Writing Around the Room. The Classroom Observation score includes the General Classroom Environment and Language, Literacy, and Curriculum subscales. Finally, the Literacy Activities Rating Scale consists of the Book Reading and Writing subscales. This measure was administered twice using a pre/post test design.

The retrospective pretest was a self-report measure designed to assess changes in the ELLM teachers' classroom practices and attitudes towards literacy learning in children. The selection of a retrospective pretest instrument instead of a traditional pretest/posttest measure was made in order to address the potential issue that the teachers may be unfamiliar with some of the program concepts prior to their involvement with ELLM and that they would not be able to adequately assess themselves on indicators related to these concepts at the usual time of pretesting (e.g., beginning of program year). The retrospective pretest instead asks participants at the end of the program year to first reflect back on their behavior or attitudes as recalled before being involved in the program (e.g., Before ELLM) and to rate themselves in addition to providing a rating of their current performance and beliefs (e.g., Now). It was believed that this method provided the teachers sufficient time to understand the program elements before being asked to assess their use of them in their classroom teaching.

The retrospective pretest was administered at the annual teacher training institute held in June of each year and it specifically asked teachers to rate both before ELLM and Now their frequency of use of particular model components (e.g., read alouds, print/writing activities, Word Wall) as well as rating their beliefs about the importance of these program elements to the development of their student's literacy skills. Teachers who were in their second year of participation in ELLM, were only asked to rate their current frequency of use and perceptions of importance of the program elements so that these results could then be compared to their first year responses of "Before ELLM" and "Now".

The teacher satisfaction survey was another self-report measure designed to provide teachers with an opportunity to express their level of satisfaction with the program components and to provide input as to any changes they would like to see happen. The instrument consisted of four sections including ratings of satisfaction and importance of the program elements, their perceptions of changes observed in the classroom since ELLM was implemented and what changes to the program, if any, they wanted to occur. This measure was also administered annually at the annual teacher training institute.

Child measures. Depending on the age of the child, different age-appropriate instruments were administered. For children who were at least four years old at the time of pretesting, the Test of Early Reading Ability Third Edition (TERA-3) and the Alphabet Letter

Recognition Inventory (ALRI) were utilized. For the children who were at least three but less than four years old at the time of pretesting, the Kaufman Survey of Early Academic and Language Skills (K-SEALS) was administered. Both the TERA-3 and the K-SEALS are standardized, normed assessment instruments. The ALRI was a tool developed by the UNF ELLM project that was adopted and modified slightly.

The TERA-3 instrument consists of a picture book of items and is valid for children ages 3 years 6 months to 8 years 6 months. The instrument is comprised of the following 3 subtests that yield an overall composite Reading Quotient score. These include the following:

- **Alphabet subtest** – measures knowledge and use of letters
- **Conventions subtest** – measures arbitrary aspects of English print (e.g., how to hold book, where to begin reading on a page, punctuation)
- **Meaning subtest** – measures variety of ways in which children comprehend print (e.g., signs, logos, relationships among words, paraphrase and interpret stories)
- **Reading Quotient** - best indicator of child's overall reading ability or predictor of future reading ability

The K-SEALS measure also consists of a picture book of items and is valid for children aged 3 years to 6 years 11 months. It is comprised of three subtests that yield two language scale subtotal scores (i.e., expressive and receptive skills) and an overall composite score. The subtests and subtotal scores are as follows:

- **Vocabulary:** Naming or pointing to object or actions shown in pictures, identifying object from descriptions.
- **Numbers, Letters & Words:** Naming or pointing to numbers, letters or words; counting; solving number problems.
- **Articulation:** Pronouncing or naming of common objectives or actions.
- **Expressive Skills:** Naming objects, numbers, letters, words or actions shown in pictures.
- **Receptive Skills:** Responding to verbal instruction by pointing to objects, numbers, letter, words or actions, identifying objects from their descriptions.

The ALRI measure consists of upper and lower case letter flashcards. It yields a total score that represents the number of letters out of 52 letters that were correctly recognized. The UNF ELLM project presents the letters to children in one consistent random order however, upon adoption by the Hillsborough County ELLM program, a second random order of letters was created to minimize the likelihood of possible order effects. Each child was randomly assigned one of the two orders for presentation at pretest and where possible this same order was again utilized at the time of posttesting.

Family measure. Because children's literacy skill development cannot be solely attributed to the efforts of the classroom teacher but rather may also be the result of the family's involvement, a pre/post family survey was developed to assess the frequency of literacy-related activities that may occur at home. This survey attempted to measure changes

over time in the occurrence of basic literacy activities (e.g., reading aloud, writing/drawing, library visits) at home.

Data Collection

With the exception of the TIS data and two self-report teacher measures, all of the other measures were administered using a pre/post test design wherein the pretest data was collected in August/September (November in Year 1 due to timing of project start-up) and post test data collected in April/May. The TIS data were collected throughout the program year and the two teacher measures were administered at the teacher training sessions that were held in June following each year of the program. All of the child measures were administered in the morning typically between the hours of 8 a.m. and 11:30a.m. and only to children whose parents/caregivers had granted permission for them to participate.

During these first two years of the project, the literacy coaches participated in all facets of the data collection including the distribution of the teacher self-report measures, coordination with the child care center directors for the collection of the family surveys, and direct administration of the child, classroom and teacher observation measures. Due to the high number of children to be measured, the coaches were assisted in this data collection by members of the ELLM evaluation team and several other Children's Board staff. All data collectors, including the literacy coaches, were provided training on the administration of the child measures and testing guidelines/protocols for the protection of human study participants.

Findings

The data collected in Years 1 and 2 of the project were analyzed however the primary focus of this report is on the findings from Year 2 given that Year 1 was of shorter duration due to project implementation. Where appropriate, the findings will be compared across years. The findings are separated into three sections that present the analyses and results for the teacher measures, child measures and family measure. At the conclusion of each of the findings, a "Key Result" summary has been added to assist readers and program staff to easily and accurately articulate and disseminate the findings.

Summary of Descriptive and Statistical Analyses of Teacher Measures

The ELLM teachers were assessed through a variety of observational and self-reported measures. The coaches observed and rated teachers on a recurring basis through the use of the TIS measure in conjunction with free text journals and at pretest and posttest sessions with the ELLCO. Teachers also completed two self-report instruments, a retrospective pretest that assessed their incorporation of literacy activities into their classroom instruction and an overall program satisfaction survey. Results from each of these measures will now be presented.

What were teacher participation and attrition rates? Prior to conducting any analyses on the teacher measures, the sample sizes and attrition rates for ELLM teachers were examined. A total of 38 teachers participated for some duration of time during the second year of ELLM. Twenty-four of the teachers were present for the entire year, seven teachers left during the program year and were replaced by seven others who participated for only a portion of the year. This represents an overall attrition rate of 22% (i.e., 7 of 31 teachers) in teachers'

participation in ELLM. However, of the 24 second-year teachers who participated for the entire year, 70% were returning teachers from the first program year.

What was the distribution of teachers' scores on the TIS measure? Each of the ELLM classroom teachers were observed by the coaches throughout the year and measured with the TIS instrument on both their frequency of use and performance in using transitions and activities. At varying intervals the coaches observed and recorded teachers' frequency of use of transitions and the sources of the activities they used with the children in their classroom. Observations of teachers' transitions were dichotomously recorded in terms of use or no use. Teacher's activity sources were recorded as originating from the literacy packs, resource centers/coach or other sources. There were a total of 203 observations for the year for the 26 classroom teachers. The observed frequencies and percents of each are presented in Table 1.

Variable	N	Percent
Transitions		
Transitions Used	105	52%
No Transitions	98	48%
Total	203	100%
Activity Source		
Literacy Pack	48	24%
Resource Center/Coach	31	15%
Other	124	61%
Total	203	100%

Table 1. Observed Frequency of Use by Teachers

As shown in the table for transitions, teachers were observed by the coaches to use transitions slightly more frequently than not using them. For the three activity sources, teachers' were observed using activities from the literacy packets more often than activities obtained from the resource center or coach, however, activities from other sources were those most frequently observed.

In addition to recording the observed frequency associated with these elements of the model, the TIS measure also includes two ratings of teachers' performance using a five-point Likert scale that ranges from 1=Inadequate performance to 5=Excellent performance. The first rating relates to the coaches' assessment of the teacher's ability to effectively move children from one activity to another. Since the promotion of the use of transitions (e.g., musical or non-musical) represents an important classroom behavior management technique, teachers' abilities to effectively incorporate these into their classroom teaching can serve as an indicator of teachers' understanding of and compliance with this strategy. The second rating on the TIS relates to the coach's assessment of the teacher's understanding of the literacy activity they are using in the classroom.

Although a total of 203 ratings were recorded for each of the two variables over the year, these assessments occurred periodically and therefore the frequency and timing of these observations were not necessarily consistent across the set of teachers. A total of 24 of the teachers received between two and 15 observation ratings from the coaches throughout the year with an average of approximately 8 ratings ($M=7.71$, $SD=3.42$) over the year. In order to examine teachers' performance on these variables over time given the variability in the

frequency of assessment, a total of five ratings (two from the fall session and three from the spring session) were randomly selected and analyzed for the 20 teachers who had received at least five observation ratings over the year. The overall means and standard deviations were computed across all teachers' sets of 5 ratings to determine the average teacher performance on both variables. For transitions, the overall mean was 2.59 (SD=.317) which equates to between poor and adequate performance in the use of transitions. For the activities variable, the overall mean was slightly higher at 2.88 (SD=.223) but still within the range of poor to adequate performance.

To test whether teachers' mean performance on these variables statistically changed over time, the nonparametric Friedman Test was employed instead of repeated measures ANOVAs given the small sample size. Results of these analyses indicated that there were only statistically significant mean differences (i.e., highly unlikely that these differences occurred as a result of chance) in the set of five transition ratings of the teachers over the year ($\chi=10.28$, $p<.05$) and not for the activity source variable. The combined teacher means on this variable however, were not indicative of incremental increased performance over time and instead revealed that performance at Time 2 ($M=3.10$; e.g., generally late fall) received the highest ratings followed by Time 1 ($M=2.65$), Time 3 ($M=2.45$), Time 5 ($M=2.40$) and Time 4 ($M=2.35$).

Key Result: ELLM teachers' effectiveness in using transitions varied inconsistently from poor to adequate performance at different points over the year and represents an area for improvement.

What were the main themes related to teachers' performance as recorded in coaches' journals? The literacy coaches kept free-text journals on their observations of the teachers for use in their feedback sessions during the period of January-March 2004. Qualitative analysis techniques (e.g., content analysis) were employed to look for main themes or recurring issues that could be identified. The data was analyzed globally across all teachers first and then separately by first and second year teacher designations. The main findings focused on teaching skills in regards to group size activities, transitions, behavior management and children's responsiveness to activities and learning. In addition, the data highlighted the importance of teachers' ability to closely monitor the developmental stage of their students.

In general, several similar themes seemed to emerge for both first and second year teachers. These issues highlighted the importance of utilizing small group activities for greater effectiveness (e.g., enhances transitions, promotes self-esteem, reduces behavior management issues). Both groups of teachers also need to be more aware of the developmental stage of the child particularly in regard to their selection of age-appropriate activities that are also appropriate in duration (i.e., awareness of attention span).

Journals for second year teachers indicated a need to allow for greater freedom of movement for children to create a more appropriate learning environment. It was also noted that these teachers seemed to understand the philosophy and concepts (perhaps with the exception of the differences between "sound knowledge" and "phonemic awareness") but needed to develop their implementation skills to be more effective overall. Additionally, these teachers need to take advantage of opportunities for oral development and remember that teacher interaction and conversations with their students are crucial for language development. Journals for first year teachers identified a need for the development of different teaching

techniques and in particular to work on making learning more fun (e.g., long reading times and too many questions are unproductive).

Key Result: While first year ELLM teachers require additional assistance in acquiring basic literacy teaching skills, second year teachers' need to focus on implementation skills and, in general, all teachers' classroom effectiveness could be improved through greater use of small groups and developmentally appropriate activities.

How well did teachers perform as measured by the ELLCO? A total of 18 ELLM teachers received both pretest and posttest scores on the ELLCO. This measure includes three observation instruments that yield total scores comprised of ratings on the Literacy Environment Checklist, the Classroom Observation score and the Literacy Activities Rating Scale score. Given that this was the first year using the ELLCO, prior to conducting statistical analyses it was necessary to gain an understanding of overall ELLM teacher performance in comparison to a benchmark or standard to assist in the interpretation of the scores obtained. The ELLCO technical appendix (Smith & Dickinson, 2002) reported data collected from several research studies conducted in New England that were used for purposes of establishing the psychometric properties of the instrument. These data were used as a benchmark for comparison with the ELLM teacher's scores. For purposes of this comparison, the ELLM means were computed using a sample size of 36, which treated each teacher's pretest and posttest scores as independent samples. The published sample was comprised of 255 preschool teachers in classrooms that were primarily in Head Start programs. Because this comparison was merely for benchmarking overall teacher performance, statistical analyses were not computed on these means. The comparisons between the samples for each of the observation instruments are presented in Figures 1-3.

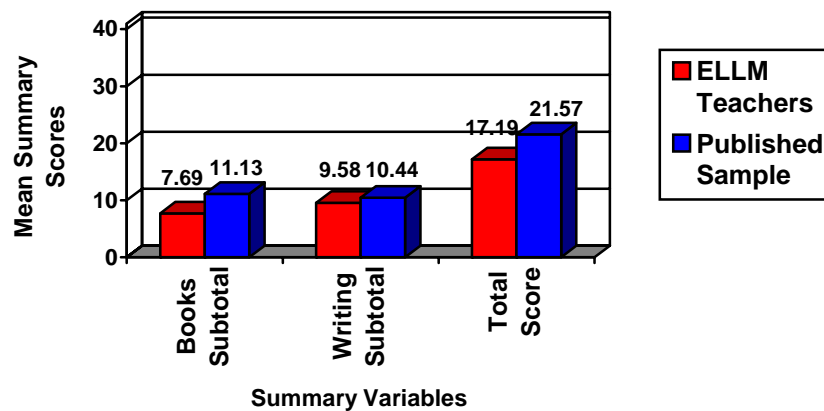


Figure 1. Mean ELLCO Literacy Environment Checklist Summary Scores for ELLM Teachers and Published Sample Teachers

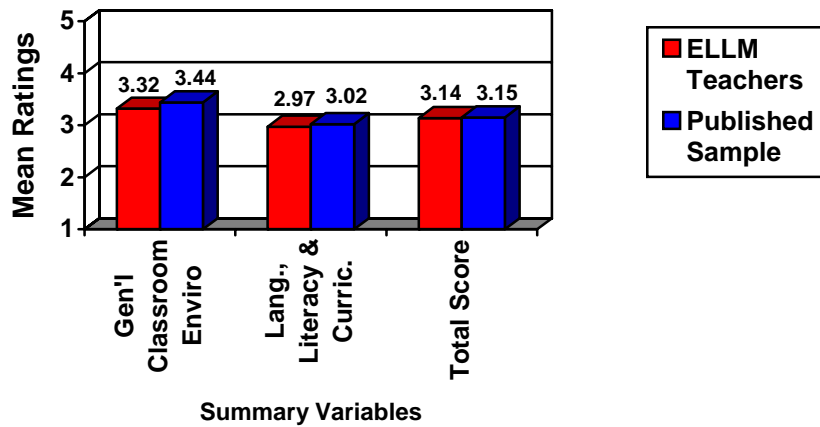


Figure 2. Mean ELLCO Classroom Observation Ratings for ELLM Teachers and Published Sample

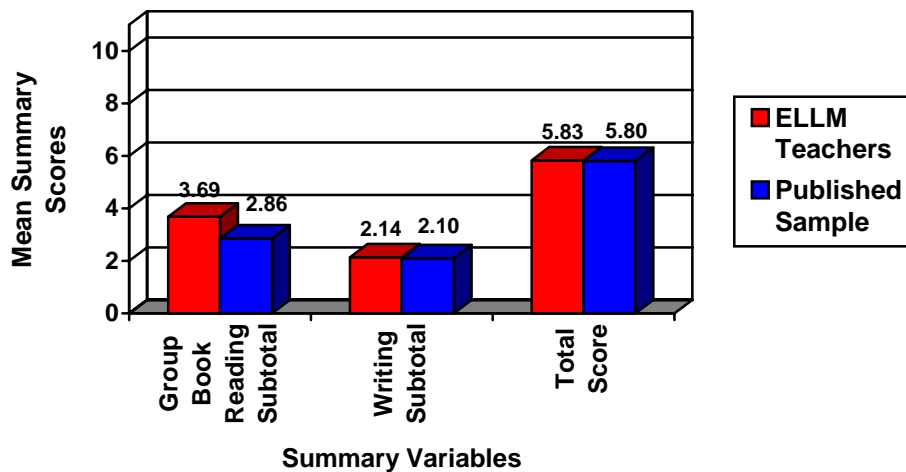


Figure 3. Mean ELLCO Literacy Activities Rating Scale Summary Scores for ELLM Teachers and Published Sample

The figures above illustrate that the ELLM teachers' performance, as measured by the ELLCO subscales, appeared to be generally comparable to those obtained from the teachers in the published studies. Performance on the Literacy Environment Checklist seemed to display the greatest degree of disparity between the samples with the ELLM teachers appearing to exhibit lower performance particularly for scores received on the Books Subscale. The two groups appeared to have relatively comparable performance on the Classroom Observation measurements and for the Literacy Activities Rating Scale, ELLM teachers seemed to slightly outperform the teachers in the published sample.

Upon completion of this benchmarking exercise, ELLM teachers' mean performance from the time of pretest to that of posttest was examined. To investigate for statistical evidence of this change, non-parametric Wilcoxon Signed Ranks Tests were employed for each of the subtests and scale scores for the sample of 18 teachers (Note: ELLM teacher pretest and posttest scores were treated as repeated measures for these analyses). Results are presented in Table 2.

Variable	Pretest	Posttest	Difference Score
Environment Checklist Summary Scores			
Books Subtotal	6.67	8.72	2.05*
Writing Subtotal	8.56	10.61	2.05*
Total Score	15.22	19.17	3.95*
Classroom Observation Ratings			
General Classroom Environment	3.24	3.39	0.15
Language, Literacy & Curriculum	2.89	3.05	0.16
Total Score	3.07	3.22	0.15
Literacy Activities Rating Scale			
Group Book Reading	3.61	3.76	0.15
Writing Subtotal	1.89	2.41	0.52*
Total Score	5.50	6.18	0.68
Note. * $p < .05$			

Table 2. ELLM Teacher's Mean ELLCO Pretest, Posttest and Difference Scores

These results indicated that while all mean scores were in the direction of showing improvement from pretest to posttest, only statistically significant gains were found for the Literacy Environment Checklist Total Score ($z = -2.646$, $p < .01$) including both its Books Subscale ($z = -2.69$, $p < .01$) and Writing Subscale ($z = -2.332$, $p < .05$) scores and the Writing Subscale of the Literacy Activities Rating Scale ($z = -1.998$, $p < .05$).

Key Result: ELLM teachers' average ELLCO scores were generally comparable to the published sample and statistically significant gains were shown over time for their average scores on the Environment Checklist and the writing subtotal of the Literacy Activities Rating Scale.

What is the degree of association between the TIS and the ELLCO? To establish whether similar constructs of teacher classroom performance were being evaluated with the two teacher measures, ELLM teachers' mean transition and activity ratings on the TIS were compared with performance on the Classroom Observation measure of the ELLCO. This portion of the ELLCO was selected for comparison because the test developer indicated that it was designed to specifically address the need to evaluate early language and literacy experiences or classroom features that are known to support literacy development in children (Smith & Dickinson, 2002). Pearson correlation coefficients were computed to assess the degree of association (i.e., relatedness) between the mean transition and activity ratings obtained from the TIS and the mean subscale and total scores on the Classroom Environment measure of the ELLCO. Correlation scores can range continuously from zero to plus/minus

one with scores at the zero end of the continuum representing little or no degree of relatedness up to one where they are essentially the same. Correlations other than zero can be positive or negative where negative correlation scores indicate that there is an inverse relationship (i.e., as one variable increases the other decreases) and positive correlations represent relationships in similar directions (i.e., as one variable increases the other also does or vice versa). The results of these analyses are presented in Table 3.

	TIS Transition Rating	TIS Activity Rating	ELLCO Classroom Environment Subscale	ELLCO Language, Literacy & Curriculum Subscale	ELLCO Classroom Observation Total
TIS Transition Rating	1.00	.894**	.727**	.617*	.712**
TIS Activity Rating	.894**	1.00	.855**	.699**	.828**
ELLCO Classroom Environment Subscale	.727**	.855**	1.00	.796**	.959**
ELLCO Language, Literacy & Curriculum Subscale	.617*	.699**	.796**	1.00	.934**
ELLCO Classroom Observation Total	.712**	.828**	.959**	.934**	1.00

Note. ** correlation is significant at $p < .01$, *correlation is significant at $p < .05$

Table 3. Correlation Matrix for TIS and ELLCO Scores

The above correlation matrix shows a high level of association between the subscales of the measures both within each instrument and across the two instruments. All of the correlations obtained were positive (e.g., direction of relatedness) and statistically significant (i.e., not likely to have occurred by chance) suggesting that the instruments are measuring the same construct.

Key Result: The revised TIS tool and the ELLCO were shown to be complimentary measures thereby providing an opportunity for interim measurement with the TIS between the ELLCO pre and posttest administrations.

How often did teachers report using literacy activities before and during their participation in the ELLM program? Following the completion of Year 2, teachers completed a Retrospective Pretest that assessed their daily use of specific literacy activities in the classroom. Teachers who had just completed their first year in the program reported on their use of these activities both before they started in ELLM and after one year of participation in ELLM (see Figure 4). Teachers who completed their second year in ELLM only reported on their daily use of literacy activities during their second year of participation. These responses were then compared to the responses they provided after their first year in ELLM (ratings of before and after one year in ELLM). The results for this latter group are presented in Figure 5. It is important to note however that because of the small sample sizes, statistical analyses

were not conducted and therefore it is not known whether any of the observed differences in the graphs are statistically significant and represent meaningful differences.

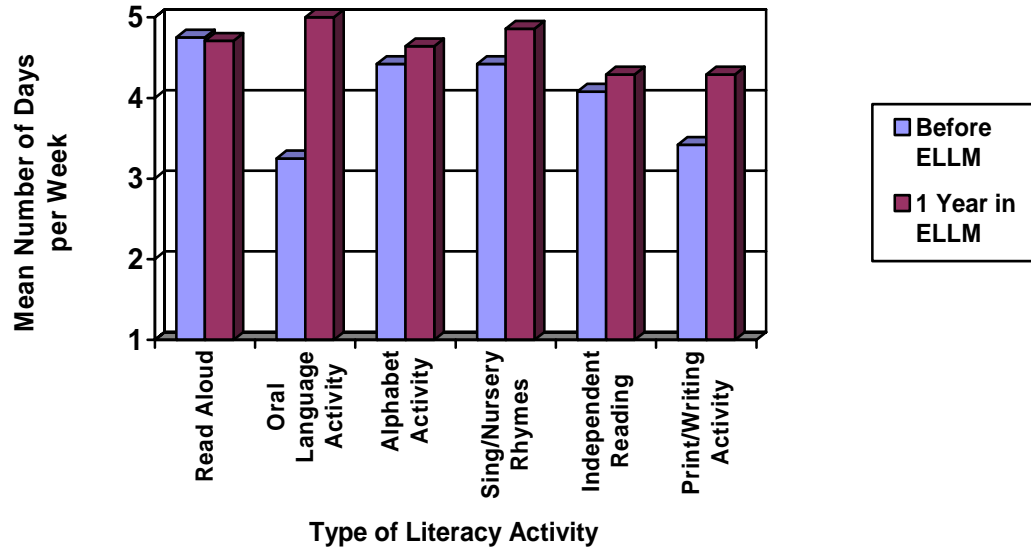


Figure 4. Mean Number of Days/Weeks 1st Year Teachers Reported Using Literacy Activities Before/After 1 Year in ELLM (N=7)

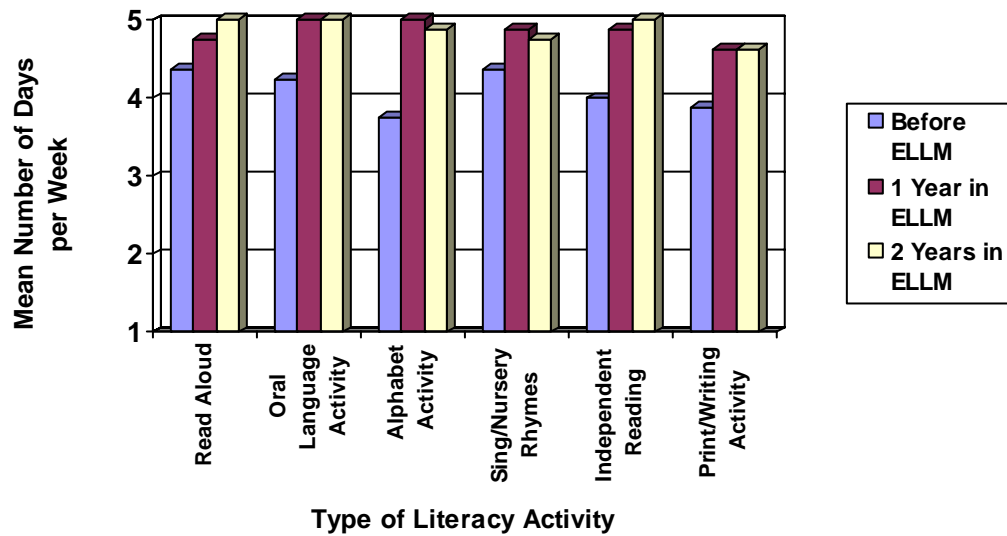


Figure 5. Mean Number of Days/Weeks 2nd Year Teachers Reported Using Literacy Activities Before/After 1 and 2 Years in ELLM (N=8)

As seen in Figure 4, the teachers who completed their first year of participation in ELLM reported increases in their daily use of all of the literacy activities except read alouds at the completion of their first year in the program compared to before they were in ELLM. It appeared that the largest observed increase in frequency of reported use was for oral language activities that changed from an average use of approximately three days per week before ELLM to everyday after participating in ELLM.

For teachers who completed their second year of participation in ELLM, Figure 5 suggests reported increases in frequency of use for all literacy activities from before ELLM participation to the end of their first year of participation. For this group, the largest observed mean increase in reported use was shown for alphabet activities. At the completion of their second year of participation in ELLM, the mean reported frequency of use appeared greater for all activities compared to before their participation in ELLM and increased or remained stable from the end of the first year for all activities except alphabet and singing/reciting nursery rhymes which seemed to decrease slightly.

Key Result: In general ELLM teachers reported and increased frequency of literacy activities in the classroom after the completion of one or two years of participation in ELLM.

How important did teachers rate literacy activities for the development of children’s literacy skills before and during their participation in the ELLM program? As part of the Retrospective Pretest, teachers were asked to rate the importance of specific literacy activities in the development of children’s literacy skills. Teachers rated the importance of these activities both before their participation in ELLM and after one or more years of participation. Ratings were made using a four-point Likert Scale ranging from 1=Not Important to 4=Very Important. The results for the first year teachers are presented in Figure 6 and in Figure 7 for the second year teachers.

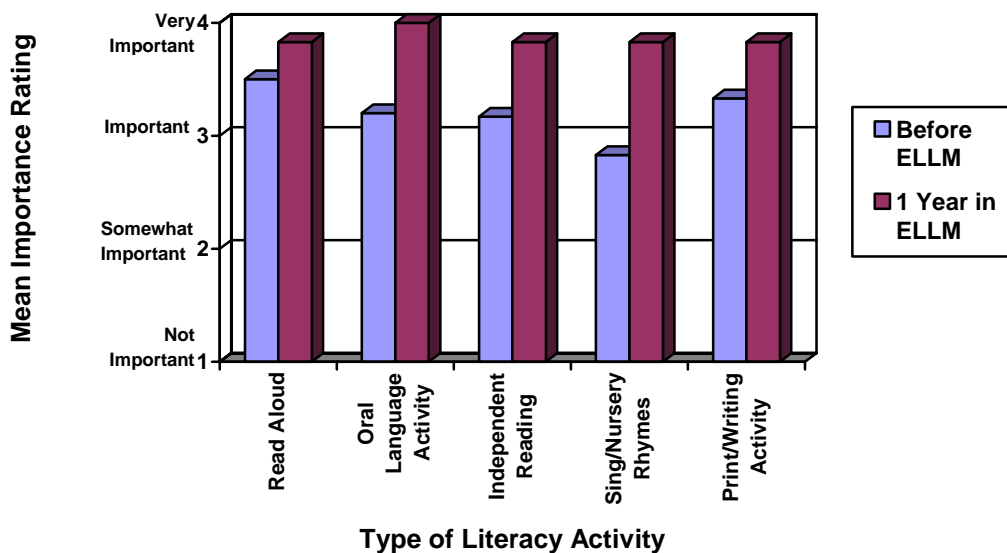


Figure 6. First Year Teachers’ Mean Importance Rating of Literacy Activities Before and After 1 Year in ELLM (N=7)

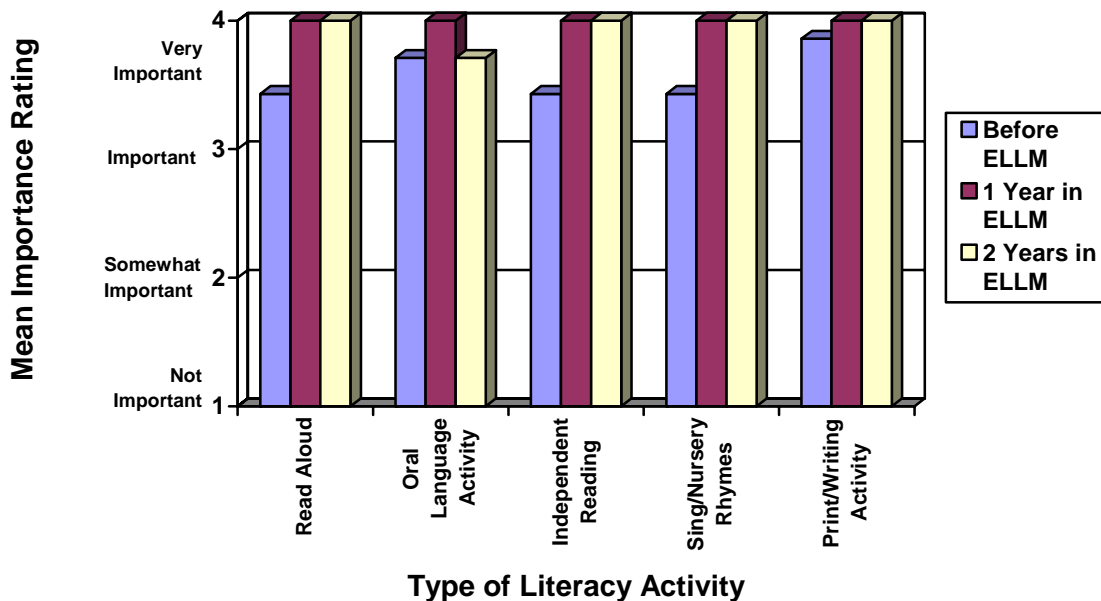


Figure 7. Second Year Teachers' Mean Importance Rating of Literacy Activities Before and After 1 and 2 Years in ELLM (N=8)

As seen in Figure 6, increases appeared in teachers' mean ratings from before participating in ELLM to after one year of participation regarding the importance of all of the literacy activities in children's literacy development. The largest increases in mean importance ratings over time seemed to be observed for activities involving singing/reciting nursery rhymes. Figure 7 shows similar increases in mean importance ratings for the second year teachers in relation to their before and after one year of participation ratings. At the end of Year 2, all mean importance rating gains seemed to be maintained with the exception of oral language activities which appeared to decrease slightly.

Also included in the Retrospective Pretest were several questions related to teachers' use of a Word Wall (e.g., wall area that has alphabet letters with related words and pictures) in their classroom for literacy instructional purposes. While only 40% of the first year teachers indicated that they had a Word Wall prior to their participation in the ELLM program, all teachers indicated they had one in their classroom after the completion of their first year in ELLM. Similar results were recorded for the second year teachers at the completion of their first year of participation in ELLM. Additionally, teachers reported on the frequency of use of the Word Wall in their teaching and the children's use of it during print/writing activities both before participating in ELLM and again at one or two years afterwards. Frequency ratings were provided using a four point Likert Scale that ranged from 1=Not at all to 4=Very Often. The mean ratings for both groups of teachers (i.e., first and second year teachers) are presented in Table 4.

	First Year Teachers		Second Year Teachers	
	Teacher's Use of Word Wall	Children's Use of Word Wall	Teacher's Use of Word Wall	Children's Use of Word Wall
Before ELLM	2.25	2.00	1.86	1.71
1 Year After ELLM	3.33	3.17	3.50	3.75
2 Years After ELLM	N/A	N/A	3.86	3.57

Table 4. Mean Frequency Ratings of Word Wall Utilization for First and Second Year ELLM Teachers

Table 4 shows the mean reported increases in frequency of use of the Word Wall by both the teachers and the children for both groups of teachers from before ELLM to one year after participating in ELLM. While the second year teachers reported mean ratings for the frequency of use of the Word Wall in classroom teaching appeared to continue to increase in their second year of participation, their mean ratings of the children's frequency of use of the Word Wall seemed to decrease slightly from Year 1 to Year 2.

Key Result: ELLM teachers reported an increase in their awareness of the importance of literacy activities after one or two years participation in ELLM and all teachers reported having a Word Wall by the end of their first year of participation in ELLM.

How satisfied were teachers with the ELLM program? At the end of the ELLM program year, teachers completed a satisfaction survey that assessed their level of satisfaction with various components of the program. These ratings were made on a five point Likert scale that ranged from 1=Not at all satisfied to 5 = Very satisfied. The mean ratings for all teachers in addition to the first and second year only teachers are presented in Table 5.

Program Component	All Teachers	1 st Year Teachers Only	2 nd Year Teachers Only
Coach's Assistance in Modifying Materials to Classroom	4.79	4.71	4.86
Coach's Feedback Sessions	4.77	4.71	4.83
Activities in Literacy Packets	4.73	4.86	4.63
ELLM Books	4.67	4.57	4.75
Resource Center	4.64	4.50	4.75
Coach's Help in Locating Resources	4.62	4.43	4.83
Monthly Teacher Tips	4.60	4.71	4.50
Coach's Modeling of Literacy Activities	4.54	4.43	4.67
Coach's Observation of Teaching Sessions	4.54	4.57	4.50

Table 5. Mean Satisfaction Ratings of ELLM Program Components for First and Second Year Teachers

As indicated in the table, overall, the teachers seemed to be quite satisfied with the various components of the ELLM program. The highest mean satisfaction ratings for all teachers combined appeared to be for the assistance teachers received from the coaches in modifying the materials to meet their individual classroom needs followed by the coach’s feedback session. First year teachers reported being most satisfied with the activities in the literacy packets followed by the individual tailorings of the program to meet their needs. Teacher comments indicated an appreciation and high level of satisfaction with the program (e.g., “I am very satisfied and very happy with the program”; “I am really grateful to ELLM and all the help they have provided for me this year. I wouldn’t have finished the year without them”). One first year teacher reported learning more from observing the literacy coach modeling the activity.

Key Result: Teachers reported a high level of satisfaction with the components of the ELLM program.

How important are the ELLM program components to teachers? As part of the satisfaction survey, teachers also rated the importance of various components of the ELLM program using a four point Likert scale that ranged from 1=Not Important to 5=Very Important. Results of these ratings are presented in Table 6.

Program Component	All Teachers	1st Year Teachers Only	2nd Year Teachers Only
ELLM Books	4.00	4.00	4.00
Literacy Packets	3.87	3.71	4.00
Coach’s Observation of Teaching	3.86	3.71	4.00
Resource Center	3.77	3.80	3.75
Monthly Teacher Tips	3.73	3.71	3.75
Coach’s Modeling of Literacy Activities	3.71	3.67	3.75
Stipends for Training Events	3.67	3.71	3.63
Teacher Get Togethers	3.54	3.17	3.86
Quarterly Stipend	3.53	3.43	3.63

Table 6. Mean Importance Ratings of Program Components for 1st and 2nd Year Teachers

As seen in the table, it appeared that based on the means for all teachers that the books provided by the ELLM program were of greatest importance to them followed by the literacy activity packets. The second year teachers’ ratings seemed to indicate that the literacy packets and coach’s observation sessions were equally as important as the books, however, it appeared that the first year teachers rated the resource center as second most important followed by what appeared to be equal ratings for the literacy packets, coach’s observation sessions, monthly teacher tips and stipends for training. Teacher comments were again very positive toward the program (e.g., “Great”; “All very helpful and informative”) and several first year teachers were particularly appreciative of the stipend, which they reported helped them stay in the field and offset their very low wages.

Key Result: In general, ELLM teachers indicated that the books and accompanying literacy activity packets were among the most important of the program components.

What changes, if any, do teachers want in the frequency of occurrence for ELLM program components? The teacher satisfaction survey also asked teachers to rate the degree of change, if any, that they desired in the frequency of occurrence of several of the components of the ELLM program. Teachers rated these components on a five-point Likert scale that ranged from 1=Much less to 5=Much more. Mean ratings on these items are presented in Table 7.

Desired Changes	All Teachers	1st Year Teachers Only	2nd Year Teachers Only
Assistance with Planning Family Activities	3.83	4.20	3.57
Coaching	3.67	3.50	3.83
Teacher Get Togethers	3.58	4.00	3.29
Assistance with Other than Literacy Classroom Strategies including Curriculum	3.45	3.80	3.17
Monthly Teacher Tips	3.38	3.50	3.29
Resource Center	3.36	3.50	3.29
Literacy Related Activities (including packets)	3.25	3.00	3.50

Table 7. Mean Ratings of Desired Changes in the Frequency of Program Components for 1st and 2nd Year Teachers

Table 7 suggests that the teachers’ mean ratings of these ELLM components indicated that for many of the components teachers, on average, reported being satisfied with the frequency of occurrence, or as in the case of assistance with planning family activities and coaching sessions, wanted them to occur a little more frequently. Mean ratings and teacher comments helped confirm that the first year teachers appeared to report a clear desire for more frequent teacher get togethers and assistance with planning family events (e.g., “I would like to attend more teacher get-togethers”; “I would like more informative teacher get-togethers with less singing and dancing at them”) whereas the second year teachers seemed to want slightly more coaching sessions and literacy related activities (e.g., “Maybe we can learn more with numbers and how to write letters”; “Its all been a life saver for this year, never will I forget it! Thanks again and again”).

Key Result: On average, ELLM teachers indicated little desire to change any of the program components although perhaps more frequent coaching sessions and assistance with family activities.

What changes have teachers noticed since participating in ELLM? Finally, teachers were asked to rate to what extent if any, they had noticed specific changes in their classroom since their participation in the ELLM program. These included rating potential changes in their teaching strategies, the classroom physical environment, their interaction with children, their job satisfaction and their children’s interest in reading and books. Ratings were made on a four point Likert scale that ranged from 1=No change to 4=Significant change. Results are presented in Table 8.

Observed Changes	All Teachers	1 st Year Teachers Only	2 nd Year Teachers Only
Children’s Interest in Reading/Books	3.83	3.60	4.00
Teaching Strategies	3.75	3.40	4.00
Interaction with Children	3.67	3.20	4.00
Classroom Physical Environment	3.64	3.25	3.86
Job Satisfaction	3.64	3.00	4.00

Table 8. Mean Observed Change Ratings for First and Second Year ELLM Teachers

The findings suggest that teachers reported being aware of observable positive improvements in their attitudes and behavior and that of their students and classrooms as a result of their participation in the ELLM program. In all of these areas, second year teachers’ mean ratings seemed to indicate a greater degree of discernable change than first year teachers. Teacher comments from both first and second year teachers were indicative of noticeable changes in both teacher behavior (e.g., “I am able to teach more”; “I have gained so much knowledge”; “I notice myself using literacy techniques when asking questions or informal interactions”) and children’s behavior (e.g., “Kids read more and learn more vocabulary; “My class of kids really enjoy everything”).

Key Result: All ELLM teachers and second year teachers to a greater extent, indicated being aware of noticeable positive improvements in their teaching, classrooms and students.

Summary of Descriptive and Statistical Analyses of Child Measures

Based on the age of the child at the time of pretesting, one of two standardized measures was administered (i.e., KSEALS for 3 year olds, TERA-3 for 4 or 5 year olds) and then repeated at posttesting. Children who were at least 4 years old at pretesting were also administered the ALRI measure. For each of the measures that were administered to the children (i.e., TERA-3/ALRI or KSEALS), pretest, posttest and difference (change) scores were computed individually for each child. Given that these measures were to serve as a proxy for measuring program (i.e., teacher) effectiveness rather than for diagnostic purposes to assess individual children, only aggregate analyses were conducted on these sets of scores. Results from these analyses will now be presented for all the child measures.

What were the participation rates of children in ELLM pre and post testing? Prior to conducting any statistical analyses on the children’s data collected, it was necessary to examine the sample sizes for each of the measures and to create sets of matched samples for children who participated at both pretest and posttest sessions. Table 9 shows these sample sizes.

Measure	Pretest	Posttest	Matched
TERA-3	149	93	87
ALRI	129	66	51
KSEALS	99	58	53
TOTAL	248	151	140

Table 9. Number of Children Participating in Year 2 ELLM Pre and Posttesting

As shown in Table 9, as expected, attrition occurred in the classrooms over the program year and therefore the sample sizes at posttest were considerably smaller than at pretest. The attrition rate presented in the table was 39%. Attrition rates in the first year were considerably lower at 19%. Also noted in the table is that there were fewer ALRI scores than TERA-3 scores even though typically all children tested with the TERA-3 measure would also be administered the ALRI. This discrepancy was attributed to instances when the ALRI was not administered due to the child showing signs of fatigue/distraction or distress or as a result of it being erroneously omitted during the testing session.

Key Result: Child attrition rates were at 39% from pretesting to posttesting which represents approximately four out of every 10 children not being available for posttesting.

Did attrition rates have an effect on children's mean scores? The attrition rate primarily reflects the high degree of mobility of children in the ELLM centers, however, it also encompasses children who may have still been enrolled in the ELLM program and were absent from the classroom at the various testing administration times. Without knowing the cause of the high attrition rates, it was necessary to refute the notion that any significant differences that may be present in children's performance over time was attributable to initial differences in abilities of the children who remained in ELLM through posttesting and those that left or were not available at the time of posttesting. This would help ensure that any differences in scores found between pretesting and posttesting for the matched sample was not merely the result of the poorer performing children leaving the ELLM program prior to posttesting. To accomplish this, pretest scores for the matched sample (i.e., children with both pretest and posttest scores) were compared to scores for children who only participated in the pretesting phase.

Analysis of Variance techniques were employed to examine for mean differences between the groups. The results for all scores on all measures (TERA-3, ALRI, KSEALS) indicated that there were no significant differences in the pretest scores of children in the matched (pre-post) and non-matched (pretest only) groups. Thus, children's attrition in ELLM was shown to be statistically unrelated to their initial performance on the measures. Preliminary analyses of the Year 1 data had suggested possible attrition effects in the TERA and ALRI (upper case letters only) pretest scores.

Key Result: There were no significant average differences in pretest scores for children who were available for posttesting and those that were not available.

What was the demographic composition of matched sample of participating children?

An examination of the matched sample of 140 children (i.e., those that pretested and posttested) revealed that it was comprised of 52% females and 48% males. The average overall age of the sample at the time of pretesting in August/September 2003 was 4.5 years old (SD=1.044 years) and ranged from 3.0 years to 7.5 years old. The average age at posttesting in April, 2004 was 5.17 years old (SD=1.06).

How did children perform on the measures at pretest and posttest? To assess children's performance on the measures, the mean and standard deviation was computed for

the pretest, posttest and difference scores for the matched sample of children. Table 10 presents the results for the aggregate and subtest standard scores of the TERA-3 and KSEALS and the total number of upper and lowercase letters recognized on the ALRI.

Measure	Pretest Mean (Standard Deviation)	Posttest Mean (Standard Deviation)	Difference Score Mean (Standard Deviation)
TERA-3 (n=87)			
Reading Quotient ^a	93.54 (17.56)	106.62 (13.54)	13.08 (16.64)
Alphabet Subtest	9.63 (3.54)	12.13 (3.05)	2.49 (3.21)
Conventions Subtest	8.78 (2.58)	10.36 (2.62)	1.57 (2.86)
Meaning Subtest	9.34 (2.98)	10.31 (2.12)	0.97 (2.88)
ALRI (n=51)			
Mean Number of Letters Recognized	26.02 (17.54)	43.51 (11.49)	17.48 (12.20)
KSEALS (n=53)			
Composite ^a	98.53 (11.20)	105.72 (14.00)	7.19 (14.62)
Vocabulary Subtest	95.28 (9.82)	102.13 (14.66)	6.85 (14.54)
Numbers, Letters and Words Subtest	101.72 (12.74)	107.36 (14.25)	5.64 (16.78)
Expressive Skills Subtest	98.60 (10.89)	105.19 (13.20)	6.58 (13.25)
Receptive Skills Subtest	97.62 (12.44)	103.72 (13.30)	6.09 (14.00)
Note. ^a Standard scores are based on a distribution with a mean of 100 and a standard deviation of 15.			

Table 10. Mean Pretest, Posttest and Difference Scores for Children on the TERA-3, ALRI and KSEALS (Year 2)

As seen in the table, on each of the measures children's mean performance increased from pretesting to posttesting. To assist in the interpretation of the mean standard scores for the TERA-3 and KSEALS, the national norm averages provided by the test developers (Reid, Hresko & Hammill, 2001; Kaufman & Kaufman, 1993, respectively) were considered. (Note: The ALRI is not a nationally normed instrument.) As noted in the table, the statistical average is 100 for both the TERA-3 Reading Quotient and KSEALS Composite standard scores. The average range of performance (e.g., half of all children scores) is between 90 and 110 for the TERA-3 Reading Quotient and between 90 and 109 for the KSEALS composite score. As indicated in the table, the pretest and posttest mean standard scores for both measures were within these national average ranges.

To help interpret children's performance on the ALRI, scores were converted to a proficiency level proportion and compared to those obtained from The Early Childhood

Longitudinal Study – Kindergarten (ECLS-K), a national study of proficiency in alphabet letter recognition of children entering kindergarten (U.S. Department of Education, 2001a; see Figure 8). Proficiency in alphabet letter recognition was defined in this study as the ability to recognize between 40 and 52 letters. The study identified 66% of children entering kindergarten as meeting this level of proficiency (U.S. Department of Education, 2001b).

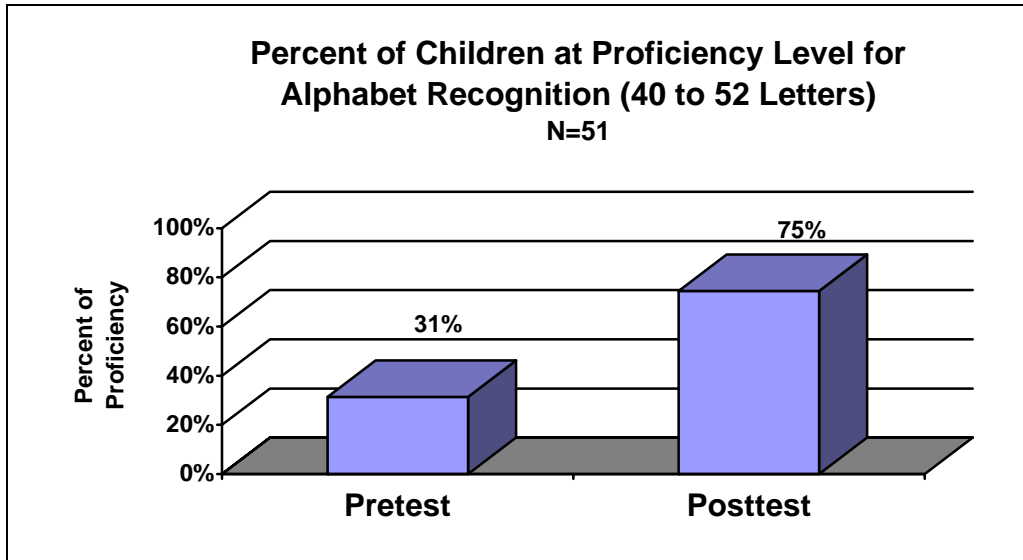


Figure 8. Percent of ELLM Children in Year 2 Attaining Proficiency Level for Alphabet Recognition (40 to 52 Letters; N=51)

As seen in Figure 8, at the time of pretesting (approximately 1 year earlier than enrollment into kindergarten), only 31% of the ELLM children achieved the defined criteria for proficiency, however, by posttesting (a few months prior to kindergarten enrollment), 75% of the children, or 9% more than the national average of 66%, obtained this proficiency level in alphabet recognition. Moreover, these Year 2 results showed improvement from those obtained in first year wherein 29% of the children were proficient at pretest and 60% at the time of posttesting.

Prior to testing the significance of the demonstrated change in performance exhibited over time, scores obtained from all children who participated in the ALRI (i.e., not just the matched sample) were examined for the presence of potential order effects. Multivariate Analyses of Variance (MANOVAs) were used to test for these potential differences in the pretest and posttest scores for the number of upper and lowercase letters recognized. Although a significant order effect was shown in the first year of the program in children's upper and lowercase pretest scores, the second year results revealed no evidence of an order effect. It should be noted however that the sample sizes of the order groups were markedly unequal at the time of pretesting (approximately two-thirds more order M than order V was utilized) but due to attrition, were balanced at posttesting. This circumstance could have impacted the ability to detect a possible order effect in the pretest scores.

Key Result: Children's average pretest and posttest scores on all measures were within the national average ranges suggesting that children in ELLM classrooms were on track.

Were the differences observed in children's pretest and posttest scores statistically significant? As shown in Table 10, children demonstrated gains in their mean scores on all measures from pretest to posttest administration. To determine whether these differences were statistically significant, repeated measures ANOVA techniques were employed for each of the mean scores on each measure separately. The results of these analyses are presented in Table 11.

TERA-3		
Dependent Variable	F(1,86)	R ²
Reading Quotient	53.78***	.39
Alphabet Subtest	52.67***	.38
Conventions Subtest	26.31***	.23
Meaning Subtest	9.81**	.10
ALRI		
Dependent Variable	F (1,50)	R ²
Number of Upper and Lowercase Letters Recognized	104.72***	.68
KSEALS		
Dependent Variable	F(1,52)	R ²
Composite	12.82**	.20
Vocabulary Subtest	11.76**	.18
NLW Subtest	5.99*	.10
Expressive Skills Subtest	13.08**	.20
Receptive Skills Subtest	10.04**	.16
Note. *** p<.001, **p<.01, *p<.05		

Table 11. Results of Repeated Measure ANOVAs for TERA-3, ALRI and KSEALS Scores

As seen in the table, all of the results indicated that the gains made by the children from pretesting to posttesting sessions were statistically significant (and not the result of chance occurrence) and, in almost all cases, the effect sizes exhibited indicated the robustness of these findings. These results again show improvement from those obtained in Year 1 wherein the gains exhibited in the TERA Reading Quotient and all subtests except Meaning were statistically significant however only the gains on the Composite and Receptive Language subtest scores for the KSEALS reached significance (all p<.05). It is important to note however, that given the limitations in the study design (e.g., absence of randomized control group design) the performance gains shown in either year cannot necessarily be attributed solely to children's participation in the ELLM program and instead could be the result of maturation and typical developmental progress (e.g., learning) that may occur over time.

Key Result: There were meaningful gains shown in children's scores on all measures however design limitations prohibit this performance to be attributed only to their participation in the ELLM program.

Did children's performance vary by gender? To further explore the nature of the demonstrated gains in the measures shown by children over time, mean differences in

performance as a function of the child's gender was explored. To assess this, multivariate ANOVAs were conducted separately on each of the measures to determine whether there were any significant differences in the pretest, posttest or difference scores between the girls and boys in the matched sample. The results of these comparisons indicated that gender was not a factor in children's performance on the measures as there were no identified significant differences between boys' and girls' performance on any of the measures.

Key Result: Children's performance on any of the measures was not shown to vary by gender.

Was children's performance impacted by teachers' degree of participation in the ELLM program? Given that the ELLM program is directly targeting the classroom teachers, it was of interest to determine whether the level of teachers' participation in the program was related to children's performance on the standardized measures. Three aspects of teacher participation were identified and included the variables of teacher turnover rates, overall length of participation in ELLM and their attendance at the summer training institutes. Each of these variables was examined separately in relation to children's performance on the measures.

Teacher turnover rates were examined first to assess the impact on the matched sample of children. This revealed that four of the four year old teachers (children measured with TERA-3 and ALRI) and two of the three year old teachers (children measured with KSEALS) left the classroom during Year 2. While this turnover likely affected all of the children in these classrooms (e.g., disruption and instability), it only impacted the scores from 19 children in the matched sample (i.e., 8 TERA-3/ALRI, 11 KSEALS). Given this small sample of scores, statistical analyses were not conducted on this variable.

The variable related to the length of a teacher's participation in the ELLM program was categorized into groups of teachers who participated since program inception (two years of participation), those that participated for the entire second year (one year of participation) and those that participated for only part of the second year (less than one year of participation). For the four year old classroom teachers, this categorization resulted in only three teachers being in the partial year only group and so for statistical purposes, children's scores in these teacher's classrooms were combined with those whose teachers had one full year of participation. Analysis of variance techniques were employed to determine if children's pretest, posttest or difference scores for the Reading Quotient or standard subscale scores on the TERA-3 differed as a function of the length of their teacher's participation in ELLM. Only one significant difference in the scores was revealed which indicated that children's standard alphabet subtest scores at posttest were significantly higher ($F(1,84)=7.28, p<.01$) for the children whose teachers participated in ELLM for two years compared to those whose teachers participated in ELLM for one year or less. There were no significant differences observed between the groups for children's ALRI pretest, posttest or difference scores.

Teachers in the three year old classrooms (children measured by KSEALS) were categorized into the three original groupings of two years participation, one year participation, less than one year participation although with an overall sample size of only 53 matched children with KSEALS scores, each of the groups were relatively small (n=10, 31, and 12, respectively). Again, ANOVAs were used to examine for significant differences in children's pretest, posttest or difference composite or standard subtest scores but the findings did not detect the presence of any significant differences in the scores. This is likely the result of the

small sample sizes of the groups and the relatively higher standard deviations indicated in the mean test scores.

Finally, the variable of teacher training attendance was examined in relation to children's performance on the measures. Teachers were categorized into two groups including those teachers who attended both summer training institutes and those that attended one or neither of the two summer trainings. While it may appear that this variable is measuring the same construct as the length in program, it should be noted that the latter group did not differentiate between teachers who attended the initial summer training at the inception of the ELLM program (i.e., 2002) or the one offered the following summer (i.e., 2003). Thus, it was possible that teachers who had been in the program since inception but did not attend both trainings could be included in this group.

Sets of univariate ANOVAs were conducted on the pretest, posttest, and difference scores of the children's performance on the Reading Quotient and standard subtests to examine for differences related to their teacher's training status. The findings revealed that for the posttest scores on the alphabet subtest, children whose teachers had attended both of the trainings performed significantly better (i.e., higher posttest scores; $F(1,85)=7.77, p<.01$) than children whose teachers attended only one or neither of the summer trainings. Similar findings were observed for the children's Reading Quotient posttest scores but just missed reaching statistical significance at the $p<.05$ level (i.e., $F(1,85) = 3.89, p=.052$). Comparable sets of ANOVAs were conducted for children's scores on the ALRI and KSEALS measures although these findings revealed no significant differences in any of the children's scores in relation to their teacher's participation levels in the trainings.

Key Result: The sample size was not sufficient to test for teacher turnover effects however children's TERA alphabet subtest scores were significantly greater for children whose teachers had participated in ELLM and the training sessions for more than one year.

Summary of Descriptive and Statistical Analyses of Family Measure

Because the ELLM program considers the family to be an important factor in a child's emergent literacy and language development, a pre-posttest family survey was designed to capture the level of literacy activities infused at home (e.g., library visits, writing/drawing). Even though there was a fairly high rate of completion of the pretest (i.e., 171 families), the return rate of the posttest surveys was extremely low (i.e., 31 families) resulting in an 82% attrition rate. This rate concurs with that obtained in the first year although there was higher total participation at both pretest and posttest (i.e., 310 and 55, respectively). Families rated the frequency of occurrence of literacy activities on a 5 point scale ranging from 1=Never to 5 = Very Often. Mean results of these ratings are shown in Figure 9.

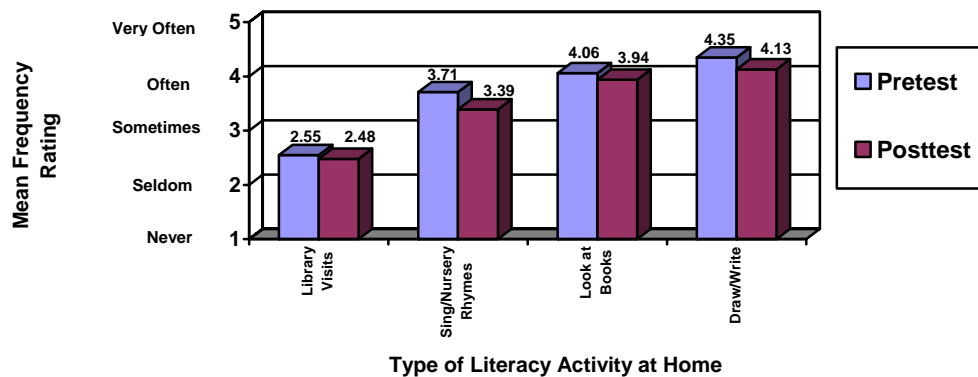


Figure 9. Mean Family Frequency Ratings by Type of Literacy Activity at Home

As shown in the figure, it appeared that families reported slightly less frequent occurrences of literacy activities in the home on the posttest than pretest. To test whether these mean differences from pretest to posttest were statistically significant the nonparametric Wilcoxon Signed Ranks tests were utilized. Results from these analyses showed that there were no significant mean differences in the frequency of literacy activities in the home over the program year. Given the small sample sizes it is likely that the greater variability in the means contributed to the lack of reliable findings. However, the posttest survey included an item that specifically asks families to report on whether they have seen any changes in their child’s interest in reading and books since they started in ELLM. The results on this item revealed that 22 of the 30 families (73%) reported that they had seen these changes. In Year 1, 90% of the families who completed posttests indicated similar observed behavioral changes.

Families who responded “yes” to this question were asked to provide additional open-ended information about these changes. Responses from Year 2 families included comments that helped confirm these positive observed changes (e.g., “reads more and writes a lot”; “eager to read and go to the library and choosing books”; “always reads books and asks me to read to him”; “he enjoys reading to me and he will initiate the reading”; “greater ability to understand, comprehend, retain it”; “he is spelling the words he sees”). Given the discrepancy in the quantitative and qualitative findings, it seems that a larger sample would be necessary to more effectively determine the outcome on these literacy related variables.

Key Result: Seventy-three percent of the families who responded indicated that they observed increases in their child’s interest in reading and books however, there was insufficient data to meaningfully test the carryover effects of ELLM in the home.

Discussion

This evaluation of ELLM represents an examination of the accomplishments that occurred primarily in the second year of the program (2003-2004), as this constituted the first full year of program implementation. The evaluation explored a variety of measured variables that served as indicators of the three project goals (i.e., teacher, child, family).

Limitations

There are several limitations that should be considered in the interpretation of the findings. First and foremost, the evaluation did not include a randomized control group study design whereby children were assigned to ELLM or control group (no ELLM intervention) classrooms. Given this limitation it is not possible to attribute any significant gains in the children's scores directly to the ELLM program in that it is equally likely that they may be the result of typical developmental progress (i.e., learning) that occurred over the program year. Secondly, for some of measures there were small sample sizes due to the low number of program participants (e.g., teachers) or as a result of attrition or non-participation at the time of posttesting. These small sample sizes either precluded the ability to perform statistical tests at all or in some cases required the use of nonparametric statistical tests that are less powerful in detecting potential program effects. Therefore, some program results may have gone undetected (e.g., teacher characteristics on children's performance). Finally, there were some variations in data collection (e.g., who collected data, how often it occurred) that may have contributed to increased variance in the scores that may also have rendered it more difficult to find program effects particularly within small samples. Despite these limitations though, the evaluation yielded some interesting findings and with methodological and design improvements over time these may prove to be reliable, robust program results.

Implications of Key Findings and Recommendations Regarding Teacher Participation and Performance

Teacher participation rates and attrition. As indicated previously, there was an overall attrition rate of 22% in teachers' participation in ELLM wherein 70% of the teachers who participated for the entire second year were returning teachers from the first year. National turnover rates for childcare staff have been estimated to range from 25% to 40% (Center for the Childcare Workforce, 2004). Compared to these national statistics, ELLM teacher turnover rates are lower than national averages. This slightly higher degree of stability shown in ELLM teachers may serve as a global indicator of the impact of the program on the teachers. Additionally, since ELLM began, there has been anecdotal evidence of the program's effect on turnover as several teachers who have left an ELLM center have actively sought employment in another ELLM center in order to remain with the program. Turnover continues to represent a significant obstacle for the achievement of high quality childcare services and an intervention such as ELLM that can help to even modestly reduce it provides a real benefit for the community and those directly impacted, including most importantly, the children served.

Turnover represents a challenge in several ways for the ELLM program. Teachers entering ELLM during the course of the ongoing program year are at a deficit in regards to their knowledge and understanding the model particularly since they have not attended the annual teacher training session. Additionally, the literacy coach and teacher relationship takes time to develop through the establishment of rapport and mutual trust. Teacher turnover can be a real setback for the program that ultimately will have an impact on the literacy progression of the children. Given the importance of maintaining stability in the classroom, the ELLM program should continually promote the benefits of the program to participating teachers and seek input from them on any added incentives that can be infused into the program to help facilitate retention. Additionally, the literacy coaches should consider providing ELLM orientation sessions on the model at varying intervals throughout the year to

assist in reducing the learning curve for teachers who are new to the program and to provide additional opportunities for relationship building.

Observed teacher performance in the classroom. Several different measures were used to assess the teachers' classroom performance in literacy instruction and fidelity in following the ELLM program principles. The TIS instrument measured both the use of transitions and activities in the classroom. The results for transitions showed that the coaches observed teachers using transitions between activities slightly more than half of the time however, their ratings suggested that teachers were not implementing them very effectively as mean ratings ranged between poor and adequate performance. Additionally, over the set of five randomly selected ratings, significant mean differences were shown however they were not indicative of incremental progress over time. Other TIS results showed that literacy coaches observed the teachers as most frequently using activities from sources other than the literacy packets or the resource center/coach and, similar to transitions, their mean effectiveness ratings in implementing these activities into the classroom were in the poor to adequate range.

The major themes present in the literacy coaches' free-text journals concurred with the TIS results wherein teachers seem to grasp the basic literacy concepts but were often lacking the skills necessary to effectively implement them into their classroom teaching. In particular, attention was needed to the age-appropriateness of selected activities and to their delivery of the material (e.g., length of time, group size, fun). Based on these corresponding findings it seems that in general, the teachers may need less concentrated focus on the content of what they are teaching but instead need greater assistance in developing the skills necessary to implement it more effectively. This is clearly an area that the literacy coaches may need to devote additional modeling and instruction/feedback on for the teachers and future evaluations can investigate for evidence of improvement.

Teacher performance was also measured within the structure of the classroom environment through the use of the ELLCO instrument. ELLM teachers' global performance was first compared to a geographic sample wherein the results appeared relatively similar between the samples. Further analyses of the ELLM teachers' scores revealed that mean performance over time (from pretest to posttest) was indicative of improvement however given the small sample sizes, these gains were only shown to be statistically significant (i.e., not the result of chance occurrence) for the Writing subscale of the Literacy Activities Rating Scale (e.g., instances of children writing or teachers assisting children with or modeling writing) and the Literacy Environment Checklist including the Books and Writing subscale scores (e.g., book area, selection, and use, writing materials and writing around the room).

To assess whether the TIS and the Classroom Observation measure of the ELLCO were measuring similar constructs of teachers' support of children's language and literacy development, these scores were correlated to determine their degree of relatedness. Results showed a high level of association between scores suggesting that the instruments were measuring the same construct. However, given that the correlation between the two TIS mean ratings (.894) was higher than the six correlations between the TIS and the ELLCO scores also suggests that there is utility in administering both measures to obtain additional granularity in measuring this construct. Thus, it seems warranted to continue to observe teacher performance with multiple measures to corroborate findings in general while also providing specificity of particular facets of their teaching performance. It is recommended however that future evaluations explore the use of independent assessors trained in administering the ELLCO instrument instead of the literacy coaches to help increase the reliability of the scores obtained

on this measure. Additionally, it would be prudent to establish some standardization of frequency of assessment as well as interrater reliability (e.g., peer review, independent assessment) on the TIS or similar observation tools used by the literacy coaches to ensure appropriate scoring and administration of these measures.

Teacher self-assessment of performance in the classroom. In addition to being observed by the literacy coaches, the teachers were asked to rate themselves and the program through two self-assessment instruments, the retrospective pretest and the satisfaction survey. The retrospective pretest permitted teachers the opportunity to learn and implement the ELLM program over the year and then to assess their behavior both before participating in ELLM and afterwards. Teachers first rated the frequency of incorporating six basic literacy activities (e.g., read aloud, alphabet activity) into their classroom teaching prior to and after participating in ELLM. Results suggested that for all teachers after a minimum of one year of participation in ELLM the mean reported frequency of use of all of the literacy activities had increased to at least four days per week.

In addition to the frequency of use, teachers were also asked to rate their perceptions about the importance of these literacy activities in children's literacy development. Similar results appeared wherein after one year in ELLM, mean ratings for each variable were at or close to the top of the range (i.e., 4=Very Important) and seemed to be maintained after the second year. Because of the high levels of reported frequency of use and the importance ratings of the literacy activities after one year of participation, there was less room for improvement to be reported in the second year of participation (e.g., ceiling effects). This result may question the utility of collecting this information beyond the initial year of participation in the program. Examination of these variables over an additional year may yield information about whether these seemingly apparent gains are maintained for teachers who continue to participate in the program and if so it may prove unproductive to continue to measure them over time.

Other information obtained from the retrospective pretest showed a 60% increase (i.e., all teachers) in the number of teachers reporting having a Word Wall in their classrooms after one year in ELLM. Additionally, results appeared to show noteworthy increases in their perceptions about both their and the children's use of the Word Wall during print/writing activities. However, as with the other items on this instrument it is unclear as to whether these gains will continue to be maintained at similar levels beyond two years of participation in ELLM.

The other self-report measure completed by the teachers included an annual satisfaction survey wherein they were asked to provide assessments about their general satisfaction with the ELLM program components, the importance of these components, any desired changes in the frequency of occurrence of them and their awareness of any behavioral changes that may have occurred for them or their students over the year. The results from the teacher ratings and comments seemed to indicate a high level of satisfaction with the elements of the ELLM program. The teachers, in general, and the second year teachers in particular, reported the highest satisfaction ratings for the individualized attention and support that was provided to them by the coaches. Similar to the retrospective pretest, mean ratings reached the maximum value of the scale (i.e., 5; ceiling effects) for even the first year teachers thereby leaving little room for improvement to be shown for the second year of ratings.

Interestingly, for the ratings on the importance of the program components, teachers ranked the books and literacy packets they received as being the most important program

element. This finding suggests that learning the skills to effectively infuse literacy development into the classroom is a necessary but not sufficient condition for success without the tools needed to effectively deliver the literacy lesson. Other results showed that teachers' ratings of the frequency of occurrence of program components also alluded to a generally high level of satisfaction with what they were receiving although there was the suggestion that coaching sessions and assistance with planning family activities could occur slightly more frequently. Finally, teachers reported being aware of positive changes in the children and their own classroom behavior and, as would be expected these appeared to a greater degree by the second than first year teachers.

Implications of Key Findings and Recommendations Regarding Children's Participation and Performance

Children's participation rates and attrition. A total of 248 children participated in the pretest session of the standardized assessment testing and 151 children were again present at the time of posttesting. This yielded a matched sample of 140 children. This sample was fairly equally balanced between boys and girls who ranged between 3 and 7.5 years old with a mean age of 4.5 years old at the time of pretesting. The attrition rate between pretest and posttest was computed to be 39%, which is notably greater than the 19% attrition shown the prior year. This higher attrition rate was most likely due to the timing of the pretest data collection which, during the first year due to startup, occurred in November when classroom enrollment was more stable whereas the second year it occurred in August/September when enrollment was more likely to fluctuate. The timing of the pretesting was changed to coincide with the beginning of the program year to ensure that it occurred prior to the start of coaching sessions. However, if future enrollment continues to be unstable until later in the fall thereby contributing to increased attrition rates, the timing of the testing may need to be reconsidered to maximize the number of children likely to be present at posttesting. Despite the high rate of attrition displayed, statistical analyses on the children's pretest scores did not reveal any attrition related effects between the pretest-only group and the pretest-posttest group.

Children's performance at pretest and posttest. Mean pretest, posttest and difference (i.e., change) scores were computed for the matched sample of children on each of the measures completed. Positive mean difference scores were obtained for all scores on each of the measures indicating gains in performance from pretest to posttest. Statistical analyses confirmed the reliability of these results. Additionally, there was no statistical evidence for the presence of order effects between mean scores on the two orders of item presentation for the ALRI measure.

Benchmark comparisons showed that TERA-3 Reading Quotient and KSEALS Composite pretest scores were within the national norm averages suggesting that the children were not entering the program at deficit levels compared to a national sample of their peers. Benchmarks for the ALRI were computed in terms of proficiency in letter recognition and results indicated that posttest scores for children in the ELLM program actually exceeded the national averages for children entering kindergarten. These results regarding children's performance represent important insights to be shared with the ELLM teachers in an effort to help alleviate any negative expectations regarding the children's abilities to learn and excel in these areas.

Several other variables were examined in relation to the children's performance on the measures to determine if they had any effect on the results. These included the child's gender

and several indicators of the teachers' level of participation in the program. Specifically, these included their turnover rates, overall length of participation in the program and their attendance at the summer training institutes. Statistical analyses were employed to examine for evidence of significant mean differences in relation to these variables. Results indicated that only the posttest alphabet subtest scores on the TERA-3 were significantly higher for children whose teachers had participated in the program for two years compared to one year and for teachers who attended both of the summer trainings. No other significant differences were revealed for any of the other scores. It is unclear the extent to which the small sample sizes on the teacher variables contributed to the inability to detect any other effects or if there are other potential indicators that would yield better results. This possibility should be explored in future evaluations.

Implications of Key Findings and Recommendations Regarding Families' Participation and Extent of Involvement

Families' participation rates and attrition. A total of 171 families completed the family survey pretest however only 31 of these families completed a posttest survey. This represents an attrition rate of 82%. Clearly, steps need to be taken to increase the families' level of participation and investment in completing the posttest survey in order to obtain an adequate sample size of matched families.

Families' performance at pretest and posttest. Families rated the frequency of literacy activities that occurred at home both at the beginning and end of the ELLM program year. Statistical analyses did not uncover any significant mean changes in ratings over time however 73% of the families reported that they had seen positive changes in their child's interest in reading and books since they started in ELLM. Because of the extremely high attrition rate the matched sample was very small and therefore likely contributed to the lack of statistically significant results. Obtaining a larger or more representative sample of families in the future would help determine the effectiveness of the family survey as a method for examining potential changes in literacy activities at home or whether a different approach (e.g., phone interviews, focus groups) could yield more reliable information with better participation rates.

Next Steps

Several next steps are proposed to continue to improve the evaluation efforts in support of the program. Since the program is currently in its fourth year of operation, some of the recommendations posed in this report have been considered already however it is important to convey them within the context of the findings revealed in this report.

- The study design limitations represent a significant barrier in fully testing the effectiveness of the model on children's performance. The attainment of a matched sample of control centers (i.e., centers not receiving the ELLM program) would provide the opportunity to parse out the portion of gains that is likely attributable to typical developmental progression by the children from that which may be directly an effect of the program. Given that a true randomized control study design is not feasible in a real-world applied childcare setting (i.e., children cannot be randomly assigned to receive ELLM or not), this quasi-experimental design would at least permit stronger conclusions about program effects than currently available without any comparison.
- The difficulty of obtaining control centers in the past has been related to the lack of incentive for the center to participate in the testing administration. One possible alternative is

to place these programs on a waitlist to receive ELLM services in the following or future year(s). The difficulty with this strategy however is that at some point it will increase the number of classrooms requiring coaching (i.e., coach's caseloads) on a regular basis. Other incentives besides receiving the full program complement may include financial incentives to the center or materials (books, literacy activities, access to resource centers) or a phased-in approach of gradually receiving more services over several years (e.g., financial, books/materials, trainings/workshops/continuing education, coaching). Sustainability efforts will be needed to determine when, if ever, ongoing ELLM teachers will step-down from the full complement of program services to enable new teachers to be phased into the program.

- The presence of small sample sizes primarily the result of attrition either prohibited statistical analyses from being conducted or hindered the ability to detect program effects when they may have occurred. A number of strategies for reducing attrition have been proposed in the previous section (e.g., timing of test administration, data collection methods used, additional incentives) that should be considered to continue to improve the likelihood of obtaining reliable results about the program's real degree of effectiveness.
- The evaluation measures/methods should continue to be refined to: (a) be consistent with best practices or industry standards for comparability, (b) advance the attainment of valid and reliable results through proper training and administration of the instruments, (c) eliminate unnecessary or ineffective data being collected (e.g., ceiling effects), and (d) be responsive to the needs and questions posed by the program to promote data-driven decision making for continued program improvement.

References

Center for the Childcare Workforce (2004). *Current Data on the Salaries and Benefits of the U.S. Early Childhood Education Workforce*. Washington, D.C.: American Federation of Teachers Educational Foundation.

Kaufman, A.S. & Kaufman, N.L. (1993). *Kaufman Survey of Early Academic and Language Skills Manual*. Circle Pines, MN: American Guidance Service Inc.

Reid, D. K., Hresko, W. P., & Hammill, D. D. (2001). *Test of Early Reading Ability Third Edition Examiner's Manual*. Austin, TX: Pro-ed.

Smith, M.W., & Dickinson, D.K. (2002). *User's Guide to the Early Language & Literacy Classroom Observation Toolkit Research Edition*. Baltimore, MD: Brookes Publishing Co.

U.S. Department of Education, National Center for Education Statistics (2001a). *The Early Childhood Longitudinal Study, Kindergarten Class of 1998-1999: Datafiles & Electronic Codebook*. NCES2001-028 {CD-ROM}. Also available at <http://www.ed.gov/pbus/edpubs.html>.

U.S. Department of Education, National Center for Education Statistics (2001b). *Entering Kindergarten: A Portrait of American Children When They Begin School: Findings from the Condition of Education*, Nickolas Zill & Jerry West, NCES2001-035. Washington, D.C.: U.S. Government Printing Office. Also available at <http://nces.ed.gov/pubs2001/20001023.pdf>.