

# INCH BY INCH, ROW BY ROW GONNA MAKE THIS GARDEN GROW\*:

## Classroom Quality and Language Skills in the Abbott Preschool Program

### Year One Report, 2002-2003 Early Learning Improvement Consortium

Released March 26, 2004

Cynthia Esposito Lamy, EdD, Rutgers University  
Ellen Frede, PhD, New Jersey Department of Education and The College of New Jersey  
Holly Seplocha, EdD and Janis Strasser, EdD, William Patterson State University  
Saigeetha Jambunathan, PhD and Jo Anne Juncker, EdD, New Jersey City University  
Heidi Ferrar, EdD, Montclair State University  
Lorraine Wiley, EdD, Rowan University  
Ellen Wolock, EdD, New Jersey Department of Education

\*From The Garden Song by David Mallett

## EXECUTIVE SUMMARY

### **What is the Abbott Preschool Program?**

- On May 21, 1998, the New Jersey Supreme Court landmark decision in *Abbott v. Burke* mandated that three- and four-year-old children in the 30 highest poverty districts in the state receive a high-quality preschool education. Both the preschool program and the 30 districts are now commonly referred to as “Abbotts.” New Jersey’s Abbott preschool program is ranked one of the highest in the nation for the level of quality, the resources committed to it and the proportion of children served (NIEER, 2004). Abbott districts provide full-day, full-year preschool programs to all eligible three- and four-year-olds. Through a Department of Education (DOE) and Department of Human Services (DHS) partnership, these classrooms now combine a DOE-funded six-hour, 180-day component with a DHS-funded wrap-around program that provides daily before- and after-care and summer programs.

### **How many children are currently served?**

- The goal of the Abbott mandate for preschool is to prepare children to enter kindergarten with the skills and abilities necessary to succeed in school. The key to reaching this goal is to ensure that programs are high-quality and that all children are included through strong outreach efforts. To this end, the Court permitted services to be provided through public schools, Head Start agencies or child-care programs, recognizing the value of each and reducing duplication of services. Expansion has been rapid. Since the first year of the program capacity has increased by over 23,000 children. The state has increased capacity by over 40% in the last two years. During the 2003–2004 school year, the fifth year of Abbott preschool implementation, the 30 Abbott districts enrolled over 38,000 three- and four-year-old children in preschool out of a possible universe of 54,000 children. Funding for over 42,000 three and four year old children has been approved for the 2004-2005 school year.

### **What is the Early Learning Improvement Consortium?**

- In an effort to assess the implementation of the preschool program, the Department of Education – Office of Early Childhood Education (DOE-OECE) formed the Early Learning Improvement Consortium (ELIC). The ELIC is a multi-year initiative in which participating New Jersey institutions of higher education assist the DOE and the Abbott districts in identifying the particularized needs of preschool children and programs by collecting and analyzing data on children and classrooms.

### **What does the ELIC investigate?**

- The primary purpose of the research conducted by the ELIC is to obtain information to improve policy and practice. In the fall of 2002, the ELIC administered tests of oral language development and early literacy skills to randomly selected kindergarten students statewide. This gives a general picture of children’s “readiness” to succeed in school. In the following winter and spring of 2003, faculty from the universities conducted structured classroom evaluations on 13 percent of the Abbott preschool classrooms to provide information on classroom practices likely to influence child learning.

### **What are the general findings of the study?**

- Despite rapid expansion, the majority of classroom scores are in the minimal to good range with 13% scoring good to excellent. The evidence suggests that as a result of preschool, children are entering kindergarten with improved language skills but still below average. There are promising indications of early literacy readiness abilities. Although the increases in program quality and children's language and literacy abilities are modest, they indicate that the classrooms are improving and are having a positive affect on children. The results reveal important areas for further improvement. Classroom quality is not high enough and, in some cases, is quite inadequate. The DOE and the early childhood education staff in the Abbott districts and contracting centers have been concentrating professional development and other initiatives on improving overall classroom quality, and increasing classroom supports for children's early literacy and mathematics skills, in particular.

### **What promising trends are there in classroom quality?**

- In 2002-2003 classroom quality scores are highest in those areas most likely to influence child learning, and these are the areas where the most growth in quality is occurring in the last few years. Teachers create a warm, nurturing environment and are improving in use of preventive management techniques that help children develop self-regulation and solve social problems. In particular, teachers are providing children with more opportunities to develop oral language, including rich vocabulary and complex sentence structure. Classrooms are better equipped in general but particularly with books, and teachers read more frequently to children. In classrooms, where children's heritage language is other than English, more teachers are supporting the child's home language in the classroom.

### **Does teacher experience influence the quality rating?**

- Improvement in classroom quality is happening at the same time as rapid expansion of the program. This rapid expansion has resulted in a large proportion of new and inexperienced teachers, many of whom are receiving their coursework in early childhood education during their first year of teaching. This lack of experience is clearly a factor in the quality and, indeed, we find a relationship between the teacher's experience and the classroom scores.

### **Do district characteristics such as size and level of poverty influence the quality?**

- Large districts and those with high poverty rates have lower quality scores. Expansion has been most dramatic in these districts and is likely one reason these districts have experienced the least growth in quality. More of the new and uncertified teachers are in the large districts.

### **Is the quality of the physical environment affecting the quality scores?**

- One important concern in the quality data is the overall low scores in classroom space and furnishings. One clear reason for this lower score is the lack of gross motor play space and equipment in many of the urban centers and schools. The average score on those specific items is below minimal. In the past two years, the DOE has been funding playground construction and equipment in many centers to help alleviate this

problem. Children who are spending 6 to 10 hours in a preschool setting need many opportunities to exercise and play outdoors. In addition, the number of preschool facilities being built with state construction funds is increasing dramatically but strategies for renovating existing facilities must be implemented to improve this aspect of the program. It is hoped that the results of the 2003-2004 observations will show the effects of these initiatives and the provision to new classrooms of almost \$20,000 in materials and equipment funds and to existing classrooms of over \$4,000.

### **What growth is found in kindergarten entry scores?**

- Analyses of the PPVT and TVIP scores show that children's oral language skills at kindergarten entry have increased as the preschool program has expanded and improved but are still well below the national average. It is important to note that regardless of whether the child attended preschool, all kindergarten children were included in the sampling pool. In 2002, approximately one third of the kindergarten children had not attended preschool and even more had not attended two years of preschool. Of those who had attended preschool, the quality of the program they experienced was lower than current programs. Given this, the results probably underestimate the effects of the current Abbott preschool program.
- A promising finding is the strong result on the Get Ready to Read screening tool. Although this likely overestimates the children's abilities, since it was designed for slightly younger children, no similar instrument exists for kindergarten age children. The skills measured, such as letter knowledge and linguistic awareness, are still highly relevant for kindergarten. Children's scores on this instrument reveal that the majority of kindergartners in Abbott districts are entering formal school with many of the early literacy skills necessary to become successful readers.

### **What initiatives at the state level are relevant to these findings?**

- Guidelines for facilities construction were developed by a task force representing the Department of Education, the Department of Human Services (DHS), child-care providers, Head Start agencies and others. These guidelines have informed amendments to the NJ Administrative Code (N.J.A.C.) 6:19-3. Facilities that meet these guidelines are currently being constructed or designed in almost every district. These will provide the quality environments which result in higher scores on classroom assessments but, more importantly, better programs for children.
- The Preschool Teaching and Learning Expectations: Standards of Quality (DOE, 2004) were revised based on the latest research in best practices for developmentally appropriate education for three-and four-year-old children. The revised Expectations consist of examples of high-quality teaching practices along with learner outcomes within each learning domain and offer significant assistance to the classroom teacher for planning instruction. The Expectations have received three favorable reviews from national organizations. Another project completed during the 2002-2003 school year to enhance the quality of Abbott preschool programs was the development of the
- Abbott Preschool Program Implementation Guidelines. The guidelines are recommendations based on the latest research and expert opinion and guide district's planning and implementation.

- In order to increase quality in the Abbott preschool classrooms, the Office of Early Childhood Education (OECE) offers a comprehensive, year-long training for master teachers who mentor and coach over 5,000 teachers and assistant teachers in the Abbott districts. The year-long course was designed to more clearly define the master teacher role and to ensure that master teachers have the skills they need to foster change and improve classroom quality.
- Typically, in the past, private child-care centers have suffered from a high rate of teacher turnover and a lack of well-trained teaching staff, thus limiting the ability to provide a high-quality program. In 2002–2003, in order to provide the high-quality programs that the court mandated, private provider teachers that received the proper training and held the appropriate certifications, received salaries comparable to in-district teacher salaries. Upgrading teacher salaries and qualifications will have the effect of stabilizing employment in the centers and creating a pool of well-trained and experienced teachers.
- High-quality educational programs undergo a continual cycle of gathering evidence about programs in order to make informed decisions toward improvement. The Self-Assessment Validation System (SAVS) is a system designed to guide the district through a systematic self-appraisal of its preschool program and to aid in program improvement. The SAVS is intended to highlight strengths of district programs and to alert districts to areas in need of improvement, which will inform program improvement.
- In addition to collecting and reporting on district data of children and classrooms, the ELIC and the OECE planned and developed the Early Learning Assessment System (ELAS), a performance-based assessment system administered by teachers during regular classroom activities. The ELAS is based on the latest research on development and learning in young children. With the ELAS, teachers learn how to observe children in the natural preschool environment on a regular basis and collect samples of work and record observations. This collected work is used to adjust the learning environment based on information about the children and to serve as a means of evaluating the skills of young children in Abbott districts on a statewide basis.

As The Garden Song by David Mallett begins “Inch by inch, row by row, gonna make this garden grow,” all of the teachers, administrators, advocates, parents and others working in Abbott preschool are sowing seeds and nurturing, scaffolding and otherwise supporting the growth of both this program and more importantly the children who participate. This report indicates that we are making progress – progress in outreach and enrollment, progress in raising teacher’s qualifications, progress in the quality of the classrooms and progress in children’s abilities to succeed in school. We have more hard work to do to fulfill the promise of Abbott preschool but the children deserve nothing less than our most dedicated efforts.

## Introduction

On May 21, 1998, the New Jersey Supreme Court landmark decision in *Abbott v. Burke* mandated that three- and four-year-old children in the 30 highest poverty districts in the state receive a high-quality preschool education. Both the preschool program and the 30 districts are now commonly referred to as “Abbotts.”

The Abbott preschool program is distinguished by several characteristics recognized recently in a report on preschool nationwide. New Jersey’s Abbott preschool program ranked as one of the highest in the nation for the level of quality, the resources committed to it and the proportion of children served (NIEER, 2004). Abbott districts provide full-day, full-year preschool programs to all eligible three- and four-year-olds. Through a Department of Education (DOE) and Department of Human Services (DHS) partnership, these classrooms now combine a DOE-funded six-hour, 180-day component with a DHS-funded wrap-around program that provides daily before- and after-care and summer programs. In total, the full-day, full-year program is available ten hours per day, 245 days a year.

The school day portion of Abbott preschool programs is staffed with one teacher and one aide and may not exceed 15 children. By September 2004, all teachers must hold a Bachelor’s degree and appropriate certification. This assures that teachers have the necessary expertise to work with young children. Districts provide one curriculum specialist to mentor teachers for every 10 – 20 classrooms, depending on classroom teacher experience. Health and social services are an integral part of the preschool program, and in addition to district social workers, child-care centers provide one family worker for every 45 children to ensure parents and children obtain referral to necessary services.

The goal of the Abbott mandate for preschool is to prepare children to enter kindergarten with the skills and abilities necessary to succeed in school. The key to reaching this goal is to ensure that programs are high-quality and that all children are included through strong outreach efforts. To this end, the Court permitted services to be provided through public schools, Head Start agencies or child-care programs, recognizing the value of each and reducing duplication of services. During the 2002–2003 school year, the fourth year of Abbott preschool implementation, the 30 Abbott districts enrolled 36,465 three- and four-year-old children in preschool out of a possible universe of 54,000 children, at a cost of approximately \$379 million. Thirty-one percent were served in school-based programs, seven percent in federally funded Head Start centers, and 62 percent in private child-care centers. Funding for over 42,000 three and four year old children has been approved for the 2004-2005 school year. Figure One below gives enrollment by year based on annual district reports. These numbers do not include children with disabilities.

### Preschool Children Served in Abbott Districts, 1998-2005

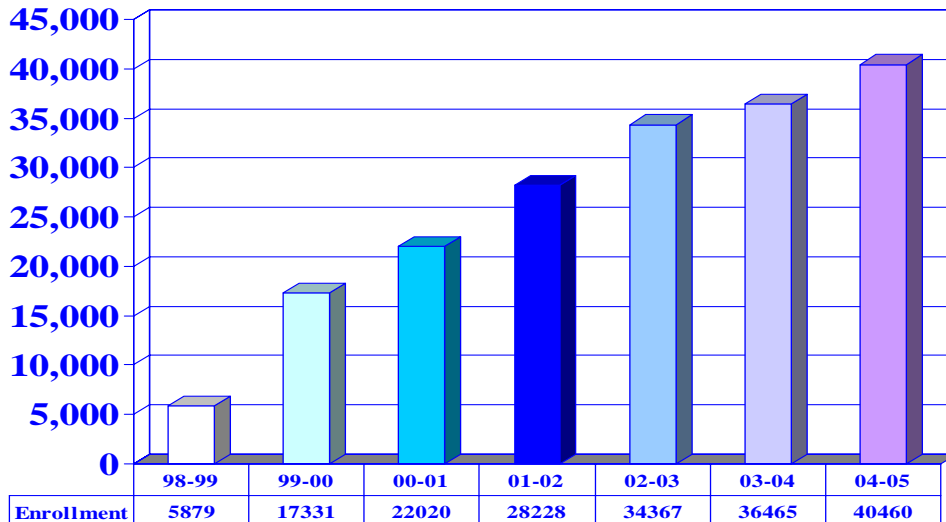


Figure 1.

Note: These figures do not include children with disabilities.  
The enrollment for 04-05 is projected based on district plans.

Through scholarships provided by the Department of Human Services, adequate salaries for teachers funded by DOE, and through the addition of more classrooms, the percent of teachers with bachelor's degrees has increased dramatically. According to district report as of October 15, 2003, over 92% of the teachers had a bachelor's degree or higher, up from an estimated 38% in 1999 (Barnett, Tarr, & Frede, 1999).

#### The Early Learning Improvement Consortium

The long-term benefits of preschool participation are clear in the research literature (Barnett, 1998). Children who participate in preschool have the following characteristics later in life:

- Increased abilities in mathematics and literacy.
- Greater success in school, graduation rates from high school and attendance at college.
- Reduced retention in grade and special education placement.
- Lower participation in crime.
- Improved employment as an adult and increased contribution to the community through volunteer work.

What is equally clear is that these benefits only accrue when the preschool program is high quality (Frede, 1998).

In an effort to assess the implementation of the preschool program, the Department of Education – Office of Early Childhood Education (DOE-OECE) formed the Early Learning Improvement Consortium (ELIC). The ELIC is a multi-year initiative in which participating New Jersey institutions of higher education assist the DOE – OECE and the Abbott districts in identifying the particularized needs of preschool children and programs by collecting and analyzing data on children and classrooms. In the fall of 2002, the ELIC administered tests of oral language development and early literacy skills to randomly selected kindergarten students in each district. This gives a general picture of children’s “readiness” to succeed in school. In the following winter and spring of 2003, faculty from the universities conducted structured classroom evaluations on 13 percent of the Abbott classrooms. The primary purpose of this data collection was to obtain information to inform policy and practice.

This report presents the findings of the first year of the ELIC study. In addition, comparisons to similar information collected by the Center for Early Education Research in 1999–2000 and 2000–2001 are made to measure change over time (Barnett, Tarr, Esposito-Lamy & Frede, 2002). The remainder of this report presents procedures and findings. Technical information on instrumentation, procedures and data analysis are included in the appendices.

### Methods

Classroom quality was measured using three instruments: the Early Childhood Environment Rating Scale – Revised (ECERS-R; Harms, Clifford & Cryer, 1998), the Support for Early Literacy Assessment (SELA; Smith, Davidson & Weisenfeld, 2001) and the Preschool Classroom Mathematics Inventory (PCMI; Frede, Dessewffy, Hornbeck & Worth, 2001). The ECERS-R, a widely used tool, was chosen to provide a comprehensive look at classroom quality and to allow comparison of New Jersey’s scores to scores from other states and other large projects. The SELA provides more specific information on classroom practices that support children’s early language and literacy skills. The PCMI, a new instrument, focuses on the materials and methods used in preschool classrooms to support and enhance children’s math skills.

A random sample of 310 preschool classrooms across the 30 Abbott districts, 11-14% of classrooms in each district, stratified by public school or private provider, were observed between January and April of 2003. For further information on the psychometric validity of the classroom observation instruments and the administration of these measures see Appendix A.

Kindergarten children’s receptive language skills and pre-reading skills were measured using three instruments: the Peabody Picture Vocabulary Test-III (PPVT-3) (Dunn & Dunn, 1997), the Test de Vocabulario en Imágenes Peabody (TVIP) (Dunn, Padilla, Lugo & Dunn, 1986) for Spanish-speaking children; and the Get Ready to Read (Whitehurst & Lonigan, 2000). Receptive language skills were measured using the PPVT-III or the TVIP in a randomly selected sample of 2355 Kindergartners across the 30 Abbott districts. The Get Ready to Read Instrument was administered to a subset of



1251 children in 21 districts. The tests were individually administered to the children during November and December, 2002. Children who spoke languages other than English or Spanish, or who had been retained in kindergarten, were not included in the sample. More information on the child language and literacy measures can be found in Appendix A.

Findings for Classroom Quality

ECERS-R

The ECERS-R rates classroom quality on a 7-point Likert scale, indicating a range of quality from inadequate (1) to excellent (7). The seven ECERS-R subscales are as follows: Space and Furnishings, Personal Care Routines, Language-Reasoning, Activities, Interaction, Program Structure, and Parents and Staff. In 2002-2003, the average ECERS-R score across all sample classrooms is 3.96. This score falls midway between minimal and good quality. The classroom total ECERS-R scores in this sample range from 1.6 to 6.3. Nearly 13% score 5 or better, placing them in the good to excellent quality range. However, 17% of the sample classrooms score below 3, indicating minimal to inadequate support for children’s cognitive and social development. See Figure 2 below.

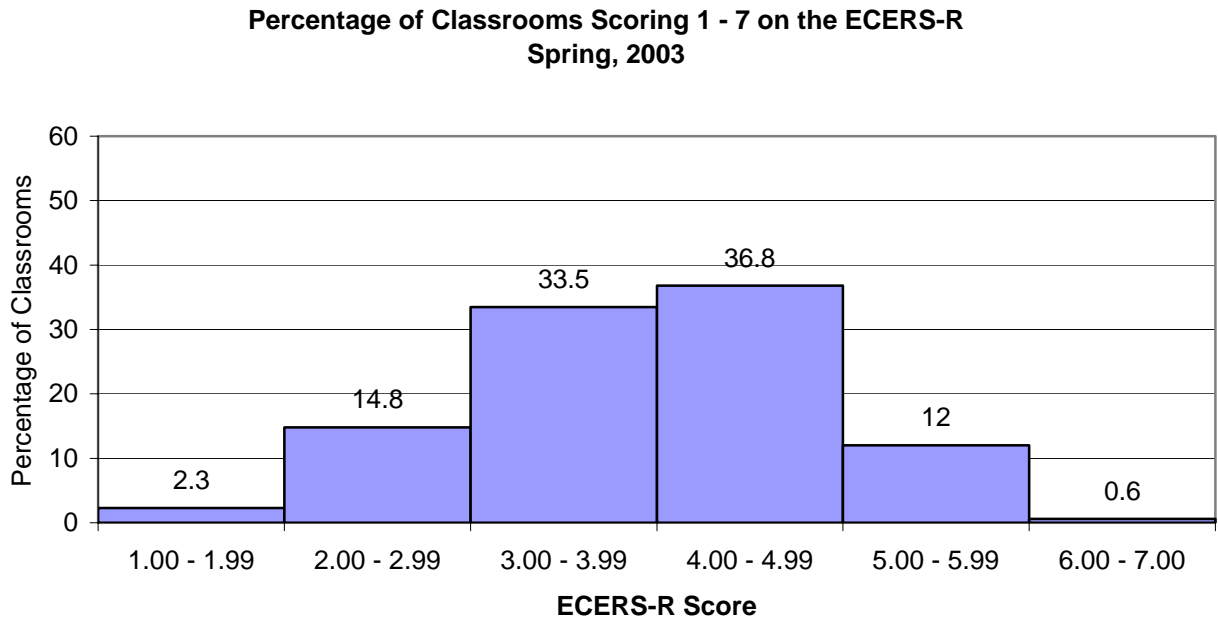


Figure 2.

The seven subscales of the ECERS-R measure different aspects of classroom quality. Table 1 below reports the ECERS-R average subscale scores and the average total score in 2003.

**Table 1.**  
**ECERS-R Total and Subscale Scores across 30 Abbott districts**  
**Spring, 2003**

<u>ECERS-R subscale</u>	<u>Average Score</u>	<u>Range</u>	<u>Standard Deviation</u>
Space and Furnishings	3.76	1.38 – 6.63	1.00
Personal Care	3.69	1.00 – 7.00	1.35
Language & Reasoning	4.27	1.00 – 7.00	1.30
Activities	3.37	1.20 – 7.00	0.94
Interactions	4.92	1.00 – 7.00	1.60
Program Structure	4.04	1.00 – 7.00	1.57
Parents and Staff	4.37	1.00 – 6.83	1.03
Overall ECERS-R score	3.96	1.60 – 6.30	.94

The subscale scores range from a low of 3.37, just above minimal, for Activities to a high of 4.92, very close to good, for Interactions. The wide ranges in each subscale show that across all areas of classroom quality – basic environment and care giving to intellectually challenging and intentional teaching practices – there are classrooms that are completely inadequate and others that are excellent or approaching excellence. For the most part, classrooms score better than minimal but less than good. Four subscales particularly relevant to educational effectiveness of the program: Language and Reasoning, Activities, Interactions and Program Structure. Abbott preschool classrooms score substantially better on two of these subscales than they do on the instrument as a whole. These measure the extent of practices such as the following:

- Language and Reasoning – good classrooms have a wide selection of books and other language materials are also found in the classroom, such as a felt board, recorded stories, games with words and pictures, puppets, and small figures. Adults read to children daily in this classroom, both formally, during whole group times, and informally during free play. The communication encourages children to develop verbal and written skills. Staff encourages this by modeling slightly more complex language than the children produce and by expanding on the ideas that children present. Adults and children can be heard discussing concepts including same/different, matching, sequence, one-to-one correspondence, classification, size, spatial relationships while children are playing with materials or working on an interesting task.
- Activities – In high scoring classrooms, materials that enhance learning and play are accessible to children for a substantial portion of the day. A “substantial portion” is defined as at least one-third of the time that most of the children are in daily attendance. A wide variety of materials are stocked on well-organized shelves, low so that children can reach without trouble or too much waiting. Materials that encourage development across all subject areas (math, science, art, music, technology) are available.
- Interactions – Effective teachers support learning by creating a positive environment for the exchange of ideas by helping, encouraging, appreciating, and

- extending what children are doing. They use preventive management techniques, by setting up the environment to reduce the chance of conflict, redirecting potential problem behavior, modeling good social skills including cooperation and compromise, and by consistently reacting to negative interactions with non-punitive support of the development of children's social skills.
- Program Structure - The classroom schedule provides a balance of structure and flexibility during the day, with some activities child-initiated and some teacher-directed, and a substantial portion of the day used for play activities. Free play is supported with plenty of materials and with adult supervision and interaction, which is understood as an opportunity for educational scaffolding. Transitions are smooth with no long waiting, and variation is made in the schedule to meet individual children's needs. Whole group time is limited to no more than 20 minutes, with most activities and routines done in small groups, and educational interactions taking place with small groups or individual children as well as with the whole group together. Additionally, children have many opportunities to select a group to be with during the day.

In analyzing items in the Activities subscale we find evidence that the lower score here reflects the fact that children are not given an opportunity to use materials for a "long enough" portion of each day. Further information about each ECERS-R subscale is reported in Appendix B.

### Change Over Time in ECERS-R Scores

The Center for Early Education Research (CEER) at Rutgers University collected child and classroom data during the 1999/2000 school year in 20 Abbott districts, and continued to collect data in subsequent years until the ELIC was formed. Thus, we can compare ECERS-R scores over four years. Analyses were performed using data (n=514) from the 1999/2000 and 2002/2003 school years, for the subset of districts for which data was collected in the first year. We find statistically significant improvements in scores for the Language and Reasoning, Activities, and Interactions subscales of the ECERS-R. Although the average total score also rose from 3.86 to 3.92 (this is different from the 3.96 figure stated previously for 2003, since it is from a sample of 20 districts, not the whole sample), this change is not statistically significant (at  $p < .05$ ).

The average Language and Reasoning subscale score rose from 3.74 in 1999 to 4.16 in 2003. The average Activities score rose from 3.18 to 3.37, and the average Interactions score rose from 4.47 to 4.83. The average Personal Care and Parents and Staff subscale scores were significantly lower in 2003, however, with the average Personal Care subscale score going from 3.98 to 3.71, and the average Parents and Staff subscale score going from 4.59 to 4.34. The average score for the Space and Furnishings subscale was unchanged over four years (3.73 in 1999 and 3.70 in 2003). See Figure 3 below. See Appendix C for statistical information.

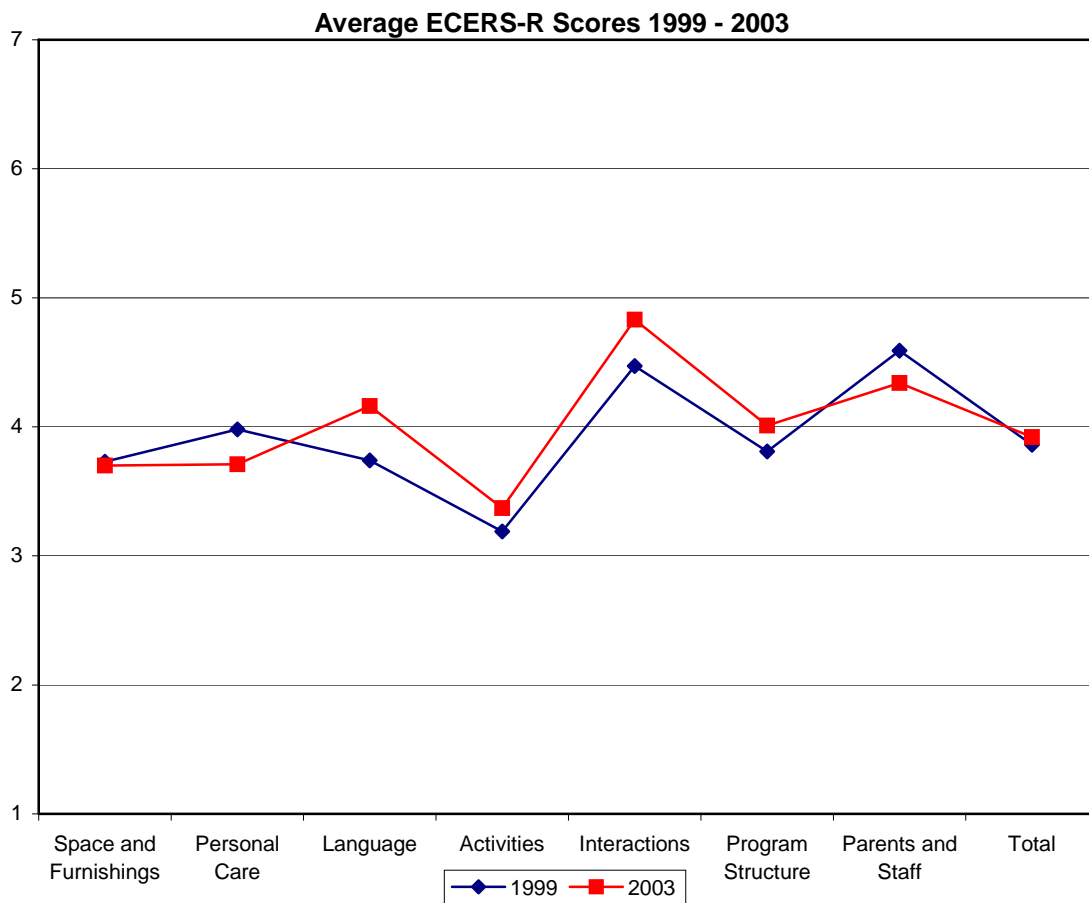


Figure 3.

### The Influence of Facilities versus Teacher Process

The ECERS-R is an “environmental” rating and assesses the materials available and the facilities, as well as the teachers use of these, and effective teaching strategies. To examine the extent to which facilities issues may be keeping classrooms from attaining higher quality scores, the items specific to facilities were used to create a “facilities influence” score. The items included in the facilities score are Items 1, 2, 7 and 8 – Indoor Space, Furnishings for Work and Play, Outdoor Space and Gross Motor Equipment. These items measure whether the space is large enough for the number of children enrolled, whether it is in good repair, whether there is enough furniture for everyone, whether the outdoor space is safe and accessible to the children, and whether there is enough safe and appropriate playground equipment. Analyses indicate that the average facilities factor score in 1999 is 3.73, while the average score in 2003 is 3.66, down slightly but not a statistically significant change.

The analyses of facilities influence scores indicate that facilities limitations tend to depress overall ECERS-R scores. However, facilities improvements are difficult to make overall when there is a need to expand and use all adequate facilities. At the same time, much effort has been put forth over the past few years at state, district and program levels to raise the qualifications of teachers for the purpose of improving classroom quality. The question arises whether aspects of classroom quality that are controlled by teachers has improved.

To address this question, items directly related to teacher process were included in a “teacher influence score” – items 5, 6, 9, 16, 17, 18, 29, 30, 31, and 32. These items measure criteria such as the teacher’s interactions, methods of discipline, and use of language with the children, along with how the teacher uses the classroom space and what materials and activities the teacher allows the children to use during the day. Analyses indicate that the average teacher influence score in 1999 is 4.18, while the average teacher influence score in 2003 is 4.42, a statistically significant and programmatically meaningful increase. See Appendix C for more detailed statistical information.

### Other Influences on Quality Scores

#### The Influence of Teacher Experience

The effort to rapidly expand preschool programs to serve all three and four-year-olds in the Abbott districts has led to the recruitment of many new teachers. First year teachers comprise 9% of our sample of teachers in 1999 and 15% in 2003. These new teachers, while benefiting from their recently completed education, nevertheless have little practical experience in the classroom. To address the question of whether the influx of so many first year teachers is having an affect on quality scores, analyses comparing the scores of first year teachers and non-first year teachers were compared, using 1999 and 2003 data.

Results indicate that while there is a difference in scores across all subscales, the differences in teacher influence score and the total scale score between first year and non-first year teachers in 1999 were negligible. By 2003, the difference in scores between these groups is highly significant for the total score, the teacher process score and for those subscales that are strongly related to teacher process – Language and Reasoning and Interactions. Teachers in their first year of teaching score significantly lower than teachers with more experience. See Figure 4 below and Appendix C for more detailed statistical information.

**Differences in Scores Between First Year and Non-First year Teachers, 1999 - 2003**

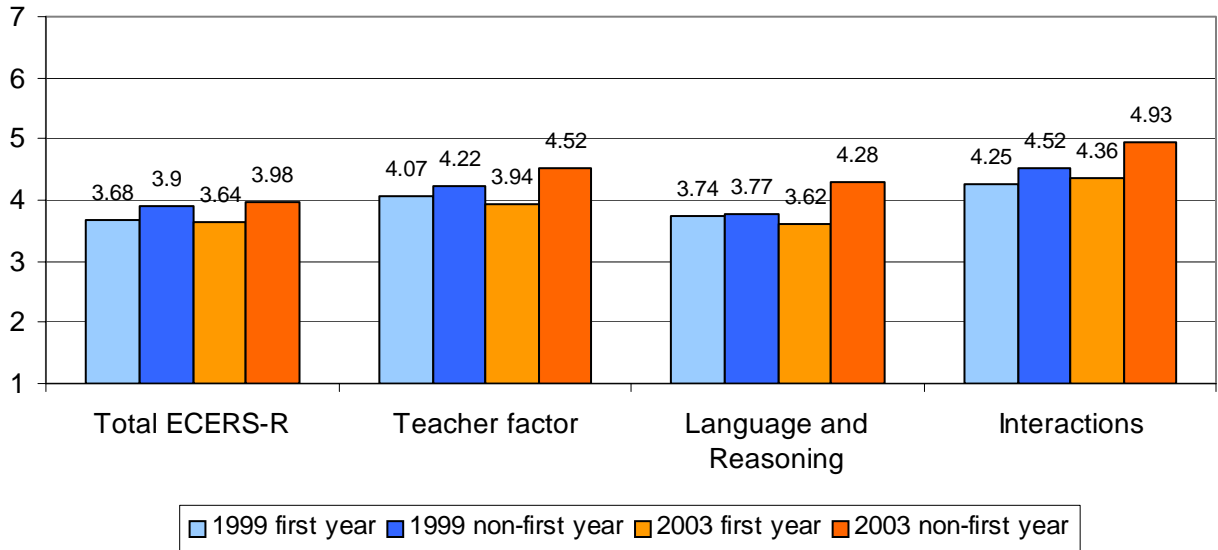


Figure 4.

### The Influence of Teacher Credentials

Research has shown that better educated teachers, especially those who also have specialized training to teach young children, provide a higher-quality preschool classroom environment than teachers without such education and certification (NIEER, 2003). Analyses were conducted to determine the influence of teacher education on classroom quality scores, accounting for the influence of teacher experience. Results of analyses indicate that teacher education had a highly significant affect on scores in 1999. In 2003 no statistically significant difference was found. This is most likely due to the requirement that Abbott preschool teachers complete their Bachelors degree and hold P-3 certification by September, 2004. Thus, there are very few teachers in our 2003 sample without these qualifications. Moreover, of the 11% of sample teachers without a Bachelors degree in 2003, 80% report being enrolled in a program leading to certification. See Figure 5 and Appendix C for more detailed statistical information.

### Teacher Education and ECERS-R Scores, 1999 - 2003

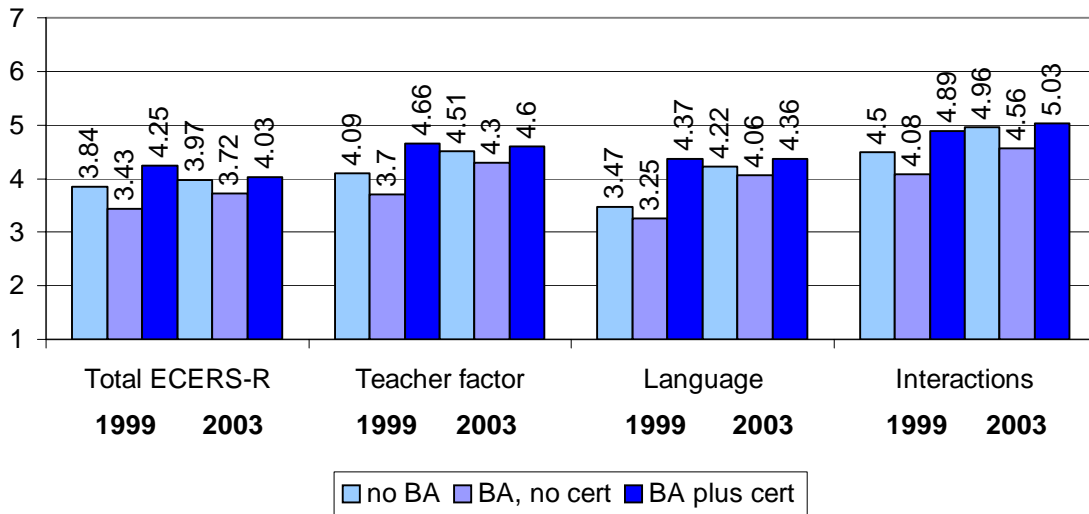


Figure 5.

### The Influence of District Characteristics

District-level characteristics such as the size of the district and the extent of poverty could be related to scores. Larger districts may face more difficult logistical issues in hiring and training teachers. Districts with higher levels of poverty among families may require more supports for families and children. In these analyses, district size is indicated by three categories based on the number of three- and four-year-olds enrolled in preschool. District poverty is indicated by the percentage of children receiving free or reduced price lunch. District poverty is divided into two groups (below 66% poverty and above 66% poverty).

Results of analyses indicate that classrooms in larger districts score significantly lower than classrooms in small districts on Space and Furnishings, Personal Care, Language, Interactions and Parents and Staff subscales; and on the facilities and teacher influence scores and the total ECERS-R scale. Districts with a higher percentage of families in poverty score significantly lower on the Interactions and the Parents and Staff subscales, and on the Teacher influence score. Figure 6 below displays the amount of unit change in the ECERS-R score related to the variable. See Appendix C for more detailed statistical information.

It is important to note that these analyses investigate differences in classroom quality scores by factors associated with districts, but not simply by district. It is not helpful to discuss differences in classroom quality scores by district, since many factors that impact the scores, including the factors mentioned above, are statistically confounded with district.

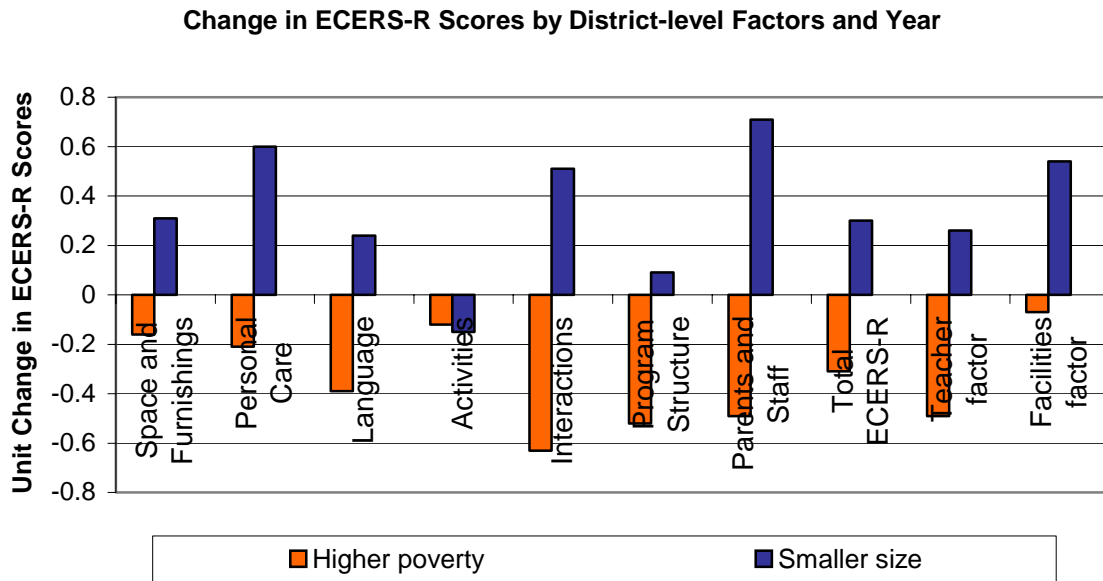


Figure 6.

A Comparison of Abbott ECERS-R Scores with Other Places and Projects

How does the classroom quality in New Jersey’s Abbott districts compare to the classroom quality found in other states and in other large research projects, as measured by the ECERS-R? Figure 7 below displays the average ECERS-R score found for the Abbott districts in the 2002/3 school year along with the average scores found for several other projects. Abbott district scores are roughly comparable to the other average scores.

**A Comparison of New Jersey's Abbott Preschool Program ECERS-R Scores with ECERS Scores of Other Projects**

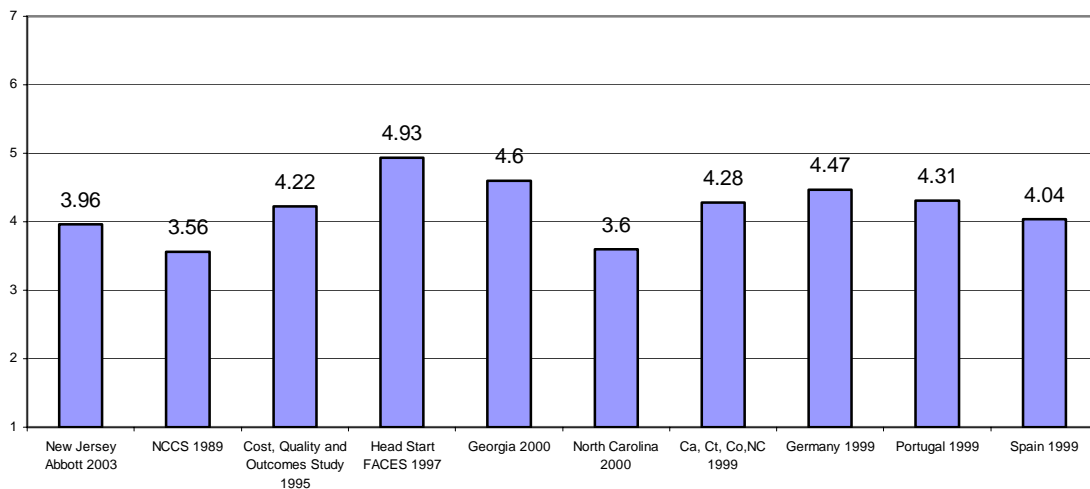


Figure 7.



## Supports for Early Literacy Assessment (SELA)

The content of the SELA comes from research in early language and literacy development and the joint position statement of the IRA/NAEYC (1998). The position statement is based on careful review of relevant research on effective classroom practices that enhance language and literacy development in the early years. The SELA measures the classroom environment and teaching practices that lead to early literacy and language development. The average total SELA score is 2.86 with a standard deviation of .78. On a scale of 1 to 5, 1 representing very low quality and 5 representing high quality, or the ideal, this score indicates that the average Abbott preschool classroom can be characterized as having mediocre support available for children's language and literacy development. Scores range from 1.00 to 5.00, with about 10% of classrooms at or near the ideal (a score of 4 to 5), and almost 12% scoring in the low to poor quality range (a score of 1 to 2). See Figure 8 below.

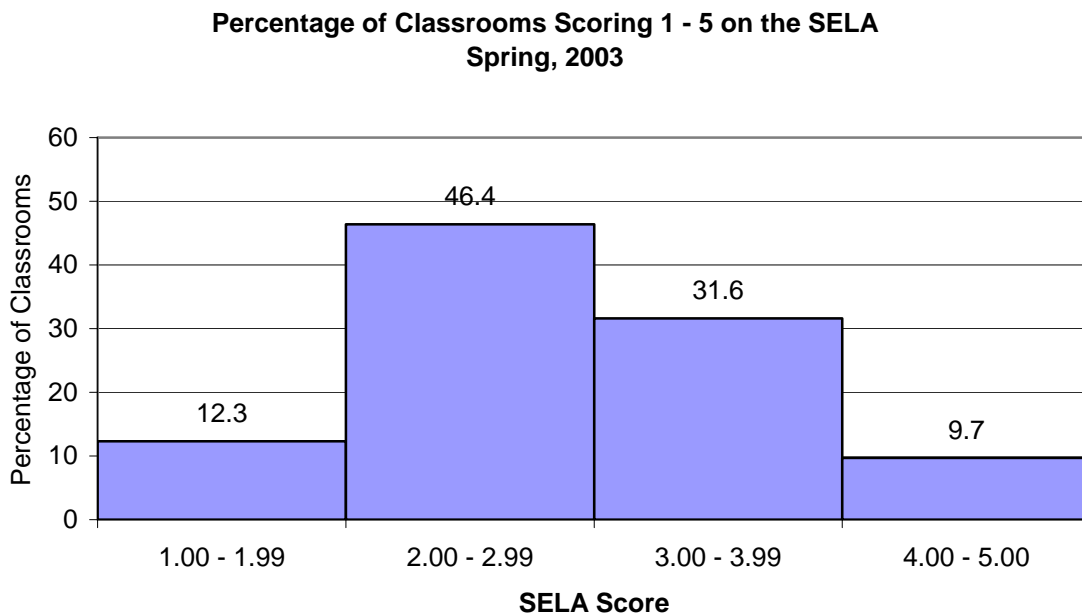


Figure 8.

The six subscales of the SELA measure different aspects of classroom quality as it is related to children's language and literacy development. Table 2 below lists the SELA average subscale scores and the average total score in 2003.

**Table 2.**  
**SELA Scores Across 30 Abbott Districts**  
**Spring, 2003**

SELA Subscale	Average Score	Range	Standard Deviation
Literate Environment	3.15	1.0 – 5.0	.87
Language Development	3.16	1.0 – 5.0	.90
Knowledge of Print/ Book Concepts	2.82	1.0 – 5.0	1.28
Phonological Awareness	1.99	1.0 – 5.0	1.21
Letters and Words	2.65	1.0 – 5.0	1.07
Parent Involvement	2.39	1.0 – 5.0	1.09
Native Language	2.35	1.0 – 5.0	1.24
Total SELA score	2.86	1.0 – 5.0	.78

The two highest-scoring subscales are the Literate Environment and the Language Development subscales, at 3.15 and 3.16 respectively. The Literate Environment subscale consists of 5 items which measure whether physical environment and materials of the classroom supports of children’s literacy skills: the use of print in the classroom, the quality of the book area and the use of books throughout the classroom, the availability of writing materials and the presence of items to support literacy in the dramatic play area (such as menus, price tags, paper and pencil). The Language Development subscale consists of 4 items which measure the teacher’s encouragement of children’s use of language through the teacher’s extension of children’s ideas, introduction of new vocabulary and linguistic structures (A child exclaims, “Look! I made a bridge!” The teacher responds, “You put supports in three places to hold up the three parts of your bridge.”), development of activities in the classroom whereby children can build knowledge and use language, and the sharing of books with children. For both subscales, about 21% of classrooms score at or near the ideal.

The lowest scoring subscale is Phonological Awareness, with an average score of 1.99. This subscale is comprised of only one item that measures the extent to which the teacher draws attention to the sounds that children hear in words – for instance, “Hey! Your name is Jenna and your sister’s name is Janell – both begin with the /j/ sound!” Nearly 50% of sample classrooms score a 1 on this item, while just over 11% score at or near the ideal.

#### Change Over Time in SELA Scores

The Center for Early Education at Rutgers University included the SELA during classroom observations for the 2001/2002 school year. Therefore, we can examine the change in SELA scores between 2001/2002 and 2002/2003. As with the analyses of change over time in ECERS-R scores, the subset of districts for which data were

collected during the first year is used in these analyses, which is why the average scores are slightly different from those reported above.

Current SELA scores indicate that much work is yet to be done; however, the change in scores over the past two years indicates that important progress is being made. The average total SELA score rose from 2.42 in 2002 to 2.82 in 2003, a statistically significant increase representing a substantial improvement in classroom support of children’s literacy development. In fact, average scores across all areas of the SELA rose substantially, and every increase is statistically significant. See Figure 9 and Appendix C for more detailed statistical information.

While all areas of literacy support as measured by the SELA improved, the greatest increase is found in the scores for Knowledge of Print/Book Concepts, rising from 2.21 in 2002 to 2.77 in 2003, more than half a scale point. Scores for the Letters and Words subscale also rose more than half a point, from 2.06 in 2002 to 2.58 in 2003. Scores for Phonological Development, although improved, remain low.

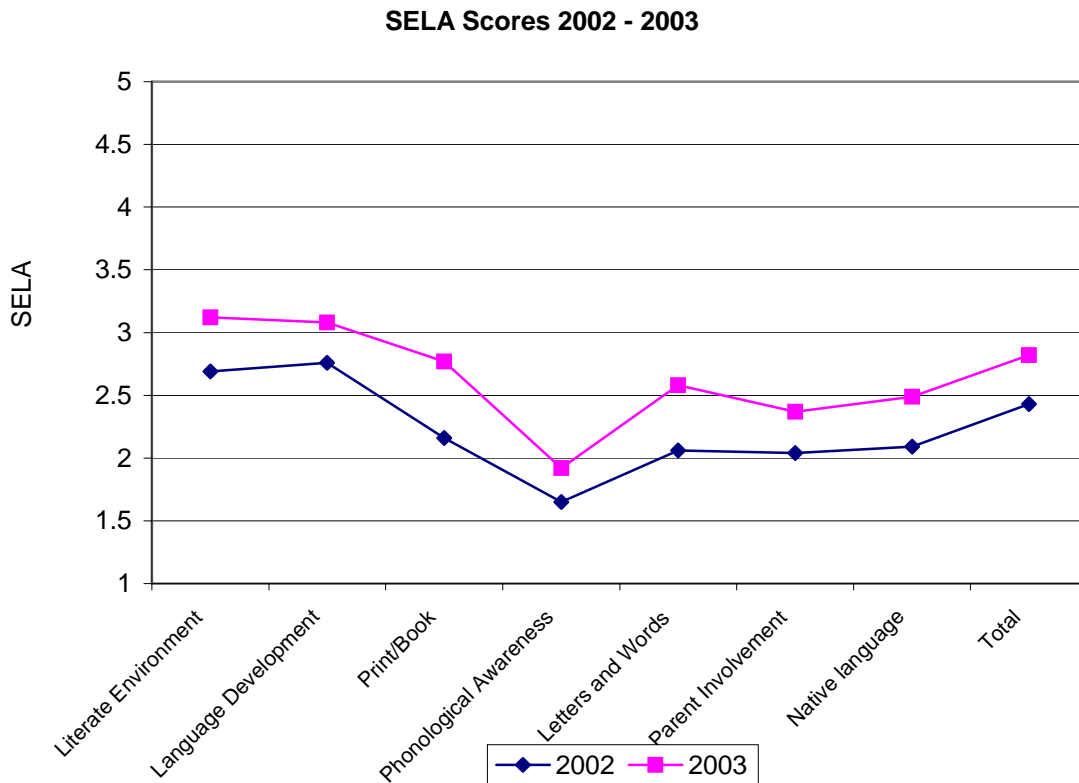


Figure 9.

## Preschool Classroom Mathematics Inventory (PCMI)

Classroom support for the development of children's early mathematical skills is measured using the Preschool Classroom Mathematics Inventory (PCMI) (Frede, Dessewffy, Hornbeck & Worth, 2001). This tool rates the materials and strategies used in the classroom to support children's early mathematical concept development, including counting, comparing, estimating, recognizing number symbols, classifying, seriating, geometric shapes and spatial relations. The standards of the National Council of Teachers of Mathematics and the National Association for the Education of Young Children inform the measure, which is comprised of 11 items on a 5-point scale, from low quality (1) to high quality (5).

The average PCMI total scale score across the sample classrooms is 1.91 (standard deviation is .65), indicating that the average Abbott preschool provides very limited support for children's mathematical skill development. The vast majority of classrooms score 1 or 2 on this scale. Note that this is the first large-scale use of a tool to measure the extent to which math knowledge and skills are supported in Abbott preschool classrooms. See Figure 10 below.

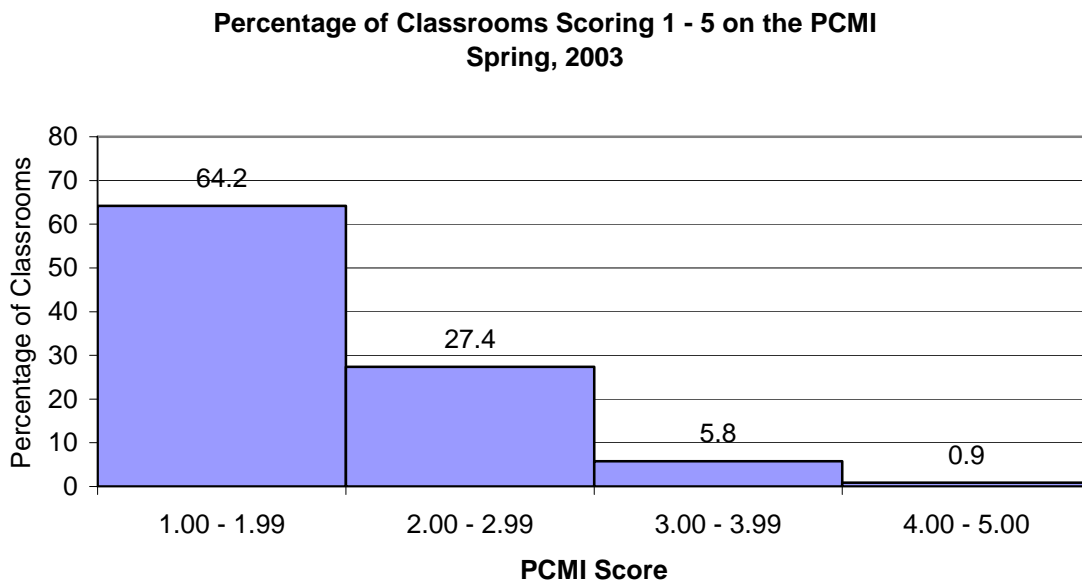


Figure 10.

The three subscales of the PCMI measure different aspects of classroom quality as it is related to children's mathematical skill development. Table 3 below reports the PCMI average subscale scores and the average total score in 2003.

**Table 3.**  
**PCMI Scores Across 30 Abbott Districts**  
**Spring, 2003**

PCMI Subscale	Average Score	Range	Standard Deviation
Materials	2.41	1.0 – 5.0	.90
Numeracy and Other Mathematical Concepts	1.68	1.0 – 4.57	.64
Parent Involvement	1.53	1.0 – 5.0	.95
Total PCMI Score	1.91	1.0 – 4.68	.65

The average total and subscale scores on this measure are quite low. The scores for the Materials and the Numeracy subscales (2.41 and 1.68 respectively) suggest that while materials that support children’s mathematical concept development need to be more available in the classrooms, considerable professional development is needed for teachers to develop the methods and activities to use along with the materials to appropriately support and scaffold children’s concept development in the classroom. It is revealing that some classrooms actually scored close to perfect on this instrument, indicating that it is possible to accomplish.

What does a high-scoring classroom on the PCMI look like? There are plenty of math-related materials to be found:

- Many diverse items for children to count – small manipulatives such as “counting bears”, rocks, buttons, blocks, small cars;
- Many examples of numerals posted for a purpose around the room, on posters, and on cash registers, number puzzles, and games, and in books.
- There are materials that encourage the understanding of one-to-one correspondence, such as pegs and pegboards, jars and lids, puzzles, Memory cards.
- Measuring can be attempted using tape measures, scales, and measuring cups, with many things to measure nearby.
- There are many items to sort in the classroom – collections of shells, beads, dishes, small manipulatives, that children can sort by function, texture, color, shape, and size.
- The children’s understanding of seriation is encouraged with materials that can be put in order from smallest to largest.
- Finally, there are plenty of geometric shapes found in the classroom – geoboards, parquetry blocks, stencils, and things with which children can create shapes – clay, yarn, construction paper and glue.

Teachers support children’s early mathematical skills by incorporating mathematical concepts throughout the day, in routines, in small groups and with individual children. For instance, children are encouraged to count each other and determine how many are absent; to count the number of napkins needed at the table; to estimate whose mound of

play dough is biggest or how many cars can fit on the track. Teachers model mathematical terminology – “When you combine the two triangles you create a new shape - a rectangle with four sides and four corners.” Teachers encourage children to compare, measure, classify and seriate, and call attention to shapes, spatial relationships, and patterns.

### Findings for Kindergartners’ Receptive Language Scores

A sample of 2355 kindergarteners across all 30 Abbott districts were assessed using the PPVT-3 (Dunn & Dunn, 1997), and also with the TVIP (Dunn, Padilla, Lugo & Dunn, 1986) when the children spoke Spanish as their first language. In cases where both the PPVT-3 and the TVIP were employed (n=287, about 12% of the sample), the highest score is used. To estimate average scores for the Abbott preschool programs, analyses employed weights based on the number of children in each district.

The PPVT and TVIP are nationally normed with an average standard score of 100 and a standard deviation of 15. In a typical American population, one would expect about 50% of children to score below the mean of 100, about 20% to score below 85 which is one standard deviation below the mean, and about 3% to score below 70, two standard deviations below the mean. In 2002-2003, the average standard score for the Abbott kindergarten sample is 87.25 (with a standard deviation of 15.17) about 13 points lower than the national average score. Using the best scores from the PPVT-3 and the TVIP for the Spanish-speaking children, there is no statistically significant difference in scores between Spanish and English speaking children in 2003.

Differences among districts in children’s scores. Average scores vary across districts. Kindergartners in smaller districts and in districts with lower poverty levels score significantly higher than children in larger districts with higher levels of poverty. In fact, kindergartners living in the smaller Abbott districts score about 8 points higher on average than children living in the largest districts in the state, while children living in the lowest poverty Abbott districts score nearly 5 points higher than children living in the highest poverty districts. There is some overlap between district size and level of poverty, and there are other factors that impact scores (e.g. percent attending preschool).

Improvement in scores over time. The Center for Early Education Research at Rutgers University collected child assessment data during the 1999/2000 and 2000/2001 school years in a subset of the Abbott districts. Across 14 districts with data from both years, the average receptive language score increased by 2.62 points, from 84.54 to 87.16, a statistically significant difference.

### Findings for Children’s Early Literacy Skills

A new measure of early literacy skills was administered to assess skills kindergartners in 21 of the Abbott districts in the fall of 2002, along with the measures of receptive language. Get Ready to Read (GRTR; Whitehurst & Lonigan, 2000) is a 20-item instrument that measures multiple aspects of children’s pre-reading skills: book

knowledge, print knowledge, letter knowledge, letter-sound correspondence, emergent writing, and three aspects of linguistic awareness – initial phonemes, rhyming, and compound words. The range of possible scores is 0 to 20. GRTR was given to a sample of 1251 English-speaking kindergarteners. Unweighted data are used in these analyses since not all districts are represented.

The average score for the sample is 15.22 with a standard deviation of 3.26. The authors divide the scores into categories indicative of very weak to very strong pre-reading skills, as described in Table 4 below. Over 41% of sample Abbott kindergarteners score between 17 and 20 on this test, which indicates very strong pre-reading skills, while over 38% score between 13 and 16, indicating strong pre-reading skills. It is important to note that the labels for the categories (e.g. very strong skills) are based on four-year-old children’s scores, and may be too generous for the current sample of kindergartners. However, given the current sample kindergartner’s PPVT-3 scores, which, though improved are still far below average, it is informative to find that these children’s pre-reading skills as measured by the GRTR are higher than might be expected.

**Table 4.**  
**GRTR Scoring Categories and Percentage of Children Scoring**  
**Fall, 2002**

<u>Range of Scores</u>	<u>Category</u>	<u>Percentage of Children Scoring</u>
0 – 6	Very weak skills	1.4%
7 – 9	Weak skills	4.7%
10 – 12	Average skills	14.1%
13 – 16	Strong skills	38.2%
17 – 20	Very strong skills	41.5%

#### Discussion

This research was conducted to inform policy and practice by describing preschool classroom characteristics and children’s skills at kindergarten entry. Overall, the majority of classroom scores range from tolerable to good with 13% scoring good to excellent. The evidence suggests that as a result of preschool, children are entering kindergarten with improved language skills but still below average. There are promising indications of early literacy readiness abilities. Although the increases in program quality and children’s language and literacy abilities are modest, they indicate that the classrooms are improving and are having a positive affect on children. The results reveal important areas for further improvement. Classroom quality is not high enough and, in some cases, is quite inadequate. The DOE and the ECE staff in the Abbott districts and contracting centers have been concentrating professional development and other initiatives on improving overall quality and literacy and mathematics, in particular.

## Promising Trends in Classroom Quality

In 2002-2003 classroom quality scores are highest in those areas most likely to influence child learning, and these are the areas where the most growth in quality is occurring in the last few years. Teachers create a warm, nurturing environment and are improving in use of preventive management techniques that help children develop self-regulation and solve social problems. In particular, teachers are providing children with more opportunities to develop oral language, including rich vocabulary and complex sentence structure. Classrooms are better equipped in general but particularly with books, and teachers read more frequently to children. In classrooms, where children's heritage language is other than English, more teachers are supporting the child's home language in the classroom.

This improvement in classroom quality is happening at the same time as rapid expansion of the program. Rapid expansion has resulted in a large proportion of new and inexperienced teachers, many of whom are receiving their coursework in early childhood education during their first year of teaching. This lack of experience is clearly a factor in the quality and, indeed, we find a relationship between the teacher's experience and the classroom scores. Although in previous years the relationship between teacher credentials and quality was strong, it is no longer evident in 2003. With most teachers in the sample meeting the requirements, it is difficult to discern a difference. In addition, those uncertified teachers who are successful in completing college coursework are the ones more likely to remain and also more likely to be successful teachers with higher scoring classrooms. Rapid expansion may also be pressing districts to use less than ideal facilities.

Expansion has been most dramatic in the largest and the highest poverty districts. This may be one reason that these districts have experienced the least growth in quality. More of the new and uncertified teachers are in the large districts. An additional possible explanation may be the greater complexity of administration in large districts with large numbers of contracting centers.

One important concern in the quality data is the overall low scores in classroom space and furnishings. One clear reason for this lower score is the lack of gross motor play space and equipment in many of the urban centers and schools. The average score on those specific items is below minimal. In the past two years, the DOE has been funding playground construction and equipment in many centers to help alleviate this problem. Children who are spending 6 to 10 hours in a preschool setting need many opportunities to exercise and play outdoors. In addition, the number of preschool facilities being built with state construction funds is increasing dramatically but strategies for improving existing facilities must be implemented to improve this aspect of the program. It is hoped that the results of the 2003-2004 observations will show the effects of these initiatives and the provision to new classrooms of almost \$20,000 in materials and equipment funds and to existing classrooms of over \$4,000.



## Growth in Kindergarten Entry Scores

Analyses of the PPVT and TVIP scores show that children's oral language skills at kindergarten entry have increased as the preschool program has expanded and improved but are still well below the national average. It is important to note that regardless of whether the child attended preschool, all kindergarten children were included in the sampling pool. In 2002, approximately one third of the kindergarten children had not attended preschool and even more had not attended two years of preschool. Of those who had attended preschool, the quality of the program they experienced was lower than current programs. Given this, the results probably underestimate the effects of the current Abbott preschool program.

A promising finding is the strong result on the Get Ready to Read screening tool. Although this likely overestimates the children's abilities, since it was designed for slightly younger children, no similar instrument exists for kindergarten age children. The skills measured, such as letter knowledge, and linguistic awareness, are still highly relevant for kindergarten. Children's scores on this instrument reveal that the majority of kindergartners in Abbott districts are entering formal school with many of the early literacy skills necessary to become successful readers.

### Initiatives of the DOE Relevant to These Findings

A number of initiatives, implemented since 2002, are designed to increase the quality of facilities and teaching practices.

#### Facilities Guidelines and Construction

With the Abbott mandate for preschool education, the 30 Abbott districts looked to their own available space in-district, in child care centers, and in Head Start programs to accommodate the preschool children. Guidelines for facilities construction were developed by a task force representing the Department of Education, the Department of Human Services (DHS), child-care providers, Head Start agencies and others. These guidelines have informed amendments to the *NJ Administrative Code (N.J.A.C.)* 6:19-3. Facilities that meet these guidelines are currently being constructed or designed in almost every district. These will provide the quality environments which result in higher scores on classroom assessments but, more importantly, better programs for children.

As noted above, the DOE has funded playground construction and provided ample materials and equipment funds to increase the quality of the classrooms. Guidance has been developed on the type of materials and equipment that every classroom and playground should possess.

In addition to informing districts of the opportunity to include some private providers who own their own buildings in the district's Long Range Facility Plan, we have encouraged districts to accommodate providers in district buildings. We will also be offering programs to interested providers on how to obtain construction loans to renovate and expand existing facilities.

## Preschool Teaching and Learning Expectations: Standards of Quality

During the summer of 2002, a task force, consisting of representatives from community organizations throughout New Jersey that support the interests of young children, met to revise the Preschool Teaching and Learning Expectations: Standards of Quality. The revisions were based on the latest research in best practices for developmentally appropriate education for three- and four-year-old children. The revised expectations consist of examples of high-quality teaching practices along with learner outcomes within each learning domain and offer significant assistance to the classroom teacher for planning instruction. The Expectations have received three favorable reviews from national organizations.

## Abbott Preschool Program Implementation Guidelines

Another project completed during the 2002-2003 school year to enhance the quality of Abbott preschool programs was the development of the *Abbott Preschool Program Implementation Guidelines*. Representatives from the Governor's office, DOE, DHS, districts, community child-care providers, Head Start agencies, professional education organizations, advocacy groups, parents, institutes of higher education and other community organizations formed work groups to analyze research and brainstorm and develop recommendations that provide guidance on all aspects of the preschool program and enable Abbott districts to fully execute the Abbott mandate. The guidelines are recommendations based on the latest research and expert opinion and are intended to inform the district's Early Childhood Program Aid Three-year Plan and Budget.

## Master Teacher Training

Vital to the success of the preschool program is the quality of curriculum and teaching. In order to increase quality in the Abbott preschool classrooms, the Office of Early Childhood Education (OECE) offers a comprehensive, year-long training for master teachers who mentor and coach over 5,000 teachers and assistant teachers in the Abbott districts. The year-long course was designed to more clearly define the master teacher role and to ensure that master teachers have the skills they need to foster change and improve classroom quality. Specifically, the master teacher training focuses on three areas: (1) in-depth training in curriculum, including the research-based guidelines; (2) assessing classroom quality through the use of structured program evaluation instruments such as the Early Childhood Environmental Rating Scale-Revised (ECERS-R), Supports for Early Literacy Assessment (SELA), and Preschool Mathematics Inventory (PCMI); and (3) coaching and mentoring strategies for adult learners. Together, this training provides the master teacher the necessary information and skills to train Abbott teachers in standards for curriculum and classroom quality.

The master teacher seminar culminated in a statewide conference in May 2003. Small groups of master teachers worked together to provide workshops on topics relevant to the *Preschool Teaching and Learning Expectation: Standards of Quality*. Over 500

participants attended the conference. Master teachers that successfully completed the course earned the status of Preschool Professional Development Fellow.

Professional development and networking for master teachers continues to be offered with focus on specific quality components such as transition practices, handling challenging behaviors, early literacy and math, and assisting English language learners.

#### Equalizing Teacher Qualifications and Pay

Typically, in the past, private child-care centers have suffered from a high rate of teacher turnover and a lack of well-trained teaching staff, thus limiting the ability to provide a high-quality program. In 2002–2003, in order to provide the high-quality programs that the court mandated, private provider teachers that received the proper training and held the appropriate certifications, received salaries comparable to in-district teacher salaries. Upgrading teacher salaries and qualifications will have the effect of stabilizing employment in the centers and creating a pool of well-trained and experienced teachers.

#### Implementing Methods for Program Evaluation and Improvement

The DOE – OECE strives to work collaboratively with districts as partners in the endeavor to establish and improve preschool programs. One aspect of that partnership is providing districts with information and leadership that assists in program improvement. The department has initiated the following strategies for accountability leading to program improvement. Numerous smaller initiatives have also been implemented.

Self-Evaluation Validation System. High-quality educational programs undergo a continual cycle of gathering evidence about programs in order to make informed decisions toward improvement. To this end, the OECE brought together stakeholders throughout New Jersey, chosen for their expertise in the field of early childhood education, to develop a self-evaluation tool for Abbott districts. The Self-Assessment Validation System (SAVS) is a system designed to guide the district through a systematic self-appraisal of its preschool program and to aid in program improvement. The SAVS is derived from the *NJ Abbott Preschool Program Implementation Guidelines*, as well as the *Guidelines for Appropriate Curriculum Content and Assessment in Programs Serving Children Ages 3 through 8* (National Association for the Education of Young Children and the National Association of Early Childhood Specialists in State Departments of Education). During June of 2003, OECE staff met with the 30 Abbott districts to explain the purpose and the process of the SAVS. The SAVS has two phases: Phase I – evaluating the program as is and establishing a plan for improvement that included revising the operational plan by October 2003; and Phase II – evaluating improvements made, which will culminate with a validation team visit in district by June 2004. The SAVS is intended to highlight strengths of district programs and to alert districts to areas in need of improvement, which will inform program improvement.

Early Learning Assessment System. In addition to collecting and reporting on district data of children and classrooms, the ELIC and the OECE planned and developed the Early Learning Assessment System (ELAS), a performance-based assessment system administered by teachers during regular classroom activities. The ELAS is based on the latest research on development and learning in young children and is fully aligned with the *Preschool Teaching and Learning Expectations: Standards of Quality*. With the ELAS, teachers learn how to observe children in the natural preschool environment on a regular basis and collect samples of work and record observations. This collected work is used to adjust the learning environment based on information about the children and to serve as a means of evaluating the skills of young children in Abbott districts on a statewide basis. The ELAS has been piloted in 7 districts and will be expanded to all of the districts in the 2004-2005 school year.

As The Garden Song by David Mallett begins “Inch by inch, row by row, gonna make this garden grow,” all of the teachers, administrators, advocates, parents and others working in Abbott preschool are sowing seeds and nurturing, scaffolding and otherwise supporting the growth of both this program and more importantly the children who participate. This report indicates that we are making progress – progress in outreach and enrollment, progress in raising teacher’s qualifications, progress in the quality of the classrooms and progress in children’s abilities to succeed in school. We have more hard work to do to fulfill the promise of Abbott preschool but the children deserve nothing less than our most dedicated efforts.

## Appendices

## Appendix A

### Measures

#### Measures of Classroom Quality

##### Early Childhood Environment Rating Scale - Revised (ECERS-R)

Program quality was assessed by trained observers using a standardized measure of preschool classroom structure and process, the Early Childhood Environment Rating Scale – Revised (ECERS-R) (Harms, Clifford & Cryer, 1998). This measure has been used extensively in the field and has well-established validity and reliability. The validity of the measure is supported by high correlations between both the scale items and ratings of items as highly important by a panel of nationally recognized experts, and between scale scores and ratings of classroom quality by experts. Internal consistency as measured by Cronbach’s alpha is reported by the authors to be adequate, ranging from .81 to .91.

Classroom quality is rated on a 7-point Likert scale, indicating a range of quality from inadequate (1) to excellent (7). The seven ECERS-R subscales are as follows: Space and Furnishings, Personal Care Routines, Language-Reasoning, Activities, Interaction, Program Structure, and Parents and Staff. Average subscale scores are calculated, as well as a total scale score averaged across all 43 items in the scale.

##### The Supports for Early Literacy Assessment (SELA)

The extent to which the classroom environment is supportive of children’s literacy development is measured with the Supports for Early Literacy Assessment (SELA) (Smith, Davidson & Weisenfeld, 2001). This measure is revised for use by this project with the deletion of 5 items that overlap with the ECERS-R. The revised measure includes 16 items on a scale from 1 to 5, low quality (1) to high quality (5) for the support of early literacy development. Six subscales are: The Literate Environment, Language Development, Knowledge of Print/Book Concepts, Phonological Awareness, Letters and Words, and Parent Involvement. This measure is informed by the Early Childhood Environment Rating Scale - Revised (Harms, Clifford & Cryer, 1998), the High/Scope Program Quality Assessment (High/Scope Educational Research Foundation, 1999) and the NAEYC publication Learning to Read and Write (Neuman, Copple & Bredekamp, 1999). Internal consistency among scale items as measured by Cronbach’s alpha on the current sample is excellent at .92.

##### The Preschool Classroom Mathematics Inventory (PCMI)

The classroom support for the development of children’s early mathematical skills is measured using the Preschool Classroom Mathematics Inventory (PCMI) (Frede, Dessewffy, Hornbeck & Worth, 2001). This tool measures the materials and strategies used in the classroom to support children’s early mathematical concept development, including counting, comparing, estimating, recognizing number symbols, classifying, seriating, geometric shapes and spatial relations. The standards of the National Council of Teachers of Mathematics and the National Association for the Education of Young Children inform the measure, which is comprised of 11 items on a 5-point scale, from

low quality (1) to high quality (5), and has two subscales, Materials and Numeracy and Other Mathematical Concepts. Internal consistency among the test items as measured by Cronbach's alpha on the current sample is excellent at .93.

### Measures of Children's Language Assessment

#### PPVT/TVIP

The Peabody Picture Vocabulary Test-III (PPVT-3) (Dunn & Dunn, 1997) measures receptive vocabulary attainment, which is a well-known correlate of cognitive development and school success. The test is designed for persons aged 2 1/2 through 90+ years and was standardized on a stratified national sample. The Spanish version of the PPVT, the Test de Vocabulario en Imágenes Peabody (TVIP) (Dunn, Padilla, Lugo & Dunn, 1986), measures receptive vocabulary for Spanish speakers.

For the 1999/00 data collection period, the TVIP was used in place of the PPVT-3 when it was determined by the assessor that the child was more appropriately assessed in Spanish. For the 2002/3 data collection period, the TVIP was used in addition to the PPVT-3 for all children who speak any Spanish, and the highest score of the two tests for each child is used.

#### The Get Ready to Read (GRTR)

The Get Ready to Read (Whitehurst & Lonigan, 2001), a new measure of reading readiness was presented to the children following the completion of the PPVT and/or TVIP. The GRTR was validated on a sample of 342 middle- and lower-class children of preschool age. It is highly correlated with other measures of emergent literacy including the Developing Skills Checklist (CTB/McGraw Hill, 1990), and is reported by the authors to be predictive of children's subsequent reading success in second grade. Internal consistency of the items is good (Cronbach's alpha = .78, as reported by the authors). Twenty items measure the following aspects of reading readiness: book knowledge, print knowledge, letter knowledge, letter-sound correspondence, emergent writing, linguistic awareness – initial phonemes, linguistic awareness – rhyming, and linguistic awareness – compound words. While the measure is meant to be presented to children during preschool, the decision was made to present it to Abbott kindergarteners based on the low scores for receptive vocabulary which had been previously found in samples of Abbott kindergarten children.

The GRTR was presented to children in 21 of the 30 Abbott districts. Since the sample is not distributed over all the Abbott districts, the data do not reflect the scores of children in the Abbott districts as a whole. Additionally, since the sample of Spanish-speaking children is small, and since the validity of the instrument is suspect in Hispanic samples (Whitehurst, 2003), the scores of Spanish-speakers are not included in analyses.

### Training and Inter-rater Reliability

University staff working in conjunction with the Department of Education – Office of Early Education staff were responsible for collecting child and classroom information from the field. Classroom observers and child assessors were hired mainly from the pool of graduate students in good standing available to the universities from

departments of education, psychology or child development. All classroom observers and child assessors were trained on each measure, on methods of conducting classroom observations, and on professional etiquette. Upon completion of training, classroom observers came to reliability on each instrument with an experienced observer three times, attaining at least 80% agreement allowing for a scoring difference of one on all measures before observing on their own. For example, the average inter-rater reliability coefficient for the 2003 data collection period across the NIEER classroom observers is .94 for the ECERS-R, .98 for the SELA and .96 for the PCMI, all allowing one score away. Using the more conservative reliability procedure of requiring exact score matches, the inter-rater reliability coefficient is .81 for ECERS-R, .70 for the SELA and .77 for the PCMI. Child assessors were accompanied by an experienced assessor who shadow-scored their first few assessments to assure inter-rater reliability.



## Appendix B

### ECERS-R Subscale Descriptive Data

#### The Space and Furnishings Subscale

The average score for the Space and Furnishings subscale is 3.76, lower than the total average. This subscale measures the quality of the indoor and outdoor space available to children during the day. The majority of classrooms (68.5%) score in the minimal to good range of quality (3 to 5), while 20% score in the inadequate to minimal range (1 to 3) and just over 10% score a 5 or better. See Figure 1 below.

**Space and Furnishings Subscale - Percentage of Classrooms Scoring 1 - 7**

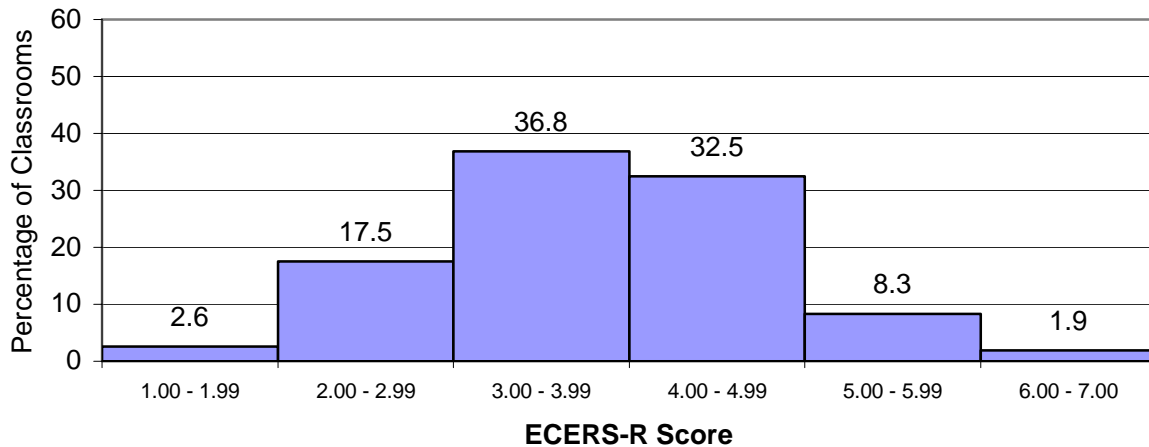


Figure 1.

A classroom scoring high on this subscale has ample room, natural lighting and good ventilation that can be controlled from the classroom. The furniture is child-sized, in good repair and is convenient to use. An easel, work bench or a sand/ water table is in use. A cozy area, not used for active physical play, has some soft furnishings and many soft toys, and is available for a substantial portion of the day. There are at least five well-equipped interest centers that children can use by themselves, with materials rotated. The classroom space is thoughtfully planned so that quiet and active play does not interfere with each other and egress is not through play centers. The walls of the classroom hold displays of children's work, at their eye level, and mostly relate to current activities or themes. This work should look quite individual, not cookie-cutter copies of a teachers' example – and should include some 3-dimensional work with clay or wood, for example.

There is a space or two for children to play alone or with a friend, protected from the group. This space is usually quite creative and is set-up with a few toys or activities that one or two children may use.

Outdoor space is safe and ample, and there is some indoor space to be used for gross motor play when weather is inclement. This space is easily accessible to the children, organized for several different types of activities to go on at once (such as play with wheel toys, ball play and hopscotch), with enough equipment that children do not have too long a wait for a favorite toy. Stationary and portable equipment stimulate skills on different levels, for instance, there are balls of different sizes, wheel toys with and without pedals, climbing structures of different levels of difficulty. The outdoor space has some convenient features such as a water fountain and close bathrooms.

The Personal Care Subscale

The average score for the Personal Care subscale is 3.69, below the total scale average. This subscale measures the extent to which classroom environment is conducive to the children’s health and safety. Over 30% of sample classrooms score 3 or below, in the minimal to inadequate range. However, 19% score a 5 or better. See Figure 2 below.

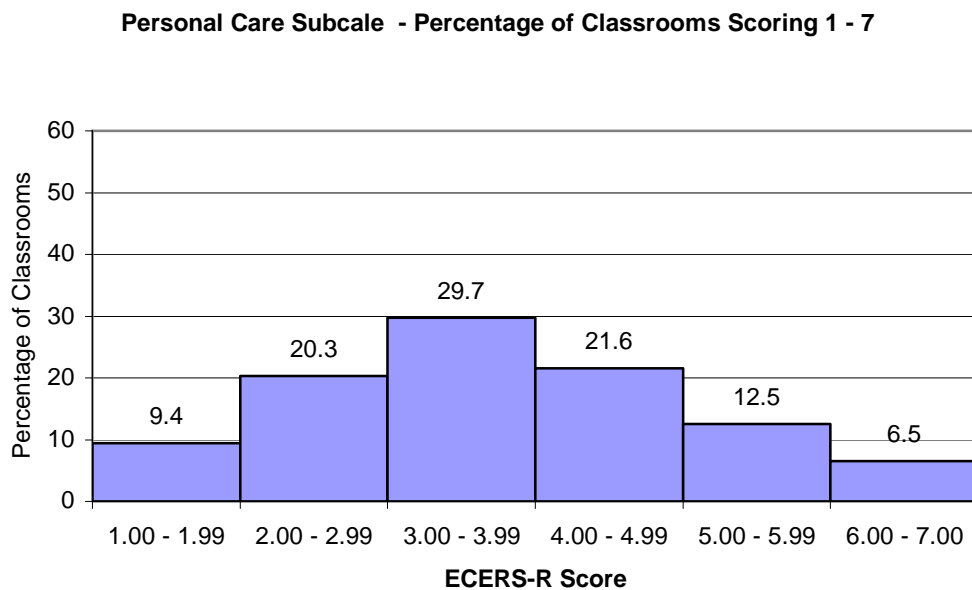


Figure 2.

A high-scoring classroom on the Personal Care subscale is characterized by close attention to the personal needs of the children. Beginning when the children enter the program in the morning, they are each greeted individually and are helped to become involved with an activity in the classroom. If a parent accompanies the child, the parent is greeted warmly and information is exchanged between the adults. Meals and snacks are nutritious, and are scheduled appropriately for young children. Meal times are pleasant, with conversation among everyone and plenty of practice and support for the development of children’s independent use of utensils and dishes. Children help set the table, serve themselves and clean up. Nap times are scheduled appropriately for children

and are flexible, with space conducive to rest and quiet activities available for non-nappers or early risers.

Toileting is sanitary and pleasant with plenty of hand-washing by children and adults. The facilities are convenient for children to use independently, but there is enough adult supervision to ensure the maintenance of sanitary conditions. Toilets and sinks must be child-sized, and if not in the classroom then very near the classroom. Health concerns are taken seriously – children have the proper attire for the weather, their noses are wiped and soiled clothing is changed. Adults are cognizant of the possible spread of germs, with hand-washing a common routine. In excellent classrooms, children are taught about health-related issues such as cleanliness and nutrition, and how to manage such things themselves. Toothbrushes are used at least once a day in all day programs, and are kept in sanitary condition.

### The Language and Reasoning Subscale

The average Language and Reasoning subscale score is higher than the total average score, at 4.27. This subscale measures the quality of the selection of books and other language-related materials in the classroom, along with the quality of the communication between adults and children to support children’s use of language and reasoning skills. Nearly one third of classrooms score in the good to excellent range (5 to 7), and over half the sample classrooms score in the minimal to good range (3 to 4). However, 56% of classrooms scoring a 4 actually score a 4.5 or better, indicating that they are very close to scoring in the good quality range. See Figure 3 below.

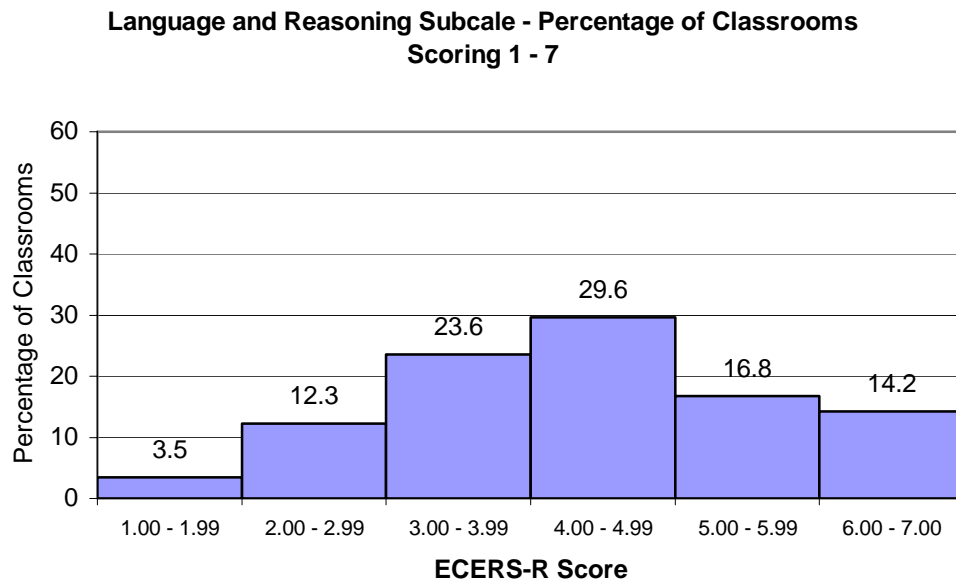


Figure 3.

A high-scoring classroom for this subscale has a wide selection of books. A wide selection includes many different types of books and stories – perhaps big picture books, rhyming books, number books, poetry, fictional stories, biographies, books about animals, weather, holidays, books in other languages. Other language materials are also

found in the classroom, such as a felt board, recorded stories, games with words and pictures, puppets, and small figures. Adults read to children daily in this classroom, both formally, during whole group times, and informally during free play. Books are rotated to relate to current themes.

The communication between adults and children in this classroom is meant to encourage children to use their verbal and written skills. To this end, staff will engage in conversation with children during whole group, small group and individual times, encouraging children to respond in more complex ways using new vocabulary. Staff encourage this by modeling slightly more complex language than the children produce and by expanding on the ideas that children present. For instance, if a boy says, “Look! I made a car!”, the adult might respond, “Wow! That shiny, red car looks like it could win a race with those big wheels!” Adults and children can be heard discussing concepts including same/different, matching, sequence, one-to-one correspondence, classification, size, spatial relationships and so forth while children are playing with materials or working on an interesting task.

### The Activities Subscale

The ECERS-R subscale with the lowest average score is the Activities subscale, with an average score of 3.37. This subscale measures the quality and quantity of the materials that the children may use in the classrooms and the amount of time the children have access to the materials. It covers materials and access to materials for dramatic play, art, music, math, science, fine motor skills and computers. Plentiful, high-quality materials are especially important for young children, who learn primarily through *doing* - and doing requires having materials to do something with. This score indicates that in the average classroom, access to materials is minimal. The frequencies of scores indicate that over one-third of classrooms score in the inadequate to minimal range on this subscale, while only about 5% of sample classrooms score in the good to excellent range. See Figure 4 below.

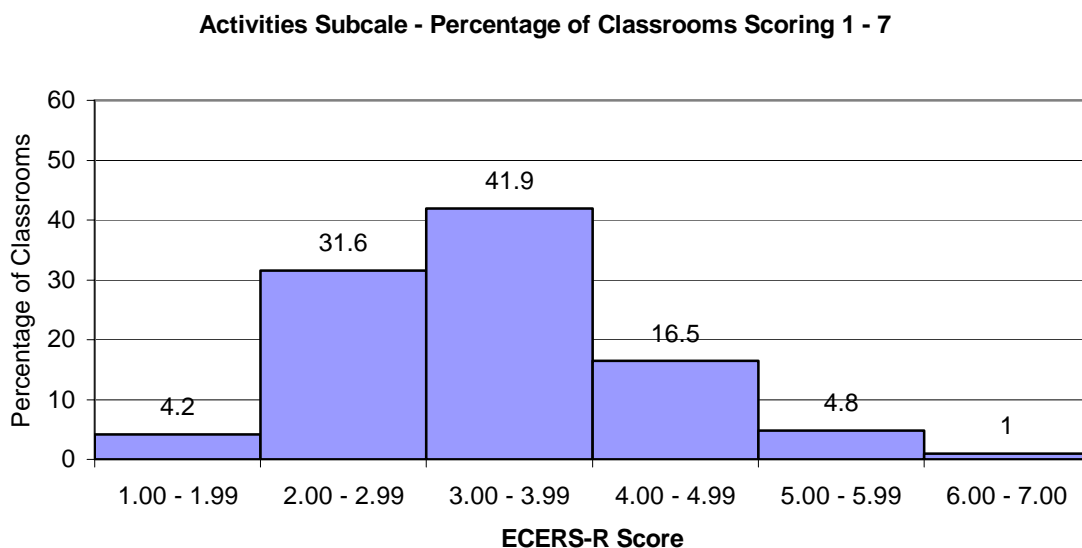


Figure 4.

What would classrooms look like if they scored high on the Activities subscale? In these classrooms, materials are accessible to children for a substantial portion of the day. A “substantial portion” is defined as at least one-third of the time that most of the children are in daily attendance. This time does not have to be in one block but should be apportioned appropriately throughout the day. A wide variety of materials are stocked on well-organized shelves, low so that children can reach without trouble. There are enough materials so that each child can choose something interesting, and can get the exact thing he or she wants without too long of a wait.

These classrooms provide a rich array of play experiences for children through the provision of many and varied materials. There are plenty of fine motor items such as small building toys, puzzles, beads for stringing and small figures. Art supplies allow for much individual expression, and include the typical crayons, markers, paints and brushes of different sizes and colors, but also clay, playdo, felt, wood, glue, safe scissors, interesting collage materials like feathers and, pom-poms, and other fun materials that can inspire creativity and assist fine motor development. Many musical instruments and recordings are available along with dancing props – some recordings should be in the languages of the children. There are several different kinds of building blocks in an activity center big enough to allow three children to build something together, along with accessories such as small vehicles, people, animals and road signs.

Sand and water play are available, with many shovels, sifters, funnels, molds, buckets, measuring cups and the like. Added features are sometimes introduced, such as bubbles in the water or rice instead of sand. Dramatic play materials are available for a variety of themes such as house, travel, pets, post office, restaurant, beach, or hairdresser; stories read in class or class trips are also used as the basis for dramatic play. Materials include the basics such as clothing, dolls and dishes, but there is much more, including items representing other cultures, and materials to support dramatic play outdoors.

Plenty of varied science materials are accessible to the children, including collections of natural objects, living things to care for such as plants, fish or small mammals, games through which children learn about the natural world, and other interesting materials such as magnets, magnifying lenses, tuning forks and the like. Math materials are plentiful and include things to count, measure and weigh along with tape measures, balance scales and rulers; geometric shapes in different sizes and colors to sort or to create patterns; number games such as lotto and dominoes, and other materials to play with written numbers such as magnetic numbers or number puzzles. For both math and science, the class is involved in a longer-term project which requires special input from the teacher or another adult, like observing and graphing change or comparing the frequency of some occurrence across days. For instance, during autumn the teacher might help children graph the outdoor temperature and compare that graph to a graph of the number of classmates who wore hats to school. And for all types of materials, things are rotated on /off the shelves as children’s interest wanes or as new classroom themes emerge.

### The Interaction Subscales

The ECERS-R subscale with the highest average score is the Interaction subscale with an average score of 4.92. This score indicates that on average interactions between adults and children, and among children, approach good quality in the Abbott preschool classrooms. Moreover, nearly 68% of sample classrooms score in the good to excellent range on this subscale, and less than a quarter fall below 4. See Figure 5 below.

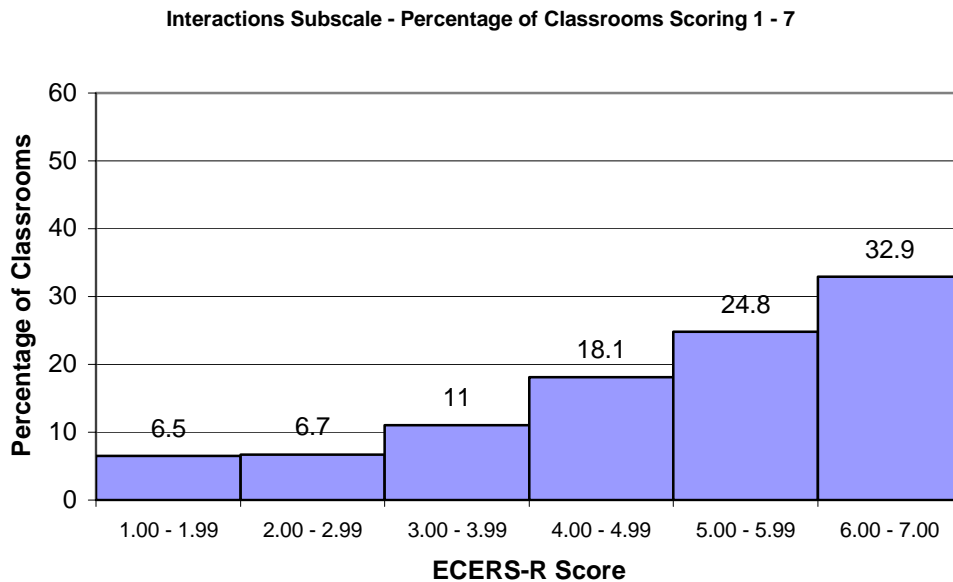


Figure 5.

What do interactions look like in classrooms that score high on the Interactions subscale? Teachers in these classrooms support learning by creating a positive environment for the exchange of ideas by helping, encouraging, appreciating, and extending what children are doing. They use preventive management techniques, by setting up the environment to reduce the chance of conflict, redirecting potential problem behavior, modeling good social skills including cooperation and compromise, and by consistently reacting to negative interactions with non-punitive support of the development of children’s social skills. This creates a positive emotional climate in which children have the best chance for optimal learning.

### The Program Structure Subscale

The average Program Structure subscale score is slightly higher than the total average score, at 4.04, placing the average classroom in the middle of the minimal to good range. This subscale measures the extent to which children have an appropriate schedule of activities and groupings during the day, are given time to play with materials, and are provided for when they have special needs. Scores for this subscale are characterized by a relatively even distribution across the scale. Nearly one quarter of classrooms score below 3, placing them in the minimal to inadequate range, while just over 30% score a 5 or better. See Figure 6 below.

**Program Structure Subscale - Percentage of Classrooms  
Scoring 1 - 7**

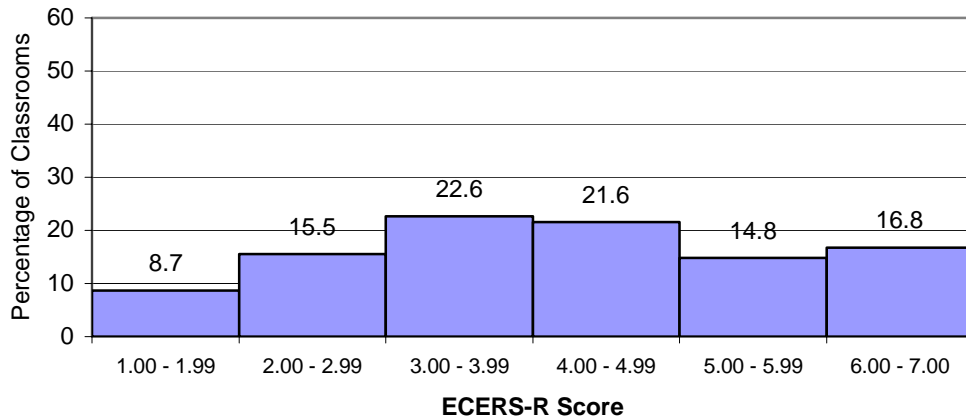


Figure 6.

What does a high-scoring classroom on the Program Structure subscale look like? The classroom schedule provides a balance of structure and flexibility during the day, with some activities child-initiated and some teacher-directed, and a substantial portion of the day used for play activities. Free play is supported with plenty of materials and with adult supervision and interaction, which is understood as an opportunity for educational scaffolding. Transitions are smooth with no long waiting, and variation is made in the schedule to meet individual children's needs. Whole group time is limited to no more than 20 minutes, with most activities and routines done in small groups, and educational interactions taking place with small groups or individual children as well as with the whole group together. Additionally, children have many opportunities to select a group to be with during the day.

Children with disabilities in the classroom are integrated into the group, with modifications made to the environment if needed, and most intervention taking place within the regular activities of the classroom. Parents and staff are actively involved in following through with activities recommended by other professionals, and there is much sharing of information about how the program is working for the child. In this sample, just over 20% of classrooms have at least one child with a disability. Three percent serve only children with disabilities.

The Parents and Staff Subscale

The Parents and Staff subscale is the next highest scoring subscale across our sample, with an average score of 4.37. This subscale measures provisions for the personal and professional needs of the program staff including professional development opportunities, evaluation, planning time, and space for private and work-related items. Provisions for the parents of the children who attend the program are also rated here, including the extent of information sharing between parents and program staff, and opportunities for parental involvement in program activities.

The distribution of scores across the scale approximates a normal curve, with nearly equal percentages of classrooms scoring above and below the mean. Nearly 30% score in the good to excellent range and nearly 62% score in the minimal to good range. See Figure 7 below.

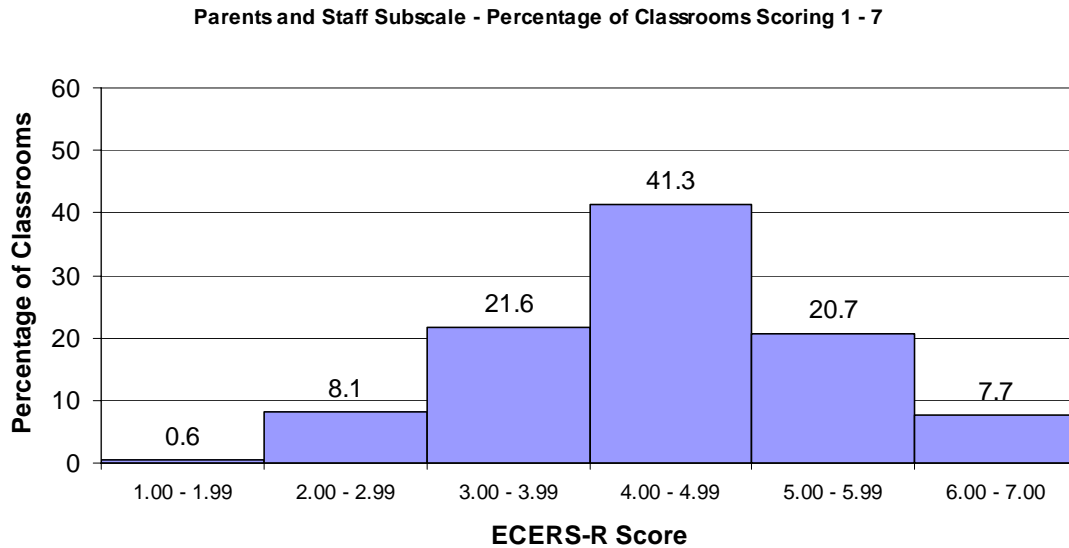


Figure 7.

A classroom scoring high on the Parents and Staff subscale has convenient and secure storage space for the personal belongings of the staff, along with adult-only restrooms, a separate lounge area for the staff with comfortable furniture, and several staff breaks scheduled flexibly during the day. There is ample space for classroom supplies and files, separate and well-equipped office administration space, and space enough for individual or group conferences that can be held separately from the children's activities. Teachers communicate with one another during the day so that the day runs smoothly, and there is planning time for staff working together. The supervision and evaluation of staff is based on frequent observations, includes self-evaluation and a formal written evaluation at least yearly, and is helpful. Areas in need of improvement are addressed with an action plan that includes professional development activities. The program will have a thorough orientation for new staff on subjects such as discipline, activities for children and relationships with parents, and an on-site library for professional development materials. The program runs monthly staff meetings that include development activities, and in-service training opportunities.

Parents are encouraged to participate in their child's program with a variety of alternatives, much respectful and positive communication between program staff and family, and the opportunity to evaluate the program yearly. Parents are given information about the program in writing, including administrative procedures and program philosophy. Additionally, parents have a decision-making role in the program, for instance they may be members of a governing board or council.



## Appendix C

### Data Analyses

#### Analyses on change in ECERS-R scores over time

A MANOVA by COHORT (coded 0 for 1999/00 and 1 for 2002/3) was used to investigate the difference in scores. The total sample size is 514. The sample size for 1999 is 262, and the sample size for 2002 is 252. Note the degrees of freedom (df) is always (1,513). Note that the direction of effect for the COHORT beta weights are reversed to indicate positive change when scores increase over time. Results are as follows:

<u>ECERS-R Subscale</u>	<u>B(p)</u>	<u>F(p)</u>	<u>Adjusted R<sup>2</sup></u>
<u>Space and Furnishing</u> COHORT	-.03(.74)	.11(.74)	.00
<u>Personal Care</u> COHORT	<b>-.28(.03)</b>	<b>4.59(.03)</b>	.01
<u>Language and Reasoning</u> COHORT	<b>.43(.00)</b>	<b>12.84(.00)</b>	.02
<u>Activities</u> COHORT	<b>.19(.04)</b>	<b>4.48(.04)</b>	.01
<u>Interactions</u> COHORT	<b>.36(.01)</b>	<b>6.68(.01)</b>	.01
<u>Program Structure</u> COHORT	.20(.17)	1.86(.17)	.00
<u>Parents and Staff</u> COHORT	<b>-.24(.01)</b>	<b>6.03(.01)</b>	.01
<u>Total</u> COHORT	.06(.48)	.51(.48)	.00

The mean subscale and total scores along with standard deviations for 1999 and 2003 are as follows:

<u>ECERS-R Subscale</u>	<u>Mean (standard deviation)</u>	
	<u>1999</u>	<u>2003</u>
Space and Furnishings	3.73(1.1)	3.70(1.02)
Personal Care	3.98(1.56)	3.71(1.34)

Language	3.74(1.38)	4.16(1.31)
Activities	3.19(1.05)	3.37(.95)
Interactions	4.47(1.59)	4.83(1.60)
Program Structure	3.81(1.80)	4.01(1.56)
Parents and Staff	4.59(1.21)	4.34(1.02)
Total	3.86(1.07)	3.92(.93)

Analyses using facilities and teacher process scores

A MANOVA by COHORT was used to investigate differences in scores between 1999 and 2003. Note the degrees of freedom (df) is (1,513) for both.

<u>ECERS-R Subscale</u>	<u>B(p)</u>	<u>F(p)</u>	<u>Adjusted R<sup>2</sup></u>
<u>Teacher Process score</u> COHORT	<b>.24(.03)</b>	<b>7.45(.03)</b>	.01
<u>Facilities score</u> COHORT	-.07(.52)	.67(.52)	.00

The mean scores along with standard deviations for 1999 and 2003 are as follows:

	<u>Mean (standard deviation)</u>	
	<u>1999</u>	<u>2003</u>
Facilities score	3.73(1.28)	3.66(1.25)
Teacher process score	4.18(1.27)	4.42(1.27)

Analyses on first year/non-first year teachers

Multivariate Analysis of Covariance (MANCOVA) by FIRSTYR (coded 0 for no and 1 for yes) was used to investigate the difference in scores between first year teachers and non-first year teachers. The COHORT variable was included in analyses to determine whether year of data collection influenced the difference in scores. Note that the direction of effect for the COHORT beta weights are reversed to indicate positive change when scores increase over time. Results are as follows:

<u>ECERS-R Subscale</u>	<u>B(p)</u>	<u>F(p)</u>	<u>Adjusted R<sup>2</sup></u>
<u>Space and Furnishing</u>			.00
FIRSTYR		3.51(.06)	
Not first year	.28(.06)		
COHORT	-.02(.81)	.06(.81)	
<u>Personal Care</u>			.01
FIRSTYR		.69(.41)	
Not first year	.17(.41)		
COHORT	<b>-.28(.03)</b>	<b>.4.68(.03)</b>	
<u>Language and Reasoning</u>			.03
FIRSTYR		<b>5.27(.02)</b>	
Not first year	<b>.43(.02)</b>		
COHORT	<b>.43(.00)</b>	<b>12.81(.00)</b>	
<u>Activities</u>			.01
FIRSTYR		3.31(.07)	
Not first year	.25(.07)		
COHORT	<b>.19(.03)</b>	<b>4.56(.03)</b>	
<u>Interactions</u>			.02
FIRSTYR		<b>4.36(.04)</b>	
Not first year	<b>.46(.04)</b>		
COHORT	<b>.38(.01)</b>	<b>6.92(.01)</b>	
<u>Program Structure</u>			.00
FIRSTYR		1.34(.25)	
Not first year	.27(.25)		
COHORT	.18(.24)	1.37(.24)	
<u>Parents and Staff</u>			.02
FIRSTYR		2.76(.10)	
Not first year	.26(.10)		
COHORT	<b>-.24(.02)</b>	<b>5.86(.02)</b>	
<u>Total</u>			.01
FIRSTYR		<b>4.50(.04)</b>	
Not first year	<b>.29(.04)</b>		
COHORT	.07(.47)	.54(.47)	

ECERS-R Subscale	B(p)	F(p)	Adjusted R <sup>2</sup>
<u>Teacher process score</u>			.02
FIRSTYR		<b>5.70(.02)</b>	
Not first year	<b>.42(.02)</b>		
COHORT	.26(.03)	5.03(.03)	
<u>Facilities score</u>			.00
FIRSTYR		1.31(.25)	
Not first year	.20(.25)		
COHORT	-.09(.45)	.58(.45)	

To more closely examine the effects of first year teachers on ECERS-R scores with the influx of many new teachers, one-way ANOVAs by FIRSTYR were run twice, once on 1999 data and again on 2003 data. Degrees of freedom (df) is always (1,244) for 1999 and (1, 248) for 2003.

Group sample sizes are as follows:

	<u>1999</u>	<u>2003</u>
First year	21	38
Non-first year	223	210

Results of analyses are as follows:

	<b>Results with 1999 scores</b>	<b>Results with 2003 scores</b>
ECERS-R Subscale	F(p)	F(p)
Space and Furnishings	.38(.54)	<b>3.83(.05)</b>
Personal Care	.76(.39)	.13(.72)
Language	.02(.90)	<b>8.39(.00)</b>
Activities	.96(.33)	2.15(.12)
Interactions	.57(.45)	<b>4.22(.04)</b>
Program structure	.45(.50)	.96(.33)
Parents and Staff	.56(.45)	2.57(.11)
Total	.86(.35)	<b>4.19(.04)</b>
Teacher process score	.28(.60)	<b>6.83(.01)</b>
Facilities score	.27(.60)	1.11(.29)

The mean scores along with standard deviations for first year and non-first year teacher groups, for 1999 and 2003 are as follows:

ECERS-R Subscale	Mean (standard deviation)	
	1999	2003
Space and Furnishings		
First yr	3.61(.82)	3.41(.94)
Non-first yr	3.76(1.12)	3.76(1.02)
Personal Care		
First yr	3.73(1.59)	3.65(1.28)
Non-first yr	4.03(1.53)	3.73(1.36)
Language		
First yr	3.74(1.24)	3.62(1.32)
Non-first yr	3.78(1.38)	4.28(1.28)
Activities		
First yr	2.99(.80)	3.16(.99)
Non-first yr	3.23(1.07)	3.43(.94)
Interactions		
First yr	4.25(1.73)	4.36(1.65)
Non-first yr	4.52(1.57)	4.93(1.57)
Program structure		
First yr	3.61(1.66)	3.79(1.74)
Non-first yr	3.88(1.80)	4.06(1.53)
Parents and Staff		
First yr	4.41(1.49)	4.10(.91)
Non-first yr	4.62(1.16)	4.38(1.04)
Total		
First yr	3.68(.94)	3.64(.95)
Non-first yr	3.90(1.06)	3.98(.92)
Teacher process score		
First yr	4.07(1.29)	3.94(1.33)
Non-first yr	4.22(1.25)	4.52(1.24)
Facilities score		
First yr	3.62(1.02)	3.46(1.17)
Non-first yr	3.77(1.31)	3.69(1.27)

### Analyses on teacher education and certification

Multivariate Analysis of Covariance (MANCOVA) was used to investigate the relative effects of teacher experience and education/certification on ECERS-R scores. The analysis was run twice, once for the 1999 data and again on the 2003 data. Teacher experience is operationalized by the FIRSTYR variable (coded 0 for no and 1 for yes). The teacher education variable is categorical (coded 1 for no BA, 2 for BA no certification, 3 for BA and certification). The degrees of freedom are always (1,215) for FIRSTYR and (2,215) for BACERT for 1999 data and are always (1, 238) for FIRSTYR and (2,238) for BACERT for 2003 data.

#### 1999 Results

<u>ECERS-R Subscale</u>	<u>B(p)</u>	<u>F(p)</u>	<u>Adjusted R<sup>2</sup></u>
<u>Space and Furnishing</u>			.07
FIRSTYR		.51(.48)	
Not first year	.18(.48)		
BACERT		<b>9.78(.00)</b>	
No BA	<b>-.44(.01)</b>		
BA no cert	<b>-.83(.00)</b>		
<u>Personal Care</u>			.02
FIRSTYR		.04(.85)	
Not first year	.07(.85)		
BACERT		<b>3.21(.04)</b>	
No BA	-.07(.75)		
BA no cert	<b>-.72(.02)</b>		
<u>Language</u>			.12
FIRSTYR		.01(.94)	
Not first year	-.02(.94)		
BACERT		<b>15.97(.00)</b>	
No BA	<b>-.90(.00)</b>		
BA no cert	<b>-1.12(.00)</b>		
<u>Activities</u>			.04
FIRSTYR		.55(.46)	
Not first year	.18(.46)		
BACERT		<b>6.06(.00)</b>	
No BA	<b>-.37(.02)</b>		
BA no cert	<b>-.62(.00)</b>		

<u>ECERS-R Subscale</u>	<u>B(p)</u>	<u>F(p)</u>	<u>Adjusted R<sup>2</sup></u>
<u>Interactions</u>			.02
FIRSTYR		.02(.89)	
Not first year	-.05(.89)		
BACERT		<b>3.98(.02)</b>	
No BA	-.39(.09)		
BA no cert	<b>-.80(.01)</b>		
<u>Program Structure</u>			.05
FIRSTYR		.05(.82)	
Not first year	.10(.82)		
BACERT		<b>6.64(.00)</b>	
No BA	-.43(.10)		
BA no cert	<b>-1.22(.00)</b>		
<u>Parents and Staff</u>			.07
FIRSTYR		.00(.95)	
Not first year	.02(.95)		
BACERT		<b>9.80(.00)</b>	
No BA	<b>-.50(.00)</b>		
BA no cert	<b>-.88(.00)</b>		
<u>Total</u>			.08
FIRSTYR		.17(.68)	
Not first year	.10(.68)		
BACERT		<b>10.49(.00)</b>	
No BA	<b>-.41(.01)</b>		
BA no cert	<b>-.82(.00)</b>		
<u>Teacher process score</u>			.08
FIRSTYR		.03(.85)	
Not first year	-.05(.85)		
BACERT		<b>10.55(.00)</b>	
No BA	<b>-.56(.00)</b>		
BA no cert	<b>-.96(.00)</b>		
<u>Facilities score</u>			.05
FIRSTYR		.64(.43)	
Not first year	.24(.43)		
BACERT		<b>6.46(.00)</b>	
No BA	<b>-.39(.03)</b>		
BA no cert	<b>-.96(.00)</b>		

2003 Results

<u>ECERS-R Subscale</u>	<u>B(p)</u>	<u>F(p)</u>	<u>Adjusted R<sup>2</sup></u>
<u>Space and Furnishing</u>			.01
FIRSTYR		2.20(.14)	
Not first year	.28(.14)		
BACERT		.81(.45)	
No BA	-.11(.59)		
BA no cert	-.21(.22)		
<u>Personal Care</u>			.00
FIRSTYR		.00(.99)	
Not first year	.01(.99)		
BACERT		1.79(.17)	
No BA	.48(.08)		
BA no cert	-.07(.74)		
<u>Language</u>			.02
FIRSTYR		<b>7.65(.01)</b>	
Not first year	<b>.67(.01)</b>		
BACERT		.01(.99)	
No BA	-.03(.91)		
BA no cert	.02(.92)		
<u>Activities</u>			.00
FIRSTYR		1.69(.19)	
Not first year	.23(.19)		
BACERT		.06(.94)	
No BA	.02(.91)		
BA no cert	-.05(.76)		
<u>Interactions</u>			.00
FIRSTYR		2.69(.10)	
Not first year	.49(.10)		
BACERT		.16(.86)	
No BA	-.00(.99)		
BA no cert	-.15(.58)		
<u>Program Structure</u>			.01
FIRSTYR		.76(.38)	
Not first year	.26(.38)		
BACERT		.03(.97)	
No BA	-.01(.97)		
BA no cert	.06(.83)		



<u>ECERS-R Subscale</u>	<u>B(p)</u>	<u>F(p)</u>	<u>Adjusted R<sup>2</sup></u>
<u>Parents and Staff</u>			.02
FIRSTYR		.55(.46)	
Not first year	.14(.46)		
BACERT		2.64(.07)	
No BA	-.11(.59)		
BA no cert	<b>-.39(.02)</b>		
<u>Total</u>			.01
FIRSTYR		2.39(.12)	
Not first year	.27(.12)		
BACERT		.42(.66)	
No BA	.03(.86)		
BA no cert	-.13(.40)		
<u>Teacher process score</u>			.01
FIRSTYR		<b>5.98(.02)</b>	
Not first year	<b>.58(.02)</b>		
BACERT		.01(.99)	
No BA	-.00(.99)		
BA no cert	.03(.89)		
<u>Facilities score</u>			.00
FIRSTYR		.31(.58)	
Not first year	.13(.58)		
BACERT		1.15(.32)	
No BA	-.07(.79)		
BA no cert	-.32(.13)		

Analyses on the effects of district factors on ECERS-R scores

Multivariate Analysis of Covariance (MANCOVA) was used to investigate the relative effects of district size and poverty on ECERS-R scores. Data collection year is also included in these analyses to determine whether improvement over time in ECERS-R scores might be impacted by district size or poverty.

The district size variable DISTSIZE operationalizes the number of 4-year-olds enrolled in preschool. It is categorical (coded 1 for small, 2 for medium and 3 for large). A small district is defined as having less than 500 four-year-olds enrolled; a medium district has between 500 and 1000 four-year-olds enrolled; and a large district has more than 100 four-year-olds enrolled). The district poverty variable POVGRP operationalizes the percentage of children receiving free or reduced price lunch. It is categorical (0 for lower poverty, 1 for higher poverty). The lower poverty group includes those districts with less than 66% of children receiving free or reduced price lunch, the higher poverty group includes the districts with 66% or more children receiving free or reduced price lunch. Note that the direction of effect for the COHORT beta weights are reversed to indicate positive change when scores increase over time.

<u>ECERS-R Subscale</u>	<u>B(p)</u>	<u>F(p)</u>	<u>Adjusted R<sup>2</sup></u>
<u>Space and Furnishings</u>			.01
COHORT	-.01(.89)	.02(.89)	
DISTSIZE		<b>3.31(.04)</b>	
Smallest	<b>.31(.05)</b>		
Medium	-.14(.28)		
POVGRP	.16(.26)	1.30(.26)	
<u>Personal Care</u>			.03
COHORT	<b>-.27(.03)</b>	<b>4.53(.03)</b>	
DISTSIZE		<b>3.96(.02)</b>	
Smallest	<b>.60(.01)</b>		
Medium	.17(.32)		
POVGRP	.21(.25)	1.32(.25)	
<u>Language and Reasoning</u>			.04
COHORT	<b>.41(.00)</b>	<b>11.92(.00)</b>	
DISTSIZE		1.31(.27)	
Smallest	.18(.37)		
Medium	.24(.13)		
POVGRP	<b>.39(.03)</b>	<b>5.05(.03)</b>	
<u>Activities</u>			.00
COHORT	<b>.19(.04)</b>	<b>4.38(.04)</b>	
DISTSIZE		.55(.58)	
Smallest	-.15(.31)		
Medium	-.06(.59)		
POVGRP	.12(.34)	.91(.34)	

<u>ECERS-R Subscale</u>	<u>B(p)</u>	<u>F(p)</u>	<u>Adjusted R<sup>2</sup></u>
<u>Interactions</u>			.05
COHORT	<b>.37(.01)</b>	<b>6.93(.01)</b>	
DISTSIZ		2.45(.09)	
Smallest	<b>.51(.03)</b>		
Medium	.09(.63)		
POVGRP	<b>.63(.00)</b>	<b>9.65(.00)</b>	
<u>Program Structure</u>			.01
COHORT	.21(.17)	1.90(.17)	
DISTSIZ		.16(.85)	
Smallest	.09(.71)		
Medium	-.06(.77)		
POVGRP	<b>.52(.02)</b>	<b>5.64(.02)</b>	
<u>Parents and Staff</u>			.08
COHORT	<b>-.22(.02)</b>	<b>5.43(.02)</b>	
DISTSIZ		<b>11.16(.00)</b>	
Smallest	<b>.71(.00)</b>		
Medium	-.05(.72)		
POVGRP	<b>.49(.00)</b>	<b>12.57(.00)</b>	
<u>Total</u>			.02
COHORT	.07(.44)	.60(.44)	
DISTSIZ		2.12(.12)	
Smallest	<b>.30(.05)</b>		
Medium	.01(.94)		
POVGRP	<b>.31(.02)</b>	<b>5.82(.02)</b>	
<u>Teacher Process score</u>			.03
COHORT	<b>.24(.03)</b>	<b>4.54(.03)</b>	
DISTSIZ		1.07(.34)	
Smallest	.27(.15)		
Medium	.09(.53)		
POVGRP	<b>.49(.00)</b>	<b>9.26(.00)</b>	
<u>Facilities score</u>			.02
COHORT	-.05(.67)	.18(.67)	
DISTSIZ		<b>5.52(.00)</b>	
Smallest	<b>.54(.00)</b>		
Medium	-.12(.44)		
POVGRP	.08(.63)	.23(.63)	

Change in SELA scores between 2002 and 2003

A MANOVA by COHORT (coded 0 for 2002 and 1 for 2003) was used to investigate the difference in scores. The total sample size is 434. The sample size for 2002 data is 182 and the sample size for 2003 data is 252. Note the degrees of freedom (df) is always (1,433). Note that the direction of effect for the COHORT beta weights are reversed to indicate positive change when scores increase over time. Results are as follows:

<u>SELA Subscale</u>	<u>B(p)</u>	<u>F(p)</u>	<u>Adjusted R<sup>2</sup></u>
<u>Literate Environment</u> COHORT	<b>.45(.00)</b>	<b>29.51(.00)</b>	.06
<u>Language Development</u> COHORT	<b>.32(.00)</b>	<b>15.74(.00)</b>	.03
<u>Knowledge of Print/ Book Concepts</u> COHORT	<b>.56(.00)</b>	<b>23.30(.00)</b>	.06
<u>Phonological Awareness</u> COHORT	<b>.25(.02)</b>	<b>5.80(.02)</b>	.03
<u>Letters and Words</u> COHORT	<b>.52(.00)</b>	<b>29.85(.00)</b>	.05
<u>Parent Involvement</u> COHORT	<b>.34(.00)</b>	<b>12.53(.00)</b>	.01
<u>Total</u> COHORT	<b>.41(.00)</b>	<b>35.47(.00)</b>	.07

The mean scores along with standard deviations for 1999 and 2003 are as follows:

<u>SELA Subscale</u>	<u>Mean (standard deviation)</u>	
	<u>1999</u>	<u>2003</u>
Literate Environment	2.68(.77)	3.12(.89)
Language Development	2.76(.80)	3.08(.88)
Knowledge of Print/ Book Concepts	2.21(1.09)	2.77(1.26)
Phonological Awareness	1.67(.87)	1.92(1.18)
Letters and Words	2.06(.83)	2.58(1.07)
Parent Involvement	2.04(.86)	2.37(1.09)
Total	2.42(.57)	2.82(.78)

## References

- Barnett, W. S. (1999). Long-term effects on cognitive development and school success. In W. S. Barnett and S.S. Boocock (Eds.), Early Care and Education for Children in Poverty: Promises, Programs and Long-Term Results. Albany, NY: State University of New York Press.
- Barnett, W. S., Tarr, J. & Frede, E. (1999). Early Childhood Education Needs in Low Income Communities: Final Report of an Assessment of Young Children's Educational Needs and Community Capacity in New Jersey's Abbott Districts. Rutgers University, New Brunswick, NJ: CEER.
- Barnett, W. S., Tarr, J., Esposito Lamy, C. & Frede, E. (2002). Fragile Lives, Shattered Dreams: A Report on Implementation of Preschool Education in New Jersey's Abbott Districts. Rutgers University, New Brunswick, NJ: CEER.
- Dunn, L. M. & Dunn, L. M. (1997). Peabody Picture Vocabulary Test-Third Edition (PPVT-3). Circle Pines, MN: American Guidance Services.
- Dunn, Padilla, Lugo & Dunn, (1986). Test de Vocabulario en Imágenes Peabody (TVIP). Circle Pines, MN: American Guidance Services.
- Harms, T., Clifford, R. & Cryer, D. (1998). Early Childhood Environment Rating Scale – Revised (ECERS-R). New York, NY: Teachers College Press.
- Frede, E. (1998). Preschool program quality in programs for children in poverty. In W. S. Barnett and S.S. Boocock (Eds.), Early Care and Education for Children in Poverty: Promises, Programs and Long-Term Results. Albany, NY: State University of New York Press.
- Frede, E., Dessewffy, M., Hornbeck, A. & Worth, A. (2001). Preschool Classroom Mathematics Inventory. Unpublished instrument.
- NIEER, (2004). The State of Preschool: 2003 State Preschool Yearbook.
- Smith, S., Davidson, S. & Weisenfeld, G. (2001). Support for Early Literacy Assessment (SELA). Unpublished instrument.
- Whitehurst, G. J. & Lonigan, C. J. (2000). The Get Ready to Read. Washington, DC: National Center for Learning Disabilities.