Smart Start and Preschool Child Care Quality in NC:

Change Over Time and Relation to Children's Readiness

March 2003

A report by the FPG-UNC Smart Start Evaluation Team

FPG Child Development Institute
The University of North Carolina at Chapel Hill

Smart Start and Preschool Child Care Quality in NC: Change Over Time and Relation to Children's Readiness

Smart Start Evaluation Team
FPG Child Development Institute
The University of North Carolina at Chapel Hill
March 2003

Donna Bryant, Kelly Maxwell, Karen Taylor, Michele Poe, Ellen Peisner-Feinberg, and Kathleen Bernier wrote this report. The authors appreciate the suggestions made by several colleagues at FPG who reviewed the report, including Don Bailey, Tal Black, Peg Burchinal, Dick Clifford, Debby Cryer, Diane Early, Jim Gallagher, and Mary Ellen Voegler-Lee. We also thank the data collectors: Becki Brinson, Dana Broach, Kim Heath, Nikki Jaeger, Joanne Liddle, Cyndee Lohr, Leanne Morris, April Points, and Tracey Tardiff. Thanks also to Steve Magers for programming and Satsuki Scoville for report design. Finally, we are sincerely grateful for the help of the preschool directors, teachers, parents, and children whose cooperation made the study possible.

Suggested Citation: Bryant, D., Maxwell, K., Taylor, K., Poe, M., Peisner-Feinberg, E., and Bernier, K. (2003). Smart Start and Preschool Child Care Quality in NC: Change Over Time and Relation to Children's Readiness. Chapel Hill, NC: FPG Child Development Institute.

This research was funded by the N.C. Department of Health and Human Services, Division of Child Development.

This and other reports from the Smart Start Evaluation Team may be found on the web: www.fpg.unc.edu/smartstart/.

Smart Start and Preschool Child Care Quality in NC: Change over Time and Relation to Children's Readiness

A Report by the FPG-UNC Smart Start Evaluation Team

Executive Summary

March, 2003

The primary goal of Smart Start is to ensure that all children enter school healthy and prepared to succeed. Based on extensive evidence that child care quality can positively affect children's learning, one of the main ways that Smart Start has tried to achieve the readiness goal is by improving the quality of children's experiences in early care and education programs. Smart Start has funded a variety of technical assistance (TA) activities to improve child care including on-site technical assistance, quality improvement and facility grants, teacher education scholarships, license upgrades, teacher salary supplements, and higher subsidies for higher child care quality or higher teacher education levels. These activities have been designed to improve child care quality and thereby expected to positively affect children's readiness for school.

This study included 110 preschool child care programs that were part of previous observational studies of NC child care quality between 1994 and 1999. The centers were located in 20 partnerships that entered Smart Start in the first, third, or fourth years of funding and were in a variety of geographic settings--urban and rural; Piedmont, East and West. We measured the quality of classroom practices and the center's level of participation in Smart Start-funded TA activities in the past year. From these classrooms we assessed 512 preschool children on their language, literacy, numeracy, and social-emotional skills.

Three main conclusions can be drawn: (1) Between 1993 and 2002, child care quality in this sample steadily and significantly increased, (2) Participation in Smart Start-funded activities was significantly positively related to child care quality, and (3) Children who attended higher quality centers score significantly higher on measures of skills and abilities deemed important for success in kindergarten than children from lower-quality centers.

While this study cannot identify which Smart Start TA activities have been most effective at improving quality, it does show that Smart Start-funded activities are significantly related to preschool classroom quality. In addition, this study replicated our earlier finding that a center's level of current participation in Smart Start-funded activities was related to classroom quality, but previous participation was not. A policy implication of this finding is that continuous quality enhancement efforts may be necessary to sustain higher levels of classroom quality. Although a significant increase in preschool classroom quality has been documented, a large proportion of preschool child care in NC is not yet at the high level of quality that is necessary to promote good outcomes for children.

Classroom quality was significantly, positively related to children's outcomes, over and above the effects of gender, income, and ethnicity. Children from poor and non-poor families were equally influenced by quality, providing support for quality improvement programs in all kinds of settings, serving all kinds of children, not just targeted to those who are poor. Children from poor families are more likely to have lower kindergarten readiness skills and thus be in greater need of positive early childhood experiences; however, all children benefit from improved programs.

Smart Start and Preschool Child Care Quality in NC: Change over Time and Relation to Children's Readiness

A Report by the FPG-UNC Smart Start Evaluation Team

North Carolina's Early Childhood Initiative, Smart Start, was established in 1993 as a partnership between state government and local leaders, service providers, and families to better serve children under six and their families. State funds are distributed to community partnerships, non-profit corporations established specifically for the purpose of supporting early care and education, family support, and health activities. The first round of twelve partnerships (18 counties) were awarded Smart Start funds in 1993 and have been called the "pioneer" partnerships. Subsequent rounds of partnerships were awarded funds each year from 1994 to 1997, until all NC counties were part of a Smart Start partnership. All 100 counties in North Carolina have received Smart Start funds since 1997, either as a single-county partnership or as part of a multiple-county partnership. Funding for Smart Start reached $220 million in 2001, but has been reduced to $190 million in fiscal year 2002-03.

The primary goal of Smart Start has been to ensure that all children enter school healthy and prepared to succeed. Smart Start's approach requires that local community partnerships plan how best to meet their own community's needs, improve and expand existing programs for children and families, and design and implement new programs. Although each partnership decides how best to meet the needs of its children and families, all work to improve the quality of early childhood education, including center-based care. By legislative mandate, partnerships spend at least 70% of their funds on child care. Statewide, about half of this amount is spent on child care subsidies for poor or working class families and about half is spent on child care quality improvement activities, both in centers and family child care homes. Activities to improve child care include on-site technical assistance (TA), quality improvement and facility grants, teacher education scholarships, teacher salary supplements, license upgrades, and higher subsidies for families to purchase higher child care quality.

This report focuses on the relation between Smart Start and center-based, preschool child care quality and children's readiness for kindergarten, addressing three main questions:

1. Has the quality of child care improved over time?
2. Does center participation in Smart Start-funded activities predict quality?
3. Do preschool children attending higher quality child care programs have better skills than children attending lower quality programs?

Previous Smart Start Evaluation studies have addressed the first two questions about quality. Preschool child care quality improved over the first six years of Smart Start and centers' level of participation in Smart Start-funded activities was significantly related to quality (Bryant, Maxwell, & Burchinal, 1999; Bryant, Bernier, Peisner-Feinberg, & Maxwell, 2002). In the new study described in this report, we collected child care observations and interviews from centers in 18 partnerships in 2002 to see if quality improvement as well as the relationship between Smart Start participation and quality continued.

Regarding the third question about the relation between classroom quality and children's outcomes, several studies have shown that overall classroom quality is related to cognitive, language, social, and emotional outcomes for children, both in the short-term and the long-term. For example, recent findings from the Cost, Quality, and Outcomes Study, a longitudinal study of child care center quality and children's long-term outcomes, found that children who attended higher quality preschool classrooms had fewer problem behaviors, better math skills, and better cognitive and attention skills through second grade (Peisner-Feinberg, Burchinal, Clifford, Culkin, Howes, Kagan, & Yazejian, 2001). In addition, the positive influences of better child care quality were even more pronounced for children at greater risk, particularly in their behavior problems and math skills. Other studies have shown that children in higher quality preschool classrooms exhibit greater competence with peers in preschool (Howes, Phillips, & Whitebook, 1992) and in kindergarten (Howes, 1990). Data from the North Carolina Head Start Quality Research Center showed that Head Start children in higher quality classrooms were rated by their teachers as having better social skills and fewer problem behaviors than children in lower quality classrooms (Bryant & Peisner-Feinberg, 2000; Peisner-Feinberg, 2000).

An earlier Smart Start Evaluation study partially addressed Question 3 by examining children's school readiness skills, but did not link school readiness to child care quality. Specifically, the 1999 study documented that children attending child care programs that participated intensively in Smart Start-funded improvement efforts were significantly more ready for kindergarten than their peers who attended non-participating child care programs (Maxwell, Bryant, & Miller-Johnson, 1999). Independent assessments and kindergarten teacher ratings showed that these children were half as likely to have language delays or behavior problems when they entered kindergarten. By including observations in classrooms, the current study fully addresses the third question of the relation between classroom quality and children's skills and knowledge.

Study Description

Sample - Centers

Data for this study were gathered from samples of child care centers in 12 pioneer partnerships (Round 1) and in 8 partnerships that entered Smart Start in 1996 or 1997 (Round 3 or 4). All centers in the current study had participated in at least one previous Smart Start Evaluation child care quality study and many had participated two or three times before. Table 1 summarizes the number of centers participating in the previous studies by year. We did not gather child care quality data from partnerships that began receiving Smart Start funding in Rounds 2 or 5. The evaluation team has conducted a wide range of studies using partnerships from all rounds, but never all rounds in the same study. We think a sample of 20 partnerships is sufficient to answer the questions posed in this study.

Table 1. **Number of centers visited in each year of the quality studies**

Study Year	Round 1	Rounds 3 & 4
1994	184	n/a
1996	188	n/a
1997	n/a	112
1999	135	85
2002	68	42

Child care in the pioneer partnerships (Round 1) has been most extensively studied because these counties were the first to participate in Smart Start. As Table 1 shows, in 1994 we

visited 184 child care centers in Round 1. Half were randomly selected from the county's list of licensed centers and half were selected specifically because they were participating in Smart Start-funded activities. Analyses of major center characteristics, including quality, showed no differences in results based on sampling strategy, so we did not use this sampling strategy again and collapsed these groups in our analyses. In 1996 we revisited all the 1994 centers that were still operating and added an additional random sample to achieve a total sample of 188 centers. In 1999, all centers visited in 1996 were asked to participate again. The results of these three previous studies of NC child care quality have been reported (Bryant, Maxwell, Burchinal, & Lowman, 1997; Bryant, Maxwell, & Burchinal, 1999; Bryant, Bernier, Peisner-Feinberg, & Maxwell, 2002).

The evaluation team first visited child care centers in Rounds 3 and 4 in 1997 as Smart Start was just beginning in these partnerships. We visited 112 centers in 1997 from 8 counties that started receiving Smart Start funds in Rounds 3 and 4. From 4 of these partnerships, we asked all child care centers to participate; in 4 larger partnerships we randomly sampled centers. The data collection procedures replicated those that had taken place in the Round 1 partnerships. In 1999, all centers in Rounds 3 and 4 that were visited in 1997 were asked to participate in another observation and interview.

The participation rate for these earlier studies was 75% in 1994, 64% in 1996, 75% in 1997, and 79% in 1999. These are satisfactory rates relative to other child care studies, and equal to or higher than participation rates in two often-cited child care observation studies with large samples (the Cost, Quality, and Outcomes Study, 1995; and the NICHD Early Child Care Research Network, 1996).

For the 2002 data collection phase, we did not have the resources to visit all previously visited centers. To assure a range of classroom quality in the sample, we divided all the centers observed in 1999 into quartiles based on their quality scores. We asked 100% of the centers in the top and bottom quartiles to participate and we randomly selected 50% of the centers in the middle two quartiles to participate. Because scores in the middle two quartiles were tightly clustered and because we randomly sampled among them, the sample included a range of quality scores. Of the 152 centers selected for participation, 13 did not serve 4-year-olds, so we did not include them. Of the remaining 139 centers, 110 agreed to participate (79%). The type of centers included 30% independent, 28% Head Start, 19% church-sponsored, 5% franchise, 2% developmental day, 1% public preschool and 15% other. The sample was not specifically drawn to be representative of the state distribution of types of centers, but the full range of child care center types did participate in the study.

To determine whether poor quality programs were more likely to drop out of the study, we compared the previous quality scores of centers in the sample in 1996, 1999, and 2002 with those not in the sample in those years. No differences were observed in the 1996 or 1999 data collection phase, however the 1999 mean quality score of the 2002 sample was significantly higher than the 1999 mean score of the centers not seen in 2002. This indicates that some of the lower quality centers in 1999 did not participate in the 2002 study. The data analysis methods were designed to take this into account so that our conclusions regarding change over time are valid.

Sample - Children

Child care directors from the 110 participating centers sent recruitment letters home to the parents of all children who were expected to attend kindergarten in the fall of 2002. Children whose parents consented were included in the study, up to 6 per classroom. If more than 6 children had consents, data collectors assessed those who had been at the center the longest time, randomly picking half boys and half girls. In total, they assessed 512 children. The average length of time children had attended their center was 23 months; 57% had attended their center for over a year. Boys made up 50.8% of the sample. Slightly more than half of the sample (54.7%) was White; 32.4% were Black, 2% Hispanic, 4.7% Native American, 2.2% Asian, and 4.1% Multiracial or Other. About half of the children (53.3%) were reported to be receiving a child care subsidy, the definition we used for poverty. Of the 274 children from poor families, 43.4% were White, 42.7% were Black, 2.6% Hispanic, 5.1% Native American, 1.5% Asian, and 4.7% were Multiracial or Other.

Procedures and Measures

One randomly selected preschool classroom was visited in each center between January and March 2002. Trained research assistants collected observational data and interviewed the classroom teacher and center director with measures described below. We provided participating directors and teachers with a $30 gift certificate for their help. Between April and June of 2002, the data collectors revisited each classroom to assess the children in a one-on-one session that lasted about 30 minutes.

Child care quality measure. The *Early Childhood Environment Rating Scale* (ECERS, Harms & Clifford, 1980) is a well-established measure of child care quality that assesses seven general areas: personal care routines, furnishings and display for children, language-reasoning experiences, fine and gross motor activities, creative activities, social development, and adult needs. Scores on each of 37 items can range from 1 to 7, with the overall mean score used as a global measure of the developmental appropriateness or quality of the classroom. To be consistent with other research, the adult needs items were not included in the overall classroom quality scores. An overall score from 1 to 2.9 is considered poor quality; scores from 3 to 4.9 are considered medium quality; and scores of 5 or greater are considered good to excellent quality. Although a revised version of the ECERS was published in 1998, we continued to use the original ECERS so that results can be compared over time on the same measure. Before data collection, observers were trained to an inter-rater agreement standard of at least 85% within one point. Inter-rater reliability during data collection (based on 10% of the observations) was 82.3% within 1 point or exact.

Director interview. Data collectors interviewed center directors to obtain information about center characteristics and services. This interview included a list of 14 different Smart Start-funded TA activities, most of them related to quality improvement, that the director or teaching staff might have participated in during the past year. Table 2 lists the categories of activities. The director interview was conducted after the classroom observation, so the observers' classroom ratings were based only on what they saw in the classroom and not biased by knowledge of center activities related to quality improvement. For each study year, we created a Smart Start participation index for each center by summing the total number of quality improvement activities.

Child assessments. Children's knowledge and skills were assessed by a number of measures, including several that were used in the NC School Readiness Assessment in 2000 (Maxwell, Bryant, Ridley, & Keyes-Elstein, 2001). The child care teacher was asked to rate children's social skills and problem behaviors (Social Skills Rating System, Gresham & Elliott, 1990). Teachers of 89% of the children (455 of 512) returned these forms. Language and math skills were assessed during one-on-one activities with the children, including the Peabody Picture Vocabulary Test III (PPVT-III) (Dunn & Dunn, 1997), the Applied Problems subtest of the Woodcock-Johnson (1989), a literacy assessment (Concepts About Print, Zill & Resnick, 1998), and 4 tasks that asked children to count and to identify letters, numbers and colors. Appendix A describes the measures in more detail. Before assessing study children, data collectors were trained on the administration of the child measures, practiced the measures several times, and then were observed by an experienced trainer to assure that they followed correct administration procedures.

Table 2. **Smart Start-funded activities**

1. Increased subsidies for higher quality care
2. Funds for teachers to attend college
3. On-site technical assistance
4. Quality improvement grants to upgrade license level (materials, facilities)
5. Teacher salary supplements
6. Support to achieve national accreditation
7. Training workshops (CPR, classroom practices, outdoor play, etc.)
8. Developmental screenings
9. Transportation for children
10. Enrichment activities in the classroom (story teller, art teacher)
11. Teacher substitutes
12. Support to improve services for children with disabilities
13. Lending library of appropriate materials
14. Subsidies – not tied to higher quality

Results

Child Care Quality Over Time

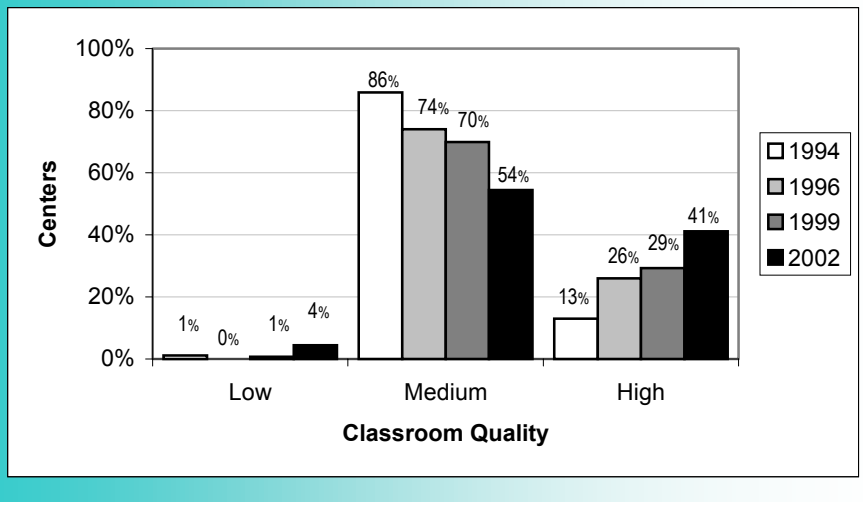
The first question addressed by this study was whether child care quality in this sample of NC centers improved over time. With Smart Start focusing considerable attention and resources on child care quality, increases would be expected. Using the overall ECERS score as the measure of quality, a statistical technique (linear model for mixed effects) was used to examine whether ECERS scores changed over time. The ECERS classroom quality scores significantly increased over time ($p < 0.001$). This analysis indicated that both the overall quality of the centers sampled each year and the quality of the individual centers significantly increased over the observed time period. Table 3 presents the mean ECERS scores, standard deviations, and ranges for centers in Round 1 and Rounds 3 & 4 by year, showing this positive increase over time.

Table 3. **Mean ECERS scores by study year**

Study Year	Round 1				Rounds 3 & 4			
	N	*Mean*	*SD*	*Range*	*N*	*Mean*	*SD*	*Range*
1994	180	4.25	0.64	2.5 - 6.3				
1996	188	4.52	0.69	3.0 - 6.3				
1997					112	4.37	0.81	2.6 - 6.4
1999	133	4.59	0.74	2.5 - 6.2	84	4.36	0.74	2.5 - 5.6
2002	68	4.73	0.93	2.6 - 6.8	42	4.76	0.96	2.8 - 6.6

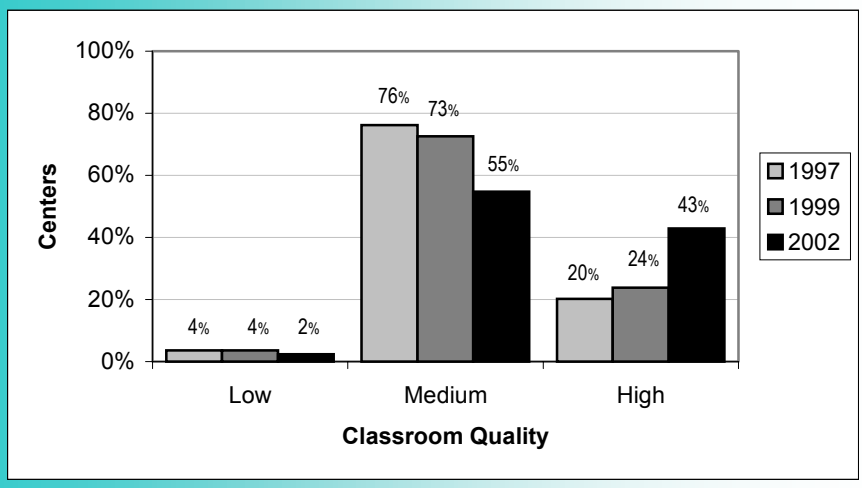
For both samples of centers (Round 1 and Rounds 3 & 4), the first observation--made at the beginning of Smart Start in their respective counties--yielded mean ECERS scores between 4.25 - 4.37, while the most recent ECERS scores, after 4 to 7 years of Smart Start, yielded mean scores of about 4.75. This is an increase of over ½ standard deviation, an effect size of 0.64 for Round 1 and 0.58 for Rounds 3 & 4. In educational research, interventions that can achieve this level of improvement are considered to be effective. (For comparison, Cohen (1988) considered an effect size of 0.5 as a "medium" effect and 0.8 as a "large" effect.)

Figure 1. **Quality of NC preschool child care: Round 1 counties**



Figures 1 and 2 illustrate the increased quality over time in Round 1 partnerships and Rounds 3 & 4. Because of the higher attrition of low quality

Figure 2. **Quality of NC preschool child care: Round 3 & 4 counties**



centers from 1999 to 2002, these figures include only centers seen at least twice. A score of 5 on the ECERS is considered "good" and a 7 is "excellent," so evaluators often use 5 as a threshold for "high" quality. For Round 1, the percentage of centers scoring at 5 or above increased steadily from 1994 to 2002. For Rounds 3 & 4, the percentage of centers in the highest category in 1999 was about the same as in 1996, but a large increase was present by 2002. In the first observation period (1994 for Round 1 centers and 1997 for Rounds 3 & 4), centers in the East had somewhat lower quality scores than centers in the Piedmont and West; however, at all later observation times, no regional differences were noted.

Relation Between Smart Start Participation and Child Care Quality

The second question was whether participation in Smart Start-funded activities was related to child care center quality. Table 4 presents the mean number of Smart Start-funded activities that centers participated in by study year and round. On average, centers participated in about 4 or 5 Smart Start-funded activities, with a range from 0 to 12. The exception is that in their baseline year of Smart Start, centers in Round 3 and 4 partnerships participated in a mean of 1 activity.

Table 4. **Mean number of Smart Start activities by study year**

	Round 1				Rounds 3 & 4			
Study Year	*N*	*Mean*	*SD*	*Range*	*N*	*Mean*	*SD*	*Range*
1994	184	3.91	2.76	0-11				
1996	188	4.93	2.87	0-12				
1997					110	1.03	1.69	0-7
1999	135	5.94	2.89	0-12	84	4.42	2.69	0-12
2002	68	4.69	2.08	0-12	42	4.38	2.39	0-9

Table 5 presents the relationship between a center's participation in Smart Start-funded activities and classroom quality based on how long the partnerships had been participating in Smart Start. This analysis shows that the number of Smart Start activities was not significantly related to classroom quality in the first year of Smart Start (1994 for Round 1 and 1997 for Rounds 3 & 4), but participation was related to classroom quality at each later

observation time (2 years, 5 years, and 8 years later). In addition, the strength of the relationship between quality and Smart Start participation increased over time. The positive effect of participation was the same for both Round 1 and Round 3 & 4 samples.

Table 5. **Relationship between number of Smart Start activities and quality**

Length of time in Smart Start	Effect Strength	Standard Error	Significance Level
During startup year	0.003	0.016	0.8431
After 2 years	0.039	0.013	0.0038
After 5 years	0.055	0.016	0.0008
After 8 years	0.085	0.033	0.0116

We also tested whether previous Smart Start participation (in 1999) predicted child care quality in 2002 and it did not. This analysis replicated our earlier finding that concurrent participation in quality improvement activities is important for supporting child care quality and that extensive previous participation does not guarantee that a center's *current* quality is high.

The Relation Between Child Care Quality and Children's Outcomes

The final research question was whether preschool classroom quality is related to children's skills and abilities. Means and standard deviations for the 11 child outcomes are presented in Table 6. The PPVT and the Woodcock Johnson are standardized measures, thus the average child in the United States will score about 100. The means in Table 6 indicate that the average child in this sample scored lower than the national average on the main measures of vocabulary (PPVT) and numeracy (WJ Applied Math). However, the average child in this study knew the rote skills of counting and naming letters fairly well. In the social development domain (also standardized measures), children in this sample were at the national average on positive social behaviors and a little above the national average in their problem behavior scores (that is, they had somewhat more behavior problems).

Table 6. **Child outcome measures**

Variable	*Mean*	*SD*	*Range*
Language and Literacy			
PPVT Receptive Language	94.9	14.85	51 - 136
Number of colors named or found	9.7	1.01	0 - 10
Number of letters named	13.5	9.93	0 - 26
Print Awareness	0.4	0.50	0 - 7
Book Knowledge	2.7	1.41	0 - 5
Story Comprehension	0.7	0.44	0 - 2
Numeracy			
WJ Applied Math	93.9	15.48	32 - 132
Highest number counted	22.2	21.16	1 - 40
Highest number counted with one-to-one correspondence	18.8	12.51	2 - 40
Social and Emotional			
SSRS Social Skills	101.7	13.46	56 - 131
SSRS Problems Behaviors	103.3	14.64	84 - 143

In our analyses, we dropped or modified some measures. The Naming Colors task had limited variability because most children knew all 10 colors so we did not use it. The two measures of counting were highly correlated; so we only used Counting with one-to-one Correspondence. Print Awareness and Story Comprehension had skewed distributions so we created binary scores for each measure (low/high).

We estimated the relation between children's scores and their classroom quality score by using a statistical technique (Hierarchical Linear Modeling or HLM) that allows us to take into account the clustering of children within classrooms. In addition to the ECERS score and Smart Start-funded activities score, we also included 3 child variables in the analysis that research suggests effect children's skills: gender, ethnicity/race, and poverty (defined as receiving a child care subsidy). For the continuous outcomes, we fit a general linear model. For the binary outcomes (Print Awareness, Story Comprehension), we fit a logistic model.

Appendix B includes the complete results from the HLM analyses. Table 7 summarizes these results showing the number of child outcomes for which each predictor was significant and the direction of the effect. Classroom quality was a significant positive predictor for 5 of the 9 child outcomes after accounting for the effects of gender, ethnicity, and poverty. Receptive Language, Print Awareness, Book Knowledge, Applied Math, and Counting One-to-One were all significantly positively related to quality. Children from higher quality centers had better skills. The effect sizes for Receptive Language and Applied Math were 0.20, which is considered a small effect in educational research.

Table 7: **Significant predictors of children's outcomes and direction of effect**

	Predictor			
Child Outcome	**High Classroom Quality**	**Boy**	**Ethnic Minority**	**Poverty**
Receptive Language	⬆ ***		⬇ ***	⬇ ***
Letters				⬇ **
Print Awareness	⬆ ***	⬇ **	⬇ *	⬇ **
Book Knowledge	⬆ **	⬇ **	⬇ **	
Story Comprehension				⬇ ***
Applied Math	⬆ ***	⬇ *	⬇ ***	⬇ ***
Counting One-to-One	⬆ ***	⬇ **		
Social Skills		⬆ ***	⬆ **	⬇ ***
Problem Behaviors		⬇ *		⬆ **

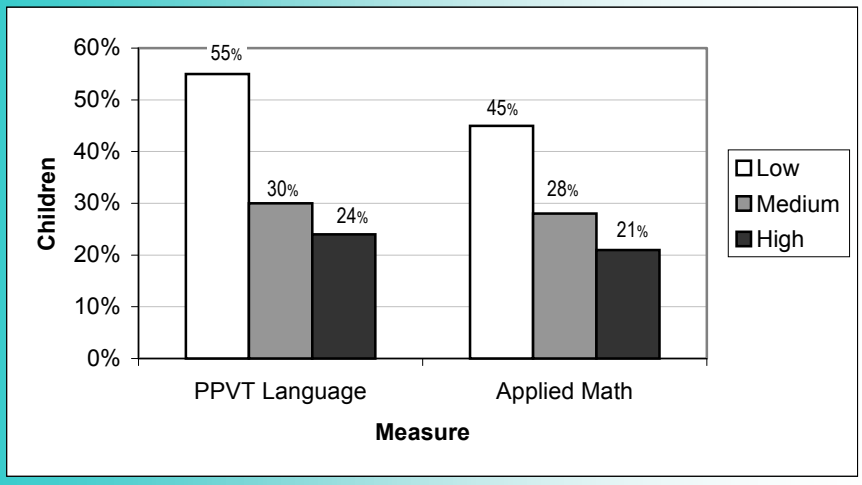
Notes: * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$.

Compared to girls, boys scored significantly lower on Print Awareness, Book Knowledge, Applied Math, and Counting, and were rated by their teachers as having fewer Problem Behaviors and higher Social Skills. Being a child of Black, Hispanic, or Other origin was related to lower Applied Math, Receptive Language, Print Awareness, and Book Knowledge scores, but higher ratings on Social Skills. Poverty was a very strong predictor of children's outcomes, related to lower scores on almost every cognitive and language measure and higher problem behavior scores.

The HLM analyses revealed only one significant interaction. For Print Awareness the effect of quality existed for boys ($\beta = 0.65$; $p = 0.006$) but not for girls. Neither subsidy nor ethnicity were found to affect the relationship between quality and any of the outcomes, indicating that the positive relation between good quality classrooms and children's outcomes was similar for poor and non-poor children and for children of various races.

Another way to describe the relation between quality and children's outcomes is to evaluate the percentage of children scoring at especially low levels on the outcome measures. Figure 3 illustrates that much higher percentages of children in the low- and medium-quality classrooms scored poorly (below 85 or, in other words, at least 1 standard deviation below the mean) in Receptive Language (PPVT) and Applied Math Problems than in the high-quality classes. The differences between low quality and medium quality seem greater than between medium quality and high quality.

***Figure 3*: Percentage of children scoring poorly by classroom quality**



Discussion

This study resulted in three main conclusions: (1) since 1994, child care quality in this sample of NC child care centers has steadily and significantly improved, (2) participation in Smart Start-funded activities was significantly positively related to child care quality, and (3) children who attended higher quality centers scored significantly higher on measures of skills and abilities that are important for school success compared to children from lower quality centers.

One of the main goals of Smart Start has been to improve the quality of children's early care and education. Smart Start leaders have focused considerable attention and funding on improving child care quality and positive effects are being seen. Local partnerships have funded a variety of strategies and activities over the years and most partnerships implement several activities in a given year. The research literature provides little guidance as to which types of technical assistance (TA) activities work best, but the overall effect of NC's Smart Start efforts has been positive.

The data from the current study do not allow us to determine whether certain types of technical assistance for quality improvement are more effective than other types. Nor can we tell whether a certain "dosage" of TA is minimally required before an effect on quality is seen. What we can conclude is that participation in more Smart Start-funded TA activities is significantly positively related to classroom quality. Given that the study used a very rough indicator of Smart Start participation (total number of activities, without regard to type or duration), we think the relationship between quality and Smart Start participation is a robust finding. Future research should undertake a more detailed and controlled study of the effectiveness of different types of Smart Start quality improvement TA activities.

The Smart Start Evaluation team recently wrote a detailed report about the strategies used by 12 partnerships that have been highly successful in increasing the proportion of high-quality programs in their communities (Taylor & Bryant, 2002). These strategies include strong leadership; strategic planning for a system of quality improvement programs; support for the education and professional development of the workforce; financial rewards for higher education and improved quality; on-site, customized technical assistance; and effective collaborations with multiple community agencies. No silver bullet or single strategy seems to have been effective, but successful partnerships have used several, coordinated approaches.

This study replicated our earlier finding that a center's level of current participation in Smart Start-funded activities was related to classroom quality, but that previous participation was not. In other words, extensive previous participation in Smart Start does not guarantee that a center's current quality is high. This finding has important policy implications. Continuous quality enhancement efforts may be needed to maintain and sustain the levels of classroom quality that will improve children's growth and development. Partially because of teacher turnover (31% in a recent statewide work force study, Russell, Lyons, Grigoriciuc, & Lowman, 2002) and partially because a large proportion of NC preschool child care is still of low to average quality, continuous efforts are still needed. Perhaps someday the early care and education system will be adequate to ensure that every child in NC has access to high quality care, but that day is not here yet.

While it is encouraging that the number of high quality classrooms in this evaluation sample has doubled since Smart Start began, almost 50% were still below the level of quality that is widely recognized as desirable for all children. The majority of child care centers provide care that is below "good" on the ECERS rating scale (a score of 5) and the average child in these centers is performing at levels below the national average. If we want NC preschoolers' skills to meet or exceed the national average, then their child care environments must improve.

Both the pioneer partnerships and the partnerships that entered into Smart Start later in the decade have seen improvements in quality and the positive effect of participating in Smart Start-funded activities. This indicates that different types of partnerships (large/small, urban/rural) have been able to achieve change. The fact that the Round 3 and 4 partnerships seem to be at about the same quality level as the Round 1 partnerships, in spite of having less time in which to achieve these gains, is possibly a sign that the pioneer partnerships were truly "pioneers." They experimented with a variety of approaches and shared the best of the approaches--those that were most effective--with the later entering partnerships. This might also explain why the centers in Round 3 and 4 partnerships participated in a smaller number of activities in their first year than did centers in Round 1. Perhaps their initial offerings to centers were more intense or spread more broadly across the partnership. Within 2 years, though, the amount of Smart Start TA activities offered in Rounds 3 and 4 was similar to that offered in Round 1.

Turning to the child outcome results, this is not the first study to show a significant positive relationship between good child care quality and positive outcomes for preschoolers, but this study replicates such results within a large sample of North Carolina preschoolers. On most measures of young children's cognitive, language and social skills, we saw a significant positive relation with classroom quality. The association between quality and outcomes was similar for children from both poor and non-poor families and for White, Black, and Hispanic children. This provides support for quality improvement programs in all kinds of settings, serving all kinds of children. All children, including children from poor families who are more likely to have fewer kindergarten readiness skills, can benefit from the richer early childhood experiences in higher quality care.

Limitations of the study. This study included only one age range of early childhood programs--preschool classrooms of 3- and 4-year-old children. Had we evaluated infant-toddler child care or family child care homes, we would likely have seen an even lower overall quality of care, as other studies have shown (e.g., Cost, Quality, & Child Outcomes Study Team, 1995; Kontos, Howes, Shinn, & Galinsky, 1995; Peisner-Feinberg, Bernier, Bryant, & Maxwell, 2000). This study also included very few programs with low quality ECERS scores (below 3 on the rating scale), possibly because low-quality programs are more likely to refuse to participate in evaluations. However, the relations we found between Smart Start participation and child care quality applied to centers across the range of quality that was included in the study and would be expected to apply to centers at the lower end of the quality continuum. In fact, had the full range of quality been represented in the study, the relations could well have been stronger.

Another limitation concerns the larger attrition rate among lower quality centers from 1999 to 2002 than in the medium and high quality groups. Of the 54 low-quality centers in 1999, 6 had closed and thus could not be observed in 2002. Of the 39 middle-quality centers in 1999, 7 had closed. Of the 29 centers that refused to participate, 8 were of high quality, 11 medium, and 10 low. The analytic strategy took into account the fact that some centers were not represented at every time point, so the results regarding the relation between participation

and quality are valid regardless. Finally, the study does not establish causality between Smart Start participation, child care quality, and child outcomes. Random assignment of centers to Smart Start and of children to centers is required to establish causality, but is not feasible for a community initiative such as Smart Start.

We mention these cautions about interpreting the encouraging results of this study only to place them in context. We believe this study convincingly demonstrates the positive links between Smart Start participation and preschool child care quality and between quality and children's readiness for school. These results also point to the need for more research on technical assistance for quality improvement to help early childhood programs know where best to invest their resources. Finally, these NC results confirm what other national studies have shown--higher quality preschool classrooms are positively associated with children's knowledge and skills. Smart Start appears to be effective in improving child care quality and children's outcomes, yet much more progress can be made.

References

Bryant, D., & Peisner-Feinberg, E. (June, 2000). *Head Start Quality and Child Outcomes: Processes and Predictors*. Presentation at Head Start's Fifth National Research Conference. Washington, DC.

Bryant, D., Bernier, K., Peisner-Feinberg, E., & Maxwell, K. (2002). *Smart Start and Child Care in North Carolina: Effects on Quality and Changes over Time*. Chapel Hill, NC: Frank Porter Graham Child Development Institute.

Bryant, D., Maxwell, K., & Burchinal, M. (1999). Effects of a community initiative on the quality of child care. *Early Childhood Research Quarterly*, *14*, 449-464.

Bryant, D., Maxwell, K., Burchinal, M., & Lowman, B. (1997). *The effects of Smart Start on the Quality of Preschool Child Care*. Chapel Hill, NC: Frank Porter Graham Child Development Center.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed.)*. Hillsdale, NJ: Lawrence Earlbaum Associates.

Cost, Quality & Child Outcomes Study Team (1995). *Cost, quality and child outcomes in child care centers, public report, second edition*. Denver: Economics Department, University of Colorado at Denver.

Dunn, L. M, & Dunn, L. M. (1997). *Peabody Picture Vocabulary Test - Third Edition (PPVT-III)*. Circle Pines, MN: American Guidance Services.

Gresham, F. M., & Elliott, S. N. (1990). *Social Skills Rating System Manual*. Circle Pines, MN: American Guidance Service.

Harms, T., & Clifford, R. M. (1980). *Early Childhood Environment Rating Scale*. Columbia, NY: Teachers College Press.

Howes, C. (1990). Can the age of entry into child care and quality of childcare predict adjustment in kindergarten? *Developmental Psychology, 26*, 292-303.

Howes, C., Phillips, D. A., & Whitebook, M. (1992). Thresholds of quality: Implications for the social development of children in center-based care. *Child Development, 63*, 449-460.

Kontos, S., Howes, C., Shinn, M., & Galinsky, E. (1995). *Quality in family child care and relative care*. New York: Teachers College Press.

Maxwell, K., Bryant, D., & Miller-Johnson, S. (1999). *A Six-County Study of the Effects of Smart Start Child Care on Kindergarten Entry Skills*. Chapel Hill, NC: Frank Porter Graham Child Development Institute.

Maxwell, K.L., Bryant, D., Ridley, S., & Keyes-Elstein, L. (April, 2001). *North Carolina's Kindergartners and Schools*. Chapel Hill, NC: University of North Carolina at Chapel Hill, Frank Porter Graham Child Development Center.

NICHD Study of Early Child Care Network. (1996). Characteristics of infant child care: Factors contributing to positive caregiving. *Early Childhood Research Quarterly, 11*, 269-306.

Peisner-Feinberg, E. S. (2000, April). Why does quality matter? In Hopmann, M. (Chair) *Findings from the Head Start Quality Research Centers*. Symposium presented at the annual meeting of the National Head Start Association, Washington, DC.

Peisner-Feinberg, E. S., Bernier, K., Bryant, D., & Maxwell, K. (2000). *Family Child Care in North Carolina*. Chapel Hill, NC: Frank Porter Graham Child Development Center.

Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L., & Yazejian, N. (2001). The relation of preschool quality to children's cognitive and social developmental trajectories through second grade. *Child Development, 72*(5), 1534-1553.

Russell, S., Lyons, J., Grigoriciuc, M., Lowman, B., (2002). *Working in Child Care in North Carolina 2001*. Chapel Hill, NC: Child Care Services Association.

Taylor, K. & Bryant, D. (2002). *Demonstrating Effective Child Care Quality Improvement*. Chapel Hill, NC: Frank Porter Graham Child Development Institute.

Woodcock, R.W., McGrew, K.S., & Mather, N. (2001). *Woodcock-Johnson III Tests of Achievement*. Itasca, IL: Riverside Publishing.

Zill, N., & Resnick, G. (1998). *Family and Child Experiences Survey* (FACES) conducted by Westat for Head Start National Study.

Appendix A: Child Assessment Battery - Spring 2002

This appendix includes a complete list of the child measures used in the study. The language and literacy measures were administered to children in a one-on-one session and the social and behavioral measures were from teacher ratings of children.

Language and Literacy

Color Names (Zill & Resnick, 1998). This subtest was adapted and used with permission from the Head Start Family and Child Experiences Survey (FACES; Zill & Resnick, 1998). The Color Names subtest is a simple color naming and identification task. Children name up to 10 colors when shown a sheet of bears of different colors. This subtest yields a raw score ranging from 0 to 10 for the number of colors named and for the number of colors named or identified.

Peabody Picture Vocabulary Test - III, Form A (PPVT; Dunn & Dunn, 1997). The PPVT was used to measure children's receptive language skills. The PPVT can be administered to individuals as young as 2 years 6 months and as old as 90 years. It consists of 204 items arranged in order of increasing difficulty; most individuals complete 60 or fewer items. Test procedures involve showing the child a picture plate and asking the child to select the picture that best represents the stimulus word presented by the assessor. The PPVT was individually administered to children. Standard scores with a mean of 100 and a standard deviation of 15 were used in the analysis. National norms were set so that 16% of children in the standardization sample had scores less than 85 and an additional 16% had scores greater than 115.

Letter Identification. Children are shown 3 pages of letters randomly ordered and including all 26 letters of the alphabet. Children are asked to name any letters they know. Scores range from 0-26.

Story and Print Concepts (Zill & Resnick, 1998). This subtest was adapted and used with permission from the Head Start Family and Child Experiences Survey (FACES; Zill & Resnick, 1998). It measures children's book knowledge, comprehension, and print awareness. For this subtest, the research assistant read to each child a book entitled "Where's My Teddy?" (Alborough, 1992, 1995) and asked 12 questions about the book and its contents. Raw scores were generated for each of the conceptual areas (i.e., book knowledge, comprehension, and print awareness). Raw scores can range from 0-5 for book knowledge, 0-2 for comprehension, and 0-7 for print awareness.

Numeracy

Woodcock-Johnson Psycho-Educational Battery – Revised (Woodcock & Johnson, 1989, 1990). The Applied Problems subtest from the Woodcock-Johnson Tests of Achievement, Form A was individually administered to children. The Applied Problems subtest consists of items designed to assess children's skills in analyzing and solving practical math problems. The Woodcock-Johnson provides norms for children as young as 24 months to adults over the age of 90 years of age. Standard scores with a mean of 100 and a standard deviation of 15 were used in the analysis. National norms were set so that 16% of children in the standardization sample had scores less than 85 and an additional 16% had scores greater than 115.

Number Identification. Children are shown a page with the numbers 1-10 in random order and asked to name them. Scores range from 0-10.

Counting and Counting Bears (adapted from Head Start Family and Child Experiences Survey Spring 1998 Assessment; used with the permission of Nicholas Zill and Gary Resnick, Westat). In these two tasks, children are asked to count and then shown a sheet of little bears and asked to count them pointing one to one. Scores range from 0 to 40 on Counting, and from 0 to 40 on Counting Bears. Children who could count above 40 were given a score of 40 and redirected onto the next task.

Social Skills

Social Skills Rating System (SSRS; Gresham & Elliott, 1990). The teacher form of the Social Skills Questionnaire was used to gather information about children's social skills and problem behaviors. Teachers completed a 30-item rating scale that measures children's social skills on a scale of 0 to 2, with a higher score indicating greater skills. Teachers also completed an 18-item rating scale that measures children's problem behaviors on a scale of 0 to 2, with a higher score indicating more problems. Standard scores with a mean of 100 and a standard deviation of 15 were used in the analysis. National norms were set so that 16% of children in the standardization sample had scores less than 85 and an additional 16% had scores greater than 115.

Appendix B: Results from the HLM Analyses

	Applied Math		Receptive Language		SSRS Social Skills		SSRS Problem Behaviors		Counting One-to-One	
	β	SE	β	SE	β	SE	β	SE	β	SE
Intercept	100.05***	4.80	99.25***	4.22	98.94***	5.10	100.95***	4.26	2.69***	0.20
ECERS	3.30***	0.97	3.14***	0.81	1.40	1.10	-1.94	1.00	0.14***	0.04
Male	-2.48*	1.18	0.35	1.03	4.65***	0.94	-2.26*	1.03	-0.16**	0.06
Subsidy	-6.33***	1.66	-6.30***	1.42	-5.05***	1.46	4.51**	1.52	-0.09	0.07
Non-white	-5.60***	1.65	-9.75***	1.44	4.08**	1.39	-2.30	1.44	-0.11	0.07
ECERS x Male	-0.31	1.26	0.28	1.10	-0.65	1.01	1.04	1.12	0.08	0.06
ECERS x Subsidy	-1.82	1.63	0.70	1.38	0.49	1.38	-1.35	1.45	0.04	0.07
ECERS x Non-white	3.09	1.56	0.02	1.35	1.48	1.41	-0.94	1.40	0.07	0.07

	Letters Named		Print Awareness		Book Knowledge		Story Comprehension	
	β	SE	β	SE	β	SE	β	SE
Intercept	1.81***	0.38	-0.30**	0.10	2.75***	0.42	1.12***	0.11
ECERS	0.14	0.08	0.39***	0.12	0.25**	0.09	0.20	0.11
Male	-0.19	0.10	-0.64**	0.22	-0.32**	0.11	-0.30	0.18
Subsidy	-0.53***	0.14	-0.85***	0.21	-0.29	0.15	-0.81***	0.24
Non-white	-0.11	0.14	-0.46*	0.20	-0.48**	0.15	-0.30	0.23
ECERS x Male	0.21	0.11	0.52*	0.24	0.13	0.13	0.20	0.18
ECERS Main Effect – Male			0.65**	0.18				
ECERS Main Effect – Female			0.13	0.16				
ECERS x Subsidy	0.03	0.13	-0.27	0.23	0.18	0.15	-0.17	0.20
ECERS x Non-white	-0.16	0.13	0.15	0.22	-0.03	0.15	-0.29	0.22

Notes: * = $p < 0.05$; ** = $p < 0.01$; *** = $p < 0.001$.