CEMETR-2004-02
OCTOBER 2004

CEME Technical Report

The Center for Educational Measurement and Evaluation

Evaluation of the Effects of Creative Curriculum on Quality and Child Outcomes in Head Start

Richard Lambert
Megan O'Donnell
Martha Abbott-Shim

A PUBLICATION OF THE CENTER FOR EDUCATIONAL MEASUREMENT AND EVALUATION

CEME
The Center for Educational Measurement and Evaluation
College of Education
UNC CHARLOTTE

Correspondence regarding this paper should be directed to:
Richard G. Lambert, Ph.D., Ed.S.
Associate Professor
Department of Educational Leadership
UNC Charlotte
Charlotte, NC 28223-0001
phone: 704.687.8867
fax: 704.687.3493
email: rglamber@uncc.edu
web: http://education.uncc.edu/rglamber

Executive Summary

Classroom quality and child outcomes were measured in Head Start classrooms across two separate sites in the southeastern United States. The classrooms were randomly assigned to treatment and control conditions. The treatment teachers received training and technical assistance to enhance the implementation of *The Creative Curriculum* ® (Dodge, Colker, & Heroman, 2002) *(CC)*. The following measures were used to examine classroom quality: Early Childhood Environment Rating Scale-Revised (Harms, Clifford, & Cryer, 1998) (ECERS-R); Caregiver Interaction Scale (Arnett, 1989) (CIS); Teacher Behavior Rating Scale (Landry et al., 2004) (TBRS); and *CC* Implementation Checklist (Dodge, Colker, & Heroman, 2003). The treatment condition yielded greater gains than the control condition across all measures at the end of the second year of the study. Stronger effect sizes were found in the treatment group at the site that maintained higher quality technical assistance. Across the full range of child outcomes, there were no differences detected between the treatment and control children.

Final Report for the Evaluation of the Effects of Creative Curriculum on Classroom Quality and Child Outcomes in Head Start

As fiscal resources continue to be inadequate to meet the educational needs of all of America's families and children, educational systems find it necessary to compete with other governmental programs for limited resources. With the current increase in accountability demands, teachers and administrators are called upon to base more educational decisions and practices on sound research evidence (Bush, 2001; Bowman, Donovan & Burns, 2001; Spodek & Saracho, 2006). For example, the No Child Left Behind legislation (US Department of Education, 2001) has mandated state accountability testing programs, the equating of aggregated child test scores with classroom and building quality, and more "evidence-based research" on educational practices.

There has been much debate over teacher directedness, appropriate amount of play, and individualization throughout the history of early childhood education in the U.S. However, comparative preschool curricula research remains scarce (Nourot 2004). Of those studies that have been done, very few have implemented random assignment, and in non-experimental studies curriculum differences are often times confounded with other program characteristics, as well as characteristics of the children attending those programs (Bowman, Donovan, & Burns 2001). There have been attempts to gather scientific evidence for the effectiveness of curriculum models or approaches in early childhood education in the past. In the 1960's and 70's research projects such as the Head Start Planned Variation Study, the Consortium of Longitudinal Studies, and the Louisville Experiment compared the child outcomes of curricular models based on differing theoretical perspectives.

The Head Start Planned Variation Study did not use random assignment; however it did introduce new curriculum models that focused on training used to increase implementation fidelity. In this particular study, the largest achievement gains were found in those models that employed direct instruction. Project Follow Through, which examined curricula models in school age children, found similar results. Larger achievement gains were found in the direct instruction models. More positive gains were also found in the emotional and social development of the children who were receiving direct instruction. The results of this study, however, continue to raise questions and dispute (House, Glass, McLean, & Walker 1978; St. Pierre, Anderson, Proper & Stebbins 1978; Stebbins, St. Pierre, Poper, Anderson & Cerva 1977; MacIver & Kemper 2002).

There were some small scale randomized studies that compared the effects of alternative models on children's learning and development with long term follow up. One of these was the High Scope Comparison Study (Schweinhart & Weikart 1997), which found that direct instruction lead to negative outcomes for social and emotional development over an extended period of time (Schweinhart, Weikart, & Larner 1986). Follow up through the age of 23 showed no decreases in delinquency and crime in comparison to earlier studies that utilized curriculum models that emphasized more child-initiated activities (Schweinhart & Weikart 1997).

As social, economic, and political conditions are placing mounting pressure on early childhood programs to produce increases in child outcomes and to justify public and private investments in the maintenance and expansion of early education and care, there is a critical need for evidence-based early childhood curricula. This need is particularly amplified by evidence illustrating how early childhood researchers began to more fully understand the complex interchanges between young children and the various contexts in which they develop (Phillips,

1996). Our understanding of both what young children can learn and do, and how supportive contexts can strengthen learning, is expanding (Bowman, Donovan, & Burns, 2000). For example, current practice in the field of early childhood education has failed to meet the challenges presented by research developments with respect to emergent literacy (Dickinson & Tabors, 2001; Dickinson, 2002).

Whereas the public is demanding more out-of-home care and economic realities make it a part of the typical American family with young children, the cost of high quality early childhood care and education is beyond the reach of many American families. The public is not only demanding more public investment in the early care delivery system but higher quality programs as well. Policy makers are often convinced of the public necessity and economic value of early childhood programs, but expect a return for the investment of public funds and a measure of accountability (Lambert, Abbott-Shim, & Sibley, 2006). Therefore, public policy is increasingly addressing educational reform, accountability, and achievement gaps; and more and more attention is focused on the role of early education in fostering school readiness. Public schools expect children to enter kindergarten ready to learn (Kagan, Moore, & Bredekamp, 1997).

The U.S. Department of Health and Human Services has begun to address these issues within their Family and Children's Experiences Survey and Quality Research Consortium research agendas (Administration on Children, Youth, and Families; 2001). The U.S. Department of Education funded two cohorts of the Preschool Curriculum Evaluation Research (PCER) grant programs. These studies have examined the impact of a variety of early childhood curricula and interventions on child outcomes and classroom quality (Lambert, O'Donnell, Abbott-Shim, & Kusherman, 2006). The research reported in this paper was funded as part of a larger study within the first cohort of the PCER grants.

The early childhood curricula under investigation in the PCER research vary in both their depth and breadth. Some of the models are add-on curricula that focus on a specific domain of learning, such as emergent literacy, preschool mathematics, or social behavior. Other researchers are examining comprehensive, integrated curricula that provide the basis for all teaching and learning in the preschool classroom. The research reported in this paper examined the implementation of *The Creative Curriculum for Preschool: Fourth Edition (CC)* (Dodge, Colker, & Heroman, 2002) and its relationship to classroom quality.

The Creative Curriculum Approach to Early Childhood Education

CC rests on a firm theoretical foundation based on the work of Maslow, Erikson, Piaget, Vygotsky, Gardner, and Smilansky (Dodge, Colker, & Heroman, 2002). It is comprehensive because it encompasses all areas of child development including the socio-emotional, physical, cognitive, and language domains. *CC* involves a highly integrated process which includes five components: "knowing how children develop and learn, creating the learning environment, knowing what content children need to learn, and understanding the teacher's role and the family's role" in providing optimal learning experiences that promote each individual child's development (Dodge, Colker, & Heroman, 2002, p. 241). This five-part curriculum framework is applied to 10 interest areas (Blocks, Dramatic Play, Toys and Games, Art, Library, Discovery, Sand and Water, Music and Movement, Cooking, Computers, and Outdoors) in a well-organized classroom learning environment.

The *CC* classroom is viewed as a community where learning takes place through positive relationships between and among children and adults. Teachers are guided to use a wide range of teaching strategies including both child-initiated learning and direct teaching and the many

teaching approaches in between. Instructional approaches are based upon the material to be taught and the individual needs of children.

The *CC* model was examined in a research study conducted by the U.S. Department of Defense Education Activity in Sure Start classrooms (Abbott-Shim, 2000). Sure Start serves the needs of primarily low income military and civilian dependents' preschool children and their families living overseas. All teaching staff (both teachers and their assistants) and their supervisors received annual training on the *CC* to ensure consistency in implementation of the model. The Sure Start children made statistically significant gains on all of the measures from the fall to the spring of the school year. The Sure Start Effectiveness Study provides some evidence that the *CC,* when implemented as intended and supported by ongoing staff development, is a curriculum model that is associated with positive child outcomes.

Approaches to Training and Technical Assistance in Early Childhood Education

The body of research literature on early childhood teacher training and staff development is also important for an ongoing understanding of the implementation of curricula. Research on the effects of curricular models is completely intertwined with the capability of individual teachers to implement the curricular models under investigation. It has been found that training improves teacher practice only if: the training participants are actively involved; the model provides teachers with both early childhood theory and practical application; the sessions are ongoing with each session building on earlier sessions; the trainer observes and provides feedback on classroom implementation; and the participants are provided opportunities to reflect on what they have learned and to share their accomplishments and challenges (Epstein, 1993). In addition, effective administrators provide systematic, in-service training on site and supervisory support for their staff's curriculum implementation (Epstein, 1993). Successful

curriculum implementation is an ongoing process. It takes thoughtful planning, staff supervision and training, ongoing monitoring, and good communication to make sure that a curriculum is consistently implemented throughout a program.

Research has shown that consultation is an effective method that can directly contribute to staff development and education (Palsha and Wesley, 1998). According to Trohanis (1980), consultation is one of the most utilized forms of technical assistance. Technical assistance is broadly defined as "a systematic process that uses various strategies involving people, procedures, and products over a period of time to enhance the accomplishment of mutual goals…" (Trohanis, 2001). Consultation is defined similarly as a problem-solving process where both a consultant and a consultee collaborate in creating solutions to mutually identified problem areas (Brown, Pryzwansky, & Schulte, 1987). According to Buysse & Wesley (2004) consultation involves establishing productive working relationships; mutually identifying needed areas of focus as well as establishing goals for change; identifying and assessing strategies in meeting those goals; and collaboratively evaluating the plan for change.

There are some key factors that contribute to developing an effective consultation model. Research suggests that higher levels of change can take place when the following occur: (1) a shared knowledge base is collaboratively formulated, (2) staff members are involved in assessing their own needs, (3) ongoing staff development occurs over an extended period of time, and (4) staff members are given the opportunity to apply their new skills and knowledge within the work setting (Bailey, 1989; Bailey, McWilliam, & Winton, 1992; Fullan, 1993; Guskey, 1986; Joyce & Showers, 1983; Winton, 1990; Winton et al., 1997). In addition to this, it has also been suggested that consultation models can be most effective when consultants are persons of diverse backgrounds from within the community. According to Palsha and Wesley (1998) this is

because these individuals have existing relationships with childcare centers and staff members and can continue providing their support and knowledge over an extended period of time.

Researchers have developed two distinct consultation models used to address needs within the centers. Wesley's (1994) innovative consultation model stressed the importance of establishing strong relationships with the consultees, which was established through frequent site visits. Emphasis was placed on the environment which included physical space, materials, people, and their interactions The In-service Education Model, developed by Palsha and Wesley (1998) also placed emphasis on collaboration between the consultant and consultee. The consultant was encouraged to specifically tailor the technical assistance plan based on the particular needs of both the consultee and classroom. In addition to this, it focused on equipping the consultees with the skills needed to identify and address ways of improving classroom quality during consultation and after it has ended.

A Review of the Literature on Teacher Change and Quality Improvement

Measures of contextual quality, in contrast to child outcome measures, can be used to demonstrate the efforts that teachers make to shape the environmental conditions and teaching practices that they control: the quality of their own classrooms as effective contexts for learning. The quality of the early childhood education has been associated with language development, cognitive growth, and social competence, however, quality has often times been reported as being poor to mediocre in most centers within the United States (Paisha & Wesley 1998). Staff education and development have been reported as being the two key discriminators between mediocre and good care (Cost, Quality, and Child Outcomes Study Team, 1995).

Although consultation models have been mentioned as being viable options for providing teachers with knowledge, skills, and support, which was mentioned above as the main

discriminator between mediocre and good care; there have been very few studies that have examined the effects of consultation on classroom quality. The first reported study to examine the effects of consultation was Wesley's 1994 article on the Innovative Consultation Model. This study yielded significant findings in six of the seven categories within the Infant/Toddler Environmental Rating Scale or ITERS (Harms, Cryer, & Clifford 1990); however, there were limiting factors involved. This model utilized consultants who were all highly educated and were from a university setting. At the conclusion of the study, it was proposed that community-based consultants from diverse backgrounds could perhaps be more effective in providing consultation.

Palsha & Wesley's (1998) follow-up study on the In-service Education Model incorporated consultants who were recruited through community agencies and who demonstrated diverse backgrounds. This study reported high satisfaction levels on a 5-point Likert Scale (5 indicating extreme satisfaction) with mean ratings of 4.4 to 4.6 for all domains. There were open-ended comments in which the consultees stated their learned ability to be able to apply what was learned in consultation to other rooms within the center (Palsha & Wesley 1998). Although these reported satisfaction levels were important, the more important evaluation method was examining the overall effectiveness that the consultation model had on improving the global quality within the classroom. The initial assessment scores were poor to mediocre, which was consistent with national data collected from the 1995 Cost, Quality and Child Outcomes Study. At the conclusion of the observation stage, none of the scores fell below minimum standards. Statistically significance differences were reported between the treatment and control groups for both the ITERS and ECERS after consultation was implemented,

supporting the notion that the model can help improve global quality in the early childhood setting.

The purpose of this study was to examine gains in classroom quality and reduction in teacher stress in treatment and control classrooms located in two separate Head Start agencies in the southeastern United States. The Treatment teachers received training and technical assistance to enhance the implementation of *CC.* The control teachers continued to implement the existing curriculum and did not receive any additional training or technical assistance beyond what their Head Start programs provided. Classroom quality was measured by nationally contracted independent observers using a variety of classroom quality measures. It was hypothesized that the treatment training and technical assistance would lead to stronger implementation of *CC,* which would result in higher classroom quality. It was also hypothesized that the *CC* training and technical assistance would be viewed as a helpful resource to the teachers and would therefore lead to fewer teachers at risk for occupational stress in the treatment group as compared to the control group.

Methods

Participants

This study was conducted within two Head Start agencies located in two separate southeastern states, North Carolina and Georgia. No sampling strategy was used to select centers from within each of the two Head Start agencies. Specific centers were purposefully chosen by the Head Start administrators and the research team based on a history of consistency among their staff and management. Whole centers were not randomly assigned. Rather, teachers within each center were blocked based on education and teacher certification status, and then were randomly assigned to treatment and control conditions within blocks. There were five treatment

and five control classrooms within each of the two Head Start agencies at the beginning of the study.

This study was conducted over two academic years and some teacher turnover took place. In year I of the study, there was teacher turnover among the lead teachers in the North Carolina site only. Two of the control teachers left the program and were replaced. In year II, there were changes in both sites. There were three changes that occurred in North Carolina. Two of the lead teachers were replaced, while one of the assistants was replaced. This turnover was concentrated in the two classrooms that were funded by More at Four, a state funded program that pays for degreed and certified teachers, but at salary levels lower than those paid by many area public schools. At the beginning of the study, one More at Four teacher was randomly assigned to the treatment condition and one to the control. Not surprisingly, these degreed and certified teachers were very difficult to retain in Head Start as they moved to the public school settings as quickly as they were offered better paying positions. Therefore, the researchers decided to drop More at Four classrooms from the study. Georgia did not have any changes within the lead teachers; however, they did have eight changes within the assistants. At the end of year two, there were four Treatment and four Control teachers in North Carolina and five Treatment and five Control teachers Georgia.

Tables 1 and 2 demonstrate the means, standard deviations, and percentages for teacher demographic variables at the end of year two for the full sample and within the two sites. There were no statistically significant differences between the treatment and control groups in total or within the two sites. Every classroom had one teacher and one aide and there were no statistically significant differences between the groups in total or within the sites regarding class size (*mean*=15.592) and teacher child ratio (*mean*=7.796). These results suggest that despite the

turnover issues in the North Carolina site, initial random assignment of the teachers along with the non-random replacement of a control teacher and the removal of the More at Four classrooms were successful strategies in achieving approximate equivalence between the groups of teachers.

Procedure

The treatment condition involved of implementation of The Comprehensive Creative Curriculum *4rd edition* (Dodge & Colker 2002), the child assessment measure for the curriculum was collected by the Treatment teachers in the fall, winter, and spring of the school year. The *CC* Implementation Checklist, the formative classroom quality measure was used by the technical assistant advisors, treatment teachers, and objective data collectors.

The control classrooms continued to implement the existing curriculum, which in both sites was a locally developed eclectic model. The control teachers continued to use the child assessment measures that were already in place in their programs and did not have an implementation checklist available to them to self-monitor the fidelity of their implementation of the local curriculum. Both programs had some limited exposure to training on a previous edition of *CC* (The Creative Curriculum *3rd edition*, 2001) several years prior to the beginning of the study.

The model of training used in this study was defined as clear and explicit delivery and discussion of *CC* content, effective participant involvement in learning opportunities, and trainer observations of the participants' implementation of the curriculum which provided the opportunity for individualized feedback and follow-up training based on observed needs. High quality technical assistance was defined as a systematic process of formative evaluation that involved classroom observation using the *CC* Implementation Checklist, feedback on the

teacher's implementation of the curriculum, and individualized assistance to enhance implementation.

The Technical Assistance Advisor's role was to support the implementation of *The Creative Curriculum* and to enhance classroom quality by fostering the teachers' self-reliance and encouraging the development of autonomy. This role includes the promotion of each teacher's self-directed learning by supporting the teachers' planning, implementation and evaluation of their own learning. In addition, they facilitated teachers' learning by locating and providing a variety of resources. The Technical Assistance Advisor supported the teachers' inquiry and self-reflection, just as the teacher supported the children's inquiry and self-reflection.

The technical assistance was focused on building long term, teacher self-reliance in the implementation of *CC*. An individualized classroom development plan guided the teachers' efforts to enhance teaching practices. In both sites, some of the technical assistance was provided by the same *CC* trainer. Each site also had their own *CC* technical assistance provider who worked both independently and jointly with the *CC* trainer. The content and amount of both *CC* training and technical assistance for the Treatment group was documented throughout the two years of the study. Each treatment teacher received four days of pre-service *CC* training and four days of in-service *CC* training during year one of the study. During year two, each treatment teacher received one day of pre-service *CC* training and four days of in-service *CC* training. In addition, the training and technical assistance that the Head Start programs provided to all of their teachers, including the study control teachers was documented.

Measures

A variety of measures were used to examine classroom quality, both formative and summative. The formative measure, the *CC* Implementation Checklist, developed by the authors

of *CC*, was used as a self-assessment tool by the treatment teachers and as a formative assessment tool by the technical assistance providers. It was also used as a summative evaluation tool by trained, blinded external observers at the beginning of the study and during the fall and spring of the second year to document implementation fidelity. These blinded external observers were sent to both treatment and control classrooms. The measure includes five sections: Physical Environment (102 items); Structure (25 items); Teacher-Child Interaction (44 items); Assessment (11 items); and Family Involvement (5 items). The content of the items is very comprehensive and focuses on both curriculum-specific aspects of practice and or a wide variety of practices that would be considered generally good early childhood practice.

The most widely used measure of environmental quality, the Early Childhood Environment Rating Scale (ECERS), (Harms & Clifford, 1980), has become a quality standard itself in the field of early childhood education. The revised version of this tool, the ECERS-R broadened the measurement of quality to include cultural diversity, family concerns, and individual children's needs. The authors field-tested the revised instrument in 45 classrooms but do not report validity information. They state that the predictive validity of the original ECERS (Peisner-Feinberg & Burchinal, 1997; Whitebook, Howes, & Phillips, 1990) and its construct validity (Rossbach, Clifford, & Harms, 1991; Whitebook et al., 1990) should hold true for the revised measure also (Harms et al., 1998). In her review of the revised measure, Paget (2001) indicates that more empirical evidence for the validity of the revised measure is needed. The authors report interrater reliability at 86% of agreement across all indicators, and item level agreement within a score of 1 point on each 7-point scale was 71%. Weighted Kappa interrater reliability scores for each item range from .28 to .90. The authors recognize the need for further research to extend the psychometric properties of the ECERS-R (Harms et al., 1998).

The CIS was designed to evaluate a caregiver or teacher interactions within a classroom. Four scales were derived through a principal components analysis: positive interaction, punitiveness/harshness, permissiveness, and detachment. Items are rated on a 4-point scale. Observers achieved a criterion of 80% agreement in the original validation study.

The TBRS addresses the teacher's instructional practices in six scales: Book Reading Behaviors (9 items); Print and Letter Knowledge (6 items); Oral Language Use (7 items); Math Concepts (2 items); Written Expression (3 items); and Phonological Awareness (2 items). This measure was under development at the time of the PCER studies and therefore limited technical information was available.

Both the ECERS-R and the CIS were collected as a pre and post test during the four year old year. Data was collected by blinded external observers from a national data collection contractor. In addition to this, the TBRS was collected as post test only during the four year old year by the same outside observers.

The Classroom Appraisal of Resources and Demands (CARD) (Lambert, Abbott-Shim, & McCarthy, 2001) has two scales, classroom Demands and classroom Resources. The teachers in this study were part of a larger study of teacher stress in Head Start and other early childhood settings (Lambert, R., O'Donnell, M., Kusherman, J., & McCarthy, C., 2006). As the sample size of teachers in the present study is small, the results of the larger study are reported herein to illustrate the measurement properties of the information yielded by the CARD. The Demands scale gives teachers the opportunity to rate the extent to which various features of the classroom context are demanding and yielded a Cronbach's alpha reliability of .941. The Classroom Resources scale gives teachers the opportunity to rate how helpful school-provided resources are in assisting with the demands of the classroom and showed a similarly high value, .950. The two

scales were not correlated (r=-.080). Subtracting standardized versions of the scales scores, Demands minus Resources, created a difference score with a reliability of .950. The scoring protocol uses a 95% confidence interval around a difference score of zero to establish cut scores for classifying teachers into one of three groups: Resources greater than Demands, Resources equal to Demands, and Demands greater than Resources. This last group is considered at risk for a stressful experience in the classroom. This three-group distinction has been shown to be useful in testing the transactional model of stress and coping among teachers (McCarthy, C. & Lambert, R., 2006).

Child outcomes were measured by an outside evaluation firm contracted by the U.S. Department of Education. The battery of child measures employed is outlined in the final report for the PCER project (Preschool Curriculum Evaluation Research Consortium, 2008). No additional site-specific child outcome measures were added beyond those used mandated for use by all sites.

Results

Classroom Quality Outcomes

Both treatment and control conditions were observed for the purpose of determining if the treatment classrooms achieved a higher level of implementation of the curriculum when compared to the control group and to determine if any classrooms achieved full implementation. Since there were both treatment and control classrooms within the same centers, it was possible that control teachers would learn certain aspects of the curriculum from the treatment teachers or through ongoing training provided by their program that focused on generally accepted early childhood practices. In addition, we expected all classrooms to have some understanding of the

components of the curriculum that were more focused on these quality teaching practices in early childhood education.

The publishers of *CC* set a minimum acceptance score of 80%, for the entire measure and within each scale, when implementing the curriculum. During the fall of the first year of the study, after pre-service training and some technical assistance, the treatment classrooms achieved on average 45.4% level of implementation, while the control classrooms achieved 33.5%, which were both below minimum standards. By the fall of the second year, after additional pre-service and in-service training and technical assistance, the treatment teachers achieved 70.9% level of implementation while the control teachers scored on average 60.8%. By the end of the second year, the treatment classrooms improved to 85.9% level of implementation and the control classrooms fell to 57.5%. Therefore the treatment classrooms achieved the publisher's minimum set standards by the end of the project. A similar pattern was observed for the teacher-child interactions section of the measure. This section contains many of the curriculum-specific aspects of practice. The treatment teachers scored 55.4% implementation during the fall of year one, rose to an average of 67.9% by the fall of year two, and ended the study at 86.9%. The control teachers scored an average of 23.1% during the fall of year one, 53.8% during the fall of year two, and ended the study at 55.1% implementation. When the scale scores were analyzed at the classroom level for achievement of the 80% correct criterion, seven of the nine treatment teachers achieved this level on the Physical Environment and Structure scales, eight of nine on the Teacher-Child Interactions scale, five of nine on the Assessment scale, and nine of nine on the Family Involvement scale. No control teachers achieved the 80% level of implementation on any of the scales except Family Involvement, where eight of the nine scored at or above 80%.

This scale examines practices that are federally monitored by the Head Start program and would therefore appear in most Head Start classrooms.

There were gains across all classroom quality measures for the treatment group; although these gains were confined to the second year of the study and the results reported in this paper are limited to the second year of the study. It is important to note that while the teachers remained constant across the second year of the study, turnover makes it impossible to examine effects across the two years of the treatment. In addition, the treatment training and technical assistance focused on the more elementary aspects of curriculum implementation during the first year of the study and high levels of implementation fidelity were not achieved until the second year of the study. All the results reported in this summary are based on site-specific analyses. However, they were sustained in cross-site analyses that considered results from all the PCER sites and therefore had higher statistical power. The reader is referred to the PCER final report for a thorough reporting of these analyses (Preschool Curriculum Evaluation Research Consortium, 2008).

For the Overall score from the ECERS-R, there were not main effects for group or change over time, but there was a statistically significant group by time interaction ($F_{(1,16)}$=5.04, p=.039). The means, standard deviations, gains, and effect sizes are represented across both sites in Table 4. The treatment group showed an increase in overall (d= .809), whereas the control group showed a decrease (d=-.068). When examined at the site level, this pattern was also present in the North Carolina site where the treatment group showed a very slight increase in (d=.144) and the control group decreased (d=-.762). This pattern was present in the Georgia site as well, where the treatment group made larger gains (d=1.401), whereas the control group

showed somewhat smaller gains (*d*=.951). The means, standard deviations, gains, and effect sizes of the overall score at the site level are reported in Tables 5 and 6.

There is limited evidence that the subscales of the ECERS-R yield information that functions as independent dimensions. However, we have chosen to report the ECERS-R scores according to the author's scoring guidelines for sub-scales.

For the Space/Furnishing score from the ECERS-R, there were not main effects for group or change over time, or the group by time interaction. The treatment group made gains in space/furnishing (*d*= .663), whereas the control group decreased in space/furnishing (*d*=-.031). The means, standard deviations, gains, and effect sizes for all of the subscales across both sites are represented in Table 4. When examined at the site level, this pattern was not present in the North Carolina site where there were gains in both the treatment (*d*=.477) and control group (*d*=.441) in the space/furnishings subscale. The means, standard deviations, gains, and effect sizes for all of the subscales in North Carolina are represented in Table 5. This pattern was present in the Georgia site, where the treatment group showed an increase in space/furnishings (*d*=.933) whereas the control group decreased on that subscale (*d*=-.327). The means, standard deviations, gains, and effect sizes for all of the subscales in Georgia are reported in Table 6.

For the Language score from the ECERS-R, there were not effects for change over time or the group by time interaction, but there was a statistically significant main effect for group ($F_{(1,16)}$=4.59, *p*=.048). The treatment group showed an increase in language (*d*= .486), whereas the control group decreased in language (*d*=-.253). When examined at the site level, the North Carolina site showed decreases in language in both the treatment (*d*=-.159) and control group (*d*=-1.493). The Georgia site, however, showed a larger gain in the treatment group (*d*=.810), and a smaller gain in the control group (*d*=.318).

For the Interaction score from the ECERS-R, there were not main effects for group or change over time, or the group by time interaction. The treatment group showed an increase in interaction (*d*= .566), whereas the control group decreased in interaction (*d*=-.058). When examined at the site level, the North Carolina site showed decreases in interaction in both the treatment (*d*=-.259) and control group (*d*=-.390). The Georgia site, showed larger gains in the treatment group (*d*=1.008), whereas the control group showed smaller gains (*d*=.321).

For the Personal Care score from the ECERS-R, there were not main effects for group or change over time, but there was a statistically significant group by time interaction ($F_{(1,15)}$=4.78, *p*=.045). The treatment group showed an increase in personal care (*d*= 1.667), whereas the control group stayed fairly constant (*d*=.088). When examined at the site level, this pattern was not present in the North Carolina site where the treatment group remained constant (*d*=.000), and the control group showed a decline in personal care (*d*=-1.104). The Georgia site showed an increase within the treatment group (d=1.175) and the control group (4.591) in the personal care subscale.

For the Activities score from the ECERS-R, there were not main effects for group, but there was a statistically significant main effect for change over time ($F_{(1,16)}$=10.23, *p*=.006) and group by time interaction ($F_{(1,16)}$=6.97, *p*=.018). The treatment group made greater increases in activities (*d*= 1.667) than the control group which stayed about the same (*d*=.088). When examined at the site level, this pattern was also present in the North Carolina site where the treatment group increased in activities (*d*=.831) and the control group decreased (*d*=-.436). This pattern was present in the Georgia site as well, where the treatment group showed a large increase in activities (*d*=3.488) and the control group displayed a smaller increase in activities (*d*=1.194).

For the Program Structure score from the ECERS-R, there were not main effects for group or change over time, but there was a statistically significant group by time interaction ($F_{(1,16)}$=8.60, p=.010). The treatment group showed a gain in program structure (d= .362), whereas the control group showed a decrease in program structure (d=-.399). When examined at the site level, the North Carolina site stayed at an almost constant level in the treatment group (d=.078), whereas the control group decreased (d=-.990). This pattern was present in the Georgia site, where the treatment group showed an increase in program structure (d=.602) and the control group showed a decrease in program structure (d=-.378).

For the Overall score from the CIS, there were not main effects for time or group by time, but there was a statistically significant main effect for group interaction ($F_{(1,16)}$=4.99, p=.041). The treatment group made greater overall gains (d = .792) than the control group which decreased in overall (d=-.142). The means, standard deviations, gains, and effect sizes for the overall scores across both sites are represented in Table 7. When examined at the site level, North Carolina's treatment group stayed about the same in overall scores (d=.077) and the control group decreased (d=-.919). In the Georgia site, the treatment group showed a large increase overall (d=2.244), whereas the control group also had small gains (d=.271). The means, standard deviations, gains, and effect sizes of the overall score at the site level are reported in Tables 8 and 9.

For the Permissiveness score from the CIS, there were not main effects for group or change over time, but there was a statistically significant group by time interaction ($F_{(1,15)}$=4.78, p=.045). The treatment group made greater declines in permissiveness (Cohen's d effect size = -.942) than the control group which increased in permissiveness (d=.568). The means, standard deviations, gains, and effect sizes for all of the subscales across both sites are represented in

Table 7. When examined at the site level, this pattern was not present in the North Carolina site where the treatment group increased in permissiveness (*d*=.332) and the control group remained the constant (*d*=.000). The means, standard deviations, gains, and effect sizes for all of the subscales in North Carolina are represented in Table 8. This pattern was present in the Georgia site, where the treatment group showed a large decline in permissiveness (*d*=-1.981) and the control group increased in permissiveness (*d*=1.132). The means, standard deviations, gains, and effect sizes for all of the subscales in Georgia are reported in Table 9.

For the Detachment score from the CIS, there was a main effect for group ($F_{(1,15)}$=7.02, *p*=.018), but there was not a statistically significant main effect for time or group by time interaction. The treatment group was lower overall and made greater declines in detachment (*d* = -.568) than the control group which increased in detachment (*d*=.373). When examined at the site level, this pattern was not present in the North Carolina site where both the treatment group (*d*=.287) and the control group (*d*=2.163) increased in detachment, however, control group's increase was more dramatic. This pattern was present in the Georgia site, where the treatment group showed at large decline in detachment (*d*=-.856) and the control group remained about the same in detachment (*d*=.097).

For the Positive Interaction score from the CIS, there was not a main effect for change over time, but there was a statistically significant main effect for group ($F_{(1,15)}$=6.45, *p*=.023) and for group by time interaction ($F_{(1,15)}$=5.01, *p*=.041). The treatment group made greater gains in positive interaction (*d* = 1.034) than the control group which decreased in positive interaction (*d*=-.163). When examined at the site level, this pattern was also present in the North Carolina site where the treatment group slightly increased in positive interaction (*d*=.163) and the control group decreased (*d*=-.906). In the Georgia site, the treatment group showed a large increase in

positive interaction (*d*=1.250), whereas the control group also had smaller gains in positive interaction (*d*=.455).

For the Harshness score from the CIS, there were not main effects for group, change over time, or group by time interaction. The treatment group made greater declines in harshness (*d* = -.291) than the control group which had a slight increase in harshness (*d*=.056). When examined at the site level, this pattern was not present in the North Carolina site where there was an increase in the treatment group (*d*=.396), whereas the control group stayed about the same (*d*=.056). In the Georgia site, the treatment group showed a decline in harshness (*d*=-.429), however, the control group showed a slightly larger decrease in harshness (*d*=-.550).

The TBRS was given as a posttest only at the end of year two. For the Features scale, there was not a statistically significant difference between the groups ($t_{(16)}$=.592, *p*=.565). However, the treatment group scored somewhat higher than the control group (*d*=.328). The means, standard deviations, differences, and effect sizes for all of the subscale scores across both sites are reported in Table 10. When examined at the site level, this pattern was also present in the North Carolina site where the treatment group showed higher scores on the features scale than the control group (*d*=.707), whereas in the Georgia site the control group scored higher (*d*=-.447). The means, standard deviations, differences, and effect sizes for all of the site specific subscale scores are represented in Tables 11 and 12.

For the Book Reading Behaviors scale, there was not a statistically significant difference between the groups ($t_{(15)}$=.000, *p*=.950). The treatment group and control group were almost equal on this subscale (*d*=.044). When examined at the site level, the North Carolina site showed a difference in favor of the treatment group (*d*=.457), whereas the Georgia site showed a difference in favor of the control group (*d*=-.447).

For the Print and Letter Knowledge scale, there was a statistically significant difference between the groups ($t_{(15)}$=2.16, p=.048). The treatment group scored significantly higher than the control group on this scale (d=1.777). When examined at the site level, this pattern was not present in the North Carolina site where the control group did better (d=-1.037). However, in the Georgia site the treatment group scored much higher than the controls (d=5.445).

For the Oral Language Use scale there was a statistically significant difference between the groups ($t_{(15)}$=3.718, p=.002). The treatment group scored significantly higher than the control group on this scale (d=2.204). When examined at the site level, this pattern was also present in the North Carolina site where the treatment group scored higher than the controls (d=.273). In the Georgia site, the treatment group also scored much higher than the controls (d=5.258).

For the Phonological Awareness scale there was not a statistically significant difference between the groups ($t_{(15)}$=.100, p=.932). The treatment group scored somewhat lower than the control group on this scale (d=-.058). When examined at the site level, this pattern was also present in the North Carolina site where the treatment group scored much worse than the controls (d=-.764). However, in the Georgia site the treatment group scored much higher than the controls (d=.702).

For the Math Concepts scale there was a statistically significant difference between the groups ($t_{(15)}$=2.280, p=.038). The treatment group scored significantly higher than the control group on this scale (d=2.428). When examined at the site level, this pattern was not present in the North Carolina site where the treatment group scored much worse than the controls (d=-1.803). However, in the Georgia site the treatment group scored much higher than the controls (d=14.747).

For the Written Expression scale there was a statistically significant difference between the groups ($t_{(15)}$=3.719, p=.002). The treatment group scored significantly higher than the control group on this scale (d=2.277). When examined at the site level, this pattern was also present in both North Carolina (d=2.464) and Georgia (d=14.747) sites where both treatment groups scored much higher than the control groups.

The teacher stress measure, the CARD, was administered to teachers and aides in all study classrooms during the spring of both years of the study. We obtained a 75% response rate in year one and an 88.9% response rate in year two. There were statistically significant differences between the groups with respect to the number of teachers at risk for stress during both years of the study. During the spring of year one, only 35.7% of treatment teachers and aides reported Demands greater than Resources, as compared with 56.3% of controls. During year two, 35.3% of treatment teachers and aides reported Demands greater than Resources, as compared to 73.3% of controls.

The Amount and Type of Technical Assistance Delivered to the Treatment Teachers

The two Technical Assistance Advisors, one in North Carolina and one in Georgia, documented each visit to a teachers' classroom by completing a Technical Assistance Report. This included the date, length of visit, teacher's name, general notes, observations and follow-up regarding the *CC* Implementation Checklist, and plans for the next visit. The number of visits, length of time for each visit, and the content of each Technical Assistance Report was summarized. The content summary noted the observation of classroom practices and/or discussion of implementation strategies. The coded categories included: learning environment; developmental assessment and use of CC.net; studies or classroom learning projects;

observations for and/or discussions with teachers about the *CC* Implementation Checklist; lesson plans and daily schedules; and modeling teaching practices.

The NC Technical Assistance Advisor had a total of 40 visits from September 2003 through May 2004 with the teachers and teaching assistants in four classrooms. The average amount of time for each visit was 5 hours. The GA Technical Assistance Advisor had a total of 115 visits during the same period of time with the teachers and teacher assistants in five classrooms. The average amount of time for each visit was almost one hour (55 minutes). Therefore, the NC teachers on average received 10 visits each for a lengthy amount of time (5 hours). Whereas the GA teachers received more frequent visits, an average of 23 visits per teacher, for much shorter time periods (1 hour).

The summary of the content for these classroom visits based on the Technical Assistance Reports is provided in Table 13. The summary of the learning environment content reflects quite different roles for the two Technical Assistance Advisors. The NC Advisor focused on the learning environment for an average of 5 times with each teacher whereas the GA Advisor included the learning environment for an average of 14 times with each teacher. A further analysis showed that the number of visits for which both Technical Assistance Advisors had a hands-on role helping to organize, label and/or repair materials was about equivalent. However, the GA Advisor had an average of six visits per teacher when she delivered resource materials to facilitate the teacher's learning and the children's studies or projects within the classroom and the NC Advisor only provided instructional resource materials on one visit out of the total 40 visits.

Both of the Technical Assistance Advisors had fairly equivalent number of visits for developmental assessment and use of the CC.net assessment data management system. The NC

Technical Assistance Advisor discussed studies or projects with only one teacher during one visit, whereas the GA Advisor discussed studies with the teachers frequently, during an average of 9 visits per teacher. The NC Technical Assistance Advisor included observations and/or discussion of the *CC* Implementation Checklist during an average five visits per teacher and the GA Advisor only focused on the checklist during an average 3.8 visits per teacher. The NC Advisor included observations and/or discussion of lessons plans and daily schedules during an average of two visits per teacher and the GA Advisor an average 3.8 visits per teacher. The GA Technical Assistance Advisor included modeling teaching practices only twice whereas the NC Advisor modeled more frequently for an average of 1.75 per teacher.

The Technical Assistance Advisors were provided the same written material and training in regard to their advisory roles. However, the qualitative data from the technical assistance reports indicated that the way in which each Advisor carried out their advisory role differed greatly. The NC Technical Assistance Advisor was didactic in her interactions with the Head Start teachers. For example, she told the teachers what learning materials were needed in the classroom and how to display the materials. A typical comment was "add unit blocks with different shapes and label shelves." The advice was usually given in reference to how it impacts their score on the *CC* Implementation Checklist such as, "the materials examined and materials added where needed so that the item could be given credit for implementation" on the checklist.

The GA Technical Assistance Advisor was facilitative and collaborative in her interactions with the Head Start teachers. For example, when a teacher indicated that the purpose of a library visit was "to enhance learning in the area of literacy," the Advisor asked "What kind of preparation would you do with the class beforehand?" The Advisor often supported the teachers' plans for children's learning projects or studies in the classroom. For

example, the teacher had read the book, *Mud Pies*, and created display charts about pies so the Advisor asked if the children "have made any pies with real mud?" Since the children had not had this opportunity, the Advisor said that she would bring "some miniature pie pans and used coffee grounds for a sensory learning experience." The GA Advisor brought resource materials on an average of six visits per teacher. When the GA Advisor provided feedback to the teacher on the *CC* Implementation Checklist, she typically asked what this meant to the teacher and how she would like to improve her classroom.

Child Outcomes

Across the entire battery of child measures, *CC* was not associated with any advantages for the treatment children. There were no statistically significant differences between the treatment and control children and these findings were sustained in the cross-site analyses as reported in the PCER final report (Preschool Curriculum Evaluation Research Consortium, 2008).

Discussion

The results of this study indicate that moderately high levels of implementation of *CC* were attained by the treatment teachers, but only after two years of training and technical assistance. Across the 19 measures of classroom quality, the treatment group yielded changes in the desired direction (or posttest scores in the case of the TBRS) which were at least an effect size of .3 greater than the control group for 16 of the measures. For the remaining three measures, the two groups were essentially equivalent. For the Georgia site, the treatment group yielded results in the desired direction which were at least an effect size of .3 greater than the control group for 15 of the measures as compared to 11 for the North Carolina site. For one of the measures in Georgia and three in North Carolina the groups were essentially equivalent. The

control group outperformed the treatment group on three measures in the Georgia site as compared to five measures in North Carolina.

It appears that more frequent visits of shorter duration, as demonstrated in the Georgia site, provided more effective technical assistance to Head Start preschool teachers. These results also suggest that the style and demeanor of the technical assistance provider can make a difference in terms of gains in classroom quality. The Georgia technical assistance provider focused more on the instructional philosophy of *CC*, gave teachers specific assistance related to *CC* teacher-child interactions and instructional strategies, used a more collaborative and facilitative style, encouraged teacher independence, and built a greater sense of trust and rapport with the teachers. In contrast, the North Carolina technical assistance provider used a much more didactic style, did not have as much rapport with the teachers, focused more on a compliance model, and emphasized the scores on the *CC* implementation checklist.

The pattern of more positive findings for the Georgia site tend to confirm earlier studies that demonstrated how a facilitative and collaborative technical assistance approach can lead to enhanced classroom quality (Palsha & Wesley, 1998; Epstein, 1993). The Georgia technical assistance provider was also more successful at implementing our intended model of consultation and remained closer to the *CC* philosophy of early childhood education. Overall, the treatment teachers and aides in both sites were less likely to be at risk for classroom stress. The *CC* training and technical assistance that the Georgia teachers received was viewed as a helpful set of resources that were useful as buffers against the demands of the classroom.

It is important to note that this study had several limitations. It is possible that the teacher turnover influenced the results in the North Carolina site as two North Carolina classrooms were dropped from the study due to repeated turnover. The study focused on a small sample of

teachers within two specific Head Start programs in the southeastern United States and the findings may not generalize to other programs or regions of the country. The teachers that remained to the end of the study were for the most part not degreed or certified and the classrooms were generally of low or moderate quality at the beginning of the study. The researchers, through a grant from the United States Department of Education, provided the relatively large amount of resources and time required to achieve full implementation of a comprehensive and integrated curricular model such as *CC*. This level of commitment may extend beyond the financial capacity of many early childhood programs.

Future research may need to focus on replicating these findings by building in planned variation in degree and certification status among teachers, perhaps using larger sample sizes and different locations around the country. It may also be important to randomize at the center level to avoid the possibility of contamination of the control group. This study was unable to pursue such a strategy because of the relatively small number of centers available and the variation in family and child demographics between them.

It is important to do this kind of rigorous research on a comprehensive integrated curricular model. *CC* is a fully integrated and comprehensive model. A study of this type illustrates that it is not easy to reach full implementation of a curriculum of this kind and suggests that two years of high quality training and high quality technical assistance may be needed to achieve full implementation of the curriculum. If a program can manage to allocate the resources needed to provide *CC* training and technical assistance, this study suggests that these processes will result in enhanced classroom quality.

Overall Conclusions

The overall results of this study raise important questions about the impact of *CC* on child outcomes. What does a finding of gains in classroom quality without simultaneous gains in child outcomes mean for early childhood policy makers? It is important for future research to examine whether comprehensive and integrative curricular approaches such as *CC* require a complex chain of events to take place in order for child outcomes to be realized. Our study suggests that teachers first need to buy into a need to totally reorient what they are doing. They then need high quality training followed by extensive technical assistance that is delivered in a very individualized and sensitive manner. It may be that after a high degree of implementation is achieved, only then can child outcomes begin to be impacted. Our research demonstrated that it may take as much as two academic years to obtain the goal of complete implementation of a comprehensive and integrative model in classrooms of low to moderate quality. If such a high level of implementation is required to realize child outcome gains, then studies of longer time duration may be needed to detect such gains.

Our results demonstrate that add-on interventions with domain-specific foci are fundamentally different in nature than comprehensive and integrative curricular approaches, and may therefore require different evaluation designs to detect their ultimate impact on child outcomes. We have learned that the process of affecting major changes in teacher behavior across all areas of classroom functioning, such as those required to fully implement *CC*, can be realized through extensive and individualized training and technical assistance. This conclusion appears to be particularly important to bear in mind when teachers have many needs and start from a position of relatively low educational attainment and low to moderate quality.

Our results also suggest that the substance as well as the style of technical assistance matters to teachers and teacher change. A sensitive, respectful, and individualized approach with

frequent contacts of short duration and meaningful assigned tasks between contacts seems to be more associated with meaningful teacher changes than a more didactic and authoritative approach. It also appears that helping teachers remain engaged in meaningful and intentional interactions with children throughout the daily schedule is important and yet challenging to realize. The next phase of research on curriculum effectiveness may need to examine the exact nature of the supporting resources available to both new and experienced teachers for each model. New teachers appeared to us to crave specific strategies and resources that help them understand exactly what to do in their classrooms. Experienced teachers seem to be much more satisfied by engaging and thought provoking training and technical assistance that helped them think about their roles as teachers in new and meaningful ways.

In our own research, we are interested in continuing to examine the effectiveness of specific coaching and mentoring strategies, particularly those that make use of data for the purpose of planning individual programs of improvement for inexperienced teachers. We are pursuing these goals through a current Early Childhood Educator Professional Development grant from the Institute for Education Sciences. We are also focusing on a system of mentoring and performance evaluation for licensed teachers working in non-public settings (child care, Head Start, etc.) in which the Pre-kindergarten / Kindergarten Teacher Performance Appraisal Instrument (PKKTPAI) which we developed is being used to inform an ongoing program of individualized teacher professional development.

References

Abbott-Shim, M. (2000). *Sure Start Effectiveness Study: Final Report*. Atlanta, GA: Quality Assist, Inc.

Administration on Children, Youth, and Families. (2001). *Head Start performance measures: Third progress report*. Washington, DC: US Department of Health and Human Services.

Arnett, J. (1989). Caregivers in day care centers: Does training matter? *Developmental Psychology, 10*, 541-552.

Bailey, D. B. (1989). Issues and directions in preparing professionals to work with young handicapped children and their families. In J. J. Gallagher, R. M. Clifford, & P. Trohanis (Eds.), *Policy implementation and P.I., 99-457*: *Planning for young children with special needs* (pp. 97-132). Baltimore: Brookes.

Bailey, D.B., McWilliam, P.J., & Winton, P.J. (1992). Building family-centered practices in early intervention: A team-based model for change. *Infants and Young Children, 5*, 73-82.

Bowman, B.T., Donvan, M.S. & Burns, M.S. (Eds). (2001). *Eager to Learn: Educating Our Preschoolers.* Washington, DC: National Academy Press.

Brown, D., Pryzwansky, W., & Schulte, A. (1987). *Psychological consultation: Introduction to theory and practice*. Boston: Allyn & Bacon.

Bush, L. (2001, July). *Ready to read, ready to learn,* White House Summit on Early Childhood Cognitive Development.

Buysse, V., & Wesley, P. W. (2004). A framework for understanding the consultation process: Stage by stage. *Young Exceptional Children, 7*(2), 2-9.

Cost, Quality, and Child Outcomes Study Team. (1995). *Cost, quality, and child*

outcomes in child care centers: Executive summary (2nd ed.). Denver: University of Colorado, Economics Department.

Dickinson, D. K. (2002). Shifting images of developmentally appropriate practice as seen through different lenses. *Educational Researcher, 31*(1), 26-32.

Dickinson, D. K. & Tabors, P. O. (2001). Beginning language with literacy: Young children learning at home and school. Baltimore, MD: Brooks Publishing.

Dodge, D.T., Colker, L.J., & Heroman, C. (2003). The Creative Curriculum® for Preschool Implementation Checklist. Washington, DC: Teaching Strategies, Inc.

Dodge, D.T., Colker, L.J., & Heroman, C. (2002). *The Creative Curriculum for Preschool (4th ed.*). Washington, DC: Teaching Strategies, Inc.

Dodge, DT. & Colker, L.J. (2001). The Creative Curriculum Developmental Continuum Assessment System. Washington, DC: Teaching Strategies, Inc.

Epstein, A.S. (1993). Training for quality: Improving early childhood programs through *systemic inservice training.* Monographs of the High/Scope Educational Foundation, no.9. Ypsilanti, MI: High/Scope.

Epstein, A.S., Schweinhart, L.J., & McAdoo, L. (1996). *Models of Early Childhood Education.* Ipsilanti, MI: HighScope Press.

Federal Register. (2001, December 17). Washigton, DC: U.S. Government Printing Office

Fullan, M. (1993). *Change forces: Probing the depths of educational reform.* London: Palmer Press.

Guskey, T. R. (1986). Staff development and the process of teacher change. *Educational Researcher, 15*, 5-12.

Harms, T., & Clifford, R. (1980). *Early Childhood Environment Rating Scale (ECERS).* New

York: Teachers College Press.

Harms, T., Clifford, R., & Cryer, D. (1998). *Early Childhood Environment Rating Scale-Revised (ECERS-R).* New York: Teachers College Press.

Harms, T., Cryer, D., & Clifford, R. M. (1990). *Infant/Toddler Environment Rating Scale.* New York: Teachers College Press.

House, E.R., Glass, G.V., McLean, L.F., & Walker, D.F. (1978). No simple answer: Critique of the "follow through" evaluation. *Harvard Educational Review,* 28 (2), 128-160.

Joyce, B., & Showers, B. (1983). *Power in staff development through research on training.* Arlington, VA: Association for Supervision and Curriculum Development.

Kagan, S. L., Moore, E., & Bredekamp, S. (Eds.) (1997). *Getting a good start in school.* Washington, D.C.: National Education Goals Panel.

Lambert, R., Abbott-Shim, M., Sibley, A. (2006). Evaluating the quality of early childhood educational settings. In B. Spodek and O. Saracho (Eds.), *Handbook of research on the education of young children*, Second Edition (pp. 457-475). New Jersey: Lawrence Erlbaum Associates, Inc.

Lambert, R., Abbott-Shim, M., & McCarthy, C. (2001). *Classroom Appraisal of Resources and Demands, preschool version.* Atlanta, GA: Head Start Quality Research Center.

Lambert, R., O'Donnell, M., Abbott-Shim, M., & Kusherman, J. (2006 April) *The effect of Creative Curriculum training and technical assistance on Head Start classroom quality.* Accepted to present at the meeting of the American Educational Research Association, San Fransisco, CA.

Lambert, R., O'Donnell, M., Kusherman, J., & McCarthy, C. (2006). Teacher stress and classroom structural characteristics in preschool settings. In R. Lambert and C. McCarthy

(Eds.), *Understanding teacher stress in an age of accountability* (pp. 105-120). Greenwich, CT: Information Age Publishing.

Landry, S. H., Gunnewig, S., Assel, M., Crawford, A., & Swank, P. (2004). *The CIRCLE Teacher Observation Rating Scale*. Unpublished manuscript.

MacIver, M.& Kemper, E. (2002). Guest editors' introduction: Research on Direct Instruction in reading. Journal of Education for Students Placed at Risk 7(2), retrieved January 5, 2006 from http://www.louisville.edu/edu/jespar/vol_7_no_2_editors.htm

McCarthy, C. & Lambert, R. (2006). The demands experienced by teachers and the resources they employ to manage stress. In R. Lambert and C. McCarthy (Eds.), *Understanding teacher stress in an age of accountability* (pp. 215-225). Greenwich, CT: Information Age Publishing.

Nourot, P.M. (2004). Historical perspectives on early childhood education. In J.L. Roopnarine & J.E. Johnson (Eds.) *Approaches to early childhood education* (pp.3-31), Englewood, NJ: Prentice Hall.

Paget, K. D. (2001). *Review of the Early Childhood Environment Rating Scale-Revised.* In J. C. Impara & B. S. Plake (Eds.), *The fourteenth mental measurements yearbook. The Buros Institute of Mental Measurements*, Lincoln: NE: The University of Nebraska Press.

Palsha, S. A., & Wesley, P. W. (1998). Improving quality in early childhood environments through on-site consultation. *Topics in Early Childhood Special Education, 18*(4), 243-253.

Peisner-Feinberg, E., & Burchinal, M. (1997). Relationships between preschool children's child care experiences and concurrent development: The Cost, Quality and Outcome Study.

Merrill-Palmer Quarterly, *43*(3), 451-477.

Phillips, D. (1996). Reframing the quality issue. In S.L Kagan & N. Cohen (Eds.), *Reinventing early care and education: A vision for a quality system.* San Francisco, CA: Jossey-Bass Publishers.

Preschool Curriculum Evaluation Research, 67 Fed.Reg. 61441-61444 (Sept. 30, 2002).

Preschool Curriculum Evaluation Research Consortium (2008). *Effects of Preschool Curriculum Programs on School Readiness* (NCER 2008-2009). Washington, DC: National Center for Education Research, Institute of Education Sciences, U.S. Department of Education. Washington, DC: U.S. Government Printing Office.

Roopnarine, J. & Johnson, J. (1987). Approaches to Early Childhood Education. Columbus, OH: Merrill.

Rossbach, H.G., Clifford, R.M., Harms, T. (1991). *Dimension of learning environments: Cross-national validation of the Early Childhood Environment Rating Scale.* Paper presented at the AERA Annual Conference, Chicago.

Schweinhart, L.J. & Weikart, D.P.(1997). The High/Scope preschool curriculum comparison study through age 23. *Early Childhood Research Quarterly, 12*(2), 117-143.

Schweinhart, L.J., Weikart, D.P., & Larner, M.B. (1986). Consequences of three preschool curriculum models through age fifteen. *Early Childhood Research Quarterly*, *1*(1): 15-46.

Spodek, B. & Brown, P.C., Curriculum Alternatives in Early Childhood Education: A Historical Perspective, in Spodek, B. (Ed.). (1982). *Handbook of Research on the Education of Young Children.* New York: MacMillan Publishing Company.

Spodek, B. & Saracho, O.N. (Eds). (2006). *Handbook of Research on the Education of Young Children.* Mahwah, NJ: Lawrence Erlbaum Assoicates.

St. Pierre, R.G., Anderson, R.B., Proper, E.C., & Stebbins, L.B. (1978). That follow through

evaluation. *Phi Delta Kappan*, 59(10), 729.

Stebbins, L.B, St. Pierre, R.G., Proper, E.C., Anderson, R.B., & Cerva, T.R. (1977). *A planned variation model. Vol. IV-A effects of follow through models.* U.S. Office of Education.

Trohanis, P. (1980). Delivering technical assistance. In R. M. Clifford & P. Trohanis (Eds.), *Technical assistance in educational settings* (pp. 27-35). Columbus, OH: Ohio State University.

Trohanis, P. (2001). Design considerations for state TA systems. The National Early Childhood Technical Assistance Center. Chapel Hill, NC: University of North C Carolina at Chapel Hill, FPG Child Development Center.

U. S. Department of Education. (2001). *No Child Left Behind Act of 2001*. Retrieved October 25, 2000, from http://www.ed.gov/nclb

Warger, C. (Ed.). (1988). *A Resource Guide to Public School Early Childhood Programs.* Alexandria, VA: Association for Supervision and Curriculum Development.

Wesley, P. W. (1994). Innovative practices: Providing on-site consultation to promote quality in integrated child care programs. *Journal of Early Intervention, 18*(4), 391-402.

Whitebook, M., Howes, C., & Phillips, D. (1990). *Who cares? Child care teachers and the quality of care in America. Final Report of the National Child Care Staffing Study.* Oakland, CA: Child Care Employee Project.

Winton, P. (1990). A systemic approach for planning inservice training related to Public Law 99-457. *Infants and Young Children, 3*(1), 51-60.

Winton, P. J., McCollum, J., & Catlett, C. (1997). *Reforming personnel preparation in early intervention: Issues, models, and practical strategies*. Baltimore: Brookes.

Table 1.
Teacher Demographics

		Both Sites		NC		GA	
		Rx	Cn	Rx	Cn	Rx	Cn
Teacher-Child Ratio	Mean	16.182	17.000	14.571	15.800	19.000	18.000
	SD	2.822	1.414	2.070	0.447	1.155	1.095
Teaching Experience	Mean	9.136	8.674	7.786	6.283	4.500	3.597
	SD	6.009	5.571	6.595	4.653	2.380	1.882
How many years have you worked at your center	Mean	4.727	4.553	4.857	5.700	4.500	3.597
	SD	5.101	3.280	6.362	4.410	2.380	1.882
Teacher's Age	Mean	41.909	35.455	42.571	38.200	40.750	33.167
	SD	9.731	9.310	11.193	13.027	7.890	4.875

Table 2.
Educational Background

	Both Sites		NC		GA	
	Rx	Cn	Rx	Cn	Rx	Cn
	%		%		%	
Do you have a CDA Credential	75.00	75.00	57.10	40.00	100.00	100.00
What is the highest degree you have completed						
High School	33.30	30.00	42.90	25.00	20.00	33.30
Technical School	16.70	20.00	14.30	0.00	20.00	33.30
AS	25.00	40.00	14.30	75.00	40.00	16.70
BS	25.00	10.00	28.60	0.00	20.00	16.70
Are you currently working towards a degree	83.30	72.70	87.50	50.00	75.00	85.70
Gender						
Female	92.30	100.00	100.00	100.00	80.00	100.00
Male	7.70	0.00	0.00	0.00	20.00	0.00
Ethnicity						
Asian	7.70	0.00	0.00	0.00	20.00	0.00
African American	84.60	83.30	87.50	80.00	80.00	85.70
European American	7.70	8.30	12.50	20.00	0.00	0.00
Other	0.00	8.30	0.00	0.00	0.00	14.30

Table 3.
Implementation Fidelity

	Fall 02 %	Fall 03 %	Spring 04 %
Total Scores			
Rx	45.4	70.9	85.9
Cn	33.5	60.8	57.5
Teacher Child Interaction Scores			
Rx	55.4	67.9	86.9
Cn	23.1	53.8	55.1

Table 4.
Implementation Fidelity

	Fall 02 %	Fall 03 %	Spring 04 %
Total Scores			
Rx	45.4	70.9	85.9
Cn	33.5	60.8	57.5
Teacher Child Interaction Scores			
Rx	55.4	67.9	86.9
Cn	23.1	53.8	55.1

Table 4.
ECERS: Both Sites

Measure		Treatment Fall	Spring	Gain	ES	Control Fall	Spring	Gain	ES
Space/Furnishing	Mean	3.778	4.194	0.416	0.663	3.583	3.556	-0.027	-0.031
	SD	0.627	0.668			0.868	1.193		
Language Score	Mean	4.389	5.111	0.722	0.486	3.833	3.389	-0.444	-0.253
	SD	1.485	0.920			1.754	1.153		
Interaction Score	Mean	4.111	5.044	0.933	0.566	3.578	3.467	-0.111	-0.058
	SD	1.647	1.009			1.922	1.889		
Personal Care Score	Mean	2.667	3.185	0.518	0.481	2.704	2.815	0.111	0.077
	SD	1.077	0.679			1.436	0.752		
Activities Score	Mean	3.373	4.378	1.005	1.667	3.172	3.268	0.096	0.088
	SD	0.603	0.494			1.091	0.689		
Program Structure	Mean	4.241	4.759	0.518	0.362	4.111	3.389	-0.722	-0.399
	SD	1.429	0.781			1.810	1.463		
Overall Score	Mean	3.642	4.356	0.714	0.809	3.406	3.320	-0.086	-0.068
	SD	0.883	0.497			1.271	0.937		

Table 5.
ECERS: North Carolina

Measure		Treatment Fall	Spring	Gain	ES	Control Fall	Spring	Gain	ES
Space/Furnishings	Mean	4.156	4.313	0.157	0.477	4.375	4.531	0.156	0.441
	SD	0.329	1.008			0.354	0.838		
Language Score	Mean	4.688	4.500	-0.188	-0.159	5.438	4.063	-1.375	-1.493
	SD	1.179	0.540			0.921	0.657		
Interaction Score	Mean	4.500	4.150	-0.350	-0.259	5.300	4.750	-0.550	-0.390
	SD	1.352	0.640			1.409	1.544		
Personal Care Score	Mean	3.125	3.125	0.000	0.000	4.042	2.875	-1.167	-1.104
	SD	1.322	0.985			1.057	0.896		
Activities Score	Mean	3.775	4.275	0.500	0.831	4.200	3.900	-0.300	-0.436
	SD	0.602	0.275			0.688	0.183		
Program Structure	Mean	5.042	5.125	0.083	0.078	5.625	4.604	-1.021	-0.990
	SD	1.064	0.722			1.031	0.809		
Overall Score	Mean	4.079	4.196	0.117	0.144	4.640	4.086	-0.554	-0.762
	SD	0.810	0.542			0.727	0.488		

Table 6.
ECERS: Georgia

Measure		Treatment Fall	Spring	Gain	ES	Control Fall	Spring	Gain	ES
Space/Furnishing	Mean	3.475	4.100	0.625	0.933	2.950	2.775	-0.175	-0.327
	SD	0.670	0.324			0.535	0.778		
Language Score	Mean	4.150	5.600	1.450	0.810	2.550	2.850	0.300	0.318
	SD	1.791	0.894			0.942	1.232		
Interaction Score	Mean	3.800	5.760	1.960	1.008	2.200	2.440	0.240	0.321
	SD	1.944	0.537			0.748	1.545		
Personal Care Score	Mean	2.300	3.233	0.933	1.175	1.633	2.767	1.134	4.591
	SD	0.794	0.435			0.247	0.723		
Activities Score	Mean	3.051	4.460	1.409	3.488	2.349	2.762	0.413	1.194
	SD	0.404	0.275			0.346	0.453		
Program Structure	Mean	3.600	4.467	0.867	0.602	2.900	2.417	-0.483	-0.378
	SD	1.441	0.767			1.278	1.064		
Overall Score	Mean	3.292	4.483	1.191	1.401	2.418	2.708	0.290	0.951
	SD	0.850	0.478			0.305	0.723		

Table 7.
CIS: Both Sites

Measure		Treatment Fall	Treatment Spring	Treatment Gain	Treatment ES	Control Fall	Control Spring	Control Gain	Control ES
Permissiveness Score	Mean	1.917	1.250	-0.491	-0.942	1.741	2.037	0.296	0.568
	SD	0.707	0.295			0.521	0.904		
Detachment Score	Mean	1.563	1.156	-0.407	-0.568	2.000	2.306	0.306	0.373
	SD	0.717	0.265			0.820	0.958		
Positive Interaction Score	Mean	2.538	3.338	0.800	1.034	2.278	2.144	-0.134	-0.163
	SD	0.774	0.609			0.821	0.682		
Harshness Score	Mean	1.708	1.472	-0.236	-0.291	2.136	1.827	-0.309	-0.289
	SD	0.812	0.423			1.071	0.676		
Overall Score	Mean	2.846	3.341	0.495	0.792	2.650	2.564	-0.086	-0.142
	SD	0.625	0.292			0.607	0.582		

Table 8.
CIS: North Carolina

Measure		Treatment Fall	Spring	Gain	ES	Control Fall	Spring	Gain	ES
Permissiveness Score	Mean	1.556	1.556	0.139	0.332	1.417	1.417	0.000	0.000
	SD	0.509	0.192			0.419	0.631		
Detachment Score	Mean	1.167	1.250	0.083	0.287	1.250	1.875	0.625	2.163
	SD	0.289	0.433			0.289	0.433		
Positive Interaction Score	Mean	2.633	2.667	0.034	0.163	2.950	2.375	-0.575	-0.906
	SD	0.208	0.058			0.635	0.350		
Harshness Score	Mean	1.519	1.630	0.111	0.396	1.389	1.417	0.028	0.056
	SD	0.280	0.740			0.501	0.292		
Overall Score	Mean	3.038	3.051	0.013	0.077	3.183	2.865	-0.318	-0.919
	SD	0.168	0.190			0.346	0.252		

Table 9.
CIS: Georgia

Measure		Treatment Fall	Spring	Gain	ES	Control Fall	Spring	Gain	ES
Permissiveness Score	Mean	2.133	1.067	-0.933	-1.981	2.000	2.533	0.533	1.132
	SD	0.767	0.149			0.471	0.803		
Detachment Score	Mean	1.800	1.100	-0.700	-0.856	2.600	2.650	0.050	0.097
	SD	0.818	0.137			0.518	1.167		
Positive Interaction Score	Mean	2.480	3.740	1.260	1.250	1.740	1.960	0.220	0.455
	SD	1.008	0.329			0.483	0.862		
Harshness Score	Mean	1.822	1.378	-0.444	-0.429	2.733	2.156	-0.577	-0.550
	SD	1.035	0.099			1.050	0.740		
Overall Score	Mean	2.223	3.051	0.828	2.244	2.223	2.323	0.100	0.271
	SD	0.369	7.529			0.369	0.683		

Table 10.
Teacher Behavior Rating Scale Scores: Both Sites

Measure		Treatment	Control	Difference	ES
Features	Mean	0.889	0.556	0.333	0.328
	SD	1.364	1.014		
Book Reading Behaviors	Mean	1.764	1.679	0.085	0.044
	SD	2.346	1.948		
Print & Letter Knowledge	Mean	3.185	2.000	1.185	1.777
	SD	1.418	0.667		
Oral Language Use	Mean	4.095	2.036	2.059	2.204
	SD	1.294	0.934		
Phonological Awareness	Mean	1.556	1.625	-0.069	-0.058
	SD	1.944	1.188		
Math Concepts	Mean	2.952	1.357	1.595	2.428
	SD	1.872	0.657		
Written Expression	Mean	3.630	1.417	2.213	2.277
	SD	1.409	0.972		

Note. *-p <.05,**-p <.01, ***-p <.001

Table 11.
Teacher Behavior Rating Scale Scores: North Carolina

Measure		Treatment	Control	Difference	ES
Features	Mean	2.000	1.000	1.000	0.707
	SD	1.414	1.414		
Book Reading Behaviors	Mean	3.929	3.286	0.643	0.457
	SD	0.665	0.436		
Phonological Awareness	Mean	0.500	1.667	-1.167	-0.764
	SD	1.000	1.528		
Math Concepts	Mean	1.500	2.095	-0.595	-1.803
	SD	0.247	0.330		
Written Expression	Mean	2.917	2.444	0.473	2.464
	SD	0.569	0.192		

Note. *-p <.05,**-p <.01, ***-p <.001

Table 12.
Teacher Behavior Rating Scale Scores: Georgia

Measure		Treatment	Control	Difference	ES
Features	Mean	0.000	0.200	-0.200	-0.447
	SD	0.000	0.447		
Book Reading Behaviors	Mean	0.000	0.714	-0.714	-0.447
	SD	0.000	1.597		
Print & Letter Knowledge	Mean	4.067	1.633	2.434	5.445
	SD	1.331	0.447		
Oral Language Use	Mean	4.914	1.486	3.428	5.258
	SD	1.062	0.652		
Phonological Awareness	Mean	2.400	1.600	0.800	0.702
	SD	2.191	1.140		
Math Concepts	Mean	4.114	0.914	3.200	14.747
	SD	1.780	0.217		
Written Expression	Mean	4.200	0.800	3.400	5.611
	SD	1.677	0.606		

Note. *-p <.05,**-p <.01, ***-p <.001

Table 13.
Quantity of technical assistance by content area and site

	NC Total Visits	NC Visits per Teacher	Ga Total Visits	Ga Visits per Teacher
Classroom environment	20	5.00	71	14.20
Assessment and *CC.net*	12	3.00	17	3.40
Studies and projects	1	0.25	45	9.00
Classroom observations	6	1.50	16	3.20
The CC Implementation Checklist	20	5.00	19	3.80
Planning and scheduling	8	2.00	19	3.80
Modeling teaching practices	7	1.75	2	0.40
Total number of visits	40	10.00	115	23.00

Note . The total number of visits for each site is not the simple sum of the number of visit by content area as more than one content area may have been addressed within a single visit.