

Quality in Inclusive Preschool Classrooms

By: [Linda L. Hestenes](#), [Deborah J. Cassidy](#), Jonghee Shim, Archana V. Hedge

This is an Author's Accepted Manuscript of an article published in

Hestenes, L., Cassidy, D., Shim, J., & Hegde, A. (2008). Quality in inclusive preschool classrooms. *Early Education and Development*, 19(4), 519-540.
doi:10.1080/10409280802230973

as published in the *Early Education and Development* 2008 [copyright Taylor & Francis], available online at: <http://www.tandfonline.com/10.1080/10409280802230973>.

Abstract:

Research Findings: Quality of care for preschool children in inclusive and noninclusive classrooms was examined in two studies. In Study 1, comparisons across a large sample of classrooms (N = 1, 313) showed that inclusive classrooms were higher than noninclusive classrooms in global quality as well as on two dimensions of quality (Activities/Materials and Language/Interactions). In Study 2, a more diverse sample of 44 classrooms (20 inclusive and 24 noninclusive) did not reveal differences on the global measures of quality but did show a difference on a measure of teacher–child interactions. Teachers in inclusive classrooms had higher quality and more appropriate interactions with all children than did teachers from noninclusive classrooms. Ratings of perceived severity of children's disabilities were not related to any of the measures of classroom quality in either study. Practice or Policy: Both studies suggest that including children with disabilities in regular preschool classrooms does not result in lower quality programs or in less adequate teacher–child interactions, particularly for children with mild to moderate disabilities. Results illustrate the importance of continued education for early childhood professionals on high-quality teacher–child interactions. Faculty in personnel preparation programs as well as policymakers need to continue to promote high-quality interactions between teachers and children.

Keywords: Education | Preschool | Teacher-Child Interactions | Inclusive Classrooms | Non-inclusive Classrooms

Article:

Introduction

The reauthorization of the Individuals with Disabilities Education Act in 1997 (Public Law 105–17) and the widespread acknowledgment of the importance of children with disabilities being placed in natural environments have resulted in more young children with disabilities being served in inclusive classrooms (Guralnick, 2001). Additionally, several factors, including the play-based focus of the usual preschool curriculum, the relative lack of standardized testing at

the preschool level, and similar developmental needs of all children (with and without disabilities), allow for easy integration of children with disabilities into regular preschool programs (Hanson et al., 2001). This type of rationale has led to approximately one half of all preschoolers with disabilities served through the Preschool Grants Program being taught in regular (inclusive) classrooms (U.S. Department of Education, 1999). The need for empirical data on the effect of child care quality on all young children in inclusive programs was addressed by the President's Commission on Excellence in Special Education (Shonkoff & Phillips, 2000; Wolery, & Bailey, 2002). However, an initial step is to determine the quality of care and education being provided in inclusive child care classrooms and to evaluate how quality is being measured in these environments.

Research results related to global child care quality for young children in inclusive and noninclusive programs are inconsistent (Bruder & Brand, 1995; Buysse, Wesley, Bryant, & Gardner, 1999; La Paro, Sexton, & Snyder, 1998), which has been a cause of concern for parents, educators, and policymakers. La Paro et al. found that segregated and inclusive classrooms did not differ on overall quality measures. Both segregated as well as inclusive settings were of moderately good quality. However, due to the small sample size, the results of this study need to be interpreted with caution. On the other hand, Buysse et al. found that 62 inclusive programs scored better on a global quality measure than did noninclusive programs. Bruder and Brand had similar results for their study in which they compared inclusive programs for toddlers with noninclusive programs. Inclusive programs observed were of higher quality than noninclusive programs.

STRUCTURAL AND PROCESS QUALITY

Although measurement of global quality in inclusive settings is critical, such measurement provides only a glimpse of the quality picture. Both structural and process quality must be thoroughly examined in order to understand the relative contribution of each to the inclusive environment. Structural variables are typically considered to be readily measured items such as teacher education, adult-child ratios, group sizes, and so on. Process quality is less easily measured and includes variables such as teacher-child interactions, child-child interactions, as well as health and safety practices (Vandell & Wolfe, 2000). Vandell and Wolfe provided a summary of the research demonstrating the relationship between structural and process quality. They reported that most research findings in the child care quality literature have indicated a significant relationship between structural characteristics and process quality. Furthermore, high-quality interactions between teachers and children are more likely to occur in classrooms where teacher-child ratios are low and teachers have higher levels of education (National Institute of Child Health and Human Development [NICHD] Early Child Care Research Network, 1996, 2000a; Phillips, Howes, & Whitebook, 1992).

For example, Phillipsen, Burchinal, Howes, and Cryer (1997) examined the associations between structural and process quality in 100 child care centers in each of four states (California,

Colorado, Connecticut, and North Carolina) by using scores on process indicators of quality—including the Early Childhood Environment Rating Scale (ECERS; Harms & Clifford, 1980), the Caregiver Interaction Scale (Arnett, 1989), and the Teacher Involvement Scale (Howes & Stewart, 1987)—as well as structural indicators. The authors found strong associations between process quality and structural quality. Structural quality factors such as lead teacher wages, classroom ratio, as well as center enrollment were strong predictors of process quality (teacher sensitivity and responsiveness). Centers with teachers who had at least some college education, lower classroom ratios, and higher pay scored higher on the ECERS as well as on the sensitivity and responsiveness factors of the Caregiver Interaction Scale.

In a similar study, Burchinal, Cryer, Clifford, and Howes (2002) examined the associations between caregiver sensitivity (process quality) and caregiver education (structural quality) in 553 child care classrooms based on Early Childhood Environment Rating Scale–Revised (ECERS-R; Harms, Clifford, & Cryer, 1998) and Caregiver Interaction Scale scores. Classrooms with teachers who had a bachelor's degree scored significantly higher on the ECERS-R and Caregiver Interaction Scale than classrooms with teachers with less education. Moreover, children in classrooms of teachers with a bachelor's degree showed significantly better language skills than did children in classrooms with teachers with less education. Pianta et al. (2005) also reported that global quality of pre-kindergarten classrooms was lower when teachers did not have a bachelor's-level education in early childhood, when teachers held traditional beliefs (had an adult-centered perspective), and when children enrolled in the classroom came from very poor families.

An NICHD Early Child Care Research Network (2002) study also reported direct associations between structural quality and process quality. This large-scale study found that structural quality directly affected process quality, and process quality in turn influenced children's outcomes. The results revealed positive associations between caregivers' training, child–staff ratio, and the quality of caregiving behaviors (e.g., sensitivity, stimulating, etc.). Caregiving quality was found to be related to children's cognitive competency. Phillips, Mekos, Scarr, McCartney, and Abbott-Shim (2000) found that classroom quality, especially the quality of teacher–child interactions, was positively related to structural factors such as teacher training and education, parent fees, teachers' wages, and teacher–child ratio, and was negatively related to group size.

PROCESS QUALITY AND CHILD OUTCOMES

The relationship between process quality, in particular, and children's development has been clearly demonstrated. Vandell and Wolfe (2000) found short-term and long-term effects of process quality on children's development: Preschoolers in child care centers with better process quality had fewer reported behavioral problems and performed better in math through second grade; these effects held even for children with less educated mothers. Peisner-Feinberg et al. (2001) reported that classroom practices were significantly related to children's cognitive and

language skills through kindergarten (e.g., math and reading). Beller and colleagues found that two of the ECERS (Harms & Clifford, 1980) subscales, Language Reasoning and Adult Needs, were predictors of developmental outcomes (Beller, Stahnke, Butz, Stahl, & Wessels, 1996). Despite such convincing evidence, questions remain about the most salient features of process quality and how to most effectively measure these features, particularly in classrooms that include children with disabilities.

TEACHER–CHILD INTERACTIONS IN INCLUSIVE SETTINGS

For teachers in inclusive classrooms, the challenge of high-quality interactions is even greater as they strive to be responsive to the needs of all children with and without disabilities. The types of interactions and conversations that are conducted with children with and without disabilities influence all facets of children's development, including their ensuing interactions with peers. Measurement of teacher–child interactions seems particularly important in understanding this dimension of process quality. Researchers in the field are called upon to study this important aspect of inclusion (Odom, 2000).

In classrooms that include children with disabilities, teachers' interaction patterns appear to be somewhat different from their interactions with typically developing children. Teachers are generally observed to be more directive and less child centered (not supportive of child-initiated activities) in their interactions with children with disabilities (Goodman, 1992). Results of another study found that teachers who were highly responsive and moderately directive in their behavior were more successful in engaging children with disabilities in meaningful activities in the classroom (Mahoney & Wheeden, 1999). Teachers' differing styles of interaction patterns with children with disabilities has been an issue of debate in the field.

Several studies have revealed that teachers use more directives with children with disabilities than with typically developing children (Chow & Kasari, 1999; File, 1994; Hestenes, Cassidy, & Niemeyer, 2003; Quay, 1991; Stipek & Sanborn, 1985). File's research indicated that teachers in inclusive preschool classrooms were more directive (“Fill up the cup”; asking closed questions) of the cognitive experiences of children with disabilities than of the cognitive experiences of typically developing children. Also, teachers were more likely to support cognitive play than social play behaviors. Indeed, support of social play (play with peers) was relatively infrequent (only 2%). Furthermore, Quay reported that teachers were more negative toward children with disabilities than toward typically developing children. A study by McWilliam, Scarborough, and Kim (2003) found that elaborations and information giving were associated with children's engagement and that interactions targeted at individual children with disabilities produced more engagement on the part of the children than did group-targeted interaction.

Studies of inclusive classrooms have suggested that teachers may be more involved with children with disabilities than with other children (Brophy & Hancock, 1985; Chow & Kasari, 1999; Hundert, Mahoney, & Hopkins, 1993), although their involvement is mixed in terms of its

appropriateness. For example, Chow and Kasari found that at the beginning of the school year in inclusive classrooms, teachers initiated more negative and task-related interactions with children with disabilities than with their typical peers. However, at the end of the school year, teacher interactions with the children with disabilities were similar to those with the typically developing children. Research has also indicated that teacher presence is predictive of more interactions between preschool children with and without disabilities in inclusive classrooms (Hestenes & Carroll, 2000). The teacher's role and involvement with young children is clearly a key aspect underlying process quality in inclusive classrooms.

THE CURRENT STUDY

The current article discusses two studies (Study 1 and Study 2) that examined the differences in quality between inclusive and noninclusive classrooms and the relationship between structural and process quality, with a particular focus in the second study on teacher–child interactions. The goals of the first study were to compare, using a large data set, differences between inclusive and noninclusive classrooms in terms of structural dimensions of quality and global quality. The specific research questions examined in this study were as follows: (a) Do inclusive and noninclusive classrooms differ on a global measure of quality (ECERS-R)? (b) Do inclusive and noninclusive classrooms differ on ECERS-R subscales and identified factor scales (Activities/Materials and Language/Interactions)? (c) Do structural components of quality (classroom ratios, teacher's level of education, teacher's experience, and course hours in special education) differ across inclusive and noninclusive classrooms? (d) How does severity of children's disabilities relate to different quality measures (ECERS-R and identified factor scales)?

The aforementioned research questions were also examined in the second study, which had a more diverse (with respect to the level of quality) sample of inclusive and noninclusive classrooms. Additional research questions examined the relationship between global assessment of quality (ECERS-R) and a measure of teacher–child interaction (the Teacher Child Interaction Scale [TCIS]). We also investigated the differences between the subscales of the TCIS across the inclusive and noninclusive classrooms. The TCIS (Farran & Collins, 1996) is a measure of process quality and includes numerous dimensions of the teacher–child interaction construct. The measure was believed to provide a more comprehensive picture (based on the literature) of the nature of teacher–child interaction. Thus, based on these measures, employed across inclusive and noninclusive classrooms, the following research questions were added: (a) Do inclusive and noninclusive classrooms differ on the amount, quality, and appropriateness of teacher–child interactions? (b) Is there a relationship between structural components of quality (classroom ratios, teacher's level of education, teacher's experience, and course hours in special education) and teacher–child interactions?

STUDY 1

Method

Participants and measures

Data were collected from 1,313 preschool classrooms in programs across the state of North Carolina. Assessors completed the ECERS-R (Harms et al., 1998) in each classroom as part of the North Carolina Star Rated License process. During this study, licensed programs received from one to five stars depending on points earned across three areas: compliance history, staff education, and program standards. Programs that wanted to earn more points in the program standards category chose to have a measure of global quality completed using the Environment Rating Scales. In preschool classrooms the ECERS-R was used. This 43-item observational instrument is completed during a 3- to 4-hr observation. Each item is rated from 1 (*inadequate*) to 7 (*excellent*). The scale contains seven subscales: Space and Furnishings, Personal Care Routines, Language–Reasoning, Activities, Interaction, Program Structure, and Parents and Staff. Because average scores are computed for state regulatory purposes without the use of the Parents and Staff subscale, this subscale was also removed from analyses in the current study.

Because the programs that were assessed were attempting to earn more points in the program standards category, most of the classrooms observed for this study were from higher quality programs. Approximately 20% to 25% of programs in North Carolina have assessments completed. Each assessor receives extensive training and must achieve and maintain at least an 85% interobserver reliability in order to complete an assessment. Once trained, the procedures require that each assessor have his or her reliability checked after every seventh assessment (if the assessor's reliability is between 85% and 89%) or after every tenth assessment (if the assessor's reliability is 90% or better). In addition to completing the ECERS-R, assessors gather information on teacher–child ratio and group size during the observation. Demographic information on the teachers as well as the teachers' perceptions of the level of severity of disability for any children with disabilities in the classroom are gathered during a teacher interview. Directors of programs provide information on children's age, gender, and type of disability, if applicable.

Procedures

Licensing consultants from the Division of Child Development in North Carolina contact the North Carolina Rated License Assessment Project (NCRLAP) if they have a program that is interested in earning more stars on their Rated License. The assessors from the NCRLAP set up an appointment with the program to complete the assessment. On the day of the visit, one of the assessors randomly selects the classrooms that are to be observed. One third of the classrooms at each age level (infant/toddler, preschool, and school age) must be assessed, with at least one classroom from each level. Each assessor spends 3 to 4 hr during the morning observing all aspects of the classroom as defined by the ECERS-R. Assessors attempt to be as inconspicuous as possible within the classroom. They also follow the children to the outdoor learning

environments. At the end of the observation, a teacher interview is conducted for approximately 20 min to complete the scale items not able to be observed and to obtain demographic information. For this study, only data from the preschool-age classrooms (that utilized the ECERS-R) are reported.

Results

Preliminary analyses

Data from 1,313 preschool classrooms were analyzed for this study. Preliminary checks of the data showed normal distributions for all of the variables used in this study, with the exception of the number of courses taken in special education by the teachers. This variable was transformed (Tabachnick & Fidell, 1996) and reanalyzed, but the results did not change. The nontransformed variable is reported for ease of interpretation. Based on previous results of a factor analysis on the ECERS-R (Cassidy, Hestenes, Hegde, Hestenes, & Mims, 2005), two factors appear to differentiate the scale into two aspects of quality: Activities/Materials and Language/Interactions. Factors represent subsets of the total scale. Planned comparisons between inclusive and noninclusive classrooms were conducted on the overall average score, the two factors, and the seven subscales of the ECERS-R.

Table 1 provides descriptive information on teacher education, teacher experience in the early childhood field, and the number of credit hours taken in special education coursework. Teacher's level of education was coded into 1 of 12 categories from *did not complete high school* to *graduate degree*. Teacher education data was not available for 17 teachers.

Table 1. Teacher Information

Variable	N	M	SD	Min	Max
Teacher education	1,296	5.84	2.88	1	12
Teacher experience (years)	1,290	11.19	7.75	0	36
Credit hours in special education courses	1,191	5.07	10.3	0	72

In the data, there were 459 classrooms (35%) that contained at least one child with an identified disability and 854 classrooms (65%) that contained only children who were typically developing. In each of the inclusive classrooms, assessors recorded the types of disabilities for each of the

children (for up to five children) and asked teachers to rate the children's severity level from 1 (*mild*) to 3 (*severe*). These ratings were based on the teachers' perceptions of the child's severity, which stemmed from what they knew about the child's diagnosis and the behaviors that occurred in the classroom. In the 459 inclusive classrooms there were 1,183 children with disabilities, and 1,145 of the children with disabilities had severity scores recorded. The types of disabilities were classified into one of nine categories (see Table 2). The average severity rating across 1,145 of the children was 1.68, and the average severity by classroom was 1.67 ($SD = 0.57$), indicating that on average children in these classrooms were perceived by their teachers as being between mild and moderate in their disability. However, the restricted 1-to-3 scale did not allow for a great deal of variability in the severity rating. The average severity by classroom was computed by adding up each child's severity rating and dividing by the total number of children with disabilities.

Table 2. Frequency, Percentage, and Severity of Disability Type in Inclusive Classrooms

Type of Disability	Frequency	%	Average Severity Rating ^a
Attention-deficit/hyperactivity disorder or attention-deficit disorder	26	2	1.88
Autism	40	3	2.00
Behavior/emotional	27	2	1.52
Cognitive delay	46	4	1.63
Down syndrome	16	1	1.93
Other health impaired	52	4	2.06
Pervasive developmental disorder	240	20	1.77
Physical disability	78	7	2.00
Speech	658	56	1.56
Total	1,183		1.68

^aHigher values represent greater severity.

Planned comparisons of inclusive and noninclusive classrooms

One-way analysis of variance was used to determine if there were differences between inclusive and noninclusive classrooms. The results indicated that inclusive preschool classrooms ($n = 459$) were higher in global quality than noninclusive classrooms ($n = 854$), $F(1, 1311) = 42.4, p < .0001$. Planned comparisons across the factors showed inclusive classrooms scored significantly higher than noninclusive classrooms in the Activities/Materials factor, $F(1, 1311) = 18.9, p < .0001$; and the Language/Interactions factor, $F(1, 1311) = 49.5, p < .0001$. Inclusive classrooms had significantly higher scores on four of the ECERS-R subscales (Language–Reasoning, Activities, Interaction, and Program Structure) when we used a stringent p value to control for

multiple comparisons. There were no differences between the Space and Furnishings or Personal Care Routines subscales (see Table 3 for means).

Teachers from inclusive classrooms ($n = 457$) had significantly more education than teachers from noninclusive classrooms ($n = 839$), $F(1, 1294) = 17.52, p < .000$ ($M_{\text{inclusive}} = 6.29$ [indicating a 2-year community college degree]; $M_{\text{noninclusive}} = 5.60$ [indicating some college]). Teachers in inclusive classrooms had more college course hours in special education than teachers of typically developing children, $F(1, 1189) = 39.1, p < .000$ ($M_{\text{inclusive}} = 7.38$ credit hours; $M_{\text{noninclusive}} = 3.87$ credit hours). Teachers in inclusive classrooms had also been in the early childhood field for significantly longer than teachers from noninclusive classrooms, $F(1, 1288) = 20.5, p < .000$ ($M_{\text{inclusive}} = 12.5$ years; $M_{\text{noninclusive}} = 10.5$ years). To determine teacher-child ratios for inclusive and noninclusive classrooms, we divided the total number of children present in the classroom by the number of adults, including teachers who were present in the classroom. Inclusive and noninclusive classrooms did not differ in the ratio of children to teachers, $F(1, 1300) = 1.59, p = .21$ ($M_{\text{inclusive}} = 6.3$; $M_{\text{noninclusive}} = 6.5$).

Table 3. Mean Scores on ECERS-R Factors, Subscales, and Average Score in Inclusive and Noninclusive Classrooms

Variable	Overall M (SD) (N = 1,313)	Range (1–7)	Inclusive M (n = 459)	Noninclusive M (n = 854)
Factors				
Activities/Materials	4.93 (1.04)	2.33–7.00	5.10	4.84**
Language/Interactions	5.56 (1.06)	1.29–7.00	5.84	5.42**
Subscales				
Space and Furnishings	4.95 (0.78)	2.13–7.00	5.02	4.91
Personal Care Routines	3.89 (0.99)	1.67–7.00	3.97	3.85
Language-Reasoning	5.41 (0.88)	1.75–7.00	5.62	5.30**
Activities	5.11 (0.91)	2.30–7.00	5.29	5.02**
Interaction	5.59 (1.18)	1.00–7.00	5.86	5.45**
Program Structure	5.67 (0.99)	1.00–7.00	5.88	5.56**
Average score	5.03 (0.70)	2.56–6.81	5.19	4.94**

Note: ECERS-R = Early Childhood Environment Rating Scale-Revised

** $p < .0001$ (comparisons between inclusive and noninclusive classrooms).

To ensure that differences in quality between inclusive and noninclusive classrooms were not due only to the effects of teacher education, experience, or special education credit hours, we reran all analyses with these variables as controls. With these control variables, inclusive classrooms were still significantly higher in quality than noninclusive classrooms. Correlations across all classrooms revealed significant associations between overall child care quality and

education ($r = .29, p < .000$), years of experience ($r = .13, p < .000$), and teacher–child ratio ($r = -.15, p < .000$).

In order to examine whether the severity of children's disabilities was related to quality measures, correlational analyses were computed. Results showed that there were no significant relationships between the average classroom severity score and the average ECERS-R score ($r = .05$), the Activities/Materials factor ($r = .06$), or the Language/Interactions factor ($r = .02$).

Discussion

The results of Study 1 indicate support for the findings of Buysse et al. (1999) and Bruder and Brand (1995) that inclusive toddler and preschool classrooms are of higher quality than noninclusive classrooms. Overall, the inclusive classrooms in the current study were higher in global child care quality and had better educated (including more hours in special education) and more experienced teachers when compared to noninclusive programs. Differences on the ECERS-R between the two groups seemed to focus primarily on subscales on which teachers had more direct control (i.e., Language–Reasoning, Activities, Interaction, and Program Structure). There were no differences between inclusive and noninclusive classrooms on the Space and Furnishings and Personal Care Routines subscales. These subscales are less likely to be influenced by the teacher and may be more dependent on programmatic decisions made by administrators or by staff consensus. This information would seem to lend support to the argument that teacher behavior is a primary contributing variable to the differences in quality between inclusive and noninclusive classrooms. However, the ECERS-R provides little detail about the specific behaviors of the teacher. For example, the Interaction subscale includes items about supervision of the children, but it only indicates the presence or absence of a behavior. For example, Item 30 (General Supervision of Children), Indicator 3.3, requires that the observer indicate whether “most supervision is non-punitive, and control is exercised in a reasonable way.” We know little about the nature of the verbalizations used in supervision and how responsive or positive the teacher was during this supervision. The Language–Reasoning and Program Structure subscales also include little information about the specifics of the teacher's influence. For example, Item 36 (Group Time), Indicator 7.2, requires that the observer indicate whether “staff engage in educational interaction with small groups and individual children as well as with the whole group.” Again, what this behavior entails is unclear; furthermore, in many cases this item would not be scored because, on the ECERS-R, if observers do not score all indicators of previous items as “pass,” then they do not score the indicators under the score of 7. Although the role of the teacher seems to be influential in the differences in quality in the two program types, because the specifics of that intervention are not clearly delineated by the ECERS-R, Study 2 examined in greater depth the nature of teacher behaviors in inclusive and noninclusive classrooms.

It is interesting to note that in this study the severity of the disability did not contribute to differences in classroom quality. Although most teachers rated the children as having mild or

moderate disabilities, the level of severity did not affect program quality. However, the majority of the children in the current study had speech delays/impairments. These types of disabilities may not impact the overall classroom climate to the extent that other types of disabilities would. Nonetheless, this information is counter to the concerns of many teachers who worry that children with disabilities will occupy a disproportionate amount of their time and impact the overall quality of their classroom. That did not prove to be the case in the current study.

It is important to note that the results of the study do not imply causality. Rather than a causal relationship between inclusion and classroom quality, it is likely that programs that include children with disabilities are those that ascribe to higher quality standards and recognize the benefits of including children with and without disabilities in their classrooms. It may also be that parents of children with disabilities select higher quality settings, or perhaps the settings improve as children with disabilities are enrolled and more individual needs are met. It is also important to note that the sample in Study 1 was a high-quality sample of programs that were attempting to achieve more stars on the North Carolina Rated License. The second study examined a more diverse sampling of quality in programs in the state and also attempted to employ a finer grained analysis of the construct of teacher–child interaction.

STUDY 2

Method

Participants

Participants for this second study came from 24 noninclusive classrooms and 20 inclusive preschool classrooms. Classrooms were from licensed child care centers in three different mid-size cities in North Carolina. A total of 72 teachers participated in the study (all female). The racial distribution was 49% Euro-American, 50% African American, and 1% Hispanic. About 70% of the teachers had a high school diploma, some college, or a 2-year college degree; 30% had a 4-year degree or some graduate coursework. Classrooms were considered inclusive if they contained at least one child with an identified disability. Classrooms included a total of 58 children with disabilities, with a mean of 2.9 and a range from 1 to 10 children per class. Each child's disability was rated from 1 (*mild*) to 3 (*severe*) by his or her teacher. A total of 51 children received these ratings. Overall, the average severity of disability for the children with disabilities was 1.96 ($n = 51$), and the average severity by classroom was 1.90 ($SD = 0.61$; see Table 4 for frequencies, percentages, and average severity ratings).

Measures and procedures

Phone calls were made to directors of programs varying in their level of quality (from one to five stars), asking if they would like to participate in the study. State-licensed child care centers were recruited from three mid-size cities in North Carolina. Interested directors were mailed letters of consent and followed up with a phone call. Once director consent was obtained,

teachers of preschool-age children from the program were asked to participate. An effort was made to obtain approximately equal numbers of classrooms that were inclusive and noninclusive.

Trained assessors completed the ECERS-R (Harms et al., 1998) in each of the 44 classrooms. Data on the ratio of the number of children to teachers were collected during the ECERS-R observation. Different trained assessors completed the TCIS (Farran & Collins, 1996) on each of the 72 teachers. The TCIS is an observational measure that was designed to code the amount, quality, and appropriateness of teachers' interactions with young children. For this measure, observations of interactions between the target teacher and children in the classroom or attempts by children to elicit attention from the teacher were recorded. The observations were conducted on two mornings during free play for 30 min each morning. During the 30-min observation, the coders kept detailed notes on the interactions that occurred, and then immediately after the observation ended they rated 11 teacher behaviors on amount, quality, and appropriateness (33 total items). The 11 teacher behaviors included physical involvement, verbal involvement, responsiveness, play interaction, teaching behavior, control of activities, directive/demands, relationship among activities, positive statements, negative statements, and goal setting. Each behavior is rated from 1 to 5, with 1 indicating *inappropriate/poor quality/none* and 5 indicating *very appropriate/high quality/frequently*. According to scale guidelines, each behavior is rated on amount, quality, and appropriateness unless the amount of the behavior is very low or nonexistent. In this case, "amount" is scored as 1 and quality and appropriateness are not scored. This scoring procedure is based on the idea that more than a minimum number of observations of a teacher behavior is needed to rate the quality or appropriateness of the behavior. Interrater reliability was established at 85% and maintained via periodic checks ($M = 92\%$; range = 86%–98%). Additional reliability analyses indicated that the TCIS had strong internal consistency across the total scale (Cronbach's $\alpha = .95$) as well as acceptable internal consistency for the subscales (Amount = .73, Quality = .93, and Appropriateness = .90). Demographic information on the teachers and children in the classrooms was also collected during a 20-min teacher interview.

Table 4. Frequency, Percentage, and Severity of Disability Type in Inclusive Classrooms

Type of Disability	Frequency	%	Average Severity Rating ^a
Attention-deficit/hyperactivity disorder or attention-deficit disorder	11	19	1.89
Autism	7	12	2.17
Behavior/emotional	3	5	2.33
Cognitive delay	1	2	
Down syndrome	3	5	2.00
Other health impaired	2	3	2.00
Pervasive developmental disorder	10	17	1.80

Physical disability	10	17	2.00
Speech	11	19	1.91
Total	58		1.96

^aHigher values represent greater severity.

Results

Preliminary analyses

The accuracy of data was checked by examining frequencies, histograms, missing data, outliers, linearity, and normality. Global quality scores on average were below the ECERS-R rating of “good” and were far lower in Study 2 than in Study 1, which had a higher quality sample.

Average scores from the seven subscales on the ECERS-R as well as a total average score on the measure were computed for each classroom. Scores for each of the 11 teacher behaviors were created by combining scores of the two observations of the TCIS. Then, summary scores on “amount” were computed by summing scores of teacher behavior under the amount area except for three items (Control Over Children's Activities, Directives/Demands, and Negative Statements). Summary scores on quality and appropriateness were created by summing all scores across 11 of the teacher behaviors under each area, respectively. In order to examine the relationship between the global quality scores and the specific measure of teacher–child interactions, correlation coefficients were computed between ECERS-R scores and TCIS scores. The results of the correlation analyses are presented in Table 5. The correlations suggested a strong relationship between the two measures, but clearly one measure did not serve as a proxy for the other.

Teacher's level of education was again coded into 1 of 12 categories from *did not complete high school* to *graduate degree*. The descriptive information on teacher education level, experience, and special education coursework across all 72 teachers is presented in Table 6. On average, teachers had slightly less than a 2-year degree and about 7½ years of experience in the field. Very few teachers had special education coursework in this sample.

Planned comparisons of inclusive and noninclusive classrooms

Unlike the first study, the results indicated that inclusive and noninclusive classrooms did not differ in their total scores on global quality (as measured by the ECERS-R), $F(1, 42) = 2.56, p = .11$; the Activities/Materials factor, $F(1, 42) = 2.2, p = .15$; the Language/Interactions factor, $F(1, 42) = 2.6, p = .11$; or the subscales (when p values were adjusted for multiple comparisons). On the TCIS, however, teachers from inclusive classrooms were rated as displaying significantly more interactions, $F(1, 42) = 5.6, p = .02$; that were of higher quality, $F(1, 42) = 7.5, p = .009$; and that were more appropriate, $F(1, 42) = 5.3, p = .03$; than those of teachers from classrooms that contained only typically developing children (see Table 7 for mean scores). There were no

differences between inclusive and noninclusive classrooms in terms of teacher education, group size, ratio, credit hours in special education, or teacher experience. In the inclusive classrooms, the perceived severity of children's disabilities was not significantly related to any of the quality measures ($r_s = -.01$ to $-.29$).

Structural dimensions associated with global quality and teacher–child interaction.

A second goal of this study was to look at structural dimensions that were related to global quality and teacher–child interaction across all classrooms. For these analyses, teacher education, experience, and hours in special education were averaged across the two teachers in each classroom. Teacher education was significantly related to overall global quality ($r = .40$, $p < .01$) as well as the two ECERS-R factors: Activities/Materials ($r = .37$, $p = .01$) and Language/Interactions ($r = .44$, $p < .01$). Education was also positively related to teacher–child interaction quality ($r = .33$, $p = .03$) and appropriateness ($r = .34$, $p = .03$), but not to the frequency of teacher–child interactions ($r = .20$, $p = .2$). Years of experience in early childhood was not related to the ECERS-R or the TCIS. In addition, there were significant negative relationships between both measures of quality and the number of children present in the classroom ($r_s = -.31$ to $-.38$, $p < .05$) and between the measures of quality and the ratio of children to teachers ($r_s = -.45$ to $-.52$, $p < .01$).

Table 5. Correlations Between Global Quality and Teacher Behaviors

Variable	Space and Furnishings	Personal Care Routines	Language–Reasoning	Activities	Interaction	Program Structure	Parents and Staff	ECERS-R Average	ECERS-R Activities/Materials	ECERS-R Language/Interactions
TCIS Amount	.478**	.526***	.445**	.353*	.533***	.287	.258	.519***	.444**	.488**
TCIS Quality	.548***	.548***	.567***	.449**	.617***	.480**	.240	.629***	.534***	.597***
TCIS Appropriateness	.560***	.553***	.616***	.517***	.624***	.558***	.304*	.672***	.574***	.633***

Note: ECERS-R = Early Childhood Environment Rating Scale–Revised; TCIS = Teacher Child Interaction Scale.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

Table 6. Teacher Information ($N = 72$)

Variable	M	SD	Min	Max
Education level	5.43	3.08	2	11
Experience (years)	7.67	7.41	0	48
Special education credit hours	1.33	3.13	0	18

Table 7. Mean Scores on ECERS-R Factors and Subscales, ECERS-R Average Score, and Mean Scores on TCIS Subscales in Inclusive and Noninclusive Classrooms

<i>Variable</i>	<i>Overall M (SD)</i>	<i>Actual Range</i>	<i>Possible Range</i>	<i>Inclusive M (n = 20)</i>	<i>Noninclusive M (n = 24)</i>
ECERS-R average score	4.41 (1.04)	2.36–6.38	1–7	4.68	4.19
ECERS-R subscales					
Space and Furnishings	4.61 (0.85)	2.75–6.25	1–7	4.88	4.38
Personal Care Routines	4.11 (1.24)	1.66–7.00	1–7	4.58	3.71
Language–Reasoning	4.36 (1.39)	1.50–7.00	1–7	4.66	4.11
Activities	4.12 (1.20)	1.30–6.60	1–7	4.18	4.06
Interaction	4.80 (1.59)	1.40–7.00	1–7	5.29	4.38
Program Structure	4.86 (1.48)	1.67–7.00	1–7	4.94	4.79
ECERS-R factors					
Activities/Materials	4.24 (1.21)	1.44–6.78	1–7	4.53	3.99

<i>Variable</i>	<i>Overall M (SD)</i>	<i>Actual Range</i>	<i>Possible Range</i>	<i>Inclusive M (n = 20)</i>	<i>Noninclusive M (n = 24)</i>
Language/Interactions	4.81 (1.45)	1.29–7.00	1–7	5.19	4.49
TCIS					
Amount	47.58 (9.64)	19–69	16–80	51.35	44.5*
Quality	69.51 (18.93)	15–109	0–110	77.25	62.79**
Appropriateness	69.61 (18.79)	19–106	0–110	76.25	63.46*

Note: ECERS-R = Early Childhood Environment Rating Scale–Revised; TCIS = Teacher Child Interaction Scale.

* $p < .05$.

** $p < .01$.

Discussion

Study 2 provides an in-depth analysis of teacher behavior using a teacher–child interaction scale that measures the amount, quality, and appropriateness of teacher behaviors. The findings contribute to our understanding of how teachers influence quality in inclusive and noninclusive classrooms. Teachers in inclusive classrooms were observed to have more interactions that were of higher quality and that were more appropriate for the age level of the child. However, global classroom quality differences were not found based on the ECERS-R scores. This may demonstrate that the TCIS is more sensitive than the ECERS-R to teacher behavior differences in this smaller sample.

Examination of the relationship between education and quality helps us understand how education impacts teachers' classroom behavior. Data from the sample showed that education was linked to the quality and appropriateness of interactions but not the frequency. Teachers with more education did not seem to have more interactions with children, but the interactions were higher in quality and were more appropriate for the ages of the children. Kontos and Wilcox-Herzog (1997) concluded in their review of the literature that quality and not quantity of interaction is most important for responsive involvement with children.

Education was not the only structural variable that influenced classroom quality. Group size and staff–child ratios were also negatively related to classroom quality. That is, the larger the group

size and higher the ratios, the lower the global quality and the less appropriate the teacher–child interactions. Although not surprising, these findings lend support to the mounting evidence for the relationship between these structural and process quality factors.

The positive correlations between the ECERS-R subscales, factors, and average global score and the different components of the TCIS are encouraging, but it is clear that one measure cannot be substituted for the other. The global measure of quality does not appear to capture the level of responsive interactions that occur between teachers and children.

GENERAL DISCUSSION

The two studies presented here provide continuing support for the critical role of the teacher in understanding classroom quality, particularly in inclusive and noninclusive classrooms. Study 1 indicates that inclusive classrooms are of higher global quality than noninclusive classrooms, whereas Study 2, with a smaller but more diverse sample, did not find differences between the groups on global quality but did find interesting differences on the quality and appropriateness of teacher–child interactions in these classrooms. The lack of a significant difference on global quality between inclusive and noninclusive classrooms in Study 2 can at least in part be explained by the smaller sample size. We would predict that with a larger sample in Study 2, differences in global quality would have emerged, but future research would need to verify this hypothesis.

What accounts for the higher quality found in Study 1 for classrooms that serve children with disabilities? We believe that education of the teacher in part explains the differences between the classroom types. In Study 1, teachers in inclusive classrooms had more education (2-year community college degree) than teachers in noninclusive classrooms (between one and two years of college). This difference in education played a role in higher global quality, although it was clear that when education was used as a control variable and the differences were still significant, the quality of the inclusive classroom was impacted by multiple components. When structural dimensions of quality were examined in relation to global quality and teacher–child interactions across the total samples in both studies, it was clear that education and the ratio of children to teachers were important components influencing quality. The experience of the teacher, however, was found to be significant in the larger and higher quality sample in Study 1 but was not related to quality in the smaller, more diverse quality sample in Study 2. Similar to previous research on the influence of teacher experience, the current studies may imply that the total number of years of experience may not be as important as the quality of those experiences. Future studies need to better differentiate the type of experience teachers have as well as the quality of the classrooms in which teachers work during those experiences.

The level of severity of disability did not impact classroom quality in either study. Although the severity level was somewhat higher in Study 2 (1.96) in comparison to Study 1 (1.68), there was no relationship to classroom quality. Whereas Study 1 reported that the majority (56%) of

children with disabilities had speech or language delays, Study 2 had a better balance, with 19% of the children having speech/language, 19% attention-deficit/hyperactivity, 17% pervasive developmental, and 17% physical disabilities. Nonetheless, even with these somewhat more serious disabilities in Study 2 there was not an overall impact on classroom quality. This unexpected finding should provide reassurance to classroom teachers and providers as they consider inclusion of a child with a disability. Although most of the children were reported to have mild to moderate disabilities, these are the typical disability types that are found in most inclusive classrooms.

Limitations of the current studies should be considered. Study 1 used a high-quality sample that may not be representative of the breadth of quality in child care facilities nationally. Because the facilities assessed were part of a voluntary quality rating system of a state regulatory process, the programs were only those that were interested in achieving a higher rating. Nonetheless, this may have actually provided a more stringent test for measuring the variables of concern. Clearly, a larger study with a more diverse range of quality would provide more generalizable results. A limitation from Study 2 was that specific interaction patterns with children with different disabilities were not studied. Depending on the disability and the severity level, children might solicit different interaction patterns from their teachers. Future research examining how teachers' behavior patterns and verbalizations change as a result of interactions with different children would provide additional insights for improving quality in inclusive classrooms. It would also be beneficial to measure more directly each child's level of severity with regard to his or her disability. Teacher's perceptions of severity are not as precise as a standardized measure that would provide more specific scores across different domains of development. A further limitation is that the quality measures used in these studies were not designed specifically to use in inclusive classrooms or to measure the unique needs of children with disabilities. There are particular issues in inclusive classrooms, including the teacher's decision to work in a classroom that includes children with disabilities, and especially children with moderate to severe disabilities, that may account for differences in interactions as well as global quality. There is an ongoing need to understand what connotes high-quality environments and interactions for children with vastly different needs and strengths. Clearly there are aspects of structure and process quality that would apply to all children, but there are certainly unique features and subtle differences that improve children's experiences and lead to better outcomes. These issues warrant further research on inclusion.

The use of both a measure of global quality and a teacher-child interaction measure in these studies seems to have provided a more comprehensive picture of quality and highlights the role of the teacher in the quality picture. A number of studies (Phillips et al., 2000; Pianta, La Paro, & Payne, 2002; Pianta et al., 2005) have used a global quality measure with a specific teacher-child interaction scale. In some of these studies, the Caregiver Interaction Scale (Arnett, 1989) and the Observation Record of Caregiving Environment (ORCE) have been used, and strong and significant correlations have been found between the measure and more global measures such as

the ECERS-R (Pianta et al., 2002). Study 2 informs the literature in this area by using a measure of teacher–child interaction that focuses not only on the amount of interactive behavior but, more important, on the quality and appropriateness of the interactions as well. Notably, the quality and the appropriateness of the interactions were the dimensions associated with teacher education. Furthermore, these studies examined inclusive and noninclusive classrooms and how they differ not only on global quality but on the quality of teacher–child interactions. Clearly, large-scale studies of inclusive and noninclusive classrooms using scales such as the TCIS are needed in order to better understand the role of teacher–child interactions in the quality equation. It would also be highly beneficial to examine how children's specific disabilities influence teacher–child interactions and other classroom dynamics. We grouped children with special needs into one general category, but clearly each type of disability in conjunction with the severity level will influence the ongoing interactions and the overall classroom environment.

Implications for Practice and Policy

One of the concerns expressed by parents and professionals when considering inclusive classrooms for young children is the fear that including children with disabilities will lower the overall quality of the program and impact teacher–child interactions. Our findings suggest that when children with mild to moderate disabilities are included in preschool classrooms, the quality is higher. As mentioned previously, neither of the current studies allows us to determine causality, so it may be that (a) teachers with higher quality interactions choose to work in inclusive classrooms; or (b) when children with disabilities are enrolled in the classrooms, the interaction patterns improve. Regardless of the direction of influence, it is clear that the teacher's role and involvement in the classroom are important parts of the quality of the classroom and that teacher preparation programs need to continue to emphasize high-quality interactions in coursework and practicum experiences. Early childhood teachers need to understand the multiple components that compose interactions (e.g., physical involvement, verbalizations, responsiveness, play interactions, etc.) and capitalize on the opportunities for interactions that occur during daily routines and free-play times indoors and outdoors. Emphasizing the quality and appropriateness of the interactions and not just the frequency of the interactions would further strengthen the skills of teachers of young children. Policy initiatives need to continue to emphasize the importance of low teacher–child ratios and small group sizes for all children. Maintaining high-quality interactions and high-quality programs is possible only when ratios and group sizes are manageable for teachers.

Notes

^aHigher values represent greater severity.

****** $p < .0001$ (comparisons between inclusive and noninclusive classrooms).

^aHigher values represent greater severity.

* $p < .05$.

** $p < .01$.

*** $p < .001$.

* $p < .05$.

** $p < .01$.

REFERENCES

1. Arnett, J. 1989. Caregivers in day-care centers. Does training matter?. *Journal of Applied Developmental Psychology*, 10: 541–552.
2. Beller, K., Stahnke, M., Butz, P., Stahl, W. and Wessels, H. 1996. Two measures of the quality of group care for infants and toddlers.. *European Journal of Psychology of Education*, 11(2): 151–167.
3. Brophy, K. and Hancock, S. 1985. Adult—child interaction in an integrated preschool programme. Implications for teacher training.. *Early Child Development and Care*, 22: 275–294.
4. Bruder, M. B. and Brand, M. 1995. A comparison of two types of early intervention environments serving toddler-age children with disabilities.. *Infant-Toddler Intervention*, 5: 207–217.
5. Burchinal, M. R., Cryer, D., Clifford, R. M. and Howes, C. 2002. Caregiver training and classroom quality in child care centers.. *Applied Developmental Science*, 6(1): 2–11.
6. Buysse, V., Wesley, P., Bryant, D. and Gardner, D. 1999. Quality of early childhood programs in inclusive and noninclusive settings.. *Exceptional Children*, 65: 301–314.
7. Cassidy, D. J., Hestenes, L. L., Hegde, A., Hestenes, S. and Mims, S. 2005. Measurement of quality in preschool child care classrooms: An exploratory and confirmatory factor analysis of The Early Childhood Environment Rating Scale—Revised.. *Early Childhood Research Quarterly*, 20: 345–360.
8. Chow, V. T. and Kasari, C. 1999. Task-related interactions among teachers and exceptional at-risk and typical learners in inclusive classroom.. *Remedial and Special Education*, 20: 226–232.
9. Farran, D. C. and Collins, E. N. 1996. *Teacher Child Interaction Scale* Unpublished instrument.
10. File, N. 1994. Children's play, teacher—child interactions, and teacher beliefs in integrated early childhood programs.. *Early Childhood Research Quarterly*, 9: 223–240.

11. Goodman, J. F. 1992. *When slow is fast enough.*, New York: Guilford Press..
12. Guralnick, M. J. 2001. *Early childhood inclusion.*, Baltimore: Brookes..
13. Hanson, M. J., Horn, E., Sandall, S., Beckman, P., Morgan, M., Marquart, J. 2001. After preschool inclusion: Children's educational pathways over the early school years.. *Exceptional Children*, 68: 65–83.
14. Harms, T. and Clifford, R. M. 1980. *Early Childhood Environment Rating Scale.*, New York: Teachers College Press..
15. Harms, T., Clifford, R. and Cryer, D. 1998. *Early Childhood Environment Rating Scale—Revised.*, New York: Teachers College Press..
16. Hestenes, L. L. and Carroll, D. E. 2000. The play interactions of young children with and without disabilities: Individual and environmental influences.. *Early Childhood Research Quarterly*, 15: 229–246.
17. Hestenes, L., Cassidy, D. J. and Niemeyer, L. 2003. A microanalysis of teachers' verbalizations in inclusive classrooms.. *Early Education and Development*, 15: 23–38.
18. Howes, C. and Stewart, P. 1987. Child's play with adults, toys, and peers: An examination of family and child care influences.. *Developmental Psychology*, 23: 423–430.
19. Hundert, J., Mahoney, W. and Hopkins, B. 1993. The relationship between peer interaction of children with disabilities in integrated preschool and resource and classroom teacher behaviors.. *Topics in Early Childhood Special Education*, 13: 328–343.
20. Kontos, S. and Wilcox-Herzog, A. S. 1997. Teachers' interactions with children: Why are they so important?. *Young Children*, 52(2): 4–12.
21. La Paro, K., Sexton, D. and Snyder, P. 1998. Program quality characteristics in segregated and inclusive settings..*Early Childhood Research Quarterly*, 13: 151–167.
22. Mahoney, G. and Wheeden, C. A. 1999. The effect of teacher-style on interactive engagement of preschool-age children with special learning needs.. *Early Childhood Research Quarterly*, 14: 51–68.
23. McWilliam, R. A., Scarborough, A. A. and Kim, H. 2003. Adult child interactions and child engagement.. *Early Education and Development*, 14: 7–27.
24. NICHD Early Child Care Research Network. 1996. Characteristics of infant child care: Factors contributing to positive caregiving.. *Early Childhood Research Quarterly*, 11: 269–306.
25. NICHD Early Child Care Research Network. 2000a. Characteristics and quality of child care for toddlers and preschoolers.. *Applied Developmental Science*, 4: 116–135.

26. National Institute of Child Health and Human Development Early Child Care Research Network. 2002. Child-care structure, process, outcome: Direct and indirect effects of child-care quality on young children's development.. *Psychological Science*, 13(3): 199–206.
27. Odom, S. L. 2000. Preschool inclusion: What we know and where we go from here.. *Topics in Early Childhood Special Education*, 20: 20–27.
28. Peisner-Feinberg, E. S., Burchinal, M. R., Clifford, R. M., Culkin, M. L., Howes, C., Kagan, S. L. 2001. The relation of preschool child-care quality to children's cognitive and social developmental trajectories through second grade.. *Child Development*, 72: 1534–1553.
29. Phillips, D., Howes, C. and Whitebook, M. 1992. The social policy context of child care: Effects on quality.. *American Journal of Community Psychology*, 20: 25–51.
30. Phillips, D., Mekos, D., Scarr, S., McCartney, K. and Abbott-Shim, M. 2000. Within and beyond the classroom door: Assessing quality in child care centers.. *Early Childhood Research Quarterly*, 15: 475–496.
31. Phillipsen, L. C., Burchinal, M. R., Howes, C. and Cryer, D. 1997. The prediction of process quality from structural features of child care.. *Early Childhood Research Quarterly*, 12: 281–303.
32. Pianta, R., Howes, C., Burchinal, M., Bryant, D., Clifford, R., Early, D. 2005. Features of pre-kindergarten programs, classrooms, and teachers: Do they predict observed classroom quality and child-teacher interactions?. *Applied Developmental Science*, 9(3): 144–159.
33. Pianta, R. C., La Paro, K. M. and Payne, C. 2002. The relation of kindergarten classroom to teacher, family, and school characteristics and child outcomes.. *Elementary School Journal*, 102(3): 225–238.
34. Quay, L. C. 1991. Caregivers' interactions with nonhandicapped and mainstreamed handicapped children in caregiving and instructional activities.. *Early Education and Development*, 2: 261–269.
35. Shonkoff, J. P. and Phillips, D. A. 2000. *From neurons to neighborhoods: The science of early development.*, Washington, DC: National Academy Press..
36. Stipek, D. J. and Sanborn, M. E. 1985. Teachers' task-related interactions with handicapped and nonhandicapped preschool children.. *Merrill-Palmer Quarterly*, 31: 285–300.
37. Tabachnick, B. G. and Fidell, L. S. 1996. *Using multivariate statistics.*, New York: HarperCollins..
38. U.S. Department of Education. 1999. *Programs for young children with disabilities under IDEA: Excerpts from the 21st Annual Report to Congress on the implementation of the*

Individuals with Disabilities Education Act., Chapel Hill, NC: National Early Childhood Technical Assistance System (ERIC Document Reproduction Service No. ED 444274).

39. Vandell, D. L. and Wolfe, B. 2000. *Child care quality: Does it matter and does it need to be improved?*, (Special Report No. 78) Madison: University of Wisconsin—Madison, Institute for Research on Poverty.

40. Whitebook, M., Howes, C. and Phillips, D. 1989. *Who cares? Child care teachers and the quality of care in America.*, Oakland, CA: Child Care Employee Project..

41. Wolery, M. and Bailey, D. B. Jr. 2002. Early childhood special education research.. *Journal of Early Intervention*, 25(2): 88–99.