

## Quality in Inclusive and Noninclusive Infant and Toddler Classrooms

By: [Linda L. Hestenes](#), [Deborah J. Cassidy](#), Archana V. Hedge, Joanna K. Lower

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

Hestenes, L., Cassidy, D., Hegde, A., & Lower, J.K. (2007). Quality in inclusive and non-inclusive infant and toddler classrooms. *Journal of Research in Childhood Education*, 22(1), 69-84. doi: 10.1080/02568540709594613

**as published in the *Journal of Research in Childhood Education* 2007 [copyright Taylor & Francis], available online at: <http://www.tandfonline.com/10.1080/02568540709594613>.**

### **Abstract:**

The quality of care in infant and toddler classrooms as compared across inclusive (n=64) and noninclusive classrooms (n=400). Quality was measured using the Infant/Toddler Environment Rating Scale-Revised (ITERS-R). An exploratory and confirmatory factor analysis revealed four distinct dimensions of quality within the ITERS-R. Inclusive classrooms were higher in quality on the overall scale as well as on three of the four factor-based scales. Teachers reported, on average, that children had mild to moderate disabilities. Correlational analyses indicated that neither having more children with disabilities nor having children with more severe disabilities was associated with higher or lower quality scores. Teacher education and teacher-child ratios were important predictors of quality. Information on low-scoring items on the Personal Care Routines subscale is also presented.

**Keywords:** Education | Early Childhood | Disabilities | Teacher-Child Ratio | Classroom Quality | Inclusive Classrooms | Non-inclusive Classrooms

### **Article:**

Inclusive education continues to be recommended practice for very young children with disabilities as growing numbers of infants and toddlers with disabilities are included in child care classrooms each year. In the reauthorization of the Individuals With Disabilities Education Act (IDEA), now recognized as the Individuals With Disabilities Education Improvement Act (IDEIA), federal regulation requires that early interventions "to the maximum extent appropriate, are provided in natural environments, including the home, and community settings in which children without disabilities participate" (Individuals With Disabilities Education Improvement Act of 2004). The overall percentage of infants and toddlers receiving services under Part C of IDEA increased by 40 percent between 1994 and 2000 (United States Department of Education, 2002), which highlights the fact that increasing numbers of young children with disabilities need specialized care and education in the United States. This rise in percentage indicates two different possibilities: there are either more young children being identified with disabilities who are in need of specialized services, or increasing numbers of families with young children with

disabilities are availing themselves of these services. It is likely that the increase stems from a combination of the two factors.

Not surprisingly, because the most naturalistic setting for infants and toddlers in the United States is usually the home, between 1999 and 2000 most young children (68 percent) received services in their homes; in 2000, 14 percent of infants and toddlers received services in self-contained facilities serving only children with disabilities. It is important to note that this percentage represents half as many children as were served in this type of setting in 1995. That is, states are increasingly recognizing the importance of inclusive classroom environments for infants and toddlers who are enrolled in group settings. A small percentage (9 percent) of infants and toddlers with disabilities were provided with services in settings that also served typically developing children, including family child care homes, nursery schools, or preschools (U.S. Department of Education, 2002). Overall, for children with and without disabilities, there has been an increase in the number of infants and toddlers cared for in group settings. In 2000, nearly 40 percent of children under age 3 were cared for in center-based facilities (Phillips & Adams, 2001). Although there has been relatively little research on children with disabilities in group settings, there have been a number of studies investigating the quality of child care environments for typically developing infants and toddlers and the implications of this quality for young children's development. In general, the research indicates that the quality of infant and toddler classrooms is poor.

For example, the Cost, Quality, and Child Outcomes Study reported that almost one-half of infants and toddlers were in classrooms that were poor quality, while only 14 percent were in classrooms considered to be developmentally appropriate (Helburn, 1995). According to the authors of the Infant/Toddler Environment Rating Scale (ITERS; Harms & Clifford, 1990), the global measure of quality used in the study, poor-quality care is defined as care with problems in basic sanitary conditions, particularly in the area of diapering and feeding, as well as safety-related issues and a lack of positive and supportive relationships with adults. Often, poor-quality programs also lack the materials required for physical and intellectual growth.

Results from the National Institute of Child Health and Human Development (NICHD) Study of Early Child Care indicated that 26 percent of the caregivers were not at all stimulating to the infants in their care. In addition, only 34 percent were highly positive and only 24 percent of infants were cared for by highly sensitive caregivers (NICHD Early Child Care Research Network, 1996). This is of particular concern, since research suggests that the quality of infant care is correlated with measures of cognitive development (Bayley Scales of Infant Development), language development (Sequenced Inventory of Communication Development), and communication skills (Communication and Symbolic Behavior Scales) (Burchinal, Roberts, Nabors, & Bryant, 1996). Higher quality infant care was also found to be associated with higher cognitive, language, and communication skills over time for 89 African American children at 12, 24, and 36 months of age (except for expressive language at 12 months) (Burchinal et al., 2000). Furthermore, a difference of 1 point on the Infant/Toddler Environment Rating Scale (Harms &

Clifford, 1990) resulted in a 6-point difference on the Mental Development Index (MDI), 1.0 months on the measure of receptive communication, 1.34 months on the measure of expressive communication, and 3 points on the Communication and Symbolic Behavior Scales. Children in low-quality infant care also score lower on standardized developmental measures in middle childhood (Vandell & Corasaniti, 1990). Overall, the level of quality in infant and toddler programs has not been strong in the United States, leading to unfortunate consequences for young children's development.

### *Quality of Inclusive Classrooms*

Far fewer studies have examined the quality of care that children with disabilities are receiving. Some research comparing the quality of preschool inclusive and noninclusive classrooms has found inclusive classrooms to be of higher quality (Buysse, Wesley, Bryant, & Gardner, 1999; Hestenes, Cassidy, Shim, & Hegde, 2007). Hestenes and colleagues reported that not only was the overall quality of inclusive preschool classrooms higher but that inclusive preschool classrooms were higher on both an activities/materials factor-based scale and a language/interaction factor-based scale of the Early Childhood Environment Rating Scale-Revised (ECERS-R). Teachers in the inclusive classrooms also had significantly higher levels of education and more coursework in special education. Teachers in inclusive classrooms were rated higher on their interactions with preschoolers, based on scores on the Teacher-Child Interaction Scale (TCIS). Results also indicated that no differences existed in classroom quality based on the level of severity of the children with disabilities who were enrolled (Hestenes et al., 2007). Although the classrooms in this study were generally representative of the higher quality programs in the state, there were, nonetheless, differences between inclusive and noninclusive classrooms. However, LaParo and colleagues report that the same percentage of inclusive and noninclusive classrooms met the criteria for developmentally appropriate practice, with 14 (48 percent) of the self-contained programs scoring 5 or above (developmentally appropriate) and 15 (52 percent) of the inclusive classrooms scoring 5 or above (LaParo, Sexton, & Snyder, 1998).

The Maternal and Child Health (MCH) Bureau Early Child Care Study of Children with Special Needs was modeled after the NICHD Study of Early Child Care and included 166 infants with diagnosed disabilities or who were at risk for developing disabilities (Booth & Kelly, 1999). Overall, the quality of child care, as measured by the Observational Record of the Caregiving Environment (ORCE), was significantly lower for infants with disabilities in child care centers than for infants with disabilities enrolled in child care homes or informal child care. This study revealed no overall difference in the quality of care in centers with and without early intervention services. Therefore, this factor did not contribute to the results regarding the quality of care in these facilities.

The Maternal and Child Health study also reveals that toddlers (30 months old) with disabilities in nonmaternal child care did not differ from children who stayed at home with their mothers on mental, motor, adaptive functioning, behavior issues, or attachment security factors (Booth &

Kelly, 2002). Other research indicates that the benefits of inclusion for toddlers with disabilities include improved social skills, cognitive development, communication, and play skills (Ingersoll, Schreibman, & Stahmer, 2001). In addition, toddlers with disabilities in inclusive settings performed as well as toddlers in self-contained settings one year after enrollment (Bruder & Staff, 1998). In summary, some previous research suggests that inclusive preschool classrooms may be of higher quality than noninclusive classrooms, but for infants and toddlers there is more ambiguity about the quality of inclusive settings, depending on the type of setting. As young children with disabilities are increasingly placed in center-based care with their typically developing age-mates, it is critical that the quality of these placements are understood and appropriate steps are taken to enhance these settings. In order to improve infant and toddler classrooms, professionals need to recognize the underlying dimensions of quality.

### *Measurement of Child Care Quality*

Most researchers who study child care quality utilize measures that fall into three broad categories. The first category includes structural indicators of quality, which contain variables that are most easily regulated by state agencies. Examples include teacher-child ratios, group size, wages, and teacher education. Structural indicators are often thought of as more distal measures of quality that indirectly impact children's outcomes. The second category of quality is process indicators, which include more complex and dynamic aspects of quality, such as teacher-child interaction and the implementation of curriculum. Process indicators are usually thought of as having a more direct impact on children's daily experiences. The third category is global quality, which includes aspects of both structural indicators and process indicators in one measure (Cassidy, Hestenes, Hegde, Hestenes, & Mims, 2005). A number of studies have reported strong relationships between process indicators, structural indicators, and global measures of quality. Phillipson, Burchinal, Howes, and Cryer (1997) examined the associations between structural quality and global quality for 100 child care centers in each of four states (California, Colorado, Connecticut, and North Carolina), using scores on the Infant/Toddler Environment Rating Scale (ITERS) and the Early Childhood Environment Rating Scale (ECERS). The authors found strong associations between rating scale scores (a measure of global quality) and structural quality. Centers with teachers who had at least some college education, lower classroom ratios, and higher pay scored higher in global quality. In a follow-up study of the same child care facilities, Burchinal, Cryer, Clifford, and Howes (2002) examined the associations between caregiver sensitivity and caregiver education in 553 child care classrooms (135 infant/toddler classrooms and 418 preschool classrooms), based on global child care quality scores. Classrooms with teachers who had a bachelor's degree scored significantly higher on measures of classroom quality than classrooms with teachers with less education. Moreover, children in these classrooms of teachers with a bachelor's degree showed significantly better language skills than did children in classrooms with teachers with less education. 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.

It is clear that aspects of structural, process, and global measures of quality are interrelated and collectively influence children's experiences and outcomes. Most of this research, however, has focused on infant, toddler, and preschool classrooms, with no mention of whether or not children with disabilities were included in those classrooms. There is also limited information on whether sub-dimensions of quality exist within global measures of quality, particularly for infant and toddler classrooms. In the Cost, Quality, and Outcome study, a factor analysis of the original Infant/Toddler Environment Rating Scale (Harms & Clifford, 1990) revealed three distinct factors: Interactions, Activities, and Health (Helburn, 1995). To our knowledge, there has not been a factor analysis of the Infant Toddler Environment Rating Scale-Revised, and therefore it is unclear whether additional dimensions of quality may be uncovered that would bring to light important aspects of quality that are unique to this younger age group of children.

In summary, although clear evidence exists of the value of high-quality environments for young children, the field lacks conclusive information on the quality of inclusive settings for infants and toddlers. The large sample available in the current study allowed us to explore whether there were unique dimensions of quality measured by the Infant/Toddler Environment Rating Scale-Revised (ITERS-R; Harms, Cryer, & Clifford, 2003). Thus, the first goal of this project was to determine whether distinct aspects of quality do exist within the ITERS-R by running both exploratory and confirmatory factor analyses. The second goal of this project was to investigate the quality of inclusive and noninclusive infant/toddler classrooms. This part of the study addressed three research questions: 1) Do inclusive and noninclusive classrooms differ in overall quality and on the factors delineated in the factor analyses?, 2) What variables predict higher quality in inclusive and noninclusive infant and toddler classrooms?, 3) What are the items on the ITERS-R that are consistently low-scoring and may need particular attention as we strive to improve the quality of infant and toddler classrooms for children with and without disabilities?

### *Context for the Current Study*

The current study utilizes a large sample of infant and toddler classrooms generated by the North Carolina Rated License Assessment Project. A recent trend across the United States has been the inclusion of a global measure of quality, such as the ITERS-R, as a component of the regulatory or licensing system. In North Carolina, the rating scales are used to determine a point total for program standards. Typically, only programs striving for a higher point total (and a higher star rating) undergo the assessment process resulting in a higher quality sample. Nonetheless, the large sample allows for statistical procedures that are not possible with smaller samples. The sample has a normal distribution and is representative of the state with regard to rural and urban and small and large programs.

## **Method**

### *Participants*

The Infant/Toddler Environment Rating Scale-Revised (ITERS-R; Harms et al., 2003) was completed in 466 classrooms across 82 counties from August 2003 to October 2004. The children in these classrooms ranged in age from 1 month to 42 months. On average, the youngest child in each classroom was 12 months and the oldest child was 21 months. The classrooms assessed for this project come from a variety of programs, including for-profit child care program, nonprofit child care programs, Early Head Start centers, early intervention programs, and church-sponsored programs. Highly trained assessors completed the ITERS-R in each classroom as part of North Carolina's Star Rated License process. During this time frame, child care facilities in North Carolina could earn from one to five stars, depending on the outcomes from three domains (i.e., program standards, staff education levels, and compliance history). One portion of the program standards requirement for centers that wanted to achieve a higher star rating was the completion of a global quality assessment using the Environment Rating Scales (i.e., ECERS-R, ITERS-R, FDCRS, or SACERS). Assessments were completed in one-third of the classrooms, including one classroom at each age level (infant/toddler, preschool, and school-age). Assessments were completed only in programs that were striving for the higher star ratings (typically 4 or 5 stars). Thus, this data set represents only the higher quality programs in the state.

Of the 466 classrooms, 64 were inclusive (13.6 percent), 2 contained only children with disabilities (.4 percent), and 400 contained only typically developing children (86 percent). Classrooms were considered to be inclusive if they enrolled at least one child with an identified disability and at least one child who was typically developing. Children were classified as having a disability if they had a diagnosis and were eligible for services. Across the inclusive classrooms, there were 125 children with identified disabilities, representing a wide range of diagnoses. The most frequent diagnoses were developmental delay (n=38), physical disability (n=22), speech delay or disorder (n=15), at-risk (n=12), or Down syndrome (n=11). Teachers' perceptions of the severity of children's disabilities were rated from 1 (mild) to 3 (severe), with an average rating of 1.80 across the inclusive classrooms. The number of children with disabilities in the inclusive classrooms ranged from 1 to 7, with an average of 2.0.

### *ITERS-R*

The Infant/Toddler Environment Rating Scale-Revised (Harms et al., 2003) is a 39-item observational instrument. Each item is rated from 1 (inadequate) to 7 (excellent) based on indicators, which are descriptions of quality listed below the 1, 3, 5, and 7 ratings. The scale contains 7 subscales: Space and Furnishings (5 items), Personal Care Routines (6 items), Listening and Talking (3 items), Activities (10 items), Interaction (4 items), Program Structure (4

items), and Parents and Staff (7 items). (1) Subscale scores are created by averaging across all of the items within a subscale, and the overall score is created by taking an average of all the items.

### *Procedures*

Highly trained assessors completed the ITERS-R in each classroom during a 3- to 4-hour observation session. At the end of each observation, a teacher interview was used to clarify demographic information and to complete the items that could not be observed. Interviews lasted approximately 30 minutes. In addition to completing the ITERS-R, assessors also collected background information on teachers, group size, and teacher/child ratios during their observation. Teachers' level of education was coded into one of 12 categories, from "did not complete high school" to a "graduate degree." At each level coded for teachers' education, increasing years of education was accounted for, as was each teacher's amount of child development/early childhood education.

### *Assessor Training*

Each assessor was a trained staff member of the North Carolina Rated License Assessment Project. The assessors receive extensive and ongoing training on the instrument as part of their job. Each assessor is trained to at least an 85 percent agreement level (based upon consensus scoring within one rating point) in programs that differ by level of quality, ethnicity, age, special needs, and program type. Reliability is maintained at that level via checks after approximately every sixth assessment. Highly reliable assessors (i.e., those maintaining a 90 percent agreement level over three consecutive reliability checks) are re-evaluated for reliability after every tenth assessment. Each assessor receives updated training and clarification on items every quarter and feedback at each reliability check. (2)

## **Results**

### *Factor Analyses of the ITERS-R*

The first set of analyses run for this study included a series of factor analyses on the ITERS-R. The goal of these analyses was to determine if there were unique aspects or dimensions of quality being measured within this global scale. Once unique dimensions were identified, it allowed us to more precisely determine how inclusive and noninclusive classrooms differed in quality. Since the sample size was large, we were able to run both exploratory and confirmatory factor analyses on the ITERS-R. A random selection procedure in SAS[R] was utilized to split the file into two data sets. One portion of these data (n=238) was used for exploratory factor analyses and the other half (n=228) was used for confirmatory factor analysis. In each of the analyses, items from the original sample with more than 10 percent missing and items with highly skewed distributions ( $> \pm 2.0$ ) were removed. This criterion led to the removal of 9 items. Six of the items were dropped due to missing values (Items: 17. Art; 19. Blocks; 21. Sand and water play; 23. Use of TV, video, and/or computer; 31. Group play activities; and 32.

Provisions for children with disabilities). These items contained missing values due to the indicators being scored "NA" if they did not pertain to infants or to toddlers. Therefore, many indicators for the items do not apply if the assessment is being conducted in an infant classroom. Three additional items were dropped due to high skewness values (Items 6. Greeting/departing; 7. Meals/snacks; and 27. Staff-child interaction). Comparisons of the two data sets revealed they did not differ by teacher education, teacher experience, ITERS-R subscale scores, or overall score.

**Table 2.** Measures of Fit for the 7-Factor, 4-Factor, and 3-Factor Models

<b>Measures of Fit Information</b>	<b>7-Factor Model</b>	<b>4-Factor Model</b>	<b>3-Factor Model</b>
Sample Size	97	205	228
Chi-Square Value	947.0655	511.8264	393.7655
Probability for Chi-Square	<.0001	<.0001	<.0001
Df	608	269	167
Ratio of Chi-Square/df	1.56	1.90	2.36
Goodness of Fit Index (GFI)	.68	.83	.85
Comparative Fit Index (CFI)	.64	.79	.79

*Exploratory Factor Analysis.* A principal component extraction method with varimax rotation was employed to explore the various ITERS-R components. Examination of the scree plot, eigenvalues more than 1, correlation matrices, and item loadings (greater than .40), resulted in a 3- or a 4-component solution. With varimax rotation, the 4-component solution produced three components with moderate to high internal consistency ( $[\alpha] = .83, .70, \text{ and } .75$ ) and one component with lower internal consistency ( $[\alpha] = .61$ ). However, the 4-component solution with varimax rotation revealed a conceptually clearer categorization of the items than did the 3-component solution. The first factor-based scale (Materials/Activities) contained 9 items and accounted for 15 percent of the variance, and the second factor-based scale (Safety/Organization) contained 7 items and accounted for an additional 10 percent of the variance. The third factor-based scale (Language/Interactions) contained 5 items and accounted for an additional 10 percent of the variance, while the fourth factor-based scale (Parents/Staff) contained 6 items and accounted for 8 percent of the variance. Overall, the four components in totality explained 43 percent of the variance (see Table 1 for item loadings). In addition, a principal factor extraction method with varimax rotation also was used to explore the factor solution. These analyses led to a similar pattern of factor-based scales with the exception of item 1 (Indoor space), which loaded on factor four in the prior analysis (principal component analysis with varimax rotation) but did

not load on any of the factors in the subsequent analysis. Thus, this item was deleted from the fourth factor-based scale in the confirmatory analysis.

**Table 1.** Item Loadings for the Four ITERS-R Factors in Exploratory Analyses

<b>ITERS-R Item Number and Name</b>	<b>Materials/Activities</b>	<b>Safety/Organization</b>	<b>Language/Interactions</b>	<b>Parents/Staff</b>
3. Provision for relaxation & comfort	.67	.08	.39	-.04
14. Using books	.69	.09	.16	.04
15. Fine motor	.73	.13	.19	-.12
16. Active physical play	.60	.17	.04	.38
18. Music/movement	.48	.34	.24	-.07
20. Dramatic play	.79	.00	.05	-.08
22. Nature/Science	.61	.11	.04	.15
24. Promoting acceptance of diversity	.62	-.11	.10	.11
30. Free play	.65	.13	-.05	.37
2. Furniture for routine care, play, and learning	.08	.43	.16	.24
4. Room arrangement	.04	.73	-.12	-.03
9. Diapering/Toileting	.12	.51	.25	.22
10. Health practices	.07	.61	.07	.06
11. Safety practices	.01	.69	.12	.00
25. Supervision of play and learning	-.05	.64	.26	.02
29. Schedule	.34	.44	.11	.31
12. Helping children understand language	.17	.15	.61	.06
13. Helping children use language	.23	.19	.68	.08
26. Peer interaction	.23	.41	.67	-.02
28. Discipline	.19	-.03	.60	.12
38. Supervision and evaluation of staff	-.06	.05	.45	.38
1. Indoor space	.08	.21	.11	.42
33. Provisions for parents	.12	-.07	.25	.51
34. Provisions for personal needs of staff	.07	.12	-.11	.61
35. Provisions for professional needs of the staff	-.02	.15	.16	.48
36. Staff interaction and cooperation	.22	-.01	.32	.43
39. Opportunities for professional growth	-.02	-.07	.25	.56

*Confirmatory Factor Analysis: 7-Factor, 4-Factor, and 3-Factor Model Comparison.* To evaluate the original 7-factor model (i.e., the 7 original subscales) and the 4- and 3-factor models found in the exploratory analyses, three confirmatory factor analyses using PROC CALIS in SAS[R] were conducted. Structural equation modeling techniques allowed us to confirm which model was the best fit for the data. These analyses were run on the second random sample, which contained 228 classrooms. Various measures of fit for each model are listed in Table 2. The [chi square] value for each model was statistically significant. The ratio of [chi square] to degrees of freedom is slightly above the recommended cutoff of 2.0 for the 3-factor model. The goodness of fit indices were far superior for the 3- and 4-factor models compared to the 7-factor model.

Each of the factor-based scales and combined 3- and 4-factor-based scales were significantly correlated with the ITERS-R overall mean (see Table 3). However, the correlations across the four factor-based scales revealed an interesting pattern. Factor 1 correlated the least with all the other factor-based scales. The correlations ranged between ( $r = .22$  to  $r = .27$ ), while factor 2 and factor 3 had the strongest intercorrelations ( $r = .42$ ).

**Table 3.** Correlations Between New Factor-based Scales, Total Score, and Regulatable Variables

<b>Regulatable Variables</b>	<b>Materials/ Activities Factor</b>	<b>Safety/ Organiz ation</b>	<b>Language/In teraction Factor</b>	<b>Parents/Staff Factor</b>	<b>ITERS-R Mean</b>
Materials/Activities	--	.22	.37	.26	.76
Safety/Organization	--	--	.42	.26	.66
Language/Interaction	--	--	--	.34	.68
Parents/Staff	--	--	--	--	.54
Combined Factor #1, #2, and #3 .95	-	-	-	-	-
Combined Factor #1, #2, #3, and #4 .97	-	-	-	-	-
Teacher Education (12 levels)	.17	.13	.26	.27	.26
Years Experience in Early Childhood	.02	.14	.02	.04	.07
Group Size	.09	-.19	-.20	.13	-.08
Teacher-Child Ratio	.02	-.33	-.27	-.08	-.23

After examining the goodness of fit indices, the correlations, and the individual items, we concluded that the best representation of unique aspects of quality in the ITERS-R was the 4-factor solution. Each of the 4 factor-based scales seemed to represent distinct aspects of quality that held together both conceptually and statistically.

#### *Comparisons Across Inclusive and Noninclusive Classrooms*

The second set of analyses was designed to make comparisons across inclusive and noninclusive classrooms. The two classrooms with only children with disabilities (i.e., self-contained classrooms) were removed from these analyses. Preliminary analyses were conducted to determine whether the variables used in the comparisons across classrooms were normally distributed. Average ITERS-R scores ranged from 3.00 to 6.39, with an overall average score of 4.94 and a standard deviation of .67. All variables had normal distributions, with the exception of the number of course hours in special education. This variable was highly skewed due to the low numbers of credit hours that teachers had received. Only descriptive information is presented on this variable.

Results of a One-way Analysis of Variance (ANOVA) showed that inclusive infant or toddler classrooms (n = 64) were higher in overall quality than were classrooms with only typically developing children (n = 400) ( $F(1, 462) = 15.55, p < .000$ ). Inclusive classrooms had a mean score on the ITERS-R of 5.24, while noninclusive classrooms averaged a 4.89 rating. Using the factor-based scales, the results showed that inclusive classrooms were significantly higher in quality on three of the 4 factor-based scales (see Table 4). There were no differences across classrooms on the activities/materials factor. Inclusive classrooms were also significantly higher on four of the subscale scores of the ITERS-R compared to noninclusive classrooms. There were no differences in subscale means for Personal Care Routines, Activities, and Interactions (see Table 4).

**Table 4.** Factor-based Scale Means and Subscale Score Means for Inclusive and Noninclusive Classrooms

<b>Factor-based Scales</b>	<b>Inclusive</b>	<b>Noninclusive</b>
Materials/Activities	5.31	4.93
Safety/Organization	4.19	3.82 *
Language/Interactions	6.19	5.78 *
Parents/Staff	5.84	5.50 *
<b>Subscales</b>		
Space and Furnishings	5.20	4.83 *
Personal Care Routines	3.74	3.48
Listening and Talking	6.23	5.73 *
Activities	5.23	4.89
Interaction	5.78	5.42
Program Structure	5.18	4.57 *
Parents and Staff	5.86	5.54

\* $p < .01$

Teacher education was coded into 12 categories that represented education levels ranging from some high school through a graduate degree. Teachers from inclusive classrooms had significantly more education than teachers from noninclusive classrooms ( $F(1,461) = 21.7, p < .000$ ). Teachers from inclusive classrooms averaged a 2-year degree, while teachers from

noninclusive classrooms had, on average, only received some college coursework. Surprisingly, teachers in inclusive and noninclusive classrooms took very few courses in the field of special education. Only 33 percent of all the teachers had any special education coursework; of those teachers, more than half (57 percent) had taken only one course. Ninety-four percent of teachers had taken four or fewer courses in special education. To ensure that the differences in quality between inclusive and noninclusive classrooms were not a function of higher levels of teacher education in inclusive settings, an analysis of covariance (ANCOVA) was conducted with education as a control variable. Results show that with the effects of education removed, inclusive classrooms were still higher in quality than noninclusive classrooms ( $F(2, 460) = 8.16$ ,  $p = .004$ ).

Comparisons between inclusive and noninclusive classrooms revealed that teachers did not differ significantly in their years of experience working in the field of early childhood education ( $F(1, 457) = 1.82$ ,  $p = .18$ ), or in the number of months they had been teaching the current group of children ( $F(1, 443) = 2.8$ ,  $p = .10$ ). Teacher-child ratios were collected during the second hour of the observation. Inclusive and noninclusive classrooms did not differ in their teacher-child ratios ( $F(1, 460) = 1.69$ ,  $p = .19$ ; mean inclusive classrooms = 2.69, mean noninclusive classrooms = 3.09).

#### *Severity of Disability*

To determine whether severity of disability was related to overall quality scores, a weighted score was produced for each classroom by adding together the teacher's ratings of the severity of each child's disability (1 = mild, 2 = moderate, 3 = severe). For example, if four children had disabilities in a classroom and the teacher rated two children as mild, one as moderate, and one as severe, the classroom received a severity score of seven ( $1+1+2+3 = 7$ ). Severity scores ranged from 1 to 13, with an average of 3.64. Bivariate correlations between overall quality scores in inclusive classrooms and the weighted severity scores were not significant ( $r = .09$ ,  $p = .51$ ). There also was no relationship between number of children with disabilities in the classroom and overall quality in inclusive classrooms ( $r = .02$ ,  $p = .87$ ). Therefore, neither having more children with disabilities nor having children with more severe disabilities was associated with higher or lower quality scores.

#### *Predictors of Quality*

In order to determine the best predictors of quality in this sample, multiple regression was utilized. Separate regressions were run for the total sample and for inclusive and noninclusive classrooms. Since the results were similar for the total sample and the noninclusive classrooms, only the inclusive and noninclusive analyses will be presented. Education, experience, teacher-child ratios, special education coursework, and the number of months with the current group of children were hypothesized to be potentially important predictors of quality, based upon the literature. Each of these variables was entered stepwise into the analysis for the noninclusive

sample, with average score on the ITERS-R as the dependent variable. The results showed that two variables were significant predictors in the model for noninclusive classrooms: Teacher education and teacher-child ratio (see Table 5). Due to the small sample size, only two predictors (based on the strongest correlations) were entered for the second regression on inclusive classrooms. These variables included teacher education and teacher-child ratio. For inclusive classrooms, only teacher-child ratio was a significant predictor (see Table 6).

Table 5. Prediction of Total Quality Score for Noninclusive Classrooms

Predictors	R <sup>2</sup>	Change in R <sup>2</sup>	F value	β
Teacher Education	.07	-	27.6*	.25*
Teacher-Child Ratio	.10	.03*	12.7*	-.18*

\*p < .000

Table 6. Prediction of Total Quality Score for Inclusive Classrooms

Predictor	R <sup>2</sup>	F value	β
Teacher-Child Ratio	.19	13.94*	-.43*

\*p < .000

#### *Descriptive Information on Low-Scoring Items*

Since the ITERS-R is being used for regulatory purposes in a number of states (e.g., North Carolina, Tennessee), there is growing interest for many providers of infant and toddler care about the items on the scale that receive the lowest scores. An examination of the mean scores for the items indicates that there are 7 items with averages below a 4.0 (considered less than "good" on the measure). Five of the lowest scoring items are from the Personal Care Routines subscale, which raises concern about whether or not the basic health and safety needs of infants and toddlers are being met. Furthermore, on three ITERS-R items (11. Safety practices; 7. Meals/snack; and 9. Diapering/toileting), scores for the noninclusive classrooms were below the "minimum" level (a score of 3) of quality.

#### **Discussion**

The goals of this study included discerning whether sub-dimensions of quality could be extracted from the ITERS-R (a global measure of quality) and exploring differences in quality for inclusive and noninclusive infant/toddler classrooms. These differences in quality were addressed in three research questions: 1) Do inclusive and noninclusive classrooms differ in overall quality and on the factors delineated in the factor analyses?, 2) What variables predict higher quality in inclusive and noninclusive infant and toddler classrooms?, and 3) What are the items on the ITERS-R that are consistently low-scoring and may need particular attention as we

strive to improve the quality of infant and toddler classrooms for children with and without disabilities? Each of these goals and research questions will be discussed in turn.

Four distinct factors for the ITERS-R were confirmed using structural equation modeling techniques (Materials/Activities, Language/Interactions, Safety/Organization, and Parents/Staff). Two of the factors--Materials/Activities and Language/Interactions--are similar to those reported in earlier studies (Helburn, 1995; Whitebook, Howes, & Phillips, 1989) on the ITERS. At least one study has also identified a third factor of Basic Caregiving Routines (Tietze & Cryer, 2004) for the ITERS. Perhaps the most surprising finding of the factor analysis was the fourth factor of Parents/Staff. This factor has not been found in factor analyses previously conducted on the ECERS, the ECERS-R, or the ITERS. Four of the seven items on the Parents and Staff subscale loaded on the Parents/Staff factor. The two factors (Safety/Organization & Parents/Staff) that are distinct from a recent confirmatory factor analysis on the ECERS-R (Cassidy et al., 2005) may suggest the greater role that safety/organization plays in infant and toddler classrooms as well as the distinct importance of provisions for parents and staff. Indeed, the two items relating specifically to parent and staff provisions were included in this factor. The importance of the parent role in infant and toddler classrooms has long been surmised, and so has the critical nature of caregivers in developing a secure attachment. This factor may at last give credence to the role of both teachers and parents in high-quality programs.

In addition to discerning whether sub-dimensions of quality existed in the ITERS-R, this study attempted to answer several questions regarding the level of quality in infant and toddler classrooms that include children with disabilities. In response to the first research question of the study, it is encouraging that infant and toddler classrooms that include children with diagnosed disabilities were significantly higher in quality than classrooms that did not include children with disabilities. This finding is in contrast to the study conducted by the Maternal and Child Health Bureau's Early Child Care Study of Children With Special Needs (Booth & Kelly, 1999), which reported no difference in quality between infant/toddler classrooms that provide early intervention and those that do not. In addition, it contrasts with the study by LaParo and colleagues (1998) indicating that inclusive and noninclusive classrooms met an equal number of indicators of best practice. Previous research suggested that preschool classrooms for children with disabilities were of higher quality than noninclusive classrooms (Buysse et al., 1999; Hestenes et al., 2007); to our knowledge, however, this finding has not been reported for infant and toddler classrooms with such a large data set. However, this finding does not indicate a causal relationship between enrollment of children with disabilities and higher quality. It is more likely that programs that are higher in quality may be more aware of the importance of inclusion and actively seek to enroll children with disabilities, or perhaps parents of children with disabilities seek out higher quality programs. Most importantly, the enrollment of children with disabilities did not diminish the overall classroom quality below the level of what is considered to be developmentally appropriate (a score of 5 on the 7-point scale). Furthermore, severity of disability, as perceived by the classroom teacher, was not associated with level of quality in

infant and toddler classrooms. Although the severity ratings provided by teachers fell between "mild" and "moderate," it is again encouraging that enrollment of children with more severe ratings did not result in lower quality scores. It is important to note that it is likely that the infants and toddlers with the most severe disabilities are more likely to be cared for in informal settings such as the home (Booth & Kelly, 1999). As an increasing number of infants and toddlers with more severe disabilities are enrolled in child care settings in the United States, it will be critical to confirm this finding. It also would be advantageous in future studies to measure severity level in children using a standardized diagnostic instrument as opposed to using teacher ratings. Although teachers' perceptions of the severity of disability likely have strong implications for classroom practices, a more consistent measure would be beneficial.

Another interesting result was that inclusive and noninclusive classrooms differed on all the factor-based scales, except for the Activities/Materials factor. This result suggests that all the classrooms were comparable in their provisions for classroom materials and the types of activities in which the infants and toddlers were able to engage, but what differentiated classrooms were the interactions, relationships, organization, and health/safety practices that took place within each setting. That is, the most distinguishable features between inclusive and noninclusive classrooms in this sample were not the activities and materials that are available but rather the dynamic interactions, organization of care, and the health and safety practices. It would be important to replicate this result with lower quality programs to ensure the generalizability of this finding.

Interestingly, the level of teacher education in inclusive classrooms was significantly higher than in noninclusive classrooms but did not alone account for the differences in quality between the two types of classrooms. Teachers in inclusive classrooms averaged a two-year college degree, while teachers in noninclusive classrooms had completed only "some" college coursework.

With regard to the second research question, an important, but not surprising, finding from this study is that both teacher education and teacher-child ratios are predictors of quality in noninclusive infant and toddler classrooms, while only teacher-child ratios predicted quality in the inclusive classrooms. It may be that the higher levels of education in the inclusive classrooms and the reduction in variability masked the impact of education in these classrooms. These findings are similar to previous research on the structural predictors of quality (Helburn, 1995; Whitebook et al., 1989). It is notable that there were no differences in teacher-child ratios between the classrooms that included children with disabilities and those that did not. However, the ratios in both types of classrooms were quite low (1:2.69--inclusive; 1:3.09--noninclusive), which seemed to meet the needs in both groups, since the level of quality in these classrooms was at a developmentally appropriate level (5.0). Nonetheless, ratios were an important predictor of quality for both classrooms. With increasing concern about the overall quality of infant and toddler classrooms, these findings highlight our continued need to educate early childhood teachers and to lower teacher-child ratios.

The final research question revealed that the Personal Care Routines subscale had the lowest average scores of the seven subscales for both the inclusive and noninclusive classrooms. The Personal Care Routines subscale (which includes such items as health, safety, and sanitation) means were noticeably lower (3.74 for inclusive classrooms; 3.48 for noninclusive classrooms) than scores on the other subscales. All classrooms were closer to minimal quality (a score of 3) than good quality (a score of 5). Significantly higher levels of education in inclusive classrooms did not result in developmentally appropriate health and safety practices in those classrooms. Maintaining appropriate procedures, particularly for hand washing, diaper-changing, and overall sanitation, appears to be difficult for many teachers in infant and toddler classrooms. These findings certainly indicate that children with disabilities are enrolled in higher quality programs, but that health and safety practices in ALL classrooms need significant attention. Increasing training sessions for teachers to improve their personal care routines would be a worthwhile endeavor.

The findings in this study also indicate that inclusive placements are still severely limited for infants and toddlers. In the current study, only about 14 percent of the 466 classrooms included infants and toddlers with disabilities. This finding is consistent with survey data collected by McDonnell, Brownell, and Wolery (1997) showing the lack of placements in high-quality facilities. In a study by Kochanek and Buka (1999), two-thirds of the exemplary early intervention programs were self-contained. We cannot attest to the extent to which children with disabilities were integrated into these classrooms. The level of quality that occurred during interactions--specifically between teachers and infants/toddlers with disabilities--is unclear, since we used a global measure of quality that did not measure individual interactions. We are unclear about the ability of teachers to implement high-quality inclusive practices, since the teachers in the inclusive classrooms had very little coursework in special education, as did their counterparts in noninclusive classrooms. Unlike the teachers in this study, teachers in inclusive classrooms in a study by Hestenes and colleagues (2007) had significantly more coursework in special education than did teachers in noninclusive classrooms. It would be interesting to conduct research on how the classrooms including children with disabilities differed with regard to teacher behaviors. Do teachers with more special education coursework interact with children in a manner that encourages involvement and acceptance of children with disabilities? It also would be important to examine the relationship between teacher-child ratios and appropriate engagement with children for teachers who have more education. It seems that continuing to educate the child care staff regarding the importance of inclusive environments, appropriate interactions with children with and without disabilities, and knowledge of best practice would increase the number of children with disabilities served in high-quality inclusive environments.

## **Conclusions**

It is encouraging to note that within this large sample, inclusive infant and toddler classrooms are higher in quality than noninclusive classrooms. The four factors of quality confirmed in this study suggest that differences are not due to the activities or materials in the classroom, but

instead are based on differences in language/interactions, safety/organization, and parents/staff variables. It is also clear that teacher education and staff/child ratios are key predictors of higher quality programs. Although this study is based on a large sample of classrooms, it must be remembered that these programs only represent the higher quality programs within North Carolina. Thus, this sample limits the generalizability of the results to other states and to samples of lower quality programs. Future research on lower quality infant and toddler classrooms as well as in programs that contain more children with severe disabilities would strengthen the results of this work. In addition, future work needs to continue to track the impact that the specific dimensions of quality (such as health and safety practices) have on the growth and development of infants and toddlers in child care programs. It is critical that we continue to monitor and understand the quality of the experiences that our youngest children--with and without disabilities--are having on a daily basis in child care settings.

## **References**

- Booth, C. L., & Kelly, J. F. (1999). Child care and employment in relation to infants' disabilities and risk factors. *American Journal on Mental Retardation*, 104, 117-130.
- Booth, C. L., & Kelly, J. F. (2002). Child care effects on the development of toddlers with special needs. *Early Childhood Research Quarterly*, 17, 171-196.
- Bruder, M. B., & Staff, I. (1998). A comparison of the effects of type of classroom and service characteristics on toddlers with disabilities. *Topics in Early Childhood Special Education*, 18, 26-37.
- Burchinal, M. R., Cryer, D., Clifford, R. M., & Howes, C. (2002). Caregiver training and classroom quality in child care centers. *Applied Developmental Science*, 6, 2-11.
- Burchinal, M. R., Roberts, J. E., Nabors, L. A., & Bryant, D. M. (1996). Quality of center child care and infant cognitive and language development. *Child Development*, 67, 606-620.
- Burchinal, M. R., Roberts, J. E., Riggins, R., Zeisel, S. A., Neebe, E., & Bryant, D. (2000). Relating quality of center-based child care to early cognitive and language development longitudinally. *Child Development*, 71, 339-357.
- Buysse, V., Wesley, P. W., Bryant, D., & Gardner, D. (1999). Quality of early childhood programs in inclusive and noninclusive settings. *Exceptional Children*, 65, 301-314.
- Cassidy, D. J., Hestenes, L. L., Hegde, A., Hestenes, S., & 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(3), 345-360.

- Harms, T., & Clifford, R. (1990). *Infant/toddler environment rating scale*. New York: Teachers College Press.
- Harms, T., Cryer, D., & Clifford, R. (2003). *Infant/toddler environment rating scale-revised*. New York: Teachers College Press.
- Helburn, S.W. (Ed.). (1995). *Cost, quality and child outcomes in child care centers: Technical report*. Denver, CO: Department of Economics, Center for Research in Economic and Social Policy, University of Colorado at Denver.
- Hestenes, L., Cassidy, D., Shim, J., & Hegde, A. (2007). *Global quality and teacher-child interactions in inclusive and noninclusive child care*. Manuscript submitted for publication.
- Individuals With Disabilities Education Improvement Act of 2004, Pub. L. No. 108-446, 118 Stat. 2647, [section] 632 (2004).
- Ingersoll, B., Schreibman, L., & Stahmer, A. (2001). Differential treatment outcomes for children with autistic spectrum disorder based on level of peer social avoidance. *Journal of Autism and Developmental Disorders*, 31, 343-349.
- Kochanek, T. T., & Buka, S. L. (1999). Influential factors in inclusive versus noninclusive placements for preschool children with disabilities. *Early Education and Development*, 10, 191-208.
- LaParo, K. M., Sexton, D., & Snyder, P. (1998). Program quality characteristics in segregated and inclusive early childhood settings. *Early Childhood Research Quarterly*, 13, 151-167.
- McDonnell, A. P., Brownell, K., & Wolery, M. (1997). Teaching experience and specialist support: A survey of preschool teachers employed in programs accredited by NAEYC. *Topics in Early Childhood Special Education*, 17(3), 263-285.
- National Institute of Child Health and Human Development Early Child Care Research Network. (1996). Characteristics of infant child care: Factors contributing to positive caregiving. *Early Childhood Research Quarterly*, 11, 269-306.
- Phillips, D., & Adams, G. (2001). Child care and our youngest children. *The Future of Children: Caring for Infants and Toddlers*, 11, 35-51.
- Phillips, D., Mekos, D., Scarr, S., McCartney, K., & Abbott-Shim, M. (2000). Within and beyond the classroom door: Assessing quality in child care centers. *Early Childhood Research Quarterly*, 15, 475-496.
- Phillipsen, L. C., Burchinal, M. R., Howes, C., & Cryer, D. (1997). The prediction of process quality from structural features of child care. *Early Childhood Research Quarterly*, 12, 281-303.

Tietze, W., & Cryer, D. (2004). Comparisons of observed process quality in German and American infant/toddler programs. *International Journal of Early Years Education*, 12(1), 43-62.

U.S. Department of Education. (2002). Twenty-fourth annual report to Congress on the implementation of the Individuals with Disabilities Education Act. Retrieved July 2, 2005, from [www.ed.gov/about/reports/annual/osep/2002/index.html](http://www.ed.gov/about/reports/annual/osep/2002/index.html)

Vandell, D. L., & Corasaniti, M. A. (1990). Child care and the family: Contributors to child development. *New Directions for Child Development*, 49, 23-47.

Whitebook, M., Howes, C., & Phillips, D. (1989). *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.