

Is Head Start Providing High-Quality Educational Services? “Unpacking” Classroom Processes

Gary Resnick, Ph.D., Nicholas Zill, Ph.D.
Westat, Inc.

I. Research Questions: Head Start Classrooms as Child Development Environments

1. How good is the average quality of Head Start classroom environments, as judged by trained observers using well-established early childhood rating scales? What proportion of Head Start classrooms are of poor quality?
2. How does the quality of Head Start classroom environments compare with other preschool and early childhood education programs?
3. How much variation is there in the quality of Head Start classrooms between Head Start programs? Across centers within programs? Across classrooms within centers?
4. How much variation in the quality of Head Start classrooms is accounted for by the larger characteristics of programs, such as the region of the country in which they are located and whether they are rural or urban programs?
5. How much variation in the quality of Head Start classrooms is accounted for by the socio-economic characteristics of the communities in which the programs are located? (e.g. racial and ethnic composition of participating families, average education level of parents, average income level of parents)
6. How much variation in the quality of Head Start classrooms is due to characteristics of the teachers (e.g. teacher education, years of experience, Child Development Associate diploma, membership in early education associations)?

II. What is Classroom Quality in Head Start?

1. Structure
 - Adult-Child Ratio
 - Class Size
2. Teacher Background and experience
 - Years of teaching Head Start
 - Highest education level achieved
 - Child Development Associate (CDA) or similar certification
 - Membership in an Early Education association
3. Classroom Process
 - Availability and Accessibility of Learning Materials
 - Arrangement of Classroom Space to Encourage Child Independence and Reflects the Child as an Individual
 - Planned Scheduling of Activities, Variety, Timing of Activities
 - Physical arrangement, facilities, equipment, furnishings, personal care routines

- Quality of curriculum in language, perceptual/fine and gross motor, creative and social domains, encouragement of cultural awareness, provisions for exceptional children
- Quality of Teacher-Child Interactions, Supervision of Activities, Tone in the Classroom
- Provisions for Parents and Staff, Opportunities for Staff Professional Growth
- Teacher behavior in the classroom: sensitivity, responsiveness, harshness, detachment, encouragement of independence/leadership

III. Measures of Quality

1. Structure

- *Adult-Child Ratio and Class Size*: observers counted the number of children in the classroom and the total number of adults involved in classroom activities (including volunteers) on 2 separate occasions during the Head Start day.
- *Teacher Background and Experience*: Self-report questionnaire given to the lead teacher in the classroom at the time of the observations.

2. Classroom Process

- *The Assessment Profile Scheduling scale*. This scale assesses the written plans for classroom scheduling and how classroom activities are implemented.
- *The Assessment Profile Learning Environment scale*. This scale measures the variety of learning materials available in the classroom that provide learning experiences in small muscle/manipulatives, self-help, art, drama/role play, science, math, language, nutrition/health, and diversity.
- *The Early Childhood Environment Rating Scale (ECERS)*. This measure consists of 37 scales measuring a wide variety of quality related processes occurring in the classroom, including routines, teacher-child interaction particularly in the use of language, learning activities, classroom tone, creative, dramatic, and gross and fine motor activities, equipment and furnishings, and staff and parent facilities. The ECERS items were rated on a seven-point scale, with the following anchors: (1) inadequate, (3) minimal, (5) good, and (7) excellent. An overall quality rating was then obtained by averaging the scores across all items.
- *The Arnett Scale of Caregiver Behavior*. This is a rating scale of teacher behavior towards the children in the class. It consists of 26 items that assesses the teacher's sensitivity, punitiveness, detachment, permissiveness, and the teacher's encouragement of child self-help. Separate ratings were done for the Lead and Assistant Teachers.

3. Classroom Observation Procedures

- FACES measured a variety of dimensions of program quality using reliable, well-known measures that are designed to be employed by specially trained classroom observers.
- The classroom observers spent an entire "Head Start day" in the classroom and, using standard measures, assessed various aspects of the classroom that were known indicators of quality.
- The classroom quality data were collected in the Fall of 1997 and again in the Spring of 1998 for classrooms containing target study children.

4. Inter-Observer Reliability

- Inter-rater reliability was assessed using "Quality Control" visitors trained in the observation methods, to conduct parallel, independent observations for all the quality measures.
- One classroom in 39 of the 40 programs was checked for inter-rater reliability.
- Percent Agreement (Fall 1997):
 - Assessment Profile Learning Environment & Scheduling: 91%
 - Early Childhood Environment Rating Scale: 51% Direct Hits, 86% Off by One

5. Internal Consistency of Quality Measures (Fall 1997)

- Assessment Profile Learning Environment: Alpha = .62
- Assessment Profile Scheduling: Alpha = .84
- ECERS: Alpha = .90
- Arnett Scales of Caregiver Behavior (Lead Teacher): Alpha = .98
- Arnett Scales of Caregiver Behavior (Asst. Teacher): Alpha = .93

6. Response Rates

- **Fall 1997 "Wave One"**: A total of 524 classrooms were observed out of 541, for a completion rate of *97 percent*.
- **Spring 1998 "Wave Two"**: A total of 543 classrooms were observed out of 584, for a completion rate of *93 percent*.

IV. Results

1. Overall Classroom Quality in Head Start is Good

Table 1. Distribution of Classrooms by ECERS Mean Scores (Fall 1997).

Labels	Score ⁱ	Frequency	Percent
Inadequate	1	0	0
	2	0	0
Minimal	3	4	0.8
	4	139	26.8
Good	5	278	53.7
	6	94	18.1
Excellent	7	3	0.6
TOTAL		518	100.0

ⁱ The average scores were rounded off to the whole number reflecting the closest scale point, so that a score of 6 or higher includes scores of 5.5 or greater. A score of 4 or lower includes scores of 4.49 or less.

The mean ECERS score for Fall 1997 was 4.93 (SD=.63). The ECERS provides labels for selected scale points that denotes a level of quality associated with the scale score. Thus, a score of 1 on the 7-point scale is considered “inadequate,” while 3 is given the label of “minimal quality,” 5 is labeled “good quality” and a score of 7 is termed “excellent quality.” As the table indicates, no programs scored below the “Minimal” category.

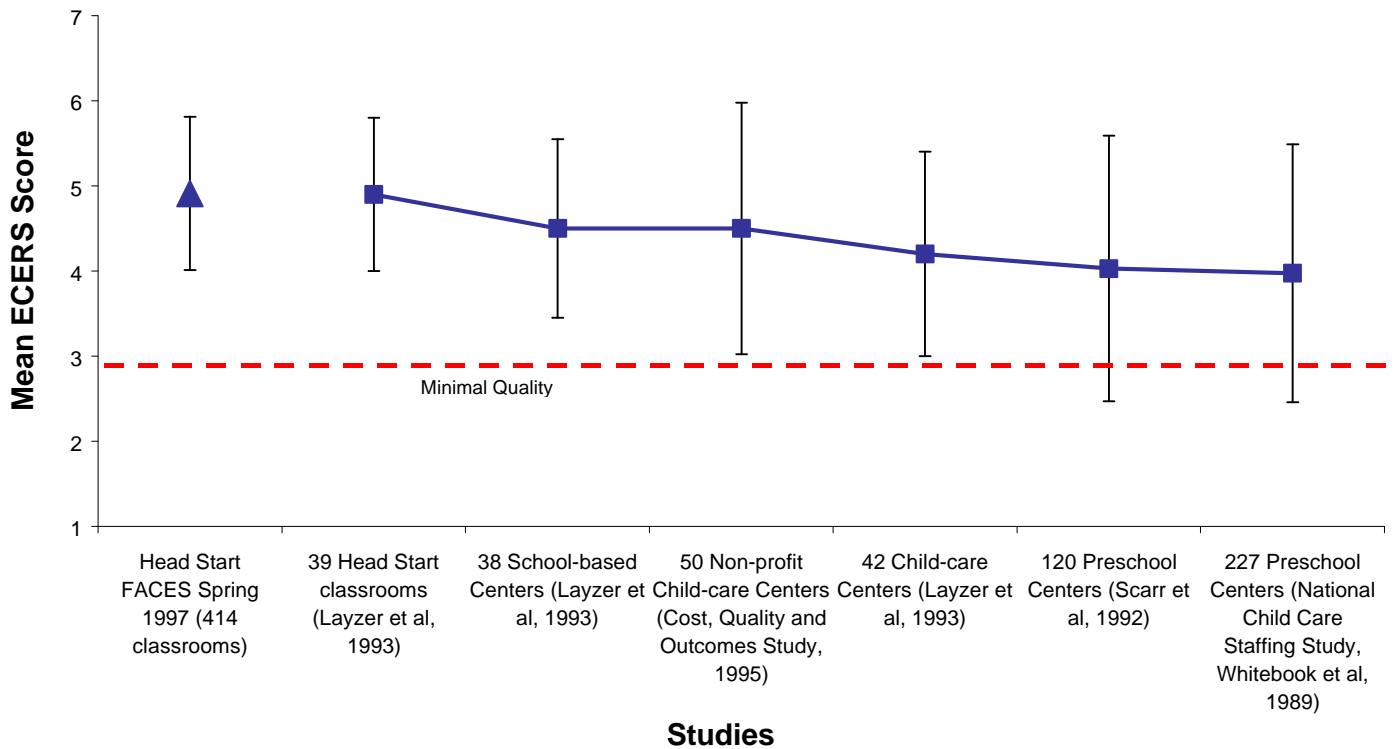
2. Classroom Quality Compares Favorably to Other Preschool Programs

Figure 1 displays a comparison of the average ECERS scores, and the range (within 2 standard deviations), between the FACES sample and previous studies of center-based preschools.

Figure 1 shows that the FACES mean of 4.9 was considerably higher than the average quality found in commercial child-care or center-based preschool programs. Also, the variability of the FACES scores was lower; none of the programs in the FACES sample fell below the “minimal” score of 3, although some were at the “3” level.

At the same time, the FACES average quality score, and variability were almost identical to those found in an earlier study that included a sample of Head Start classrooms. The lowest ECERS scores were reported by two large studies of center-based preschool programs, while school-based and non-profit child-care centers were slightly higher (but not at the same levels as the Head Start classrooms).

Figure 1. Classroom Quality Compares Favorably to Other Preschool Programs



The range represents the mean plus or minus 2 SD's. It does not include classrooms that fall outside of these boundaries.

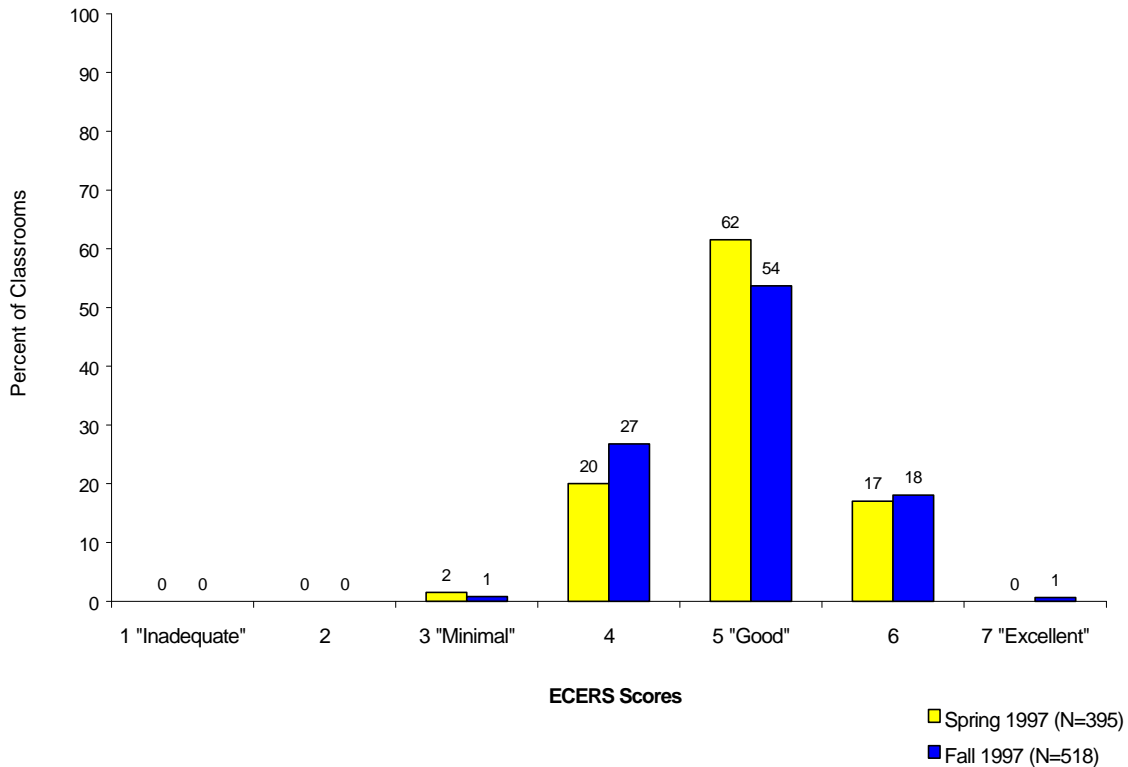
3. Classroom Quality Is Consistent Over Two Years

ECERS Scores from 395 classrooms in the Spring 1997 data collection were compared with scores from the 518 Fall 1997 classrooms. The classrooms came from within the same 40 Head Start programs. The Fall 1997 classrooms consisted of the same classrooms that were observed in the Spring 1997, plus some additional classrooms.

Figure 2 indicates that ECERS scores were consistent over the two years, within the same programs. The mean ECERS score remained the same, at 4.9. At both time periods, there were a similar number of classrooms rated "good" quality or higher. Classrooms rated "good" quality or higher (ECERS scores of 5 or above), at each time period were:

Spring 1997 78.5 %
 Fall 1997 72.4 %

Figure 2. The Percentage of Classrooms Rated Good Quality or Higher is Consistent Over Two Years.

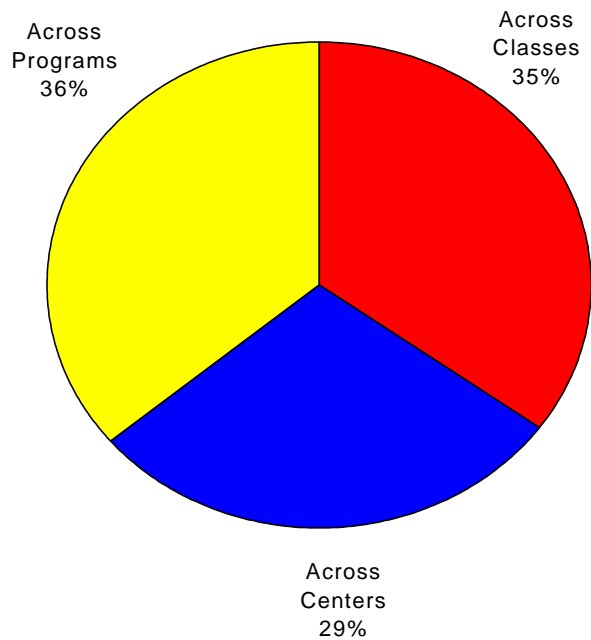


The Fall 1997 classrooms included those from the Spring 1997 as well as other classrooms within the same 40 Head Start programs. Source: Westat (1999). Analysis of data from Head Start Family and Child Experiences Survey (FACES) Spring 1997 and Fall 1997 classroom observations.

4. Classroom Quality Varies Across Programs and Centers

There was considerable variation in program quality across three levels: Head Start classrooms, centers, and programs. Classrooms are the most elemental level, where individual children experience Head Start, but classrooms are part of the larger Head Start center. A number of Head Start centers together comprise a Head Start program.

Figure 3. Percent Distribution of Variance in ECERS Mean Scores Across Head Start Programs, Centers, and Classrooms



Source: Westat (1999). Analysis of data from Head Start Family and Child Experiences Survey (FACES), Spring 1997 classroom observations.

Figure 3 shows that approximately one-third of the variation in program quality (according to the ECERS total scores) could be attributed to each of the three levels: classroom, centers and programs. Slightly more variation occurred at the classroom and the program levels than at the center level. Thirty-six percent of the variation in quality occurred at the program level and thirty-five percent of the variation in quality occurred at the classroom level, compared to twenty-nine percent attributable to variation across centers.

These findings indicate that, within a given Head Start program, classrooms in the same center tended to show more variation in quality compared with classrooms across centers (that are also part of the same program). The presence of significant variation at all of these levels supports continued efforts to improve program monitoring.

5. Classroom Quality Varies According to Socioeconomic Characteristics of the Communities in which Programs are Located

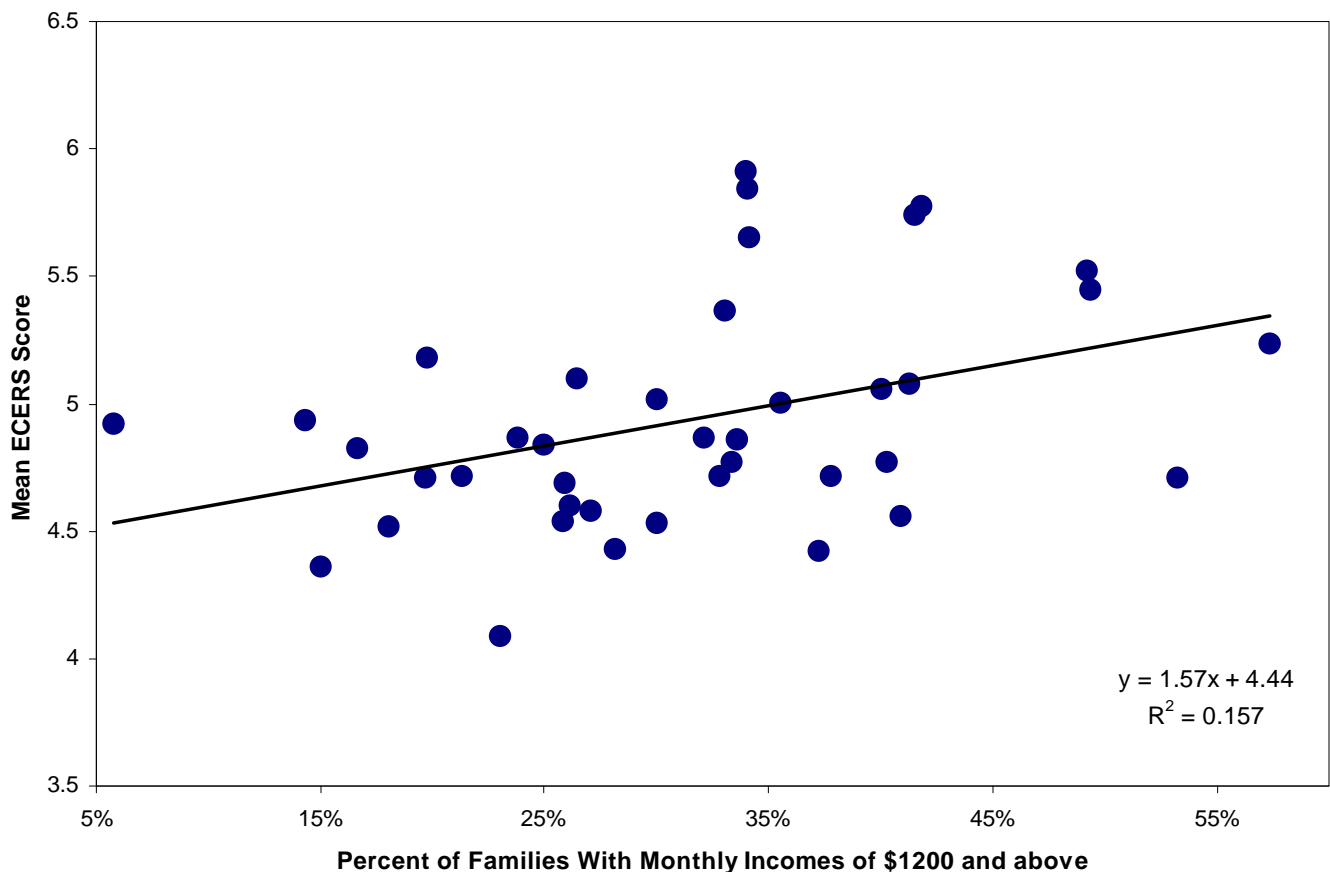
Regression analyses were conducted using average family characteristics (for sample families participating in a program) to predict to average ECERS scores for each program (using Fall 1997 data). Program averages were calculated for the following predictors:

- Average Family Income (Percentage of Families with Monthly Incomes of \$1200 or more)
- Education Level (Percentage of Parents with Some College Education or more)
- Percentage of Non-Minority Families In Program

The results are portrayed as scatterplots for the average score of the 40 programs participating in the study. The subsequent regression equation is fitted as a line through the plots.

Figure 4 indicates that classrooms with higher quality scores were in communities in which the average family income of participating families was higher. Average family income predicted approximately 16% of the variance in ECERS scores.

Figure 4. Classroom Quality Is Higher In Programs With Higher Family Income



Since the quality of language activities is an important component of the ECERS, and also may be related to children's subsequent language abilities, the same regression analysis was repeated for the Language Subscale of the ECERS, consisting of four items rating the quality of materials and instruction in receptive and expressive language, informal use of language during other activities, and the quality of activities aimed at thinking and reasoning skills.

Results indicate that average family income explains more variation in the ECERS Language subscale, compared with the Total ECERS scores. Approximately 23% of the variance in the quality of language activities in programs was explained by average family income. Again, as the average income of families in a program increases, the average process quality of the program also increases.

The average process quality using the total ECERS score was also predicted from the average percentage of families from non-minority groups who participated in the program, but to a much lesser extent. Approximately 8% of the variance in ECERS scores was explained by the percentage of non-minority families in the program. Average process quality was higher for programs when the percentage of parents from non-minority groups was higher.

When predicting to ECERS Language subscale scores, there was a slight increase in the amount of explained variation (10%). However, this predictor explained less of the variance in quality scores compared with average family income.

6. Head Start Structural Quality is High

Structural aspects of program quality, such as class size and child:adult ratio, further support the conclusion that the quality of many Head Start classrooms is good and higher than other center-based preschool programs.

Class Size (the number of children present per class)

- FACES Fall 1997 averaged 14.1 children
- FACES Spring 1998 averaged 13.8 children
- There were no significant differences across time, indicating that class size is stable in Head Start classrooms over time.
- The 75th percentile of classrooms averaged 16.0 children, while the 25th percentile averaged 12.0 children (Fall 1997 data).
- The class sizes found in FACES suggest that most Head Start programs sampled in this study meet or exceed the Performance Standards already in place.

Child:Adult Ratios (No. of children per adult)

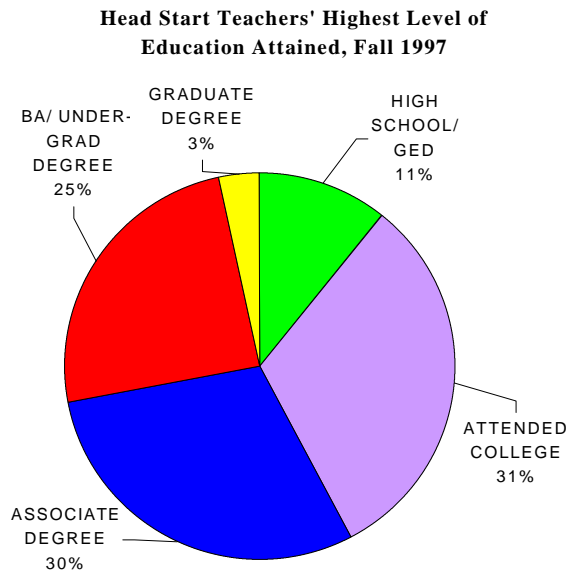
- FACES Fall 1997 averaged 6.3 children per adult
- FACES Spring 1998 averaged 6.2 children per adult
- There were no significant differences across time, indicating that child:adult ratio is stable over time in Head Start classrooms
- NAEYC accreditation standards: 8 or fewer three year olds, or 10 or fewer four year olds
- Head Start Program Performance Standards: 7.5-8.5 three year olds or fewer, or 10 or fewer four year olds
- In FACES, the 75th percentile of classrooms averaged 7.5 children per adult, whereas the 25th percentile of classrooms averaged 4.7 children per adult (Fall 1997)
- Even the worst quarter of the Head Start classrooms in the FACES sample had fewer children per adult than the NAEYC accreditation and were at the lower end of the Head Start Program Performance standards.
- However, the FACES child:adult ratios are based on the total number of adults in the classroom, including parents and other volunteers in the classroom, as long as they were actively involved in classroom activities. The Head Start Program Performance Standards and the NAEYC standards for child:adult ratio only count paid professional staff, so it is hard to compare.
- The method by which Head Start classrooms were able to have more favorable child:adult ratios was primarily through volunteer assistance, further underlining the importance of parent involvement as a contributor to overall program quality in Head Start.

These findings indicate that structural aspects of quality are important factors distinguishing Head Start classrooms from other preschool settings.

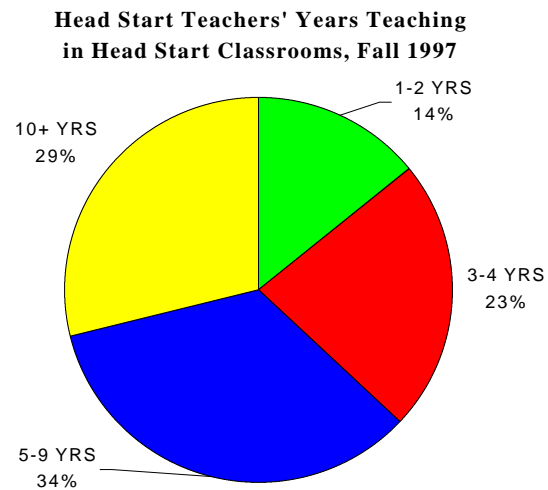
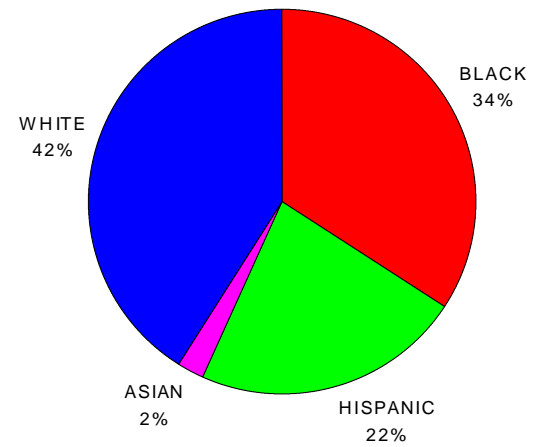
7. Head Start Teachers are Well-Trained

- The data are based on a self-report questionnaire given to the lead teacher for the classroom on the day the classroom process quality observations took place
- Head Start lead teachers are experienced and qualified to teach early childhood education
- 28.1% of lead teachers have a bachelor's degree or higher
- 76% have a Child Development Associates (CDA) diploma or an equivalent state-sponsored certificate
- Head Start lead teachers have good teaching qualifications, but lower than those of teachers in public elementary schools

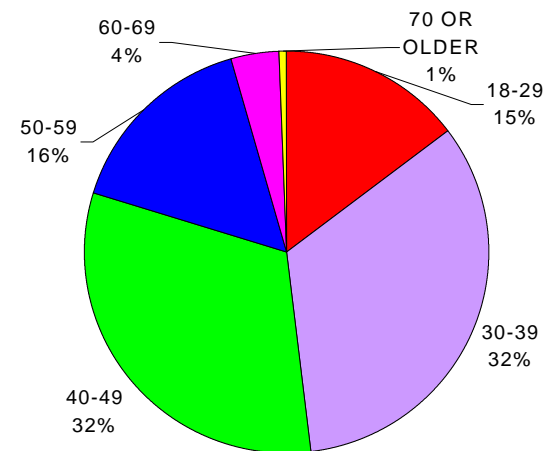
Figure 5. Characteristics of Head Start Teachers



Racial Backgrounds of Head Start Teachers, Fall 1997



Age Categories of Head Start Teachers, Fall 1997

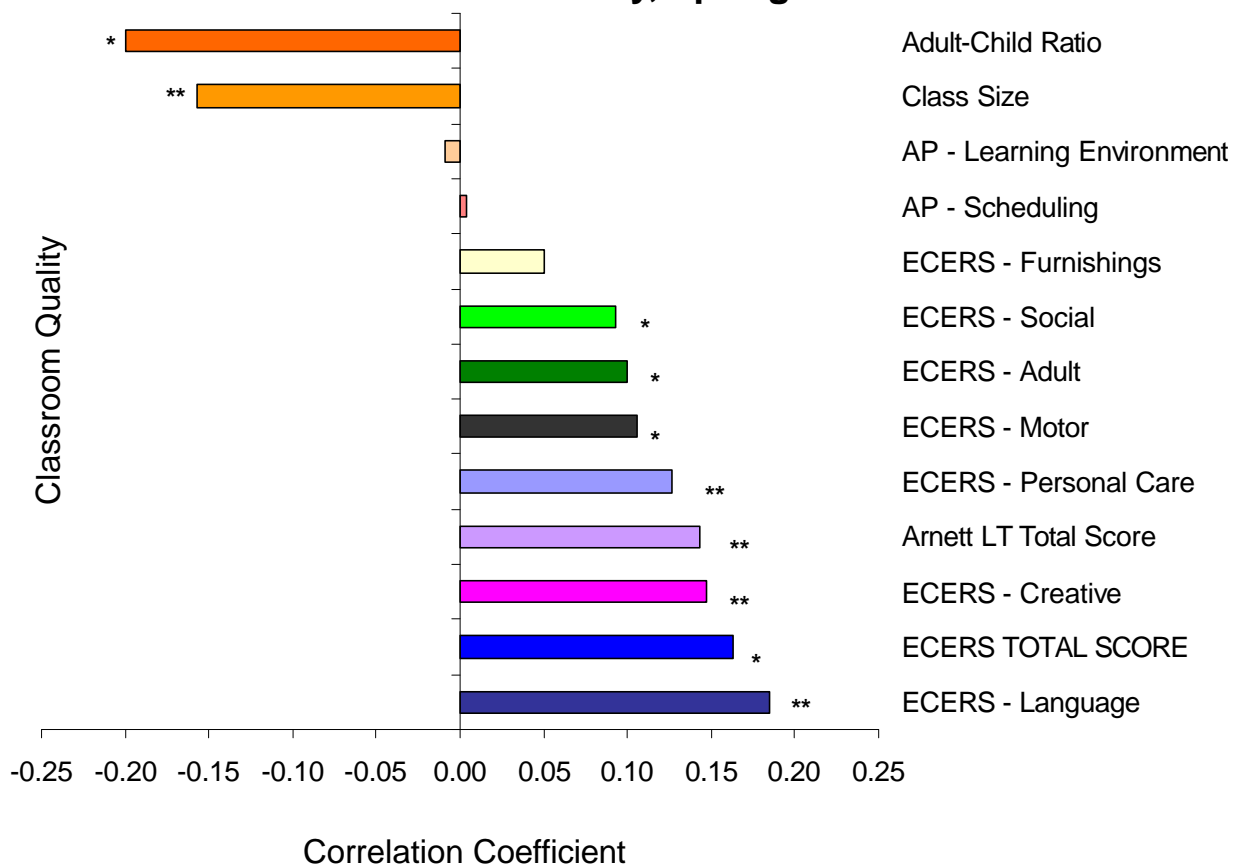


Source: Westat (1999). Analysis of data from Head Start Family and Child Experiences Survey (FACES), Fall 1997 Teacher Self-Report Form.

8. Teacher Backgrounds are Correlated with Classroom Quality Measures

- A series of simple bivariate correlations were conducted using the number of years teaching Head Start and highest education level correlated with ECERS, Assessment Profile, and Arnett scores, class size and child:adult ratio
- Results indicate significant positive correlations between highest education level and ECERS subscales and total score and the Arnett (lead teacher)
- Classrooms with higher process quality ratings and more sensitive teachers are also those with more qualified teachers
- Results also indicate significant negative correlations between teacher education and child:adult ratio and class size; classrooms with lower child:adult ratios, but also fewer children, tended to have more qualified teachers
- The only significant correlation between years teaching Head Start and quality was a negative correlation with ECERS adult subscale (provisions for adults); classrooms with fewer provisions for adults (staff and parents) tended to have teachers with more years of experience teaching Head Start

Figure 6: Highest Education Correlated with Classroom Quality, Spring 1998



Source: Westat (1999). Analysis of data from Head Start Family and Child Experiences Survey (FACES), Fall 1997 and Spring 1998.

*significant at $p < .05$; **significant at $p < .01$

9. Relationships between Program-Level Factors and Classroom Quality Using Multi-Level Modeling

- We conducted Two-Level Program Effects Models using SAS PROC MIXED (Singer, 1998), to predict classroom process quality from classroom-level and program-level factors
- Level-1 (Classroom, N=525) Predictors: Teacher Education, Years Teaching Head Start
- Level-2 (Program, N=40) Predictors: Average Family Income, Percent Non-Minority Families, Parent Education (Percent of families with at least some college), Region, Urban vs. Rural.
- Separate models were tested for each of three Level-1 Dependent Variables: ECERS total score, ECERS language subscale, Arnett scale

Results

- After taking the socioeconomic characteristics of communities into account, regional and urban-rural differences did not account for significant variation in classroom process quality
- Percent of between-program variation in classroom quality:
 - ECERS total = 33%
 - ECERS language = 22%
 - Arnett Scales = 15%
- Percent of between-program variation that can be accounted for by three measures of communities' socioeconomic characteristics:
 - ECERS total = 27%
 - ECERS language = 25%
 - Arnett Scales = 10%
- Percent of within-program variation that can be accounted for by teacher characteristics:
 - ECERS total = 6%
 - ECERS language = 5%
 - Arnett Scales = 5%

Implications of Results

- The socioeconomic characteristics of families in the communities in which Head Start programs operate are significantly related to the quality of the educational services programs provide
- After taking socioeconomic characteristics of families in the community into account, regional and urban-rural differences in quality are no longer significant
- Head Start programs are located within the communities they serve, and the characteristics of these communities may determine the level of resources available to the programs, in terms of volunteers, financial support, and the pool of experienced individuals available to teach

V. Summary of Conclusions

- Classroom quality in Head Start classrooms is good and higher than the quality of other center-based preschool programs;
- Classroom quality can be measured reliably in a large-scale survey using standard observational scales;
- Classroom quality remains consistent across two years;
- Structural aspects of Head Start classrooms (child:adult ratio, class size) are also favorable and exceed existing standards;
- Head Start teachers are well-trained and experienced, but their qualifications are lower than those of teachers in public elementary schools;
- Higher quality in classrooms is associated with better qualified teachers (those with higher education levels);
- There is more variation in classroom quality attributed to the center and the program than at the classroom level;
- Variations in quality in Head Start classrooms are significantly related to the socioeconomic characteristics of the communities that programs serve. Programs in communities with higher parent education and income levels (within the poverty range) and more non-minority families tend to have higher-quality classrooms.