Cumberland County and Greater Carlisle Area United Way  
Success by Six  
ECERS-R Training  

Barbara Carl and Richard Fiene  
Prevention Research Center  
The Pennsylvania State University  
June 2003  

Introduction  

This project consisted of both training and mentoring of staff from ten Cumberland County child care centers on the use of early childhood program quality tools that can be used to assess interactions between caregivers and children and environmental quality of early care and education settings. This number represents close to 18% of all licensed child care centers in Cumberland County.

Through this project staff from ten child care centers attended a full day workshop, as well as two days on site trainings on use of early childhood program quality assessment tools. The directors, or designated child care center staff trainers, then returned to their specific center program and conducted self-assessments of their classrooms, along with certified ECERS trainers/assessors. A pre/post test evaluation was conducted on the participants’ knowledge of environmental quality. In addition, this project provided information to directors and their staff on the importance and value of self-assessment as one of the first steps towards NAEYC accreditation. The trainers emphasized the need for consistency in self-assessment. Training consisted of the following components:

Day 1  Full-day classroom instruction on the ECERS-R and Arnett CIS, with a brief overview of the ITERS. This session also discussed the importance of consistent self-assessment. This training will be for directors and lead teachers for ten Cumberland County Child Care Center Providers (62 participant). In addition, appropriate ECERS-R and Arnett CIS resources were provided and a financial incentive was used to support attendance and participation.

Day 2  Guided field observation for the nine directors or designated center staff trainers who attended the initial training. This training was conducted on the Monday following the initial training and will be held in three neutral locations (non-participatory sites).

Day 3  Morning site visit at three participatory sites. Directors and trainers each conducted their own evaluation without the benefit of on-site interaction.
Review and consensus was conducted in the afternoon—importance of consistency stressed. Strategies for conducting self-assessments at individual centers was discussed during afternoon session.

Day 4  During the next several weeks center directors conducted an ECERS-R self-assessment of a specific room within their center. This observation was conducted in conjunction with a trainer. Ratings were reviewed with directors and scores compared. The trainer effectively modeled consensus building skills the director can use with staff when processing future assessments.

Day 5  Approximately one month later, trainers will return to centers and conduct a final environmental assessment. Results from this assessment will be compared to those of Day 4 to see if directors incorporated the suggestions made during the prior assessment. Analysis will be conducted to see if this intervention produced significant increases in the overall environmental rating score of the observed classroom.

**Workshop Training**

This training was attended by 62 child care directors and lead teachers. Participants were invited to attend based on a first come, first serve basis. The centers were informed of this project during the design phase of this Implementation Grant. All centers in Cumberland County were informed of the project through the Success By Six Child Care Newsletter. Additionally, directors were informed of the project at a monthly forum held at the United Way. Center directors requested the training be provided to their lead teachers, as well as themselves. They stated that too often they attend training and must return to the centers and reteach the information they learned. The reported this to be a challenge in ensuring that the information is understood and implemented in the same manner. By bringing as much of their staff as possible, the center as a whole would have a better foundation for utilizing the material presented. This would assist in all staff being able to process the information together and better understand how to utilize the assessment tools. The directors felt this approach would better support the concept of quality enhancement, give a stronger basis of understanding the material presented and lessen the resource burden (time, attention and resources) of each director training their staff."

A pre and post test were administered to workshop participants. A ten item questionnaire was given to attendees as part of their opening paperwork at the start of the workshop. These questions were drawn directly from the ECERS tool and chosen for their importance in the correct scoring of the measure. These pretest questionnaires were then gathered by facilitators. At the end of the workshop, participants were again given the same questionnaire as part of their evaluation of the training. As these results indicate, a transfer of learning occurred between the start of the training and the end. There was an increase in the correct response for each of the ten items, in some cases producing dramatic results. The results of these surveys were compared and are as follows.
Pre/Post Test Workshop Results

The first question on the pre/post test measure relates to art activities. In a full day child care program, at a minimum, children should have access to art materials one hour per day. The overall majority of participants were aware of this, however, there was an increase in the correct response to this question between the pre and post tests. As the chart below reflects, 95.1% of participants answered this item correctly on the pretest and a full 100% reported correctly on the posttest.

The term “substantial portion of the day” is very important in understanding the ECERS tool. In this context, the term translates to mean 1/3 of the day. This distinction is a disqualifier for most of the items in the activities section of the tool. Because of this, it is imperative for evaluators to be clear on the definition of this term. As the chart below reflects, only 34.4% of participants were clear on this term prior to the training, while 87.3% could identify the definition of the term by the end of the training.
The majority of participants were aware that children should have access to dramatic play materials for one hour daily. (Pretest responses were 90.2%, post test results were 100%).

Another fine point in the correct utilization of the ECERS tool is understanding the difference between the terms available versus accessible. For the purpose of this measure, accessible means that materials are present in the room or facility but not necessarily within the reach of children. Pretest results indicate only 3.3% of participants understood this distinction, growing to 43.6% by the posttest.

Sand play is important for preschool aged children. Many providers choose to use other substances besides sand in sand play. It is the philosophy of both the authors and trainers that food items (rice, beans and oatmeal) are not acceptable substitutes for sand as it gives children the wrong message about playing with food. In the choices listed, the only appropriate substitute for sand was birdseed. For most
participants, this was a learning item, as only 3.3% reported correctly in the pretest, versus 92.7% in the posttest.
Both the ECERS tool and Pennsylvania Department of Public Welfare require that children’s mats/cots for sleeping need to be placed a minimum of 24 inches apart. It is interesting that 63.9% answered this item correctly in the pretest, versus 67.3% in the posttest. It is curious that this is a DPW regulation and one would suspect that most directors and caregivers should already know this information. However, these pre/post tests do not reflect that knowledge.

The ECERS tool specifically identifies interlocking blocks, i.e., Legos, as fine motor materials, not blocks. In the pretest, only 41% of participants answered this correctly, versus 96.4% in the posttest.

The ECERS tool also specifically identifies that a walk is not considered gross motor activity for preschoolers. Pretest responses indicated only 23% correctly identified this answer, versus 89.1% in the posttest.
Another dramatic increase in knowledge occurred in the question regarding shaving cream play with children. Only 23% of pretest respondents were aware that this can be a toxic substance for children, versus 96.4% in the posttest.

The ECERS and ITERS tools are very clear regarding health and safety of children. Many providers seem unaware of the “rules” of handwashing. The measures are very clear that baby wipes are not an acceptable substitute for handwashing, except with very young infants who cannot independently hold up their head. The pretest measure indicated that 70.5% of respondents were aware of this versus 100% of posttest respondents.

Response to this training was overwhelmingly positive. A total of 53 participants completed the Pennsylvania Pathways Trainee Evaluation Form. Of these, 94.1% felt the level of training was just right and 100% felt they had learned something they can use in their child care work. The overall experience of the training was positive, with 90.7% of participants rating both the trainer and the training as good to excellent. Eighty-five percent of those in attendance report that they will tell others about the training.

It is important to note the willingness of participants to learn from this training. The response of participants was overwhelmingly positive. Of the original ten center Directors contacted to participate, eight followed through to completion. Two Directors
declined to participate with the follow-up training. One Director attended only the initial workshop and another participated with two days of the on-site training. They both felt that they have accredited centers and could not benefit from this type of mentoring.

Reliability Practice Site Visits With Directors

The second component with this training was site visits with Directors from nine child care centers that participated in the initial classroom training. This facet of the training consisted of three days of actual classroom observation where participants utilized the ECERS tool under the guidance of reliability certified ECERS trainers. Participants conducted these observations in three child care centers who voluntarily agreed to allow the training in their sites. The first day of this on site training was conducted using a guided facilitation model whereby trainers openly discussed the descriptors in the tool. This method was used in an attempt to ensure that participants obtained clarity of the assessment in a practical manner. The overall average of correct responses was 72.3%. A score of 85% is considered “reliable” and is a measure of clear understanding of the tool.

The second day of on site training was conducted with participants and trainers again observing on site. At this time, however, all members scored individually while on site and discussing their scores afterwards. The overall average of correct responses was 81.5%.

The third observation day was conducted individually with Directors at their own site. Directors chose the particular room in their center that they wanted to observe. The Director and trainer scored individually and then compared scores. (This information also became the pretest measure). The overall average of correct responses was 89.2%. After completion of score comparison, trainers and Directors identified ways to address deficiencies in the scores. A follow-up post test date was set at this time for the trainer to return and conduct the final post test assessment.

The increasing correct number of responses of directors compared to seasoned assessors is an important aspect of this training process. All eight directors who participated in this training to completion became reliable ECERS-R assessors. The chart below reflects the increase in effective use of the ECERS-R tool, ranging from participation in the one day workshop, followed by one on-site, guided observation, and finally, independent assessment.
These data have serious policy implications. The majority of ECERS-R training is conducted via one day workshops. As these data show, one day classroom training has little impact on the understanding of this complex assessment measure. Directors who attend only a workshop are unable to translate what they learned in that setting to clear assessment of a child care facility. To best educate and inform directors on this measure, a minimum of two days “in the field” with a well-trained assessor are suggested.

The following section reflects the pre and post test results of centers that participated in this study. This information will be provided first by identifying the overall ECERS-R score of study participants, as well as an overview of subscale items, coupled with the pre/post test results. Finally, the Cumberland County results will be compared to those of the 2002 Pennsylvania Early Care and Education Study.

**Pre/Post Test Results of Mentored Sites**

Both the mentoring and comparison groups were tested for equivalence at the beginning of the project in the pre-test data collection phase. On the ECERS-R, both groups showed no statistically significant differences at the pre-test. When the programs and caregivers were measured at the post-test, significant changes did occur. The section below provides detail of the overall increase in ECERS-R score, as well as the sub-scale measures.
Composite ECERS-R Score of Study Participants

Study participants produced dramatic increases in their ECERS-R score between the pre and post test assessment. This almost one full point increase raised the overall quality of study participants from an “adequate” score to that of “good”. Some of the subscales that comprise this mean score increased at an even larger rate and produced significant increases in scores.

Subscale Results

The Space and Furnishings subscale assesses the following items:
- Indoor space
- Furniture for care, play and learning
- Furnishings for relaxation
- Room arrangement
- Space for privacy
- Child-related display
- Space for gross motor
- Gross motor equipment

The composite post test score for this subscale was 5.57, versus the pretest score of 4.75.

The Personal Care Routines subscale assesses the following items:
- Greetings/departure
- Meals/snacks
- Nap/rest
- Toileting/diapering
- Health practices
- Safety practices
The composite post test score for Cumberland County was 5.19 versus 3.71 pretest.

The Language-Reasoning subscale assesses the following items:

- Books and pictures
- Encouraging children to communicate
- Using language to develop reasoning skills
- Informal use of language

The composite post test score for Cumberland County centers was 5.97 versus 4.84 pretest score.

The Activities subscale assesses the following items:

- Fine motor
- Art
- Music/movement
- Blocks
- Sand/water
- Dramatic play
- Nature/science
- Math/number
- Use of tv, video and/or computer
- Promoting acceptance of diversity

The composite post test score for Cumberland County centers was 5.33, versus 3.88 pretest score.
The Interaction subscale assesses the following items:

- Supervision of gross motor activities
- General supervision of children
- Discipline
- Staff-child interactions
- Interactions among children

The composite post test score for Cumberland County centers was 5.75 versus the 4.98 pretest score.

The Program Structure subscale assesses the following items:

- Schedule
- Free play
- Group time
- Provisions for children with disabilities

The composite post test score for participating centers was 6.17 versus 5.69.

The Parents and Staff subscale assesses the following items:

- Provisions for parents
- Provisions for personal needs of staff
- Provisions for professional needs of staff
- Staff interactions and cooperation
- Supervision and evaluation of staff
- Opportunities for professional growth

The composite post test score for Cumberland County centers was 5.25 versus 5.10 pretest score.
As reported graphically above, improvement in the mean score was attained for each subscale. Some subscales increase more dramatically and significantly than others. The chart below ranks subscale improvements, along with the statistically significance of these results.

<table>
<thead>
<tr>
<th>Subscale Category</th>
<th>Mean Pre to Post Test Score Improvement</th>
<th>Statistical Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Care Routines</td>
<td>+ 1.48</td>
<td>.000</td>
</tr>
<tr>
<td>Activities</td>
<td>+ 1.45</td>
<td>.003</td>
</tr>
<tr>
<td>Language and Reasoning</td>
<td>+ 1.13</td>
<td>.008</td>
</tr>
<tr>
<td>Space and Furnishings</td>
<td>+ .814</td>
<td>.004</td>
</tr>
<tr>
<td>Interactions</td>
<td>+ .776</td>
<td>.112 (Not Statistically Significant)</td>
</tr>
<tr>
<td>Program Structure</td>
<td>+ .473</td>
<td>.048</td>
</tr>
<tr>
<td>Parents and Staff</td>
<td>+ .154</td>
<td>.325 (Not Statistically Significant)</td>
</tr>
<tr>
<td>Composite ECERS-R Score</td>
<td>+ .966</td>
<td>.001</td>
</tr>
</tbody>
</table>

These increases in mean ECERS-R scores indicate the effectiveness of this targeted mentoring strategy. Mentoring obviously worked for all these programs, but it worked best in situations where the staff was most receptive. Further study should be conducted to identify characteristics of those directors whose scores increased most dramatically.

The Cumberland County child care centers who participated in this study reflect very high quality facilities. All subscales indicate quality in the “good” (5.0 minimum) scoring. The ranking of these scales, from highest to lowest, are listed below:

- Program Structure (6.17, with a range of 1.67 – 7.00, SD 1.82)
- Language and Reasoning (5.97, with a range of 3.25 – 7.00, SD 1.26)
- Interactions (5.75, with a range of 2.20 – 6.80, SD 1.60)
- Space and Furnishings (5.57, with a range of 4.38 – 6.25, SD .58)
- Activities (5.33, with a range of 3.0 – 6.10, SD 1.05)
- Parents and Staff (5.25, with a range of 4.66 – 5.83, SD .55)
- Personal Care Routines (5.19, with a range of 2.83 – 6.83, SD 1.37)

This ranking also indicates a wide variety of the quality of child care in participating sites. Keeping in mind that a score of 5.0 is considered “good”, 4.0 is “adequate” and 3.0 is “minimal”, care should be taken to further educate facilities that scored the lowest in an attempt to raise the overall quality of child care in Cumberland County centers.
**Comparison of Cumberland County Participating Control Versus Intervention Sites**

As indicated above, there are some dramatic and significant improvements in the environmental quality of the Cumberland County facilities that participated in this targeted mentoring process. To clarify further that these differences were not just a natural fluctuation of change, we shall examine both the pre and post test scores of the control sites versus those of the intervention sites. As you will see, all control sites did improve in quality, however, with no statistical significance. Those facilities that improved the most reported attending several workshop presentations over the course of the two months between the pre and post test data points. The results are as follows:

Both the control and intervention sites began this study at relatively the same starting point (4.55 for intervention, 4.33 for control). However, after completion of the targeted ECERS-R training and mentoring, the intervention score increased to 5.51, versus the control post test score of 4.63. This is a statistically significant increase ($t = .048$, df = 16).

**Subscale Results**

In reviewing the Space and Furnishings Subscale, the intervention score increased from a pretest mean score of 4.75 to a post test mean score of 5.57. This compares to the mean pretest control group score of 4.21 versus 4.58. This is a statistically significant increase ($t = .030$, df = 16).
In the Personal Care Routines Subscale, dramatic increases in the mean score were again seen in the intervention group. The pretest score was 3.71 versus the posttest score of 5.19. This compares to the control group pretest score of 3.78 versus 3.52. This represents an actual decline in scoring of .26 (versus the 1.48 increase in scoring for this same subscale by the intervention group).

This is a statistically significant increase (t = .006. df = 16).

The Language and Reasoning Subscale also saw an increase in the mean score for those who participated as the intervention group. The pretest score was 4.84 versus the post test score of 5.97. This compares to the control group mean pretest score of 4.83 versus the post test score of 5.07. While this was a dramatic increase, this was not a statistically significant increase (t = .098. df = 16).
The Activities Subscale also saw dramatic increases in the mean score for the intervention sites. The pretest mean score was 3.88 versus the post test score of 5.33. This compares to the control group, whose mean pretest score was 3.80 versus 4.01 post test. This is a statistically significant increase ($t = 0.014, \ df = 16$).

The Interactions Subscale produced very interesting results. The mean pretest score for those in the intervention group was 4.98. The Post test mean score for this group was 5.75, an increase of .77. For the control group, however, the mean pretest score was 5.5, and the post test score was 5.4. This represents a decrease of .10. It is interesting to note that even though the control group sites started out at a higher rate than those of the intervention sites (+.52), the actually reported a decrease in the post test measure. While this was a concerning decrease, this was not a statistically significant increase ($t = 0.3315, \ df = 16$).
The Program Structure Subscale indicates some differences between the intervention and control group, which are evidenced by their pretest scores. The intervention group reflected a mean pretest score of 5.70, versus the mean control group pretest score of 4.11. Posttest intervention scores indicate a mean score of 6.17, versus 5.44 in the control group. (This was not a statistically significant increase \(t = 0.207, \ df = 16\)). This subscale measures quality indicators that relate to schedule, group time and free play of children. These data indicate a significant \(-1.59\) difference in the pretest assessment, and, as such, this subscale cannot be assumed to be equal.

The Parents and Staff Subscale also produced some interesting results. The mean pretest score for intervention sites was 5.10, and the post score was 5.25. This compares to the mean pretest score of 5.34 for the control group and post test score of 5.38. While the intervention sites report both lower pre and post test scores, it should be noted that these reflect an increase of .15, versus the .04 increase for the control groups. While this was an increase, this was not a statistically significant \(t = 0.357. \ df = 16\).
As reported graphically above, control group sites, while showing some improvement in the mean score for some subscales, actually decreased in some categories. Additionally, even though some improvement was noted, the only increase which is statistically significant is that of the Program Structure subscale. The chart below ranks subscale scores, along with the statistically significance of these results.

<table>
<thead>
<tr>
<th>Subscale Category</th>
<th>Mean Pre to Post Test Score Improvement</th>
<th>Statistical Significance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Personal Care Routines</td>
<td>- .257</td>
<td>.197 (Not Statistically Significant)</td>
</tr>
<tr>
<td>Interactions</td>
<td>- .100</td>
<td>.426 (Not Statistically Significant)</td>
</tr>
<tr>
<td>Parents and Staff</td>
<td>+ .303</td>
<td>.880 (Not Statistically Significant)</td>
</tr>
<tr>
<td>Activities</td>
<td>+ .207</td>
<td>.397 (Not Statistically Significant)</td>
</tr>
<tr>
<td>Language and Reasoning</td>
<td>+ .245</td>
<td>.391 (Not Statistically Significant)</td>
</tr>
<tr>
<td>Space and Furnishing</td>
<td>+ .364</td>
<td>.197 (Not Statistically Significant)</td>
</tr>
<tr>
<td>Program Structure</td>
<td>+ 1.33</td>
<td>.009</td>
</tr>
<tr>
<td>Composite ECERS-R Score</td>
<td>+ .303</td>
<td>.089</td>
</tr>
</tbody>
</table>

Aside from the Program Structure Subscale, these results indicate that the control and intervention sites were statistically similar to one another and the improvements noted by the intervention groups cannot be attributed to chance alone. These data further support the effectiveness of this targeted mentoring strategy in increasing the overall quality of child care environments.
Comparison of Cumberland County Post Test ECERS-R Scores Versus 2002 Pennsylvania Early Childhood Quality Study

Based on the seven point scale, this score indicates that the average participating child care center in Cumberland County is functioning at a good level of care, with a mean score of 5.51. This score is also considerably higher than the 3.89 that was determined to be the statewide average in the Pennsylvania Early Childhood Quality Settings Study.

Additionally, when assessing the individual subscale categories, the center sites in this study scored significantly higher than those represented in the Quality Study. Care must be taken, however, to remember that the sample size in this study was only eight facilities and generalizations to the larger population must be carefully viewed. Also, the range of scores is great, which indicates a wide variability in the quality of care. That said, the scores obtained by the Cumberland County child care centers who agreed to participate in this study reflect quality which ranges from minimal to good.

Subscale Results

The Cumberland County centers scored dramatically higher than those in the ECQS in the Space and Furnishings subscale items (5.57 versus 3.95).
Compared to the ECQS results, participating Cumberland County centers scored significantly higher in the Personal Care Routines subscale (5.19 versus 3.39).

Compared to the ECQS results, Cumberland County centers scored significantly higher in the Language-Reasoning subscales (5.97 versus 4.13).

Compared to the ECQS results, Cumberland County centers scored higher in the Activities subscales (5.33 versus 3.13).
Cumberland County centers also scored significantly higher than the ECQS average on the Interaction subscales (5.75 versus 4.60).

In the Program Structure subscale, Cumberland County centers scored significantly higher than the ECQS average (6.17 versus 4.20).

Finally, in reviewing the Parents and Staff subscale, Cumberland County centers scored significantly higher than the ECQS average (5.25 versus 3.89).
Conclusions and Recommendations

- The ECERS-R (as well as other industry standard environmental rating scales) is a complex tool that requires adequate training if it is to be utilized appropriately. Results of the pre/post test measure conducted at the close of the initial workshop indicate that most participants were unaware of many of the details that are critical in scoring. Understanding of terms “available”, “accessible”, and “substantial portion of the day” are pivotal in achieving accurate scoring. Without proper training, child care providers and directors are unable to correctly assess their facilities.

- Both the ECERS tool and Pennsylvania Department of Public Welfare require that children’s mats/cots for sleeping need to be placed a minimum of 24 inches apart. It is interesting that 63.9% answered this item correctly in the pretest, versus 67.3% in the posttest. It is curious that this is a DPW regulation and one would suspect that most directors and caregivers should already know this information. However, these pre/post tests do not reflect that knowledge.

- The increasing correct number of responses of directors compared to seasoned assessors is an important aspect of this training process. All eight directors who participated in this training to completion became reliable ECERS-R assessors. These data have serious policy implications. The majority of ECERS-R training is conducted via one day workshops. As these data show, one day classroom training has little impact on the understanding of this complex assessment measure. Directors who attend only a workshop are unable to translate what they learned in that setting to a clear assessment of a child care facility. To best educate and inform directors on this measure, a minimum of two days “in the field” with a well-trained assessor are necessary. An additional impact of this is well-trained and “reliable” directors can impart their knowledge gained to their staff and increase the overall quality of care throughout their facility.

- These data clearly demonstrate how the mentored programs improved from the pre-test to the post-test on the Early Childhood Environment Rating Scale – Revised. This is a very important finding because the majority of mentoring projects in the past have utilized anecdotal evidence to demonstrate their effectiveness. This study, similar to those previously conducted by Fiene, 2001, clearly demonstrate the effectiveness of mentoring in quality improvement utilizing randomized trials of interventions.

- It is clear from data analysis of the control group sites that, while some fluctuation in scoring between the pre and post test data points, there were no indicated changes that reflect statistically significance (4.33 versus 4.63). There were also
no statistically significant differences between any of the subscale scores of the control group.

- The public policy implications of this study are significant in that this is a clear demonstration that a short term, targeted mentoring intervention will produce positive changes in the overall quality of child care programs. Previous research (Johnson, 1994) has indicated that increasing the number of training hours produces more developmentally appropriate behaviors in child care staff. Mentoring fits this model because it is intensive one on one intervention in which the mentor and protégé are engaged in problem solving activities to improve the overall quality of environments of child care programs.
References
