

# BEST PRACTICES FOR CONDUCTING PROGRAM OBSERVATIONS AS PART OF QUALITY RATING AND IMPROVEMENT SYSTEMS



Research-to-Policy, Research-to-Practice Brief OPRE 2011-11b

June 2011



**DISCLAIMER:**

The views expressed in this publication do not necessarily represent the views or policies of the Office of Planning, Research and Evaluation, the Administration for Children and Families or the U.S. Department of Health and Human Services.

**ACKNOWLEDGMENTS**

Dr. Hamre would like to acknowledge colleagues at the National Center for Research in Early Childhood Education ([ncrece.org](http://ncrece.org)) and the Center for Advanced Study of Teaching and Learning for their collaboration and support of this work. Dr. Maxwell would like to acknowledge colleagues at FPG Child Development Institute ([fpg.unc.edu](http://fpg.unc.edu)) for their support of this work. Both authors would like to thank Dr. Thelma Harms and Ms. Cathy Riley for reviewing the brief. The authors would also like to thank the following people for reviewing the outline of the brief: Kirsty Brown and colleagues from the Pennsylvania Office of Child Development and Early Learning, Anna Carter from the North Carolina Division of Child Development, Vicki Hawley from the University of Minnesota, Kathy Glazer from the BUILD Initiative, and Bentley Ponder from the Georgia Department of Early Care and Learning.

# Best Practices for Conducting Program Observations as Part of Quality Rating and Improvement Systems

## Research-to-Policy, Research-to-Practice Brief OPRE 2011-11b June 2011

### Submitted to:

Ivelisse Martinez-Beck, PhD., Project Officer  
Office of Planning, Research and Evaluation  
Administration for Children and Families  
U.S. Department of Health and Human Services

### Submitted by:

Bridget Hamre, Center for Advanced Teaching and Learning, University of Virginia  
Kelly Maxwell, Frank Porter Graham Child Development Institute, University of North Carolina, Chapel Hill

Contract Number: GS10F0030R

Project Director: Kathryn Tout  
Child Trends  
4301 Connecticut Ave NW  
Washington DC, 20008

### Suggested Citation:

Hamre, B. K., & Maxwell, K. L. (2011). *Best Practices for Conducting Program Observations as Part of Quality Rating and Improvement Systems*, Research-to-Policy, Research-to-Practice Brief OPRE 2011-11b. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

This Brief was developed by members of the Quality Initiatives Research and Evaluation Consortium (INQUIRE) which is designed to facilitate the identification of issues and the development and exchange of information and resources related to research and evaluation of quality rating and improvement systems (QRIS) and other quality initiatives. INQUIRE is funded by the Office of Planning, Research and Evaluation through the Child Care and Early Education Policy and Research Analysis and Technical Expertise contract with Child Trends.





## **Best Practices for Conducting Program Observations as Part of Quality Rating and Improvement Systems**

Program observation offers an important window into program quality and an opportunity to identify strengths and weaknesses of programs. Observations assess the degree to which programs are providing children with safe and stimulating environments and warm, sensitive, and stimulating interactions—the aspects of early childhood environments most closely aligned with positive developmental outcomes.

Although implementing observations can be complex as well as time and resource intensive, data from observations are an important component of quality rating and improvement systems (QRIS) and are used to inform both the ratings for the public and the improvement plans for programs. In a 2010 Compendium of QRIS, 22 of 26 QRIS included program observations as part of the rating system.<sup>10</sup> Recognizing the widespread use of program observations in QRIS, the purpose of this Brief is to highlight issues and recommendations for conducting program observations as part of a QRIS. Some of the issues in this Brief apply to all early care and education settings, including family child care homes; others apply only to center-based programs with multiple classrooms. The first section of the Brief includes considerations when selecting a quality measurement tool. The second section describes issues related to hiring and training individuals to conduct observations. The third section describes issues about planning, scheduling, and conducting observations; the fourth section covers scoring and reporting.

## Selecting an Observational Tool for Measuring Quality

Once the decision has been made to include an observational tool in a QRIS, the next step is to determine *which* measure(s) to use. Although QRIS currently include a limited number of observational tools, there are many available observational measures of quality and resources for comparing different measures.<sup>1,11</sup>

Consideration of the following issues may be helpful in selecting an observational measure.

*What is the definition of quality and how well does the measure reflect the definition?* If an observational measure is selected without a definition of quality, then the measure will implicitly define “quality.” A measurement tool ideally should reflect the explicit QRIS definition of quality. Georgia’s Department of Early Care and Learning, for example, completed a crosswalk comparison of its definition of quality early care and education with several existing measures in order to select tools for its quality improvement initiatives that best measured aspects of quality that were important to them.<sup>5</sup>

*What is the purpose or intended use of the measure?*

It is important to match the measure with the intended purpose.<sup>12</sup> If, for example, providers in the QRIS use the measure as a self-assessment tool, was the measure designed for this purpose and how easy is the measure to use? If the measure is supposed to be helpful for technical assistance (TA), how well is the measure aligned with the TA content?



*Is there evidence of the*

*measure’s ability to produce reliable and valid scores?* Reliability refers to the consistency in measurement. Can two individuals using the same measure observe the same classroom at the same time and score the classroom similarly? Validity refers to whether the tool measures what it is intended to measure. It is not good enough for a publisher to describe a tool as a measure of quality child care. There must be some evidence that the tool is measuring quality; this is usually accomplished by demonstrating that scores on the tool are related to other measures or other aspects of quality. Validity may also be demonstrated with data linking higher scores on the quality measure to better child outcomes.

*Has the measure been used in the settings of interest?* If the QRIS applies to center-based early care and education programs, family child care, and school-age programs, then the observational measures selected must be appropriate for those different settings. This can be challenging; there are many more measures that have been developed for center-based programs than for family child care and school-age programs.

The remaining sections of this Brief focus on issues that apply to conducting observations of classrooms and family child care homes, regardless of the measure selected. Because the Environment Rating Scales<sup>2-4</sup> and Classroom Assessment and Scoring System<sup>8</sup> are two of the more commonly used measures in early care and learning settings, this Brief includes information about these particular measures as well as a more general discussion of issues across all measurement types.

## Selecting and Training Staff on Observational Protocols

This section of the Brief presents recommendations regarding selecting staff to conduct observations, determining the number of observers needed, and providing appropriate training.

### *Selecting Staff for Monitoring and Evaluation Work*

For staff who will be observing classrooms for the purpose of evaluation or monitoring, it is important that they are able to learn to score reliably on the observational system. People with experience learning and using observational measures are generally preferred. Among those without such experience, it is important to determine the extent to which they will be able to make objective ratings of classrooms.

In general, it is important for observers to have some familiarity with early childhood programs. It is not the case, however, that observers need to be very experienced teachers or program directors. Experienced teachers or administrators often have strong opinions about classroom environments that may make it more difficult for them to follow the specific observational rubric of the measurement tool.

It is extremely important for observers to be free of conflict of interest with the programs they will be observing. It is difficult, for example, to have a technical assistance provider collect program evaluation data for the classrooms she serves. If the observer is a part-time employee, then it is important to determine whether volunteer or paid efforts outside of the part-time observation job could present a conflict of interest (e.g., she sells learning materials and leaves her business card with teachers whom she observes).

### *Selecting Staff for Technical Assistance and Professional Development Work*

In contrast to the selection of staff to conduct observations for evaluation and monitoring purposes, it is recommended that those who will be providing technical assistance to teachers and programs, based on the observation findings, have extensive experience in early childhood settings. This experience fosters relationships between them and the teachers with whom they are working; it also provides TA staff with a wealth of experiences from which to draw as they do their work.

Because in many cases TA staff will not be using the observational measures to determine a quality score for classrooms, potential TA staff often ask whether they need to be trained to reliability on the instruments. Based on work with many states using observational methods to provide TA, the authors of this Brief highly recommend intensive training for TA staff on the observational measures. TA staff need to learn to use the observational instrument well so that they can provide feedback to teachers and programs that is consistent with the results from the observational measure.

### *Number of Observers Needed*

The number of observers needed depends on a variety of factors that are discussed in greater detail in the section on planning and scheduling observations, such as how many observations need to be conducted, for how long, etc. However, there are a few other factors to consider.

In general, it is preferable to have a smaller number of people conduct more observations, rather than having many people conduct fewer observations. Training a smaller number of observers is more efficient and less expensive. In addition, as described below, there are substantial costs related to providing ongoing support to ensure observers are coding reliably, and these costs are magnified with more observers. Observers are more consistent in their use of an observational instrument if they use it frequently.

With that caveat in mind, some programs and states have hired full time observational staff within their agencies, while others contract the work out. Both of these are viable options, but when using contract staff it is important to ensure a high level of quality control over their work.

In most cases, observers can complete observations on one classroom/family child care program in a day. This may vary depending on the instrument and the setting in which it is being used. It is also possible that observers may be trained to use multiple observational instruments and provide ratings on multiple instruments in one day.<sup>A</sup>

It is also important to note that more people will need to be trained than will ultimately collect data. There are a variety of reasons for this including attrition and the fact that in many cases not all trainees will be able to meet the reliability threshold.

### ***Finding the Right Training for the Right Audience***

There are a variety of training options available for most observational measures. It is important to ensure that the right training is provided to the right group of people.

There are four overarching types of trainings. These types of trainings, listed from least to most intensive, are described briefly below. Table 1 provides a summary of training options currently available within each category for the ERS and CLASS, including information on length of training and cost.

- **Gain familiarity with the instrument.** States and others implementing QRIS provide trainings to give programs and teachers information about the instruments with which they will be evaluated. These trainings typically last between a half to a full day and provide attendees with an overview of the instrument. They are typically not intended to provide trainees with the capacity to use the instruments to score classrooms or to provide feedback to others about the results of an observation.
- **Using the instrument for evaluation purposes.** People who will be using the observational instrument as a determinant of rating levels need more intensive training. These trainings typically consist of an overview on the instrument development, familiarization with each of the scales and items included in the instrument, and specific information on how to score the instrument. There is also substantial time for trainees to practice using the instrument and to calibrate their scoring to the scores of authors or master coders. Trainings typically last at least two to three days. This practice may be done live or using videotapes. In some cases a reliability test is provided at the end of the training. If there is not a reliability test included in the training, it is essential to provide a test for trainees prior to collecting data. In other cases, assessors must meet a pre-identified training standard (e.g., agreement within 1 point for at least 85% of the items).

<sup>A</sup> A forthcoming study conducted for the Office of Planning, Research and Evaluation in the Administration for Children and Families, U.S. Department of Health and Human Services by Mathematica Policy Research examines how results on observational measures vary depending on the parameters used for observation (number of observers, use of multiple instruments during an observational session).

- **Providing technical assistance or professional development based on classroom observations.** Training for individuals who will be using results from the assessment to give teachers and programs feedback is typically more intensive than trainings designed for those who will be using the instrument only to score classrooms/family child care programs. It is important that those giving feedback to teachers and programs have in-depth knowledge about the instrument, including a strong understanding of each of the components of quality rated by the instrument, how ratings are obtained, and the reasons why each component of quality is important to children’s development. These individuals should also have specific training in how best to support teachers and programs in improvement.
- **Training of trainers.** Using observational measures at a large scale requires training large cadres of people on an ongoing basis. To help meet this need, some observational instruments provide training that certifies participants to train others on the instrument. These are typically quite intensive trainings, lasting at least a week and sometimes longer. Content focuses on all of the information covered in the trainings described above as well as information on and practice leading trainings.

### *Training Large Numbers of Observers*

As discussed above, using observational instruments as a part of city, county or state QRIS requires training large numbers of people. It is recommended that initial trainings are provided directly by instrument authors or other certified trainers. Over time, systems may work toward a train the trainer model, once there is sufficient experience with the instrument and local capacity to ensure that trainings are of high-quality.

### *Maintaining Reliability and Consistency of Scores among Observers*

A measure is not inherently reliable. Maintaining high levels of reliability requires continuous effort. Thus, it is not sufficient to provide one training session for observers and expect that they will be able to continue scoring reliably over time. Drawing upon the experience of researchers conducting multi-site evaluation studies using observational instruments, it is recommended that processes to provide ongoing support and calibration be implemented to ensure that the instrument is being used consistently across observers, sites, and time. Although this process is resource intensive, it is essential to ensuring the successful use of observational instruments. If programs and teachers do not have confidence that they are being assessed fairly and consistently, the whole system will be undermined.

QRIS and other initiatives using observational measures should develop a system of regularly checking reliability of observers and providing them with feedback about their scores. This may be done in a variety of ways. For example, some observational instruments offer special trainings or online resources to help keep cohorts of observers reliable. QRIS may also develop these types of training on their own. Reliability checks can also be accomplished by sending multiple raters to observe the same classroom. Evaluation protocols typically provide for two raters that conduct assessments in 10-20% of classrooms that are observed. Individual instruments have recommendations regarding the acceptable threshold for inter-rater agreement and ways in which this should be calculated. Regardless of the way QRIS check reliability, it is desirable that there is some check on reliability and provision of feedback to observers at least once a month during ongoing data collection. If results of these monthly sessions suggest observers are scoring consistently, once a month testing and support may be sufficient. If there are significant problems with reliability, providing more intensive supports may be necessary.

## Planning, Scheduling, and Conducting Observations

In this section, we discuss issues related to determining which classrooms to observe; cost issues; length, frequency and timing of observations; and using multiple observational measures. Some issues, like determining which classrooms to observe in center-based programs, apply only to those programs with more than one classroom whereas other issues, like using multiple measures, apply equally to all programs including family child care homes.

### *Number and Type of Classrooms to Observe*

Although one could argue the importance of observing each and every classroom in a program to obtain the most accurate measure of program quality, resources typically limit the number of classrooms that can be observed. Thus, one important decision in conducting observations is sampling—determining the number and types of classrooms to observe in programs that have more than one classroom. The sampling decisions generally depend on the purpose of the observation. For most QRIS, the purpose is to determine the observed quality of the program. Thus, it is important to observe each type of classroom (i.e., infant/toddler, preschool, school-age) and to observe enough classrooms to be “representative” of the program. At a minimum, one classroom of each type provides information about the quality of care for children of varying ages.

How many classrooms should be sampled and observed in order to have a “representative” sense of the program’s overall quality? Research provides some guidance regarding the number of classrooms to observe in each program. Researchers from the McCormick Tribune Center for Early Childhood Leadership collected data from Head Start and public preschool programs in Illinois to compare the accuracy of classifying programs to one of four star levels using observations from either one-third or one-half of the classrooms in each program.<sup>6</sup> They found that sampling one-half of the classrooms resulted in 72% of the programs being assigned the star level that they would have received if all of the classrooms were observed. A sample of one-third of the classrooms resulted in a 67% match. Neither infant/toddler nor school-age programs were included in this study; it may be important to sample a certain percentage of classrooms within each type (i.e., infant/toddler, preschool, school age). It is also important to consider the range of scores required at each level, recognizing that a larger sample of classrooms may be needed to accurately place a program in a small quality range (e.g., ERS score between 4.0 and 4.25) as compared to a larger quality range within a level (e.g., ERS score between 4.0 and 5.0). In the Illinois study, the levels were defined by fairly large ECERS-R ranges (e.g., Level 2 required an ECERS-R average of 3.50 to 4.24).

Researchers in Missouri also collected data about sampling.<sup>9</sup> They found that sampling one-half of the classrooms in each program resulted in 86% of the programs being assigned the tier level they would have received if *all* of the classrooms were observed. A sample of one-third of the classrooms resulted in a 75% match. It is important to note that the Missouri study included data from infant-toddler, preschool, and school-age classrooms.

Research suggests that sampling more classrooms per program is best. Determining the percentage and number of classrooms is not just science, though. The cost of conducting observations generally limits the number of classroom observations. (Costs are discussed later in this Brief.) Weighing the costs, it is not recommended that QRIS observe every classroom in programs if the purpose is solely to determine the program’s rating. However, it is clear that observing every classroom may be useful for other purposes such as providing technical assistance.

In general, the authors of this Brief suggest these guidelines for the number of observations in QRIS:

- Observe at least one classroom for each age range served by the program (e.g., infant/toddler, preschool, and school age); and
- Ideally observe one-half of the classrooms in each program.

### ***The Cost of Observations***

The cost of conducting classroom observations varies state-to-state depending, in part, on factors like the geographical distribution of programs and the market rate for employing people to conduct assessments. When determining the cost of conducting classroom observations, administrators should consider the following:

- Time in the classroom conducting the observation
- Time outside of the classroom to complete related paperwork and reports
- Time to schedule visits, if the observer is responsible for this
- Travel time
- Initial training time
- Supervision time to discuss observation issues and questions
- Time conducting inter-rater observation visits (which includes time in the classroom and time later discussing the observations)

The total cost of conducting classroom observations as part of a QRIS also depends on the number of programs and classrooms observed as well as the frequency with which programs are observed. For example, Tennessee conducts classroom observations in every licensed program every year; North Carolina conducts classroom observations in programs every 3 years (or more frequently if requested and paid for by the program) and only for those programs who have first met other criteria. Thus, the costs for conducting assessments in a state like Tennessee would be expected to be greater than for a state like North Carolina.

Based on experience with conducting observations as part of research studies, the authors of this Brief estimate that the direct cost of conducting an observation ranges from \$300 to \$500 per classroom. Readers may find it useful to review Mitchell and Ghazvini's<sup>7</sup> detailed estimate of the cost of conducting assessments as part of Florida's QRIS; it not only describes the types of costs but also estimates the costs in Florida.

### ***Length of Observations***

Most instructions for classroom observational measures include guidance on the length of time necessary to conduct an observation. The authors of the ERS recommend a minimum of 3 hours to conduct an observation. Depending on the program schedule, the assessor may require more than 3 hours to observe particular items (e.g., observing both arrival and nap may require more than 3 hours). QRIS may want to provide guidance on the maximum length of time to spend in classrooms completing an ERS to provide some consistency in ratings across assessors so that some observations are not based on 3 hours and others based on 6. The authors of the CLASS recommend a minimum observation of 2 hours (four 30-minute observation cycles), but provide a scoring sheet that allows up to 3 hours of observation (or six 30-minute observation cycles).

### ***Frequency and Timing of Conducting Observations***

In QRIS, program observations are conducted as part of a program's overall rating. QRIS vary in terms of the length of time a rating is valid. For about half of the states, QRIS ratings are valid for 1 year. Other states have longer time periods between ratings (2 or more years), and a few states vary the length of time based on the quality level of the program.<sup>10</sup> QRIS administrators may also establish criteria—other than time—that would require re-examining a program's rating. The most common trigger is a licensing violation, but other triggers could include a change in ownership or location.<sup>10</sup> If observations are included in the rating system, then observations should be conducted every time the program is reviewed and assigned/reassigned a rating.

Timing is another issue in considering when to conduct program observations. The goal is to observe and rate the quality of care provided during a "typical" day. For child care centers and after-school programs, this may mean avoiding days that include special events, like holiday parties or field trips. For family child care programs, a "typical" day may be more challenging to define because of the variability in children served. Providers, for example, may care for only one baby one day of the week and five children (infants through preschoolers) other days. Family child care homes may also serve more school-age children during summer and school breaks. It is recommended that QRIS consider what is (and is not) "typical" for child care centers, family child care homes, and after-school programs and develop guidelines for determining the conditions under which a visit should be made.

### ***Using Multiple Observational Instruments***

Can the same person complete more than one measure during the same observation time in a classroom? The answer depends on the type of assessment instruments used. Here are some considerations when determining whether one person could complete multiple instruments during the same observation period (see footnote A).

- Do the instruments measure similar constructs (e.g., language and pre-literacy) in similar ways (e.g., rating scale)? If so, then possibly.
- Can the observation period be divided into segments so that the person completes one instrument during one period and completes the other instrument during another time period (e.g., completes the first instrument based on observations from 8:00 - 8:30 and 9:00 - 9:30 and completes the second instrument based on observations from 8:30-9:00 and 9:30 – 10:00)? If so, then possibly.
- Does at least one of the instruments require the person to code particular behaviors (e.g., teacher language) during a time frame (e.g., every 30 seconds)? If so, then it will probably be difficult to complete another instrument at the same time.

A recent review of observational measures used within QRIS suggests that most states currently include only 1 classroom observational instrument. Some states, like Missouri, include both the ECERS-R and the ECERS-E. These can reasonably be completed by the same person at the same time because they both measure relatively similar aspects of quality (although the ECERS-E more specifically measures curricula) and are scored similarly on a scale of 1 to 7.

Some states are using or considering using the CLASS in addition to the ECERS-R as part of their QRIS. The ECERS-R is based on a 3-hour observation that requires the person to attend to particular events (e.g., greeting), teacher language (e.g., types of questions the teacher asks children), and activities (e.g., reading to children). The CLASS requires the person to pay focused attention to everything during a 20-minute cycle, spend 10 minutes coding, and then conduct another 20-minute observation cycle. Although it is possible to have one person complete both instruments during the same observation period, the best results may be obtained if the instruments are used on different days or two observers rate the classroom on the same day with one rater focusing on ECERS-R and the other on CLASS.

## Reporting Observational Data

Once data are collected, QRIS must make decisions about how to report observational data to programs and teachers and other stakeholders. The specifics on reporting will depend on the instrument being used and the purpose of the assessment; however, there are a few principles to consider.

First, it is important to make sure that those receiving the data have the information needed to appropriately interpret the scores. It is typically not sufficient to simply share overall summary scores, without providing a context for what those scores mean. Summative reports should explain what was observed and how that led to a particular score. It is most helpful if these reports provide specific information about what was observed on the day the scores were made, rather than simply giving generalized descriptions.

It is also important to make sure that results are shared in a way that acknowledges the purpose and method with which they were obtained. For example, if QRIS sample one classroom per age level in a center, it is important that these data are shared at the center level. Individualized feedback to a teacher should only be given if their classroom was actually observed; this individualized feedback should not be generalized to other teachers in the center.

## Summary

Program observation is a common and important component of quality rating and improvement systems. Once the decision has been made to include observation as part of a program's rating, QRIS administrators must make a series of decisions regarding how to best collect data on the observed quality of programs. This Brief is intended to provide general guidance for making these decisions. There remains much to be learned, however, about how to most effectively observe in early childhood classrooms.<sup>11</sup> Current observational measures fall short on capturing all aspects of importance to high quality early childhood programs. The emergence of new observational measures in the next few years will help provide information about a broader range of program and classroom strengths and areas for growth.

## References

- <sup>1</sup>Halle, T., Vick Whittaker, J. E., & Anderson, R. (2010). *Quality in Early Childhood Care and Education Settings: A Compendium of Measures, Second Edition*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- <sup>2</sup>Harms, T., Clifford, R. M., & Cryer, D. (2005). *Early Childhood Environment Rating Scale: Revised edition*. New York: Teachers College Press.
- <sup>3</sup>Harms, T., Cryer, D., & Clifford, R. M. (2006). *Infant/Toddler Environment Rating Scale: Revised edition*. New York: Teachers College Press.
- <sup>4</sup>Harms, T., Cryer, D., & Clifford, R.M. (2007). *Family Child Care Environment Rating Scale: Revised edition*. New York: Teachers College Press.
- <sup>5</sup>Maxwell, K. (2008, June). *Overview of measurement issues in Quality Rating and Improvement Systems*. Presentation at Quality Rating and Improvement Systems: Creating the Next Generation of QRIS conference, Minneapolis, MN.
- <sup>6</sup>McCormick Tribune Center for Early Childhood Leadership National-Louis University. (2008, Spring). *When quality counts and money matters (Research Notes)*. Chicago, IL: Author. <http://cecl.nl.edu/research/rn.htm>.
- <sup>7</sup>Mitchell, A., & Ghazvini, A. (July, 2007). *Florida's Quality Rating System: A conceptual model for estimating cost assumptions, explanations, and supporting data*. Publication 2007-Q001a. The Policy Group for Florida's Children and Families.
- <sup>8</sup>Pianta, R. C., La Paro, K. M., & Hamre, B. (2008). *Classroom Assessment Scoring System (CLASS): Pre-K version*. Baltimore: Paul H. Brookes Publishing Co.
- <sup>9</sup>Thornburg, K., Mauzy, D., Mayfield, W., Scott, J., Sparks, A., Mumford, J., Foulkes, T., & Fuger, K. L. (In Press). Data-driven decision making in preparation for large scale QRS implementation. In M. Zaslow, I. Martinez-Beck, K. Tout & T. Halle (Eds.), *Quality measurement in early childhood settings*. Baltimore: Paul H. Brookes Publishing Co.
- <sup>10</sup>Tout, K., Starr, R., Soli, M., Moodie, S., Kirby, G., and Boller, K. (2010). *Compendium of Quality Rating Systems and Evaluations, OPRE Report*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.
- <sup>11</sup>Zaslow, M., Martinez-Beck, I, Tout, K., & Halle, T. (Ed.). (2011). *Quality Measurement in Early Childhood Settings*. Baltimore: Paul H. Brookes Publishing Co.
- <sup>12</sup>Zaslow, M., Tout, K., Halle, T., & Forry, N. (2009). *Multiple purposes for measuring quality in early childhood settings: Implications for collecting and communicating information on quality. OPRE Issue Brief #2*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services.

**Table 1. TRAINING OPTIONS FOR ERS and CLASS (as of September 2010)**

<b>Training to Use Tool to:</b>	<b>ERS</b> (see <a href="http://ers.fpg.unc.edu/">http://ers.fpg.unc.edu/</a> for updated information)	<b>CLASS</b> (see <a href="http://www.teachstone.org">www.teachstone.org</a> for updated information)										
<p><b>Gain Familiarity with Tool</b></p> <p><b>Target Audience:</b> Teachers, administrators, policymakers</p>	<p><b>BASIC TRAINING (LEVEL 1)</b></p> <p><b>Description:</b> Training participants will learn: how the scales measure quality; where and when the scales are used; to become somewhat familiar with the format and content of the scales; basic instruction in scoring; and procedures for completing an observation. No assessment of participant competence is completed.</p> <p><b>Length:</b> 1 day or less</p> <p><b>Cost:</b> varies</p> <p>Arrangements, including costs, available from Environment Rating Scales Institute, Inc.</p> <p>Contact Cathy Riley: email= <a href="mailto:criley968@gmail.com">criley968@gmail.com</a>, or call ERSI, Inc. at 919-338-2639</p>	<p><b>INTRODUCTION TO THE CLASS</b></p> <p><b>Description:</b> Gain an understanding of the CLASS framework, its various uses, and the teacher-child interactions that support children’s learning during this six-hour interactive program. No assessment of participant competence is completed.</p> <p><b>Length:</b> 6 hours (2 and 4 hour versions available upon request)</p> <p><b>Cost:</b> \$225 per participant. Class size is limited to 25 participants.</p> <p><b>LOOKING AT CLASSROOMS</b></p> <p><b>Description:</b> Learn to understand and identify the interactions that help children achieve at higher levels using this web-based, self-study program. No assessment of participant competence is completed.</p> <p><b>Length:</b> self-paced, approximately 20-30 hours</p> <p><b>Cost:</b> 6 month enrollment fee</p> <table border="1" data-bbox="917 1381 1453 1659"> <thead> <tr> <th><u># of enrollees</u></th> <th><u>Fee per subscriber</u></th> </tr> </thead> <tbody> <tr> <td>1 - 4</td> <td>\$175</td> </tr> <tr> <td>5 - 24</td> <td>\$155</td> </tr> <tr> <td>25 – 99</td> <td>\$135</td> </tr> <tr> <td>100 or more</td> <td>Site license available</td> </tr> </tbody> </table>	<u># of enrollees</u>	<u>Fee per subscriber</u>	1 - 4	\$175	5 - 24	\$155	25 – 99	\$135	100 or more	Site license available
<u># of enrollees</u>	<u>Fee per subscriber</u>											
1 - 4	\$175											
5 - 24	\$155											
25 – 99	\$135											
100 or more	Site license available											

<b>Training to Use Tool to:</b>	<b>ERS</b> (see <a href="http://ers.fpg.unc.edu/">http://ers.fpg.unc.edu/</a> for updated information)	<b>CLASS</b> (see <a href="http://www.teachstone.org">www.teachstone.org</a> for updated information)
<p style="text-align: center;"><b>Observe and Score Classrooms</b></p> <p><b>Target Audience:</b></p> <p>Observers or raters who will be using the tool to provide scores</p>	<p style="text-align: center;"><b>INTRODUCTORY TRAINING</b></p> <p><b>Description:</b> Consists of the Introductory Training, plus two days of guided practice observations. Includes Level 1 and 3-hour practice observations in an early childhood classroom with an experienced group leader and debriefing to discuss scores and interpretation of scale. No assessment of participant competence is completed.</p> <p><b>Length:</b> 3 days</p> <p><b>Cost:</b> \$1025</p> <p>Available from UNC – for more information visit <a href="http://ers.fpg.unc.edu/">http://ers.fpg.unc.edu/</a></p> <p style="text-align: center;"><b>INTENSIVE SUMMER INSTITUTE</b></p> <p><b>Description:</b> Consists of Levels 1 and 2, plus two days of lecture in which a select number of scale requirements are explained. Includes: Two 3-hour practice observations in an early childhood classroom with an experienced group leader; debriefing to discuss scores and interpretation of scale requirements; lecture on the interpretation of specific indicators, and what is required. No assessment of participant competence is completed.</p> <p><b>Length:</b> 5 days</p> <p><b>Cost:</b> \$1825</p> <p>Available from UNC – for more information visit <a href="http://ers.fpg.unc.edu/">http://ers.fpg.unc.edu/</a></p>	<p style="text-align: center;"><b>CLASS OBSERVER TRAINING</b></p> <p><b>Description:</b> This two-day training helps you build an in-depth understanding of the CLASS and teaches you how to use the tool to accurately observe and code classrooms. Explicit instruction and guided practice using authentic classroom videos help you prepare for the CLASS reliability test that follows this training. Those who pass the CLASS reliability test are certified to use the CLASS to observe and code classrooms for one year, after which recertification is required.</p> <p><b>Length:</b> 2 days, plus online reliability test</p> <p><b>Cost:</b> \$670 per participant plus materials (approximately \$60/person). Class size is limited to 15 participants.</p>

Training to Use Tool to:	ERS (see <a href="http://ers.fpg.unc.edu/">http://ers.fpg.unc.edu/</a> for updated information)	CLASS (see <a href="http://www.teachstone.org">www.teachstone.org</a> for updated information)
	<p><b>TRAINING TO RELIABILITY</b></p> <p><b>Description:</b> Consists of Levels 1 and 2, plus additional days of practice scoring followed by debriefing with authorized group leaders. Includes: Four 3-hour practice observations in an early childhood classroom with an experienced group leader; debriefing to discuss scores and interpretation of scale requirements; lecture on the interpretation of specific indicators, and what is required. Instruction in how to complete reliability training based on outcomes of participants is provided. Reliability scores are documented, and letters documenting reliability scores are available if desired.</p> <p><b>Length:</b> 5 days with option of additional days available if needed</p> <p><b>Cost:</b> Based on needs of agency or individuals requiring training.</p> <p>Arrangements, including costs, available from Environment Rating Scales Institute, Inc.</p> <p>Contact Cathy Riley: email= <a href="mailto:criley968@gmail.com">criley968@gmail.com</a>, or call ERSI, Inc. at 919-338-2639</p>	

<b>Training to Use Tool to:</b>	<b>ERS</b> (see <a href="http://ers.fpg.unc.edu/">http://ers.fpg.unc.edu/</a> for updated information)	<b>CLASS</b> (see <a href="http://www.teachstone.org">www.teachstone.org</a> for updated information)
<p style="text-align: center;"><b>Provide Technical Assistance or Professional Development</b></p> <p><b>Target Audience:</b> TA or PD staff who provide feedback to teachers and programs</p>	<p style="text-align: center;"><b>IN-DEPTH TRAINING</b></p> <p><b>Description:</b> Designed especially to enhance the skills of technical assistance (TA) providers using the scales to help improve the quality of children’s programs. The 8-day training includes: pretest to determine level of pre-course knowledge; lecture by authors and associates, on the interpretation of specific indicators, the research and rationale for requirements, what is required to meet requirements; explanations about how to best use the scales in efforts to maximize the effects of coaching, mentoring and technical assistance; and post-test to measure content knowledge.</p> <p>Pre-requisite is Level 1-2 or similar training, plus experience using the scale.</p> <p><b>Length:</b> 8 days</p> <p><b>Cost:</b> \$2000</p> <p>Available from UNC – for more information visit <a href="http://ers.fpg.unc.edu/">http://ers.fpg.unc.edu/</a></p>	<p style="text-align: center;"><b>MyTeachingPartner COACHING</b></p> <p><b>Description:</b> MyTeachingPartner (MTP) is a research-based, intensive and ongoing coaching program proven to boost effective classroom interactions and improve child outcomes. MTP coaches and teachers focus their work together on the teachers’ classroom video and the CLASS.</p> <p>Choose from the following MTP options:</p> <ol style="list-style-type: none"> <li><b>1. MTP Teacher Support</b> Teachers receive direct coaching focused on the CLASS from Teachstone’s expert MTP Coaches. In addition to this year-long, one-on-one support, Teachstone provides all teacher materials, video cameras, the MTP web interface, and a webinar-based kickoff event.</li> <li><b>2. MTP Coach Support</b> Coaches receive training and support from Teachstone as they implement MTP with your teachers. In addition to extensive training and yearlong support, Teachstone provides the MTP web interface, all coach and teacher materials, and video cameras.</li> </ol> <p><b>Length:</b> 10 months (initial training for coaches is 3 days, assuming coaches are already reliable on CLASS)</p> <p><b>Cost:</b> \$6,500 per participant for MTP Teacher Support</p> <p style="padding-left: 40px;">\$8,600 per participant for MTP Coach Support</p>

Training to Use Tool to:	ERS (see <a href="http://ers.fpg.unc.edu/">http://ers.fpg.unc.edu/</a> for updated information)	CLASS (see <a href="http://www.teachstone.org">www.teachstone.org</a> for updated information)
<p><b>Become a Certified Trainer</b></p> <p><b>Target Audience:</b> Trainers</p>	<p>Not currently available.</p>	<p><b>CLASS TRAIN-THE-TRAINER</b></p> <p><b>Description:</b> Learn how to provide the Introduction to the CLASS and the CLASS Observation Training to your colleagues during this three-day event. Deepen your content knowledge, learn tips for leading successful CLASS programs, practice leading discussions and training exercises, and receive feedback on your performance from an experienced CLASS trainer.</p> <p>CLASS Observation Certification is a prerequisite, and we recommend that, after becoming certified, participants observe and code a minimum of ten classrooms before attending this program.</p> <p><b>Length:</b> 3 days</p> <p><b>Cost:</b> \$2,100 per participant (includes materials required to conduct trainings). Class size is limited to 10 participants.</p>
<p><b>OTHER TRAINING OPTIONS</b></p>	<p><b>SUMMARY REPORT WRITING</b></p> <p><b>Description:</b> For participants who have been trained to reliability. Consists of 1-2 days training on how to write accurate and credible summary reports following an assessment. No assessment of participant competence is completed.</p> <p><b>Length:</b> 1-2 days with option of additional days available if needed</p> <p><b>Cost:</b> Based on needs of agency or individuals requiring training.</p> <p>Report reviewing by ERSI staff is also available. This provides review and correction to reports written by others to ensure appropriateness of information provided.</p> <p>Arrangements, including costs, available from Environment Rating Scales Institute, Inc.</p> <p>Contact Cathy Riley: email= <a href="mailto:criley968@gmail.com">criley968@gmail.com</a>, or call ERSI, Inc. at 919-338-2639</p>	<p><b>CONSULTATION: USING DATA TO GUIDE PROFESSIONAL DEVELOPMENT DECISIONS</b></p> <p><b>Description:</b> Collaborate with experienced CLASS consultants to match teacher needs to CLASS-based professional development programs. Interpretation and reporting of your aggregate CLASS data informs decisions that support effective teaching and learning.</p> <p><b>Length:</b> Project dependent</p> <p><b>Cost:</b> Dependent on scale and nature of consultation.</p>

Note: Information for this table was obtained from websites and confirmed by authors of the measures. However, readers should consult the websites referenced in the table to obtain the most current information.