

## Use of Observational Tools in QRIS

Observational tools are used in QRIS for the assignment of ratings and as a method for supporting programs' continuous quality improvement (CQI). Commonly used observational tools are supported by a research base with established protocols for tool administration. This fact sheet provides a descriptive analysis of how observational tools are being incorporated into the 41 QRIS.

This fact sheet is one of a series about the state of quality rating and improvement systems (QRIS) in the United States. In 2016, there were 41 fully operational QRIS in the United States.<sup>1</sup> Data are from the QRIS Compendium (<http://QRIScompendium.org>), a catalog of the QRIS operating in the United States as of October 31, 2016.<sup>2</sup>

## Common Observational Tools Used in QRIS Rating Processes

Use of observational tools has become nearly universal since 2010.

- ◆ In 2010, most QRIS (76 percent) were using observational assessment as part of the QRIS rating process.
- ◆ In 2016, 38 QRIS (93 percent) incorporated observational tools into their indicators for all types of participating programs.

The environment rating scales (ERS)<sup>3</sup> and the Classroom Assessment and Scoring System (CLASS)<sup>4</sup> are the most frequently used tools as seen in Figure 1.

- ◆ Thirty-one QRIS (76 percent) use ERS and 18 QRIS (44 percent) use CLASS.
- ◆ Thirty-one QRIS use multiple tools, with CLASS and ERS being the most common combination (15 use this combination; 37 percent).
- ◆ Eleven QRIS (27 percent) use other types of assessment tools, including self-developed tools (3 QRIS; 7 percent) and the Program Quality Assessment<sup>5</sup> (2 QRIS; 5 percent).

### Environment Rating Scale Instruments

- The Early Childhood Environment Rating Scale-Revised (ECERS-R)
- The Infant/Toddler Environment Rating Scale-Revised (ITERS-R)
- The Family Child Care Environment Rating Scale-Revised (FCCERS-R)
- The School-Age Care Environment Rating Scale (SACERS)

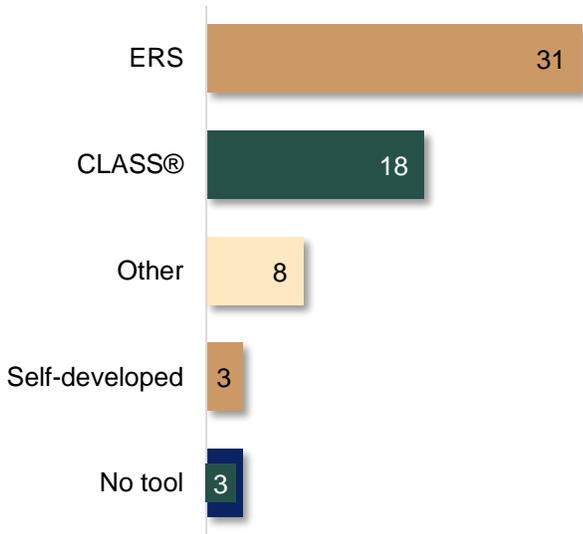
<sup>1</sup> State with a QRIS: AR, AZ, CA, CO, DE, FL (3 localities), GA, IA, ID, IL, IN, KY, MA, MD, ME, MI, MN, MS, MT, NC, ND, NE, NH, NJ, NM, NV, NY, OH, OK, OR, PA, RI, SC, TN, UT, VA, VT, WA, WI. While most QRIS operate at the state level, three represent separate counties in Florida (Duval, Miami-Dade, and Palm Beach). The California QRIS, while represented in the Compendium as one system, is implemented at the county level and does not include all counties in the state.

<sup>2</sup> Data compiled in 2014 and 2015 from <http://QRIScompendium.org> are used to show changes in the number of QRIS with specific features. These data are supplemented by an earlier version of the Compendium titled *Compendium of Quality Rating Systems and Evaluations*, which was released in 2010. Retrieved from <https://www.acf.hhs.gov/opre/resource/compendium-of-quality-rating-systems-and-evaluations>. Publicly available details about QRIS standards and implementation and interviews with QRIS administrators also inform the fact sheet.

<sup>3</sup> Harms, T., Clifford, R. M., & Cryer, D. (2005).

<sup>4</sup> Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008).

**Figure 1: Number of QRIS Using Observational Assessment Tools, 2016**



Note: N = 41 QRIS. Some are in multiple categories.

## Tools Used With Centers

Not all QRIS use observation tools with all types of participating programs, although ERS and CLASS remain the most common tools for centers.

- ◆ Of the 41 QRIS that are using observational tools in centers, 31 (76 percent) use ERS tools (either ECERS-R, ECERS-3, or ITERS-R) while 18 (44 percent) use CLASS.
- ◆ Nineteen QRIS (46 percent) use multiple tools to assess centers participating in the QRIS, with 14 QRIS (34 percent) using both ERS and CLASS.

## Tools Used With FCC

Thirty-four QRIS (83 percent) use observational tools to assess, rate, and provide quality improvement assistance to family child care homes (FCC). ERS are the most common, used by 29 QRIS (71 percent).

- ◆ Eight QRIS (20 percent) use CLASS with FCC.
- ◆ Nine QRIS (22 percent) use multiple tools; six QRIS (15 percent) use both CLASS and ERS.
- ◆ Eight QRIS (20 percent) use other types of assessment tools, including the Caregiver Interaction Scale<sup>5</sup> and self-developed tools.

<sup>5</sup> HighScope Educational Research Foundation. (2017).

<sup>6</sup> Arnett, J. (1985).

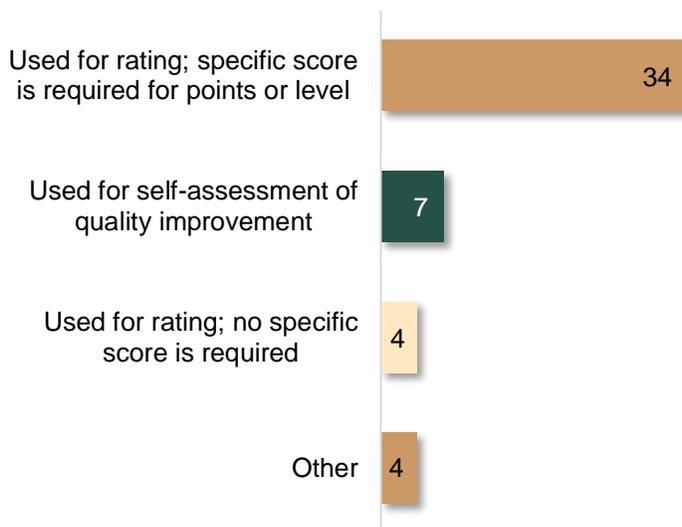
## Purpose of Observational Tools

The majority of QRIS (34; 83 percent) use observational tools in the rating process with centers by assigning points or levels to programs if they attain or exceed specific scores.

### Purpose of Observations in Centers

- ◆ Seven QRIS (17 percent) use observational tools in centers for self-assessment purposes and/or quality improvement for centers participating in the QRIS.
- ◆ Eight QRIS (20 percent) use observational tools either in the rating process with no specific score required or for other reasons.

**Figure 2 : Purpose of Observational Tools for Centers, 2016**



Note: N = 41 QRIS. Some are in multiple categories.

### Purpose of Observations in FCC

- ◆ For FCC, 29 QRIS (71 percent) use observational tools in the rating process by assigning points or levels to programs if they attain or exceed specific scores.
- ◆ Five QRIS (12 percent) use observational tools for self-assessment purposes and/or quality improvement.
- ◆ Seven QRIS (17 percent) use observational tools either in the rating process with no specific score required or for other reasons.

## Classroom Selection Processes

QRIS that use observational tools typically adhere to standardized processes for selecting which classrooms will receive observations in programs with multiple classrooms for each age group. Selection of classrooms is typically a concern for centers, which are more likely to have multiple classrooms, but it can also be applicable to larger FCC providers. The most common method used is random selection of classrooms after certain parameters are met. This is true for both CLASS (13 QRIS; 32 percent) and ERS tools (21 QRIS; 51 percent).

## Age Groups Observed

The administration of observational tools varies across systems. QRIS typically observe a subset of classrooms in centers; in some cases not all age groups are observed. These decisions are made based on administration cost, QRIS participation rules, the ages of children served, and whether tools are appropriate or validated for use with particular age groups. Most commonly, QRIS randomly select classrooms to receive observations after certain parameters are met (e.g., that a certain percentage of classrooms or at least one classroom is observed for each age group).

- ◆ For centers, 38 QRIS (93 percent) use observational tools to observe preschool-age environments and 37 QRIS (90 percent) observe infant and toddler environments.
- ◆ Nineteen QRIS (46 percent) use observational tools for school-age children in centers.
- ◆ For FCC, 35 QRIS (85 percent) observe preschool-age environments and 34 (83 percent) observe infants and toddlers.
- ◆ Fourteen QRIS (34 percent) use observational tools for school-age children in FCC.

## Observer Training

QRIS use similar methods to train observers on how to use the tools. The most common training method among QRIS that use observational tools is to have an official tool anchor to train staff in person (which is done by 30 QRIS; 73 percent). An anchor is a trained observer who meets 90 percent reliability for the ERS and 80 percent for CLASS and who can be used for reliability testing of other observers. The remaining eight QRIS (20 percent) use a combination of methods, including online training and using a CLASS trainer who is not an anchor.

## Observer Reliability

QRIS use different reliability processes to ensure that the observers administering the tools are doing so in a consistent manner. The most common method used to ensure reliability is the one recommended by the tool's author.

- ◆ For ERS, 27 QRIS (66 percent) are using the authors' recommended reliability tests, which is 85 percent agreement with consensus scores.
- ◆ For CLASS, 15 QRIS (37 percent) are using the recommended initial reliability, which is 80 percent agreement with consensus.

The most common ongoing reliability check, used by 16 QRIS (39 percent) occurs between every 6 and 10 visits.

## QRIS Showcase: Pennsylvania Keystone STARS

Pennsylvania's QRIS, Keystone STARS, is a four-level block system that has been operating since 2002. Since its inception, Keystone STARS has used environment rating scales to assess infant/toddler, preschool, and school-age classrooms in center-based and home-based programs. Selection of ERS tools happened prior to implementation of the system. The State chose the ERS tools because of their all-encompassing approach to assessment and inclusion of interactions, physical space, and materials, among many other facets of quality. In 2014, the QRIS began using the newest version of the center-based preschool tool, the *Early Childhood Environment Rating Scale, Third Edition*.

Keystone STARS incorporates the ERS tools early on in a program's initial participation. Programs are given the opportunity to attend professional development events, either online or in person, where they receive an overview of the assessment. At this point, they have an initial chance to think through which assessment areas they would like to focus on in their programs. Programs that have completed the training then go through a self-assessment process to further refine their goals. Home-based providers complete this self-assessment process by partnering with another provider to identify and meet their quality improvement goals. Center-based programs typically have teachers observe one another in their classrooms and they use this information, in collaboration with the center director, to set goals.

Programs at the one- and two star levels are only required to use the ERS for self-assessment, although they have the option to voluntarily have an assessment completed by an external observer. Those programs that do decide to voluntarily participate are given detailed reports about their programs' strengths and areas for improvement. They can then use this information to improve quality and work toward higher star levels within the system. External assessments by trained observers are required for programs going for the three and four star levels. For three star programs, a 4.0 average score across all subscales of the ERS is required, while four star programs must achieve a 5.0. Programs can receive up to two assessments in a 12-month period. This allows programs that do not meet the cut-off score on their initial assessment to receive technical assistance to improve and then try again with another assessment, all within a 1-year period.

One of the strengths of Keystone STARS assessment model is its continuous use of one assessment over a long period. This provides a wealth of data that show where programs are on a continuum of best practices over many assessment cycles. STARS staff report that access to this kind of information is valuable for many programs because it helps them identify how factors, like staff turnover or use of new curricula, affect scores. Keystone STARS is undergoing a revision of its standards as of spring 2017. The use of observational assessment tools in a future version of Keystone Stars will likely align with the overall goal of the QRIS: to remove barriers for programs to improve and maintain high levels of quality and provide a research-based process for programs to systematically improve outcomes for children.

## References

- The Build Initiative & Child Trends. (2016). A catalog and comparison of quality rating and improvement systems (QRIS) [Data system]. Retrieved from <http://qriscompendium.org/>
- Mathematica Policy Research & Child Trends. (2010). *Compendium of quality rating systems and evaluations*. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. Retrieved from <https://www.acf.hhs.gov/opre/resource/compendium-of-quality-rating-systems-and-evaluations>
- Harms, T., Clifford, R. M., & Cryer, D. (2005). *Early childhood environment rating scale*. New York, NY: Teachers College Press. Retrieved from <http://ers.fpg.unc.edu/>
- Pianta, R. C., La Paro, K. M., & Hamre, B. K. (2008). *Classroom assessment scoring system*. Baltimore, MD: Paul H. Brookes Publishing Company, Inc. Retrieved from <https://curry.virginia.edu/classroom-assessment-scoring-system>
- HighScope Educational Research Foundation. (2017). *Preschool program quality assessment*. Red-e Set Grow, LLC. Retrieved from <https://highscope.org/assessment/program/details>
- Arnett, J. (1985). *Caregiver Interaction Scale*. Educational Testing Service.

*This document was developed with funds from Grant # 90TA0002-01-00 for the U.S Department of Health and Human Services, Administration for Children and Families, Office of Head Start, Office of Child Care, and Health Resources and Services Administration, by the National Center on Early Childhood Quality Assurance in cooperation with Child Trends. This resource may be duplicated for noncommercial uses without permission.*

*The National Center on Early Childhood Quality Assurance (ECQA Center) supports State and community leaders and their partners in the planning and implementation of rigorous approaches to quality in all early care and education settings for children from birth to school age. The ECQA Center is funded by the U.S. Department of Health and Human Services, Administration for Children and Families.*

### National Center on Early Childhood Quality Assurance

9300 Lee Highway  
Fairfax, VA 22031

Phone: 877-296-2250  
Email: [QualityAssuranceCenter@ecetta.info](mailto:QualityAssuranceCenter@ecetta.info)

**Subscribe to Updates**  
[http://www.occ-cmc.org/occannouncements\\_sign-up/](http://www.occ-cmc.org/occannouncements_sign-up/)



ADMINISTRATION FOR  
**CHILDREN & FAMILIES**