THE CHILD CARE QUALITY RATING SYSTEM (QRS) ASSESSMENT

Defining and Measuring Quality: An In-Depth Study of Five Child Care Quality Rating and Improvement Systems

OPRE Report 2011-29

August 2011

Defining and Measuring Quality: An In-Depth Study of Five Child Care Quality Rating and Improvement Systems

OPRE Report 2011-29
August 2011

Mathematica Policy Research
Pia Caronongan
Gretchen Kirby
Lizabeth Malone
Kimberly Boller

Prepared for:
Office of Planning, Research and Evaluation
Administration for Children and Families
Department of Health and Human Services
370 L'Enfant Plaza Promenade, SW
7th Floor West, Room 7A011
Washington, D.C. 20447

Project Officers:
Ivelisse Martinez-Beck
Kathleen Dwyer

Prepared by:
Mathematica Policy Research
600 Maryland Ave., S.W., Suite 550
Washington, DC 20024-2512

Project Director:
Gretchen Kirby, Mathematica

Co-Principal Investigators:
Kimberly Boller, Mathematica
Kathryn Tout, Child Trends

This document was prepared under Contract #HHSP233200800394G with the Administration for Children and Families, U.S. Department of Health and Human Services. The views expressed in this publication are those of the authors and do not necessarily reflect 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. This report and other reports sponsored by the Office of Planning, Research and Evaluation are available at http://www.acf.hhs.gov/programs/opre/index.html.

Suggested citation: Caronongan, P., Kirby, G., Malone, L., Boller, K. (2011). *Defining and Measuring Quality: An In-Depth Study of Five Child Care Quality Rating and Improvement Systems*. OPRE Report # 2011-29. Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation.

OPRE
Office of Planning, Research and Evaluation
Administration for Children and Families

MATHEMATICA
Policy Research

Child TRENDS

ACKNOWLEDGMENTS

This report was produced as part of the QRS Assessment Project, funded by the Office of Planning, Research and Evaluation in the Administration for Children and Families, U.S. Department of Health and Human Services. We are grateful to our federal project officers, Ivelisse Martinez-Beck and Kathleen Dwyer, who provided valuable guidance and feedback pertaining to all aspects of this work. We also benefited from the knowledge shared by members of the QRS Assessment expert panel, which helped to shape the questions and approach for the in-depth study of quality measurement. This expert panel included Sheri Azer Fischer, J. Lee Kreader, Kelly Maxwell, Deb Swenson-Klatt, Kathy Thornburg, and Gail Zellman as well as members of the project team, Shannon Christian and Karen Tvedt from Christian and Tvedt Consulting.

We would like to thank other members of the QRS Assessment team, Kathryn Tout at Child Trends and Heather Zaveri at Mathematica, for their thoughtful contributions during various stages of the in-depth study. Other colleagues at Mathematica assisted in the preparation and production of this report. Anastasia Erbe and Chris Pefaure helped to code and organize data collected from interviews and documents. Amanda Bernhardt and Jane Retter provided editorial support. Kari Beckmann and Eileen Curley provided administrative and secretarial support.

Finally, this report would not have been possible without the generosity of QRIS administrators, staff, planners and evaluators who shared their time and insights, provided extensive documentation, and patiently fielded multiple follow-up questions from our research team. We would especially like to thank the individuals in each of the five states who were gracious with their time in arranging our site visits as well as their knowledge to support this work. These include: in Miami-Dade, Pam Hollingsworth of the Early Learning Coalition of Miami-Dade/Monroe and Jesse Leinfelder and Lisa Pittman of the Children's Trust; in lllinois, Megan Fitzgerald and Holly Knicker of the Illinois Department of Human Services; in Indiana, Melanie Brizzi and Janet Deahl of the Bureau of Child Care in the Family and Social Services Administration; in Pennsylvania, Debi Mathias, Catherine Cormany, and Philip Sirinides of the Office of Child Development and Early Learning, and Gail Nourse of the Pennsylvania Key; and in Tennessee, Barbara Wall and Gary Smith of the Tennessee Department of Human Services and Bingham Graves with the University of Tennessee-Knoxville.

THIS PAGE LEFT BLANK FOR DOUBLE- SIDED PRINTING

CONTENTS

III *(continued)*

TABLES

EXHIBIT

FIGURE

THIS PAGE LEFT BLANK FOR DOUBLE- SIDED PRINTING

EXECUTIVE SUMMARY

Quality measurement serves as a foundation for child care Quality Rating and Improvement Systems (QRIS). Understanding the variation that exists in quality measurement, the different contexts in which states make decisions, and the factors that play into their decisions can help administrators identify where their state may fall along the spectrum in practice (or the direction in which they should steer), and can help researchers design approaches that take these differences into account. Such information can also aid in identifying opportunities for moving toward some common practice and research goals.

The quality components included in a QRIS define a state's framework for measuring quality and signal to providers and parents the practices that should be included in high-quality early child care and education programs. There are commonalities in the quality categories that are included across QRIS (Tout et al. 2010), demonstrating that many states and communities are using a similar foundation upon which to build their rating systems. However, the manner in which states and localities combine and aggregate these quality categories to develop QRIS ratings has many nuances, producing rating systems with important variations that can impede direct cross-QRIS comparisons and research approaches.

Recognizing the need for information on the quality measurement practices in QRIS, this in-depth study of select QRIS was launched as part of the Child Care Quality Rating Systems (QRS) Assessment project, funded by the Office of Planning, Research, and Evaluation (OPRE) within the Administration for Children and Families.[1] We examined the approaches used by states and communities to measure quality through the QRIS by focusing on three research questions:

1. What is the variation in how select QRIS define and measure quality, and what accounts for the variation in their approaches?
2. What are the specific processes used by select QRIS to measure each component of the quality rating and determine the overall rating level?
3. What is the availability of consistent and reliable data on quality ratings within select QRIS and how are the data currently being used?

To answer these questions, we selected five QRIS for the in-depth study: Miami-Dade County, Florida; Illinois; Indiana; Pennsylvania; and Tennessee. A summary of key characteristics of the five QRIS is presented in Table ES.1.

[1] To conduct the QRS Assessment project, in 2008 OPRE contracted with Mathematica Policy Research, Child Trends, and Christian and Tvedt Consulting. The goals of this project are to (1) gather and analyze existing and new information on QRIS implementation and research to inform decision making on QRIS development and refinement and (2) build the capacity for ongoing monitoring and evaluation within and across systems.

Table ES.1. Overview of QRIS Participating in the In- Depth Study of Quality Measurement

	Miami-Dade County	Illinois	Indiana	Pennsylvania	Tennessee
QRIS Name	Quality Counts	Quality Counts	Paths to Quality	Keystone STARS	Star-Quality Child Care Program
Starting Year of Statewide Implementation	2008	2007	2008	2003	2001
Number of Rating Levels	5	4	4	4	3
Structure of Rating Levels	Combination	Building blocks	Building blocks	Building blocks	Combination
Eligible Programs					
Center-based	✓	✓	✓	✓	✓
Head Start/ Early Head Start	✓	✓	✓	✓	✓
Pre-kindergarten	✓	✓	n/a	✓	✓
Licensed FCC	✓	✓	✓	✓	✓
License-exempt Homes	✓	✓			
School-aged Programs	✓	✓	✓	✓	✓
Other			Child care ministries		
Total Number of Participating Programs	430	1,030	2,040	4,420	2,749
Percentage of Participants in Each Level	Level 1: 13% Level 2: 29% Level 3: 35% Level 4: 19% Level 5: 4%	Level 1: 10% Level 2: 30% Level 3: 60% Level 4: 1% Tier 1: 52%[a] Tier 2: 22% Tier 3: 27%	Level 1: 61% Level 2: 15% Level 3: 13% Level 4: 10%	Level 1: 46% Level 2: 29% Level 3: 12% Level 4: 14%	Level 1: 2% Level 2: 19% Level 3: 61%
Total Number of Children Served	28,000 (as of July 2010)	43,465 (as of April 2011)	75,993 (as of May 2011)	168,530 (as of June 2010)	Not available

Source: Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); QRS Data (Illinois Department of Human Services, May 2011); Keystone STARS 2010 Program Report (OCDEL, 2010); Tennessee Report Card & Star Quality Program Year 8 Annual Report (Pope and Magda 2010); Paths to QUALITY, Monthly Management Report (FSSA, 2011); Trends from Miami-Dade's QRIS (ELC,2010).

n/a = not applicable

[a] License-exempt homes have a separate three-tier system in Illinois.

Quality Definitions, Thresholds, and Rating Criteria

Each of the five QRIS looked to existing systems to inform the design of their program, but they did not adopt another existing QRIS as a whole. Instead, the design of each system was influenced by characteristics of the local early education and care market, as well as existing licensing and accreditation standards.

Selecting Quality Rating Components

Although there is some overlap in quality components included in ratings, we found considerable variation in the specificity and rigor of indicators for each component.

- **Licensing.** The role of licensing depends on the perceived rigor in the licensing requirements and the maturity of the QRIS. Licensing compliance is either a complete or partial requirement at level one, or a prerequisite for participation. Only Miami-Dade does not include licensing compliance as a requirement at any level. To ensure a level playing field for all providers, four of the five QRIS require license-exempt centers to obtain a license in order to participate in their QRIS.
- **Ratio, group size, and health and safety indicators.** QRIS requirements for child-staff ratio, group size, and health and safety were influenced by the licensing requirements in each state. Two of the five QRIS include additional requirements beyond what is required for a licensed provider for child-staff ratio and group size in their QRIS standards for center-based programs in order to bring providers in line with accreditation standards by the time they reach the highest rating level.
- **Staff qualifications.** All five QRIS incorporate staff qualifications into their ratings. Education and training are the most typical indicators used but there is wide variation in how requirements are defined at each level. For example, the number and specificity of requirements vary by position type as does the percentage of staff who must meet requirements.
- **Administration and management.** Quality indicators in the area of administration and management cover two main topics—staff management (such as staff benefits, annual professional development plans for staff, and the use of differentiated salary scales based on education and experience) and program administration (such as risk and fiscal management, program evaluation, and strategic planning). Systems vary in the specificity of requirements as well as the rating level at which each is required.
- **Family partnerships and community involvement.** All five QRIS include requirements for family partnerships and community involvement; however, they typically rely on self-reported information from providers. Communication with families is a common indicator across all five systems but there is great variation in the modes specified and frequency required.
- **Environment.** Four of the five QRIS use the Environment Rating Scales (ERS; Harms et al. 1995, 2005, 2006, 2007), citing their wide use and recognition in the field. Indiana does not use the ERS scales in their entirety, but includes some items similar to ERS items in their rating tools. The four QRIS that use the ERS integrate scores into quality ratings by setting a minimum score that providers must meet to qualify for a particular rating level in building block systems or receive a number of points in combination systems. At the highest rating level, all four QRIS require a score of 5.0 or higher which aligns with the "good" range on the scale. There is greater variation among QRIS in the minimum ERS scores required at lower levels.
- **Individualization of services**. This group of components reflects the extent to which providers tailor or individualize services to meet the needs of children and families by using child assessments, provisions for special needs, developmentally appropriate curricula, and practices that respond to and recognize cultural and linguistic diversity.

With few exceptions, standards for these components are included in the ratings at higher levels across QRIS, indicating that these features are not necessarily expected of a provider demonstrating a baseline level of quality.

- **Accreditation.** Across the five QRIS, respondents perceived that accreditation represents the high end of the child care quality spectrum. Two QRIS require accreditation to reach the highest QRIS level. While accreditation is required of providers at level four in Indiana, providers must also undergo an observational assessment and demonstrate that they meet all standards of the lower levels. Other QRIS chose not to make accreditation a requirement but use alternative ways to incorporate accreditation status into ratings. In Pennsylvania, accreditation fulfills partial requirements at level four. Miami-Dade and Tennessee QRIS do not include accreditation as a requirement at any level, but award additional points to accredited providers.

Laying Out the Structure of Quality Rating Levels

- **Methods for combining indicators.** The five QRIS combine and aggregate components using either a building block approach in which a provider must meet all of the standards required at one level before moving on to the next or combination systems, wherein a provider is rated on individual components before scores are combined to obtain an overall rating. Respondents in QRIS using a building block approach noted that this structure allows for a clear and consistent representation of how each level is defined. Conversely, planners from the QRIS that employ a combination system indicated that it was necessary to provide multiple avenues to achieve a higher rating, while still prioritizing what they felt were the most important elements.
- **Number of rating levels.** Planners' and administrators' knowledge of licensing and accreditation standards served to establish a range of quality for the QRIS to cover. They designed the levels of the QRIS to help providers progress from licensing requirements (at the base) to standards that are largely equivalent to accreditation (at the top). The number of intermediate levels in each QRIS was influenced by what planners and administrators felt were reasonable expectations in terms of improvements that providers could achieve over time and supports that could be provided to help providers make that progress.
- **Terminology for levels.** Respondents in each QRIS indicated that they devoted a substantial amount of thought and discussion during QRIS planning to what ratings should be called, such as stars or levels. Four of the QRIS use the term "stars" because the term denotes a certain level of prestige and accomplishment that they want providers to associate with being a QRIS participant. Indiana's QRIS uses the term "paths" to place more emphasis on the value of the quality improvement process.

Processes for Quality Measurement

The next step in our investigation of quality measurement was an analysis of the processes that sites implement to collect information on each component and its indicators, and to determine the overall rating level.

Pre-rating Process

- All five QRIS require attendance in overview sessions in which information is provided about the goals of QRIS, the system standards, expectations of participants, and resources and supports available to help providers at each stage of the process.
- Each site has a preparation process in place to help providers learn about QRIS standards and gauge whether they are ready to undergo the rating process. During this stage, a range of supports are offered to providers including self-study materials such as workbooks and worksheets to help understand the standards and conduct self-assessments, additional training sessions, and individualized technical assistance provided through consultation with QRIS specialists.

Gathering Evidence for Individual Components

- The first step in the formal rating process is an evidence review to determine whether a provider meets requirements for individual quality components. Across the QRIS, raters are distinct staff from QRIS specialists who perform the pre-rating and supportive roles with providers.
- Rating teams vary substantially in size due to the nature and amount of work that raters are required to do in each site.
- All five QRIS require raters to have a bachelor's degree and three require that this degree be in early childhood education or a related field. None of the five QRIS has a formal protocol for training new raters or firm guidelines for initial and ongoing reliability. However, three have developed materials to improve the consistency of the evidence review process.
- Across QRIS, raters review evidence for at least 2 and as many as 10 components for each provider. Evidence is usually obtained through direct observation, director/provider interview, document review (the most common method), or a combination of the three.
- The required evidence for some components is fairly straightforward—for example, providers need only present current certificates to demonstrate licensing compliance and accreditation status. Other components can require extensive effort or documentation. For example, demonstrating staff qualifications requires access to and review of education and training documentation for numerous individual staff members.

Conducting Assessments Using Standardized Measures

- Four of the QRIS assess the quality of the environment using the ERS. Illinois also uses the Program/Business Administration Scales (Talan and Bloom, 2004, 2009) to assess a number of other components. Indiana does not assess any components using standardized measures but includes some observational indicators in their quality rating tool. (The rest of this section focuses on ERS and thus Indiana is not included.)
- The number of assessors per QRIS ranges widely from 7 assessors in Illinois to nearly 60 assessors in Tennessee. The workload for assessors is similar across sites, with assessors conducting between 8 to 12 assessments per month.

- Assessment teams include lead assessors who supervise groups of assessors, train new assessors, and serve as anchors. Anchors' ratings serve as the benchmark upon which the ratings of other assessors are compared for consistency. Lead assessors also conduct assessments themselves, albeit with a smaller caseload.
- Assessors are required to have a minimum of a bachelor's degree, and are preferred to have this degree in early childhood education. In addition to education level, two QRIS require experience in early childhood settings and Miami-Dade requires assessors to be bilingual due to the demographics of the providers and families in their locality. Miami-Dade and Pennsylvania also gauge the writing skills of assessor candidates.
- Each QRIS has built upon publisher-provided materials and guidelines to design protocols for conducting training sessions in-house. The basic parameters of the training process are similar across sites. Few of the current assessors in the five QRIS have received direct training from the authors of the ERS. In lieu of sending all assessors for training with ERS authors, three QRIS have sent at least some of their lead assessors to receive training.

Procedures for Conducting Classroom Observations

- Information on children's ages is used to determine which measures will be used for the observation. In mixed age classrooms, sites typically use the measure appropriate for the age of the majority of children in the room.
- Three sites do not inform providers of the exact date of the visit and instead give providers a window of three to four weeks during which they can expect the assessment visit to take place, and allow providers to designate blackout dates during which they cannot be observed due to scheduling conflicts.
- Assessors typically observe one-third of the classrooms for each age group served and conduct at least one assessment for each age group. In the case of multiple classrooms, the classrooms observed are selected randomly on the morning of the assessments. There are additional guidelines for selecting classrooms for observation. For example, three of the QRIS require that at least half of enrolled children are present in a particular classroom, two QRIS exclude classrooms that are staffed by a substitute teacher, and two QRIS exclude classrooms if the teacher is new.
- To calculate facility ERS scores, Illinois and Pennsylvania take the average score across all classrooms and scales administered. Tennessee also calculates an average across classrooms. However, if any individual classroom receives an ERS score below 3.0, the entire facility assumes that classroom's score. Miami-Dade produces separate averages for each scale administered (such as an ECERS average and an ITERS average).

Assigning Component and Final Ratings

- In Indiana and Miami-Dade, component ratings are automatically calculated in QRIS databases based on data entered by the rater and/or assessor for individual indicators. In Illinois, Pennsylvania, and Tennessee, raters manually calculate ratings for each component.
- Miami-Dade, Indiana, and Tennessee have the calculation of overall ratings automated in their QRIS databases; that is, based on individual component ratings, the database

automatically calculates the overall rating. In Illinois and Pennsylvania, raters determine overall ratings manually by reviewing individual component ratings.

Data Collection, Use, and Analysis to Refine Quality Measurement in QRIS

Availability of Data on Quality Measurement Ratings

- At a minimum, each QRIS database stores information on current and historical quality ratings.
- Beyond the ratings, all QRIS databases store component-level ratings for at least some quality rating components. Three QRIS databases store indicator-level data.

Use of Data to Monitor and Evaluate QRIS

- Administrators examine distributions of quality ratings at least annually to examine how QRIS participants are progressing. Administrators also examine data at the component or indicator levels to identify areas showing substantial progress and areas where large numbers of providers tend to underperform.
- Several sites monitor the supports that providers access in preparation for the rating process. Respondents discussed plans to eventually link these data to quality improvements made over time to determine which components have required the most support from QRIS Specialists.
- Of the sites we visited, only Indiana had a study already in process to compare developmental outcomes of children in the care of providers with varying quality rating levels. Respondents in other QRIS also expressed an interest in examining relationships between quality ratings and child outcomes but noted that the cost of conducting child assessments was prohibitive.
- Respondents noted that the scarcity of time and resources and the need for better integration and more detailed information on quality and outcomes are the key challenges in using the available data for research and evaluation purposes.

Research Directions

This in-depth study describes what is conceptualized as quality and how it is measured in five QRIS. Although the five QRIS profiled in this report incorporate a greater number of components in quality ratings than earlier iterations of QRIS, there remain many unanswered questions about which quality components to include, and how, within the rating systems.

In terms of quality measurement processes, we found greater consistency in the administration of the ERS across QRIS than in the procedures for gathering evidence on other quality components or calculating ratings. Nonetheless, there continue to be threats to the reliability of standardized assessments including limitations in the number of assessors trained directly by authors of the measures and inconsistencies in the number of classrooms observed. The measures of other quality components present challenges to consistent, reliable data collection and interpretation. Multiple modes of data collection—such as observation, interview, and document review—could serve to confirm the presence of quality components (and increase reliability) but would likely introduce tradeoffs in terms of cost.

Among the five QRIS studied, there is generally greater consistency in the definitions of the quality components at the highest rating levels than at the baseline levels. Cut-off points at intermediate levels are somewhat arbitrarily determined. Whether differences between providers at each level would translate to meaningful differences in child outcomes is an open question. At the highest level, QRIS standards overlap considerably with recommendations of accrediting organizations such as the National Association for the Education of Young Children. Further research may help shed light on whether features specified for the highest level are consistent with quality thresholds that have been linked to positive outcomes for children.

I. INTRODUCTION

Quality measurement serves as a foundation for child care Quality Rating and Improvement Systems (QRIS). The quality components included in a QRIS define a state's framework for measuring quality and signal to providers and parents the practices that should be included in high-quality early child care and education programs. Descriptive studies of QRIS indicate that providers focus their efforts on improving the components that are measured by the rating system and give less attention to program areas that are not measured (Zellman and Perlman 2008; Thornburg 2008).

There are commonalities in the quality categories that are included across QRIS (Tout et al. 2010), demonstrating that many states and communities are using a similar foundation upon which to build their rating systems. However, the manner in which states and localities combine and aggregate these quality categories to develop QRIS ratings has many nuances, producing rating systems with important variations that can impede direct cross-QRIS comparisons and research approaches. A cross-QRIS descriptive study focusing on five pioneer QRIS states found that administrators used a similar (albeit limited) research base to inform the development of quality standards, but that additional factors such as feasibility and cost, as well as values and goals, ultimately contributed to decisions about the content of their ratings (Zellman and Perlman 2008).

Understanding the variation that exists in quality measurement, the different contexts in which states make decisions, and the factors that play into their decisions can help administrators identify where their state may fall along the spectrum in practice (or the direction in which they should steer), and can help researchers design approaches that take these differences into account. Such information can also aid in identifying opportunities for moving toward some common practice and research goals. Recognizing the need for information on the quality measurement practices in QRIS, this in-depth study of select QRIS was launched as part of the Child Care Quality Rating Systems (QRS) Assessment project, funded by the Office of Planning, Research, and Evaluation (OPRE) within the Administration for Children and Families.[2]

After just a decade of QRIS implementation, suggestive findings from descriptive and validation studies about the potential influence of QRIS on quality are encouraging but limited, and evidence from rigorous research about the effectiveness of QRIS on quality and child outcomes is entirely lacking. Yet, in an era of evidence-based practice, administrators are facing increasing pressure to justify the time and expense of QRIS by demonstrating quality improvement and further connecting quality improvement in programs to improved child outcomes. Going beyond QRIS to quality improvement initiatives as a whole, there is limited evidence of a large and potentially meaningful association between quality and child outcomes (Burchinal et al., 2009). However, the lack of compelling evidence may reflect gaps in the ability of research to measure what many policymakers, practitioners, and researchers alike still consider a meaningful relationship between quality and outcomes. Burchinal and colleagues suggest that the effect of high-quality care may only manifest if children receive a particular dosage (the time spent in care), or alternatively, that the measures of quality currently in use are not capturing the dimensions of the care environment that

[2] To conduct the QRS Assessment project, in 2008 OPRE contracted with Mathematica Policy Research, Child Trends, and Christian and Tvedt Consulting. The goals of this project are to (1) gather and analyze existing and new information on QRIS implementation and research to inform decision making on QRIS development and refinement and (2) build the capacity for ongoing monitoring and evaluation within and across systems.

are most closely associated with children's outcomes. Another potential challenge in capturing associations between quality measures and children's outcomes is the amount of error introduced by quality measurement procedures and tools (Raudenbush and Sadoff, 2008). These important issues are beyond the scope of this work, but are being examined by other research efforts funded by OPRE, such as the combined QRS Assessment/FACES pilot study of observational measures of classroom quality and the Child Care and Early Education Quality Features, Thresholds and Dosage and Child Outcomes (Q-DOT) Study Design.

The Assessment project, however, has focused on examining the details of quality measurement within QRIS to better understand what the summary ratings in select QRIS represent, how they vary across systems, and importantly, the particular methods by which components are measured and the summary ratings are produced. Measures that are used for accountability (as in QRIS) must meet high standards for consistency, reliability, and validity—whether they are gathered through surveys, document reviews, or observational measures (Zellman et al. 2011; Lugo-Gil et al. 2011; Zaslow et al. 2009). The QRIS rating represents the integration of multiple measures to produce a single summary rating; therefore it must meet technical standards not only for each individual component but also for the methods of determining the final rating.

Consistency, reliability, and validity in measuring quality components and determining the summary rating hold a great deal of importance from the perspective of all QRIS stakeholders. First, the QRIS rating must gain the trust and confidence of both policymakers and the public if it is to work as a consumer education tool that merits public funding. Second, the system must ensure equity across providers and the types of care settings in the rating process, particularly when financial resources and technical supports are linked to different rating levels. And third, it is vital for research purposes that the system measures what it is intended to measure in order to support strength in analysis and confidence in findings. For example, poorly done measurement could contribute to the lack of an observed association between quality levels and outcomes and lead to unsubstantiated conclusions. However, if we know that measurement is conducted systematically and with rigor, then the field can focus on assessing whether or not the measures themselves are capturing what is important, or if the link between quality and children's outcomes is weak.

Given the importance of quality measurement in QRIS to both practice and research, there is a need for more in-depth information about how quality is defined and measured. This in-depth study helps fill this need by taking a close look at quality measurement in a manner that can help assess the fidelity of implementation. This type of information lays important groundwork for future monitoring and evaluation by identifying areas in need of refinement as well as opportunities for cross-system analysis.

A. Research Strategy and Questions

The Compendium of QRS and Evaluations (Tout et al. 2010)—the first product of the Assessment project—described the definition of quality in each of 26 QRIS by developing standard categories that can be applied across all QRIS. In total, the Compendium identified 13 categories that capture the range of components used by QRIS to define quality within the rating structure (Exhibit I.1).

Exhibit I.1 Thirteen Categories of Quality Components Used by 26 QRIS

Licensing compliance	Family partnerships
Ratio and group size	Administration and management
Health and safety	Cultural and linguistic diversity
Curriculum	Accreditation
Environment	Provisions for special needs
Child assessment	Community involvement
Staff qualifications	

Source: Compendium of QRS and Evaluations (Tout et al. 2010), produced as part of the QRS Assessment project.

Using the quality categories defined by the Compendium as an organizing framework for data collection and analysis, the in-depth study on quality measurement examined the approaches used by states and communities to measure quality through the QRIS. The study was developed to address three research questions:

- What is the variation in how select QRIS define and measure quality, and what accounts for the variation in their approaches?
- What are the specific processes used by select QRIS to measure each component of the quality rating and determine the overall rating level?
- What is the availability of consistent and reliable data on quality ratings within select QRIS and how are the data currently being used?

B. Study Methods

This in-depth, qualitative study is part of a two-pronged approach into the exploration of quality measurement in QRIS. The second part of the approach is a secondary data analysis of the prevalence of quality components by QRIS level and the association of quality components with observed quality (Malone et al., forthcoming). The selection of QRIS for inclusion in both qualitative and quantitative analyses was accomplished concurrently. The analysis goals of the secondary data analysis drove the initial selection of nine QRIS for participation in 60-minute telephone interviews that focused on the coverage, access, and documentation of data collected as part of the quality rating process. Four criteria were used to select the initial nine QRIS based on characteristics from the Compendium that were indicative of the program stability and structure that could support the secondary data analysis. These characteristics included (1) active QRIS in operation for at least one year prior to 2010, (2) a building block or combination rating structure to support the comparison of levels, (3) an observational measure of environment quality, and (4) the presence of a QRIS database (or linked sources) containing overall rating levels and the possibility of component-level data. The nine QRIS included Miami-Dade, Florida; Palm Beach, Florida; Illinois; Indiana; Kentucky; Maryland; Oklahoma; Pennsylvania; and Tennessee. Between April and June 2010, researchers conducted telephone interviews with eight of the selected QRIS (all but Kentucky).[3]

[3] Kentucky was about to launch an evaluation of their own and could not participate at the time.

Selected QRIS. Based on the information gathered during the telephone interviews on data coverage, access, and documentation, we selected five QRIS for the in-depth study: Florida-Miami Dade, Illinois, Indiana, Pennsylvania, and Tennessee. Representatives from each of these QRIS reported the availability of comprehensive data on quality ratings and measures, housed within a single QRIS database (such as Miami-Dade) or a series of linked systems (such as Pennsylvania). Also, most of these systems reported using the data in QRIS management and improvement efforts. This purposeful selection was intended to identify a set of QRIS from which the research team could gather detailed and comprehensive information on the quality measurement process as well as select a subset for participation in the secondary data analysis. Three of these five QRIS—Miami-Dade, Illinois, and Tennessee—were requested to provide data for that analysis. A companion report presents the findings from that work (Malone et al. 2011). A summary of key characteristics of the five QRIS that participated in the in-depth, qualitative study on quality measurement is presented in Table I.1.

Data collection. Two researchers conducted 1 to 1 1/2 day site visits to each of the five QRIS to interview a range of respondents through a combination of individual and small group interviews. The respondents included QRIS administrators, QRIS planners and developers, staff who gather evidence for component ratings and conduct standardized assessments, staff who serve as primary liaisons to participating providers throughout the enrollment and rating process, QRIS evaluators, and architects and managers of QRIS data systems.

To guide data collection, we organized the research around four topics: (1) quality definitions, thresholds, and rating criteria; (2) use of observational measures; (3) processes for quality measurement; and (4) data collection, use, and analysis to refine quality measurement in QRIS. These topics guided the development of a master protocol as well as a series of data collection tools for use during the site visits. Applicable sections of the protocol and the data collection tools were selected for use with each respondent type with planned overlap to enable triangulation of the data in the analysis. Data collection in this way ensures that findings are based on mutually confirming lines of evidence (Yin 2009). One researcher held primary responsibility for leading each interview, while the other was responsible for taking handwritten notes during each interview. In addition, each interview was digitally recorded.

C. Analytic Approach and Study Limitations

To support the cross-site analysis, the notes from each interview conducted were transcribed and as needed, researchers referred back to the recordings. One researcher involved in each interview developed the notes and then sought confirmation on the accuracy and completeness from the other researcher present. We created codes that followed the structure of the master interview protocol and applied them to each of the individual interview notes. We used the qualitative analysis software package, Atlas.ti (Scientific Software Development 1997) to organize and code the data in line with the coding framework. This enabled comprehensive and systematic analysis of each research topic across the five sites.

Throughout the analysis, we used the 13 quality component categories developed in the Compendium as discussed above. Given the use of these categories as an organizing framework for the analysis and reporting, Table I.2 provides a summary of the categories that are included in the standards for each of the five QRIS that are the focus of this in-depth study.

Table I.1. Overview of QRIS Participating in the In- Depth Study of Quality Measurement

	Miami-Dade County	Illinois	Indiana	Pennsylvania	Tennessee
QRIS Name	Quality Counts	Quality Counts	Paths to Quality	Keystone STARS	Star-Quality Child Care Program
Starting Year of Statewide Implementation	2008	2007	2008	2003	2001
Number of Rating Levels	5	4	4	4	3
Structure of Rating Levels	Combination	Building blocks	Building blocks	Building blocks	Combination
Eligible Programs					
Center-based	✓	✓	✓	✓	✓
Head Start/ Early Head Start	✓	✓	✓	✓	✓
Pre-kindergarten	✓	✓	n/a	✓	✓
Licensed FCC	✓	✓	✓	✓	✓
License-exempt Homes	✓	✓			
School-aged Programs	✓	✓	✓	✓	✓
Other			Child care ministries		
Total Number of Participating Programs	430	1,030	2,040	4,420	2,749
Percentage of Participants in Each Level	Level 1: 13% Level 2: 29% Level 3: 35% Level 4: 19% Level 5: 4%	Level 1: 10% Level 2: 30% Level 3: 60% Level 4: 1% Tier 1: 52%[a] Tier 2: 22% Tier 3: 27%	Level 1: 61% Level 2: 15% Level 3: 13% Level 4: 10%	Level 1: 46% Level 2: 29% Level 3: 12% Level 4: 14%	Level 1: 2% Level 2: 19% Level 3: 61%
Total Number of Children Served	28,000 (as of July 2010)	43,465 (as of April 2011)	75,993 (as of May 2011)	168,530 (as of June 2010)	Not available

Source: Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); QRS Data (Illinois Department of Human Services, May 2011); Keystone STARS 2010 Program Report (OCDEL, 2010); Tennessee Report Card & Star Quality Program Year 8 Annual Report (Pope and Magda 2010); Paths to QUALITY, Monthly Management Report (FSSA, 2011); Trends from Miami-Dade's QRIS (ELC,2010).

n/a = not applicable

[a] License-exempt homes have a separate three-tier system in Illinois.

Table I.2. Inclusion of the Categories of Quality Components in the Five QRIS Participating in the In-Depth Study of Quality Measurement

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Licensing Compliance		✓	✓	✓	✓[a]
Ratio and Group Size	✓				✓
Health and Safety		✓		✓	
Staff Qualifications	✓	✓	✓	✓	✓
Family Partnerships	✓	✓	✓	✓	✓
Community Involvement		✓		✓	
Administration and Management	✓	✓	✓	✓	✓
Environment	✓	✓	✓	✓	✓
Curriculum	✓	✓	✓	✓	✓
Child Assessment	✓	✓	✓	✓	
Cultural and Linguistic Diversity	✓	✓	✓		
Provisions for Special Needs	✓	✓	✓	✓	
Accreditation	✓[b]	✓	✓	✓	✓[c]

Sources: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); site visits conducted as part of the QRS Assessment project.

[a] Providers must be licensed to participate.

[b] Accredited providers receive a "plus" designation to their quality rating.

[c] Accredited providers receive two bonus points.

The purposeful selection of QRIS for this in-depth study focused on systems that may be more fully developed than others in their practices with regard to quality measurement and data collection. As such, the findings from this study pertain to these particular QRIS and are not reflective of all states or counties implementing QRIS. Also, inherent in qualitative research is the potential bias introduced by respondents. We purposefully selected respondents who were identified as the most knowledgeable on the research topics pertaining to this study and fulfilled particular roles in the quality measurement process. We could not interview all staff involved in QRIS quality measurement, even those with particular roles. For example, we did not interview the entire team of raters or assessors due to the size of the group and limitations in staff availability. The primary contact who assisted in scheduling at each site selected the staff who would participate in particular interviews (such as a subgroup of raters or ERS assessors).

D. Roadmap to the Report

The three research questions are addressed in turn in each of the chapters that follow. In Chapter II, we describe the definitions of quality in each of the five QRIS as reflected in the quality components included and the rating levels and structure. In Chapter III, we provide details of the processes used to measure each of the quality components and assign the final rating level in each of the QRIS. In Chapter IV, we report on the availability and use of data from the quality measurement process. In the concluding chapter, we highlight themes that emerged from this study and discuss potential directions for future research.

THIS PAGE LEFT BLANK FOR DOUBLE- SIDED PRINTING

II. QUALITY DEFINITIONS, THRESHOLDS, AND RATING CRITERIA

Previous research has documented the variation that exists in the QRIS such as differences in components that make up ratings, individual indicators for each component, and how indicators are combined to produce a quality rating (Tout et al. 2010). Less information exists about how different systems came to vary to this extent.

The quality measurement process is driven by the indicators that each QRIS has selected as indicative of the construct "quality". As such, understanding how each QRIS defines quality, along with the various factors that influenced that definition, is an important first step in studying how quality is measured in the system. In this chapter, we describe states' experiences in planning, developing, and designing their systems with the goal of illustrating how information from the extant literature on child care quality combined with factors specific to local contexts have resulted in differences—both subtle and striking—in the structures and features of current QRIS.

A. Background and Context of the Planning and Development Process

Establishing the five QRIS included in this study was driven by local needs and circumstances. It is important to understand the political impetus that led to the creation of each system because this context determined who was present at the table during planning and their subsequent influence on each system's definition of quality.

Tennessee's system, the most mature of the five, was created in 2001 in response to multiple cases of provider negligence that resulted in child deaths. These incidents brought about a push to find a way to better regulate child care settings and resulted in formulation of the Child Care Evaluation and Report Card Program, along with the voluntary Star-Quality Program, which was mandated by law.

Pennsylvania's Keystone Stars program, implemented in 2002, was created after a report commissioned by the Governor's Task Force on Early Childhood Care and Education found decreases in environmental quality of child care centers and homes from 1996 to 2002 (Fiene et al. 2002). Establishing a rating system also grew out of the need for a quality indicator that could be used to identify high quality programs and facilitate integration with other early childhood initiatives in the state. For example, Head Start programs could use quality ratings to determine whether a particular provider could meet HS performance standards—a requirement for child care partners providing wraparound services. Respondents noted that although they previously relied on accreditation status as an indicator, there was a need for a finer-grained measure.

Indiana's Paths to Quality (PTQ) system, launched in 2007, was created because administrators sought a way to tie together the patchwork of investments that had been implemented in the state's early childhood system to that point and be able to examine whether those investments were bringing about desired outcomes in terms of quality improvements. Indiana Administrators and planners noted that although funds were limited and they were faced with a child care environment in which a substantial portion of providers were not required to be licensed because of religious affiliation, they were committed to establishing a system that would help families make sense of the landscape of child care options. In exploring options for the system, administrators looked to a program that had already been implemented in Allen County in 2001, and subsequently implemented in several other counties in Northeast Indiana. Given the challenges they faced, planners stressed the importance of being able to build on a homegrown system, particularly since

the earlier PTQ efforts demonstrated a positive reception from providers, with increasing rates of participation each year (Elicker et al., 2007). Planning efforts in Indiana focused on bringing existing early childhood initiatives together into an integrated system. They continue to closely link new quality projects to PTQ—not making initiatives exclusive to participants in this system but making sure that all incentives and resources provided through those projects fit within existing PTQ requirements.

Illinois and Miami-Dade County have newer systems established within the last five years. Respondents from both QRIS noted their increasing awareness of developments in QRIS at the national level as they were planning their systems. Illinois' Quality Counts system began as a tiered reimbursement system that grew out of efforts to improve their state Child Care Assistance Program (CCAP) by tying financial incentives to the provision of high quality care for subsidy-eligible families. Miami-Dade's Quality Counts system is funded through The Children's Trust, a funding stream dedicated to initiatives for children in that county. At the time of planning, Florida already had an existing statewide tiered reimbursement system tied to accreditation called the Gold Seal Quality Care program. However, there was not a program that offered supports to providers as they worked to improve quality.

Although the exact circumstances under which each QRIS was established varied, we found common goals across the five system sites (Table II.1). Each QRIS seeks to improve child care quality as a primary goal. To increase quality, systems venture to help child care providers by offering information, professional development (PD), supports, and incentives. The five QRIS intend for quality ratings to help parents recognize and access high-quality child care providers. Only one QRIS has an explicitly stated goal to improve child outcomes.

1. Participants in the Planning Process

Each of the five QRIS looked to existing systems to inform the design of their program, but they were not interested in adopting another QRIS as a whole. Respondents noted that studying existing systems helped them gain a sense of possible options and have a basis of comparison in thinking about what was feasible given their own local needs and constraints on resources. Indiana planners noted that although they found some desirable features of other systems—for example, multiple assessment visits in a year and large incentives tied to ratings—they knew such features would be too expensive to implement and sustain in their QRIS. Miami-Dade's planners recalled learning from visits and consultation with other QRIS but emphasized that it would not have been advisable to implement an existing system "off the shelf" for a variety of reasons. First, in terms of design, they did not want to implement a pilot program or limit the system to subsidized children and families, emphasizing the importance of launching a program accessible to everyone in the county. Second, it was important to include the local community in the process of designing the system.

Respondents across QRIS felt that convening key stakeholder groups was an essential step to ensure buy-in from the community and encourage participation in the eventual system. Through this process, planners gathered information about local providers, families, and children to contextualize options and better match to local needs. Table II.2 summarizes the various individuals or entities involved in the planning process. In the four QRIS that were creating statewide systems, the planning process was spearheaded by state child care administrative agencies and included staff from offices overseeing licensing efforts. All five QRIS included representatives from child care subsidy administrators in their planning efforts.

Table II.1. Goals of Five QRIS

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Quality Goals	Ensure young children throughout Miami-Dade County are provided with the best possible early learning experiences through high quality early learning programs	Increase the number of high-quality child care providers available to families	Increase the quality of care for all children	Increase the quality of care for children	Improve the quality of child care in Tennessee
Provider Goa	Provide information and support to early learning programs to improve and sustain the quality of their programs	Provider incentives and supports to improve the quality of programs service children participating in CCAP	Support	Increase	To encourage
Family Goals	Provide families with an easy to use tool to select the best program for their child	Improve parents' ability to recognize quality child care	Provide parents with a method to make informed child care choices	Increase parents' understanding and demand for higher quality care	To provide support and information to parents as they seek to secure quality child care for their children
Child Goals		Increase school readiness of children in care			

Sources: Elicker et al., 2007; Illinois Department of Human Services, 2004; Miami-Dade Quality Counts, 2009; OCDEL, 2010; Pope & Magda, 2010.

Table II.2. Participants in QRIS Planning

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Child Care Licensing		✓	✓	✓	✓
Child Care Subsidy Administrator	✓	✓	✓	✓	✓
Center-Based Providers	✓	✓		✓	✓
Family Child Care Providers	✓	✓		✓	✓
Head Start Representatives	✓	✓	✓	✓	✓
Department of Education	✓			✓	✓
Local AEYC			✓		✓
Local CCR&R	✓	✓	✓[a]		
Other TA Providers/ Partners	✓			✓	✓
Researchers/ Consultants	✓	✓			✓

Source: Site visits conducted as part of the QRS Assessment project.

[a]Includes developers of PTQ.

Indiana's planning process included a smaller group of individuals because their efforts focused on refining the original Paths to Quality program to make it more suitable for statewide implementation. In the other four QRIS that were building new programs, representatives (including program directors and union leaders) of center-based and family child care providers participated in the process. In Illinois, Pennsylvania, and Tennessee, providers served on working groups or committees created to formulate standards and requirements. Providers also offered feedback on proposed plans through focus groups and/or public meetings in two of the five QRIS.

The planning process also involved representatives from other agencies involved in the provision of early childhood services whose programs were both potential QRIS participants or partners of QRIS participants. These agencies included Head Start (five QRIS) and the Department of Education (three of five QRIS). Representatives from local child care resource and referral agencies (three QRIS), local affiliates of National Association for the Education of Young Children (NAEYC) (one QRIS), and other providers of technical assistance and PD (three QRIS) also participated to discuss existing and potential initiatives and infrastructure available to help providers move through rating levels. Finally, three of the QRIS brought in outside experts with previous experience in implementation or evaluation of large-scale child care quality improvement efforts to share substantive knowledge about child care quality indicators and strategies for monitoring them.

B. Designing the System

As plans for each QRIS began to take shape, planning groups were tasked with identifying essential components of child care quality, finding ways to quantify each component and bring all components together to produce a single rating, and deciding on an appropriate number of levels to represent reasonable improvements in quality over time. In this section, we describe each of these issues in turn, focusing on various factors that QRIS planners considered in making their decisions.

1. Selecting Quality Rating Components

Previous research has documented some commonalities in components of quality ratings as defined by different systems. The earliest iterations of QRIS all included standards for child-staff

ratios and group size, staff qualifications, and environment (Zellman and Perlman 2008). More recently, systems have also included indicators for family partnerships, administration and management, and accreditation. Additional standards pertaining to child assessment, curriculum, community involvement, cultural and linguistic diversity, and provisions for special needs have only begun to emerge (Tout et al. 2010).

We begin with a description of the components of each site's quality ratings and the indicators used to assess them. For simplicity and ease of comparison across QRIS, we will use the following component categories, based upon categories created in the Compendium, in our discussion of the different indicators of interest:

- **Licensing compliance**—indicators referring to a program's licensing status
- **Ratio and group size**—indicators provide guidelines for the number of children per caregiver and the total number of children in a classroom or home
- **Health and safety**—indicators provide guidelines for provisions to protect children's health and safety
- **Staff qualifications**—indicators specifying the educational qualifications and training of the teaching staff, program director, or family child care provider
- **Administration and management**—indicators refer to administrative procedures and structures, human resource policies, employee benefits and other provisions in place to manage staff and program operations
- **Family partnerships**—indicators refer to activities and strategies to involve and engage families
- **Community involvement**—indicators refer to practices and strategies to promote connections between the program and the community and/or help families and children connect with resources in the community
- **Environment**—indicators of activities, practices, materials, and provisions in the environment to promote children's optimal learning and development
- **Individualization of services**—indicators of a program's efforts to tailor services to children's needs through the following strategies:[4]
  - **Curriculum**—specifications about the curriculum used and the extent to which classroom activities are developmentally appropriate
  - **Child assessment**—practices that promote ongoing assessment of children's needs for the purposes of improving individual and group instruction and sharing information with parents

[4] This classification is used to combine elements for ease in access to the information as well as to provide comparable information to the secondary data analysis included in another forthcoming report of the Quality Rating System (QRS) Assessment project.

  - **Cultural and linguistic diversity**—provisions that support cultural competence and intentional practices that promote respectful interactions with diverse children and families
  - **Provisions for special needs**—practices and strategies to promote inclusion of children with special needs
- **Accreditation**—indicators refer to a program's status with regard to program accreditation by a national accrediting body

Across QRIS, not a single component enters into the rating in the same way (Table II.3). For example, although the five QRIS include indicators of staff qualifications, family partnerships, administration and management, and environment in their quality ratings for center-based programs, the level at which each indicator is required varies.

Table II.3. Quality Level at Which Components Enter Ratings for Center- Based Programs

	Quality Counts, Miami-Dade County[a]	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Licensing Compliance	--	Level 1	Level 1[b]	Level 1	Required for enrollment
Ratio and Group Size	Level 1	--	--	--	Level 1
Health and Safety	--	Level 4	--	Level 1	--
Staff Qualifications	Level 1	Level 1	Level 2	Level 1	Level 1
Administration and Management	Level 1	Level 3[c]	Level 2	Level 1	Level 1
Family Partnerships	Level 1	Level 3[c]	Level 2	Level 1	Level 1
Community Involvement	--	Level 3[c]	--	Level 1	--
Environment	Level 1	Level 1	Level 2	Level 1	Level 1
Curriculum	Level 1[d]	Level 3[c]	Level 3	Level 3	Level 1[e]
Child Assessment	Level 1	Level 3[c]	Level 3	Level 2	--
Cultural and Linguistic Diversity	Level 1	Level 3[c]	Level 2	--	--
Provisions for Special Needs	Level 1	Level 3[c]	Level 3	Level 2	--
Accreditation	"Plus" rating	Level 3	Level 4	Level 4 (not required)	2 points added to total score

Sources: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); site visits conducted as part of the QRS Assessment project.

[a]Miami-Dade and Tennessee use combination systems in which criteria are specified for each component corresponding to a certain number of points.

[b]License-exempt providers must go through the voluntary certification program (VCP) in order to participate.

[c]Providers can achieve a Level 3 rating if they are accredited or meet a cutoff score on the PAS. The PAS is only administered if a provider is not accredited or chooses to receive an assessment despite accreditation status.

[d]The curriculum component is currently being assessed and scored but points are not yet included in the final rating.

[e]Tennessee introduced their Developmental Learning component on January 1, 2010.

In general, licensing serves as the foundation on which standards in the five QRIS are built but the current role of licensing depends on the perceived rigor in the licensing requirements and the

maturity of the QRIS. As Table II.4 illustrates, one of the five QRIS requires a license in good standing as the complete requirement to receive a QRIS level one rating. In Illinois and Pennsylvania, additional requirements beyond licensing are necessary at the first rating level. Tennessee originally required licensing compliance at level one but their standards have recently been revised (in 2009) to instead consider licensing a prerequisite for participation. Administrators felt that given the relative maturity of the Tennessee Star-Quality System, it was necessary to raise the base requirements to achieve a level one rating. Of the five QRIS, only Miami does not include licensing compliance as a requirement at any level. Although the requirements at level one for staff qualifications and ratio and group size align exactly with Florida licensing requirements, the Miami-Dade QRIS also includes additional standards encompassing other quality components.[5]

Table II.4. Role of Licensing Compliance in QRIS

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Licensing Compliance Required	No	Yes	Yes	Yes	Yes
Level Where Licensing is Required	Some Level 1 requirements are identical to licensing standards	Level 1	Level 1	Level 1	Pre-requisite for
Other Requirements at Licensing Level	Additional components specified for Level 1 beyond those that are equivalent to licensing	Yes	None	Yes	n/a

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable

The tie between licensing and QRIS eligibility and/or entry was also influenced in each QRIS by the patterns in the types of care used, the licensing requirements for different care settings and early learning programs, and the need to create comparable standards across settings. Planners deliberated over how requirements would affect providers who are exempt from state licensing standards and regulations. License-exempt providers include religious institutions (churches, parochial schools, and so on), family day care providers serving a small number of children, relatives, and in-home providers who only care for a small number of children. Table II.5 shows participation rates for children receiving Child Care and Development Fund (CCDF) subsidies[6]. Although a majority of children receiving subsidies are in licensed or regulated settings across states, a considerable number of children (as many as 45 percent in Illinois) are cared for in non-regulated settings.

Table II.5. Types of Child Care Settings Attended by Children Receiving Subsidies, by State

	Florida	Illinois	Indiana	Pennsylvania	Tennessee

[5] Seven counties in Florida have their own licensing standards, but Miami-Dade is not one of them.

[6] We were unable to find similar statistics for the general population of children in each state.

	Florida	Illinois	Indiana	Pennsylvania	Tennessee
Number of Children Under 5 Years Old, 2009	1,167,892	890,818	443,195	743,681	428,145
Percentage of Children	5.5%	3.8%	4.5%	6.5%	6.4%
Subsidized Children in Licensed or Regulated Settings	91.3%	54.2%	72.8%	70.5%	90.4%
Center	81.8%	32.8%	36.2%	59.3%	77.4%
Family/Group Home	9.5%	21.4%	36.7%	11.3%	12.9%
Subsidized Children in Settings Legally Operating Without Regulation	8.6%	45.8%	27.2%	28.0%	9.6%
Center	8.1%	3.9%	23.4%	0.0%	0.3%
Care In Child's Home	0.5%	17.2%	0.2%	1.3%	0.9%
Family/Group Home	0.0%	24.7%	3.7%	26.7%	8.4%

Sources: U.S. Census Bureau: State and County QuickFacts; Office of Child Care CCDF Data Tables, 2009.

Planners were aware that families access a wide variety of settings in each of their states and felt that it was in the best interest of the system to have a wide range of providers involved in the QRIS. However, it was also important for ratings to convey comparable standards of quality at each level across provider types in order for ratings to be meaningful and useful for parents. To ensure a level playing field for all providers, four of the five QRIS require license-exempt centers to obtain a license in order to participate in their QRIS (Table II.6). Indiana requires license-exempt centers to go through a Voluntary Certification Program (VCP) in lieu of a license. The VCP program was in existence before PTQ was launched statewide but in planning for PTQ, administrators worked on revising VCP standards to align them with requirements for licensed providers at level one[7]. License-exempt homes are only eligible to participate in the QRIS in Illinois and Miami (which does not require a license). Illinois planners recognized the large contingent of license-exempt homes serving children in their state (as shown in Table II.5) and created a separate set of QRIS standards for those providers.

The QRIS can be considered a means of defining quality not just in child care settings but across a spectrum of early learning programs to include programs such as Head Start and pre-kindergarten. As a result, QRIS planners also considered the context of use and licensing requirements for these programs as they relate to QRIS eligibility and entry. Although enrollment rates in these programs are not as high as rates for all private providers combined, Head Start and

[7]Indiana also revised licensing requirements for family child care providers so that they would be required to obtain a Child Development Associate (CDA) credential within three years, and changed the definition of lead teacher to make it comparable across provider types.

public pre-kindergarten do serve a sizeable number of children in some states. For example, in Florida, 13 percent of children under age five are enrolled in public pre-kindergarten programs and in Pennsylvania, nearly five percent of children are enrolled in a pre-kindergarten program (Table II.7).

Table II.6. QRIS Eligibility and Entry Requirements for License- Exempt Child Care Providers

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
License-Exempt Centers Eligible	Yes	No	Yes	No	No
QRIS Entry Different for License-Exempt Centers	No	n/a	Must go through Voluntary Certification Program in lieu of license	n/a	n/a
License-Exempt Homes Eligible	Yes	Yes	No	No	No
QRIS Entry Different for License-Exempt Homes	No	Yes; rating standards are specified for unregulated homes	n/a	n/a	n/a

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable

Table II.7. Enrollment in Head Start and State- Funded Pre- kindergarten Programs, by State

	Florida	Illinois	Indiana	Pennsylvania	Tennessee
Number of Poor Children Under 6 Years Old, 2009	323,640	215,267	129,430	172,641	135,787
Head Start Enrollment, 2009	35,390	39,435	14,145	35,253	16,339
Approximate Percentage of Poor Children in HS[a]	10.9%	18.3%	10.9%	20.4%	12.0%
Number of Children Under 5 Years Old, 2009	1,167,892	890,818	443,195	743,681	428,145
State-Funded Pre-k Access	4-year olds only	3- and 4-year olds	None	3- and 4-year olds	3- and 4-year olds
Total Pre-k Enrollment	155,877	87,451	0	31,796	18,252
Percentage of Children Under 5 Enrolled in Pre-k	13.3%	9.8%	0.0%	4.3%	4.3%

Sources: Barnett et al., 2010; Head Start Program Fact Sheet, FY 2010; National Center for Children in Poverty, State Profiles; U.S. Census Bureau: State and County QuickFacts;

[a] This percentage is an approximation because some Head Start programs may enroll children who are age 6 or older, as well as children from families with incomes up to 130 percent of the poverty threshold.

Head Start and public pre-kindergarten programs are required to abide by separate standards set at the federal level—the Head Start Program Performance Standards (HSPPS)—or by local education agencies. However, it is up to each state to decide if licensing is required for Head Start programs. Head Start and public pre-kindergarten programs are required to obtain a license in order to participate in four of the five QRIS (Table II.8). Only Pennsylvania accepts compliance with HSPPS in lieu of a license.

Table II.8. QRIS Eligibility and Licensing Requirements for Head Start and Pre-kindergarten Programs

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Head Start Programs Eligible	Yes	Yes	Yes	Yes	Yes
QRIS requirement for licensing Different for Head Start	No	No	No	Compliance with HSPPS accepted in lieu of license	No
State Pre-k Programs Eligible	Yes	Yes	n/a	Yes	Yes
QRIS requirement for licensing Different for Pre-k Programs	No	No	n/a	No	No

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable

Ratio, group size, and health and safety indicators. Decisions to include additional QRIS requirements for child-staff ratio, group size, and health and safety beyond what is required for a licensed provider were influenced by the licensing requirements in each state. Three of the five QRIS—in Illinois, Indiana, and Pennsylvania—do not include additional requirements for child-staff ratio and group size in their QRIS standards for center-based programs. Respondents from these QRIS reported that their licensing standards set a strong base and that the cost of additional data collection was not justified. In addition, in each of these three QRIS, accreditation is integrated into the higher QRIS rating levels and ratios and group size are embedded into accreditation requirements (further discussion of the role of accreditation is discussed later in this chapter.)

Miami and Tennessee are the only QRIS that include additional provisions for ratio and group size in their rating systems. Planners in Miami did not think that Florida's licensing standards were stringent enough. For example, the required child-staff ratio for two-year-olds in Florida is 11:1—three more children per staff member than the second highest ratio licensing requirement of the QRIS studied (8:1 in Illinois). Although Miami's standards for ratio and group size at the lowest QRIS level are identical to Florida licensing standards, requirements at higher rating levels narrow disparities with other states' standards. At the highest rating level, Miami's QRIS requirements are comparable to NAEYC recommendations for ratios and group sizes, although they are at the higher end of the recommendations (Table II.9). Tennessee's licensing requirements for ratio and group size are comparable with the other three QRIS (refer to Appendix Table A.1). However, because Tennessee's Star Quality Program is administered in conjunction with their licensing program, there is no additional cost or burden associated with collecting data on ratios and group size; licensing staff conduct visits to facilities throughout the year.

Table II.9. Ratio and Group Size Requirements for Center- Based Programs by Licensing, QRIS Level, and Accreditation in Florida and Tennessee

	Florida Licensing Standards[a]	Quality Counts, Miami-Dade County (Highest Level)	Tennessee Licensing Standards	Tennessee Star Quality, (Highest Level)	NAEYC Recommendations
Child-Staff Ratio					
Infants	4:1	4:1	4:1	4:1	3/4:1
Ones	6:1	4:1	6:1	4:1	3/4:1
Twos	11:1	6:1	7:1	5:1	4-6:1
Threes	15:1	9:1	9:1	8:1	6-9:1
Fours/Fives	20/25:1	10:1	13/16:1	13/15:1	8-10:1
Group Size					
Infants	--	8	8	8	6-8
Ones	--	12	12	12	6-8
Twos	--	12	14	10	6-12
Threes	--	18	18	16	12-18
Fours/Fives	--	20	20	20	16-20

Sources: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); NCCIC & NARA, 2010; NAEYC, 2008.

[a]Seven counties in Florida have their own licensing standards, but Miami-Dade is not one of those counties.

Only Pennsylvania and Illinois set standards for health and safety in their QRIS. Keystone Stars requires that programs track illnesses and injuries, and that staff participate in annual training in topics such as first aid and child abuse. Illinois requires staff certification in CPR and first aid at the highest level. As noted in the Compendium, very few states include health and safety provisions in their rating systems likely because such measures tend to be covered by licensing requirements (Tout et al. 2010). In addition, as we will discuss in a later sections, many QRIS include observational assessments of the environment in their ratings and those measures assess health and safety practices.

Staff qualifications. All existing QRIS described in the Compendium incorporate staff qualifications into their ratings—with education and training being the most typical indicators used (Tout et al. 2010). Education and training were also common indicators in the five QRIS studied—although there is wide variation in how requirements are defined at each level (Table II.10).

First, QRIS varied in the extent of requirements defined by position type. Four QRIS have separate requirements for center directors, but only two of the five distinguish between requirements for lead versus assistant teachers. The fact that Indiana and Tennessee do not specify requirements for assistant teachers may be linked to the fact that state licensing standards do not specify requirements for this position either (refer to Appendix Table A.1).

The staff qualifications requirements for Illinois Quality Counts are aligned entirely with requirements of the Illinois Great START Program, a wage supplement program for child care practitioners established several years before their QRIS. Great START defines education level requirements by position, but Illinois Quality Counts standards only require a certain percentage of center staff to meet Great START requirements at each rating level. At the first two levels, a percentage of all staff (10 to 20 percent) must meet Great START requirements whereas at levels three and four, a percentage of teaching staff (25 to 30 percent) must meet these requirements.

Table II.10. QRIS Staff Qualifications Requirements for Center- Based Programs Across Rating Levels

	Quality Counts, Miami-Dade County	Illinois Quality Counts[a]	Indiana Paths to Quality	Pennsylvania Keystone Stars[b]	Tennessee Star-Quality
Specific Requirements by Position					
Directors	✓	--	✓	✓	✓
Lead Teachers	✓	--	--	✓	--
Assistant Teachers	✓	--	--	✓	--
Requirements for Teachers[c]					
Education Level	Percentage of teachers (25-50%)	Percentage of teachers (10-30%)	Percentage of teachers (25-50%)	Percentage of teachers (25-100%)	Percentage of teachers (10-50%)
ECE Credits or Degree	Number of credits (0-18)	Number of hours/credits (15-30)	--	Number of credits (12-30)	--
Credential (CDA, State-Awarded, CCP, Montessori, etc.)	Percentage of teachers (50-100%)	--	Percentage of teachers (25-50%)	--	Percentage of teachers (10-50%)
Continuing Education/In-Service Training	Hours per year (10-30)	--	Hours per year (15-20)	Hours per year (12-24)	Hours per year (12-18)
Experience	--	--	--	--	Years in early care/education program (0-4 depending on education level)

Source: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); site visits conducted as part of the QRS Assessment project.

[a]Based on requirements of the Great START Wage Supplement Program.

[b]Based on Pennsylvania Keys to Quality Early Learning Career Lattice.

[c]We specify lead teacher requirements for QRIS that differentiate standards by teaching position.

Four QRIS define staff qualifications requirements by the percentage of staff meeting certain criteria for level of education and in-service training. For example, at level two, Miami requires all teachers to have a high school degree, 50 percent of teachers to have a state credential, and 50 percent of teachers to complete 15 hours of in-service training per year. Indiana's level two has identical requirements for in-service training, but requires 25 percent of teachers to be credentialed. Tennessee accepts previous experience as an alternative to meeting education level requirements; at level two, 25 percent of teachers must have either three years of experience in early care/education or be enrolled in a credentialing program. Table II.10 illustrates that although QRIS tend to measure indicators using the same units, at each level there is substantial variation in the percentage of teachers required to meet each standard as well as in the type of training each teacher is expected to complete.

Administrators and planners acknowledged that staff qualifications requirements can be challenging to understand and track. Further, respondents noted that it would not have been prudent to set requirements that providers could not realistically achieve, given local systems and resources. Where possible, the QRIS drew on current state requirements and existing PD systems in selecting staff qualifications indicators. Illinois respondents noted that aligning QRIS requirements with the Great START program was a strategic move because providers were already familiar with the requirements for that program and there were already existing incentives for working toward those requirements. Further, the Great START Program is backed by state law, whereas Illinois Quality Counts is not. Pennsylvania defines staff qualification requirements according to the Keys to Quality Early Learning Career Lattice, which is used among practitioners across the state's early care and learning programs (including state-funded pre-kindergarten programs, Head Start, child care providers, and technical assistance providers).

Administration and management. Quality indicators in the area of administration and management cover two main topics—staff management and program administration. Tables II.11 and II.12 list the various indicators assessed and the rating level at which each is required for center-based programs. Overall, Miami, Illinois and Pennsylvania have more extensive indicators for this component than Indiana and Tennessee.

Illinois uses the Program Administration Scale (PAS) (Talan and Bloom 2004) as a measure of administration and management.[8] The PAS was developed by scholars at the McCormick Center for Early Childhood Leadership at National Louis University (NLU). One of the PAS developers served as a consultant on Illinois' QRIS planning work group. During planning for Illinois Quality Counts, the PAS tool had been piloted with 67 Illinois centers statewide. The PAS was designed to complement standardized measures of the child care environment (the Environment Rating Scales) by producing ratings of organizational quality on a similar seven-point scale. Planners in Illinois noted that the PAS was the only known measure that captured these constructs at that time. Their decision to include this measure in quality ratings was driven by the desire to capture dimensions of program administration at a level of detail that would inform research and evaluation efforts as well as be useful by programs for self-assessment and planning. Illinois' indicators from the PAS are more detailed and comprehensive than what is displayed in Table II.11 and II.12. However, the PAS is administered only to non-accredited providers applying for level three, and all providers applying for level four. Further, Illinois Quality Counts only requires providers to meet a PAS score threshold at those levels (4.25 at level three and 5.0 at level four) and does not specify requirements in terms of particular indicators.

Although none of the indicators is common to all QRIS, several are assessed by four of the five (Table II.11). For example, in the area of staff management, four of five QRIS require an annual PD plan for staff. Miami and Illinois require such a plan as part of a performance evaluation. Tennessee requires a PD plan for a certain percentage of staff beginning at lower quality levels. The same four QRIS require a salary scale that is differentiated based on staff experience and levels of education at higher rating levels. The provision of staff benefits, such as health insurance or paid vacation days, is also a common indicator (four of five QRIS). Staff benefits are required across rating levels in Pennsylvania (levels two to four) and Tennessee (levels one to three), with the number of benefits increasing at higher levels. Illinois' PAS measure assesses the number of benefits provided as well as

[8] The Business Administration Scale (BAS; Talan and Bloom, 2009) is used in family child care programs.

the quality of each benefit provided. For example, providers are asked whether staff have access to a retirement plan and whether and how much the employer contributes to that plan. The only benefit required in Indiana is paid planning time for the lead teacher.

Table II.11. QRIS Staff Management Indicators for Center- Based Programs Across Rating Levels

	Quality Counts, Miami-Dade County	Illinois Quality Counts[a]	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Staff Meetings	Quarterly (L3)	2x a year to twice a month (L3)	--	1 within the last 6 months (L1)	--
Staff Orientation	Staff handbook includes job descriptions (L2), policies and procedures (L4)	Written procedures; consistency of implementation (L3)	--	Staff handbook includes job descriptions (L2), policies and procedures (L3)	--
Staff Benefits Provided	--	Up to 5 benefits and amount of each; Paid planning time for teaching staff; Salary increases (L3)	Paid planning time for lead teacher (L3)	2-4 benefits (L2-L4); Paid planning time for all teachers (monthly to weekly; L3-L4)	Number (2-4; L1-L3)
Staff Differentiated Salary Scale	Yes/no (L5)	Yes/no (L3)	--	Yes/no (L3)	Yes/no (L3)
Staff Performance Evaluation	Annually (L3); Includes classroom observation (L4)	1-3 times/year; based on multiple sources of evidence (L3)	--	Annually; Includes 2 classroom observations (L3)	--
Staff Professional Development Plans	Included in performance evaluation (L5)	Included in performance evaluation (L3)	--	Completed annually (L1)	Completed annually for 50-100% of staff (L1-L3)

Source: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); site visits conducted as part of the QRS Assessment project.

Note: Notation in parentheses indicates the level at which a particular indicator is required.

[a]Comparable PAS Indicators selected for illustrative purposes. Not an exhaustive list.

All but Tennessee include indicators in the area of program administration (Table II.12). Three QRIS assess whether a risk management plan is in place although each requires it at different levels—Miami at level one, Illinois at level three, and Pennsylvania at level four. The same three QRIS require certain fiscal management practices. Miami and Pennsylvania both prescribe a projected one-year budget, evidence of record keeping, and quarterly analyses of finances. However, Miami requires all three features at level five, whereas Pennsylvania spreads out requirements over levels two and three, with more features required at level three. Other program administration indicators assessed are marketing strategies, modes and frequency of program evaluation—including soliciting input from parents and staff—and strategic planning.

Table II.12. QRIS Program Administration Indicators for Center- Based Programs Across Rating Levels

	Quality Counts, Miami-Dade County	Illinois Quality Counts[a]	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Marketing Strategies	Plan in place (L4)	Plan in place, materials used, procedures for reaching out to prospective parents (L3)	--	--	--
Risk Management Plan and Procedures	Plan in place (L1)	Plan in place, dissemination to staff, frequency of drills (L3)	--	Plan in place (L4)	--
Fiscal Management	Projected one-year budget, record keeping, quarterly analysis (L5)	Budget planning, accounting practices (L3)	--	Projected one-year budget (L2), record keeping (L2), quarterly analysis (L3), CPA review (L4)	--
Program Evaluation	--	Tools used; Frequency (L3)	Annually (L3)	Staff surveys conducted (L3)	--
Quality Improvement and Strategic Planning Activities	--	Written plan, frequency of review (L3)	Strategic plan (L3)	Quality Improvement plan (L3); Strategic plan (L4)	--

Source: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); site visits conducted as part of the QRS Assessment project.

Note: Notation in parentheses indicates the level at which a particular indicator is required.

[a]Comparable PAS Indicators selected for illustrative purposes. Not an exhaustive list.

Family partnerships and community involvement. The inclusion of family partnerships and community involvement indicators in the five QRIS denotes a growing recognition that child care providers' relationships with the families and larger communities they serve are an important component of quality. The challenge, as expressed by respondents, is that there are few standardized measures for these components and they typically have to rely on self-reported information from providers. As a result, QRIS are constrained to use indicators that they can verify by reviewing provider documentation or observe during a one-time visit.

Overall, Illinois and Pennsylvania measure the most indicators for these components (Table II.13). Illinois' PAS measure includes a range of indicators even beyond those specified in the table. However, as previously mentioned, the PAS is only administered to a subset of providers.

All of the QRIS require some form of communication with families. Miami requires providers to have at least three modes of communication beginning at level two. Indiana and Pennsylvania do not require a particular number or type of mode but do require a certain frequency of communication. Tennessee requires two forms of communication—the use of bulletin boards and a second mode that is used quarterly to monthly.

Three indicators are assessed by four of the five QRIS, yet are specified in different ways. Specifically, four QRIS prescribe a certain number of parent/teacher conferences and family

activities per year but the frequencies and the levels at which they are required varies. For example, Miami expects two conferences per year at level four, whereas Tennessee requires one per year beginning at level one. Pennsylvania's required frequencies for conferences varies by level: one at level two and two at level three. Similarly, the provision of opportunities for families to evaluate the program is required by four QRIS, but the nature of the measure ranges from a yes/no indicator (Tennessee), to specifying the frequency of the evaluation (Illinois, Indiana, and Miami-Dade), to specifying the modes for evaluation (Illinois). Other indicators are assessed by three or fewer QRIS including parent orientation, parent education and referrals, family participation in planning, and transition support.

Table II.13. QRIS Family Partnerships and Community Involvement Indicators for Center- Based Programs Across Rating Levels

	Quality Counts, Miami-Dade County[a]	Illinois Quality Counts[a]	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Communication with Families	3 or more modes (L2)	Up to 7 modes; Frequency (L3)	Frequency	Frequency (daily; L2)	Quarterly bulletin board; Frequency (quarterly to monthly; L1)
Parent Orientation	Family handbook (L1)	Written information; Guided tour (L3)	--	Parent orientation meeting (L1)	--
Parent/Teacher Conferences	2/yr (L4)	1-2/yr (L3)	--	1-2/yr (L2,3)	1/yr (L1)
Family Activities	2/yr (L3)	Types of school activities families participate in (L3)	--	1/yr (L3)	1/yr (L3)
Parent Education	--	Up to 7 family supports (L3)	--	Handouts (L2)	Handouts (L2); 1 training /yr (L3)
Transition Support Provided	Yes/no (L4)	--	--	Yes/no; types provided	--
Opportunities for Family Evaluation of Program	1/yr (L5)	Up to 2 Modes; Frequency (every 2-3 years to annually; L3)	1/yr	--	Yes/no (L3)
Opportunities for Family Participation in Planning	--	Advisory or governing board (L3)	--	Policy in place (L4)	Advisory council; Frequency (2 mtgs/yr)
Parent Referrals/Community Resources	--	Up to 7 family supports (L3)	--	List of resources (L1); Referral plan (L3)	List of resources (L3)
Community Involvement	--	Community outreach strategies; Collaboration with early childhood community (L3)	--	Outreach to agencies involved in transition activities (L3)	--

Source: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); site visits conducted as part of the QRS Assessment project.

Note: Notation in parentheses indicates the level at which a particular indicator is required.

[a]Comparable PAS Indicators selected for illustrative purposes. Not an exhaustive list.

Environment. All five QRIS studied include an observational assessment of the environment in their ratings. This was due, in part, to our criteria for selection, but also, as documented in the compendium, this is a typical feature of the QRIS (Tout et al. 2010). As Table II.14 shows, each QRIS requires providers to undergo an observational assessment regardless of rating level. The only exception is Illinois, which allows accredited centers to fully meet QRIS requirements at level three (without additional requirements). Although these assessments require an investment of time and money on the part of the QRIS as well as for individual providers, across the board respondents from all the QRIS indicated that this component provides crucial information about child care quality that would be difficult to obtain by any other method. One respondent added that the observational aspect adds credence to the process and ratings.

Four of the five QRIS use the Environment Rating Scales (ERS; Harms et al. 1995, 2005, 2006, 2007)—using the particular scale that is appropriate for the setting and age group assessed (Table II. 14). Planners selected the ERS scales for use as observational measures of the environment because of their wide use and recognition in the field. A few respondents expressed concern that the ERS scales are limited as a measure of the quality of teacher-child interactions and may place undue emphasis on certain health and safety indicators, such as hand washing. Three of the QRIS have considered the possibility of using the Classroom Assessment Scoring System (CLASS; Pianta et al. 2007) as a measure to capture additional aspects of quality such as child engagement and the quality of interactions. However, given the costs associated with adopting a new measure, such plans are still in the exploratory stages.

Indiana does not use the ERS scales in their entirety, but includes some items similar to ERS items in the rating assessment tools. Respondents in Indiana involved in planning noted that the cost associated with using the entire ERS for quality ratings was prohibitive. However, they also noted that because they did not use the entire ERS, they were able to include additional observational indicators to capture the quality of interactions and feel that their resulting measure is comprehensive and well-rounded. A comparison of observational indicators on the Paths to Quality assessment tool with the ECERS-R (Harms et al. 2005) shows more items adapted from the language and reasoning, activities, interaction, and program structure subscales than from the space and furnishings and personal care routines subscales of the Early Childhood Environment Rating Scale-Revised (ECERS-R). The PTQ tool includes 13 indicators related to classroom interactions, only 6 of which overlap considerably with ECERS-R indicators. In addition, some items on the PTQ assessment are more specific in terms of interactions. For example, the ECERS-R includes an indicator about how staff encourage appropriate social interactions with peers. The PTQ standards are phrased similarly to NAEYC criteria, referring specifically to the use of problem-solving approaches to resolve conflicts between children.

The four QRIS that use the ERS integrate scores into quality ratings by setting a minimum score that providers must meet to qualify for a particular rating level or receive a number of points. ERS scores can range from 1 to 7 (where 1=inadequate, 3=minimal, 5=good, and 7=excellent as defined by instrument developers). Table II.14 shows the minimum thresholds set by each QRIS. At the highest rating level, all four QRIS require a score of 5.0 or higher which aligns with the "good" range on the scale. There is greater variation among QRIS in the minimum ERS scores required at lower levels. Pennsylvania does not have minimum score requirements at levels one and two, although providers are required to undergo an ERS assessment and craft a written improvement plan if the assessment yields a score lower than 3.0 on any ERS subscale at level two. Miami and Illinois require a score of 3.0 at level one, whereas Tennessee requires a score of 4.0. The thresholds set at the lower levels were influenced by planners' awareness of quality levels in their county or state

Table II.14. QRIS Observational Tools for Measuring the Quality of the Environment

	Quality Counts, Miami-Dade County[a]	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Measure Used	ECERS-R[a] ITERS-R[a] FCCERS-R	ECERS-R[b] ITERS-R[a] SACERS[c] FCCERS-R[e]	Observational indicators developed from NAEYC standards and ERS items	ECERS-R ITERS-R SACERS FCCERS-R	ECERS-R[a] ITERS-R[a] SACERS[d] FCCERS-R[f]
Observational Assessment Required at All Levels	Yes	Yes, except for accredited centers at level 3	Yes	Yes	Yes
How Environment Score Feeds into Final Rating	Equivalent points for average score on each scale is doubled (for centers) or quadrupled (for FCCs) before adding to point total	Score must meet cutoff for appropriate rating level	n/a	Score must meet cutoff for appropriate rating level	Environment component star rating added to other components ratings and averaged
Minimum ERS Requirements by Level	L1: 3.0 L2: 3.5 L3: 4.0 L4: 4.5 L5: 5.5	L1: 3.0 L2: 3.5 L3: 4.25 L4: 5.0	n/a	L1: self-assessment L2: improvement plan if score < 3 L3: 4.25 L4: 5.25	L1: 4.0 L2: 4.5 L3: 5.0

Source: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); site visits conducted as part of the QRS Assessment project.

n/a = not applicable

[a]Excludes the Parent & Staff scale (items 38-43).

[b]Excludes Parent & Staff scale and item 29.

[c]Excludes items 10, 11, 34 - 36, 39, 41- 43, 45-49 supplemental.

[d]Excludes items 9 -14, 36, 39, 41- 43, 45-49 supplemental.

[e]Excludes Parent & Staff scale.

[f]Excludes items 30-32 (adult needs), 33-39 supplemental. Tennessee previously used the FDCRS but transitioned to the FCCERS-R in August 2010.

at the time of planning. Respondents in Miami and Illinois noted that they did not want to set the bar too high at the lowest level out of concern that it would discourage providers from participating.

Individualization of services. Table II.15 summarizes a group of components that reflect the extent to which providers tailor or individualize services to meet the needs of the children and families they serve. This includes the use of developmentally appropriate curricula, the use of child assessments, provisions for special needs, and practices that respond to and recognize cultural and linguistic diversity. With few exceptions, standards for these components enter in at the higher rating levels across QRIS, indicating that these features are not necessarily expected of a provider demonstrating a baseline level of quality.

The two newer QRIS (Illinois and Miami) included more indicators in these areas than the other three, illustrating that these components have only begun to receive increased attention in the field as vital aspects of quality. Respondents in the states with more mature QRIS acknowledged that they have only recently revised their standards to include at least some of these components.

The only individualization indicator currently assessed in all five QRIS is the use of developmentally appropriate curriculum. A variety of methods are employed to assess whether this standard is met. Miami compares a provider's curriculum to a list of curricula that have previously been reviewed and approved by the Early Learning Coalition, the QRIS administrator. To achieve a level four rating, providers document that they have fully implemented a developmentally appropriate curriculum by achieving scores of 4.5 or greater on the activities, program structure, and language reasoning subscales of the ERS. Illinois reviews curriculum documentation as part of the PAS assessment, to determine whether the curriculum was developed based on published professional standards (for example, by an organization like NAEYC) or individual state learning standards.

Curriculum review and approval takes time and resources. To streamline the process, three QRIS use state Early Learning Guidelines (ELGs) as a benchmark for curriculum content. ELGs, which many states have developed in recent years in conjunction with the federal Good Start, Grow Smart Initiative, identify critical skills that children should develop and acquire from birth to age 5 (NCCIC 2010). In Indiana, providers at level three are required to use a curriculum that is aligned with the Foundations to the Indiana Academic Standards for Young Children from Birth to Age 5. In Pennsylvania's Keys to Quality system, providers are evaluated on the extent to which learning standards are reflected in classroom activities—at lower levels, staff may use these standards as a resource for lesson planning, but at level four, programs are required to conduct a crosswalk of their curriculum with the standards. Finally, Tennessee introduced a new component called Developmental Learning in January 2010, requiring providers to document classroom activities pertaining to each developmental area identified in the Tennessee Early Learning or School-Age Development Standards (TN-ELDS/SADS). Linking QRIS requirements to ELGs allows for building on existing infrastructure and maximizes resources because states have invested considerable efforts into educating providers on the guidelines. In fact, Indiana, Pennsylvania and Tennessee also require training on ELGs in their quality rating criteria. Miami also has a curriculum training requirement but it is not explicitly linked to state-provided training.

In addition to the implementation of developmentally appropriate activities, high-quality programs in four QRIS are also expected to conduct assessments to monitor children's growth. These QRIS require programs to have a system in place to conduct child assessments and, in three QRIS, to use assessment results to inform planning. Indiana and Miami require observational assessments but do not identify a specific tool to be used. Illinois' PAS indicator on assessment does

Table II.15. QRIS Indicators for Individualization of Services for Center- Based Programs Across Rating Levels

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Curriculum is Developmentally Appropriate	Yes/no (L1); Fully implemented (L4)[a]	Yes/no (L3+)	Yes/no; Reflected in classroom activities (L3)	Level of integration with Learning Standards (L2-L4)	Use of developmental standards
Curriculum Training	6 hours/yr for lead teachers (L2) or all teachers (L3)[a]	n/a	Curriculum orientation session (L3)	Annually for director (L2); for teachers (L3)	Director and staff (50-100%) training
Child Assessment Process in Place	Yes/no (L4)	Yes/no; Tools used (L3)	Yes/no (L3)	Frequency (L2-L4; 1-3x/yr)	n/a
Child Assessment Results Guide Planning	Yes/no (L5)	Yes/no (L3)	n/a	Yes/no (L3); Results shared with parents (L4)	n/a
Cultural and Linguistic Diversity	Communication in other languages (L5)	Communication in other languages; Staff asks about cultural practices (L3)	Classroom materials represent diverse cultures and backgrounds (L2)	n/a	n/a
Provisions for Special Needs	Screening and referral process; Activities developed for children with potential delays (L5)	Screening and referral process; Activities developed for children with potential delays (L5)	Written plan; adaptation of space, materials, and activities (L3)	Screening and referral process (L3); Activities to meet IEP goals (L4)	n/a

Source: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); site visits conducted as part of the QRS Assessment project.

Note: Notation in parentheses indicates the level at which a particular indicator is required.

n/a = not applicable

[a]The curriculum component is currently being assessed and scored but points are not yet included in the final rating.

not specify a particular mode of assessment, but requires that the measure used be reliable. Pennsylvania has the most specific standards for child assessments—requiring the use of observational assessments at least once and up to three times a year (increasing by QRIS level) and specifying the electronic collection and reporting of child outcomes using the Ounce and Work Sampling assessments by level three.

Finally, QRIS also assess the extent to which providers make accommodations for particular subgroups of families they serve. Four QRIS assess provisions that programs offer to support children with special needs. These strategies range from having a screening and referral process in place (Miami, Illinois, and Pennsylvania), adapting space and materials for children with special needs (Indiana), and developing activities for children with potential delays (all but Tennessee). QRIS assess how programs recognize and support cultural and linguistic diversity by providing resources to communicate with families whose primary language is not English (Illinois and Miami), making an effort to learn about the cultural background of families (Illinois), and making sure that classroom materials reflect diverse cultures and backgrounds (Indiana).

Accreditation. Prior to the emergence of QRIS, accreditation was one way for child care programs and providers to distinguish themselves from others beyond complying with licensing regulations. Across QRIS, respondents perceived that accreditation represents the high end of the child care quality spectrum. According to respondents, the accredited providers who were involved in QRIS planning noted that the accreditation process does not always reap rewards for providers in terms of increased demand for care. These providers were reportedly supportive of QRIS as another means to encourage accreditation. However, it was also universally acknowledged that the accreditation process takes a considerable amount of time and resources on the part of providers. If accreditation was to be required to progress through the system, planners and administrators wanted to ensure that the system would be able to offer providers the necessary supports and resources to achieve it.

Two QRIS—Illinois and Indiana—require accreditation to reach the highest level (Table II.16). While accreditation is required of providers at level four in Indiana, providers must also undergo an observational assessment and demonstrate that they meet all PTQ standards of the lower levels. To support QRIS participants as they go through the accreditation process, Indiana linked an existing program, the Indiana Accreditation Project, to their system. The project, an initiative sponsored by the Bureau of Child Care and administered by the Indiana Association for the Education of Young Children (IAEYC), offers child care providers financial and technical assistance as they work through each stage of the accreditation process. Supports provided through this project were reported to be helpful in the original implementing counties (Elicker et al. 2007). During planning for statewide PTQ, representatives from IAEYC and the original counties worked with administrators to ensure that the same types of supports would be available for new providers participating in the statewide system.

Illinois accepts accreditation as an alternative to undergoing assessments (ERS and PAS) at level three, and requires accreditation in addition to individual requirements at level four. At both levels three and four, providers must still comply with the staff qualification requirements, but it is only at level four where providers are required to be accredited and also score above 5.0 on the ERS and PAS. Respondents in Illinois noted that they are reconsidering the exemption from observational assessments at level three because the current procedure prevents them from examining changes in quality for accredited providers over time and comparing observational measures of quality for providers at different levels.

Table II.16. Role of Accreditation in QRIS for Center- Based Programs

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Role of Accreditation in Rating	"Plus" designation to final rating	May be used as alternative to requirements at Level 3, required at Level 4	Required at Level 4	May be used to fulfill partial requirements at Level 4	Two extra points before calculation of final rating
Observational Assessment Required for Accredited Providers	Yes	At level 4 only	Yes	Yes	Yes
Other Requirements at Accreditation Level	n/a	None at Level 3. Yes at level 4, staff qualifications	Yes, must meet standards for lower levels	Yes, staff qualifications	n/a
Number of Accrediting Bodies Currently Recognized[a]	13	4	4	5	9

Source: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); site visits conducted as part of the QRS Assessment project.

n/a = not applicable

[a]Some agencies are specific to family child care or school-age care providers.

Other QRIS chose not to make accreditation a requirement but use alternative ways to incorporate accreditation status into ratings (Table II.16). Pennsylvania's system previously rated all accredited providers as level four automatically based on the assumption that those providers would meet the ERS standard at that level. Under current standards, Keystone Stars accepts accreditation to fulfill partial requirements at level four, although providers must meet requirements for staff professional development, strategic planning, and child assessment. Accredited providers must also undergo an ERS assessment and achieve a facility average score of 5.25 or higher. Respondents indicated that it was essential to conduct observational assessments of all programs for research and comparison purposes. They also noted the importance of demonstrating that the time and effort required of accredited providers is comparable to what is required of providers who are not accredited.

The Miami-Dade and Tennessee QRIS do not include accreditation as a requirement at any level, but award additional points to accredited providers. Miami-Dade Quality Counts adds a "plus" designation to the final quality rating for accredited providers, but accreditation does not actually affect the calculation of points that determines the quality rating. Tennessee currently awards two additional points to a provider's total score (across components) before taking the average score to obtain a QRIS rating. Planners had discussed requiring accreditation at the highest level but were concerned about the accreditation costs preventing providers from achieving the highest rating

levels. They were also concerned about possible changes to accreditation standards that were beyond their control and how those would reflect upon their own homegrown QRIS standards. In addition, they noted that most accrediting agencies conduct on-site visits once every three years, whereas their licensing procedures require annual visits (at least) and they wanted to maintain that protocol consistently.

Respondents emphasized the importance of conducting a review of standards and procedures set by external agencies—such as accrediting bodies or the HSPPS—for the purpose of maintaining transparency and consistency in the standards and the meaning of quality ratings. In some cases, QRIS procedures have been revised to improve alignment between external standards and QRIS criteria. For example, Illinois previously accepted compliance with HSPPS as a substitute for accreditation, but they changed this policy when they found that not each specific center location affiliated with large Head Start grantee organizations undergo an on-site review. Administrators felt strongly that each facility participating in Illinois Quality Counts should receive an on-site assessment.

Most QRIS have a list of approved accrediting bodies. For center-based programs, all QRIS recognize accreditation through the NAEYC and National Early Childhood Program Accreditation (NECPA).[9] These lists are continuously updated and reviewed. Indiana and Pennsylvania have instituted formal processes to review accrediting agencies' procedures and standards and ensure alignment with QRIS requirements. Organizations that are interested in becoming a recognized accrediting body in the state must submit an application to the state child care administration office demonstrating that they meet certain criteria. The two QRIS use similar criteria to evaluate accrediting organizations beginning with fiscal and administrative capacity, an established accrediting process and procedures to avoid conflicts of interest. Pennsylvania specifies that an organization's process must have been in place for three years or more and Indiana requires at least 100 accredited facilities. Both QRIS require an on-site assessment as part of the accrediting process and organizations are asked to document qualifications and training procedures of staff who conduct the assessments. Further, the protocol for establishing and maintaining the reliability and validity of assessment procedures and instruments must be described in depth. Finally, organizations are required to link accreditation standards to research on child care quality and demonstrate how standards are aligned with those of the state QRIS.

2. Laying out the Structure of Quality Rating Levels

Once planners had identified the different components that would be included in ratings, discussions turned to determining how the QRIS would be structured. Structural aspects of the QRIS include how individual indicators would be combined to calculate an overall rating, the number of rating levels and what would be required at each, and what the levels should be called. These structural features for each of the five QRIS are summarized in Table II.17. Understanding the structure of each QRIS and the factors considered in their design allows for a fuller appreciation of what ratings mean and the extent to which ratings are comparable across systems.

[9] Miami's Quality Counts recognizes accreditation as defined by Florida's Gold Seal program. The Gold Seal program is administered independently of Quality Counts and has its own application and review process for both accrediting agencies and providers. The Gold Seal list currently includes 13 organizations, including those recognized by Illinois and Indiana. Providers with Gold Seal accreditation receive higher reimbursements for the subsidized children they serve (Florida Department of Children and Families, 2011).

Table II.17. QRIS Rating Structure, Levels, and Terminology

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Rating Structure	Combination	Building block	Building block	Building block	Combination
Number of Levels	5	3 or 4[a]	4	4	3
Site-Specific Term for Levels	Stars	Stars	Level	Stars	Stars

Source: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); site visits conducted as part of the QRS Assessment project.

[a]License-exempt homes have a separate 3-tier system in Illinois.

Methods for combining indicators. The five QRIS combine and aggregate components in one of two ways. Illinois, Indiana, and Pennsylvania use a building block approach in which a provider must meet all of the standards required at one level before moving on to the next (higher) level. Miami and Tennessee employ combination systems, wherein a provider is rated on individual components before scores are combined to obtain an overall rating. Combination systems resemble building block systems within rating components. That is, criteria are set to meet a particular level for each component. The key difference between building block and combination systems is that in combination systems, a provider may meet criteria for a high rating on one component but a lower rating on another. In Miami, each provider receives a certain number of points for each component, those points are added up and the point ranges for star levels correspond to ranges of total points. In Tennessee, points for each component are added up and averaged. The average score is rounded up to a whole number, which becomes the provider's overall rating.

Respondents in QRIS using a building block approach noted that this structure allows for a clear representation of what each level of quality looks like. Knowing a provider's overall rating allows for easy identification of the individual criteria that a provider has met because each one would have had to be fulfilled to achieve that rating. One respondent noted that they did not want to have "too many possible ways to get to a rating." Conversely, planners from the QRIS that employ a combination system indicated that it was necessary to provide multiple avenues to achieve a higher rating, while still prioritizing what they felt were the most important elements.

Number of rating levels. Planners' and administrators' knowledge of licensing and accreditation standards served to establish a range of quality for the QRIS to cover. Prior to establishing a QRIS, many states implemented what were essentially two-level QRIS that rewarded accredited providers with higher reimbursements from the state for serving subsidy-eligible children and families. However, the sharp disparity between licensing and accreditation requirements was such that few providers successfully pursued accreditation (Mitchell 2005; Tout et al. 2010). By adding intermediate benchmarks for quality to bridge licensing and accreditation standards, administrators believed that the QRIS would be more likely to get providers in the door and offer targeted supports for quality improvement.

The variation in baseline requirements across the five QRIS reflects differences in perceptions about what were reasonable expectations for a beginning QRIS participant, based on licensing requirements and what planners knew about current levels of quality and resources in each QRIS. At the first level, the number of components assessed ranges from three in Illinois and Indiana, to eight

in Pennsylvania (Table II.18).Respondents in Miami noted their perception that that the baseline for quality as signaled by state licensing requirements was relatively low. For example, the requirements at Miami's first level are typically lower than standards set by other QRIS, particularly in the area of child-staff ratios and group size. Also, providers need not meet all the listed criteria at the lowest level, given Miami's combination system.

The number of intermediate levels in each QRIS was influenced by what planners and administrators felt were reasonable expectations in terms of improvements that providers could achieve over time and supports that could be provided to help providers make that progress. These issues were balanced against the need for some distinction between each level so that providers are required to exert some effort to progress from one level to the next. Miami's planners felt that providers were likely to need more scaffolding to progress to the highest level; thus, the QRIS has five rating levels in order to accommodate more gradual changes in requirements from one level to the next. Tennessee planners were determined to keep the program simple by limiting the system to three levels. There is a smaller gap between Tennessee's lowest and highest levels and their requirements at the lowest level, particularly for ERS scores, are higher than the other QRIS. Tennessee's system also assesses fewer components in total, making it feasible to have fewer levels. In Illinois, planners were concerned that providers might be discouraged from participating if they set requirements at the highest level too high. As a result, they set the bar slightly lower but left room to add higher levels later.

Requirements for the highest rating level are comparable across QRIS and far less disparate than requirements at the first level (Table II.19). At the highest level, the range in the number of components assessed is much narrower, with the QRIS assessing six to nine components. Further, the indicators are similar. The narrowing of disparities at the highest levels reflects the influence of nationally recognized standards of child care quality. In fact, there is considerable overlap in requirements at the highest levels and NAEYC accreditation requirements (Appendix Table A.2), even if the two QRIS that require accreditation at the highest level are excluded.

Terminology for levels. Respondents in each QRIS indicated that they devoted a substantial amount of thought and discussion during QRIS planning to what ratings should be called, such as stars or levels. Planners and stakeholders were cognizant that the symbolism accompanying this choice in terminology could affect public perceptions of what the ratings mean. In the end, decisions were made based on which aspects of the system were to be emphasized. Four of the QRIS use the term "stars" because respondents in these QRIS noted that the term denotes a certain level of prestige and accomplishment that they want providers to associate with being a QRIS participant.

Although Illinois also refers to rating levels as stars, respondents there noted concern that the prestige associated with the term is only being attributed to the highest star ratings. One respondent drew a parallel with hotel ratings, stating that "no one wants to stay at a one-star hotel." There is a concern among administrators that the stigma associated with a low star rating is leading providers to hold out of the system until they feel they are at a level that would merit a higher rating. Indiana's QRIS, the only one of the five that does not use the term "star," was designed to emphasize the value of the quality improvement process—hence, the use of the term "paths." All PTQ participants spend at least a brief period at the first level. Respondents indicated that this terminology reduces the stigma associated with being at a lower level (Table II.17).

Table II.18. QRIS Components Necessary to Meet Lowest Rating Level for Child Care Centers

QRS Component Category	Quality Counts, Miami-Dade County[a]	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality[a]
Licensing Compliance	n/a	Yes	Yes	Yes	Yes
Child-Staff Ratio[b]	Infants 4:1 Ones 6:1 Twos 11:1 Threes 15:1 Fours/Fives 20:1	Infants 4:1 Ones 5:1 Twos 8:1 Threes 10:1 Fours/Fives 10/20:1	Infants 4:1 Ones 5:1 Twos 5:1 Threes 10:1 Fours/Fives 12/15:1	Infants 4:1 Ones 5:1 Twos 6:1 Threes 10:1 Fours/Fives 10:1	Infants 4:1 Ones 5:1 Twos 8:1 Threes 9:1 Fours/Fives 13/16:1
Group Size[b]	n/a	Infants 12 Ones 15 Twos 16 Threes 20 Fours/Fives 20	Infants 8 Ones 10 Twos 10 Threes 20 Fours/Fives 24/30	Infants 8 Ones 10 Twos 12 Threes 20 Fours/Fives 20	Infants 8 Ones 12 Twos 14 Threes 18 Fours/Fives 20
Health and Safety	n/a	No requirement	n/a	Illnesses and injuries tracked	n/a
Curriculum	Use of approved curriculum	No requirement	No requirement	Copies of Learning Standards on site	n/a
Environment	ERS Score 3.0-3.49	ERS Score 3.0-3.49	No requirement	ERS self-assessment but no cutoff	ERS Score 4.0-4.49
Child Assessment	n/a	No requirement	No requirement	No requirement	n/a
Staff Qualifications[d] (% of Staff)					
Education Level/Credential	HS/GED (100%), state credential (50%)	CDA or similar, and some college (10%)	Associate's degree	Associate's degree	HS/GED or Experience
ECE Credits	Yes	n/a	Yes	Yes	Yes
Continuing Education/In-Service Training	10 hours (100%)	15 hours	n/a	6 hours	18 hours
Years of Experience	n/a	n/a	n/a	n/a	Experience or HS/GED
Family Partnerships					
Family Resources	Family handbook	No requirement	No requirement	Family handbook, information on transitions	Modes of communication
Family Activities	n/a	No requirement	No requirement	Orientation	Parent meetings, Conferences
Family Participation in Planning	n/a	No requirement	No requirement	n/a	n/a

Table II.18 *(continued)*

QRS Component Category	Quality Counts, Miami-Dade County[a]	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality[a]
Administration and Management					
Staff Management	n/a	No requirement	No requirement	PD plan, staff meetings	Staff benefits
Fiscal Management	n/a	No requirement	No requirement	n/a	n/a
Administrative Management	Risk management plan in place	No requirement	No requirement	n/a	n/a

Source: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); site visits conducted as part of the QRS Assessment project.

n/a = not applicable (component not required by QRIS)

[a]Miami-Dade and Tennessee use combination systems so we provide the required indicators to receive the minimum score for each component.

[b]Illinois, Indiana, and Pennsylvania do not specify ratio and group size requirements for QRIS but require licensing compliance for QRIS participation. For comparison purposes, we present state licensing requirements for ratios and group size.

[c]For simplicity, we compare staff qualifications for a teacher.

Table II.19. QRIS Components Necessary to Meet Highest Rating Level for Center- based Programs

QRS Component Category	Quality Counts, Miami-Dade County[a]	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality[a]
Licensing Compliance	n/a	Yes	Yes	Yes	Yes
Child-Staff Ratio[c]	Infants 4:1 Ones 4:1 Twos 6:1 Threes 9:1 Fours/Fives 10:1	Infants 3/4:1 Ones 3/4:1 Twos 4-6:1 Threes 6-9:1 Fours/Fives 8-10:1	Infants 3/4:1 Ones 3/4:1 Twos 4-6:1 Threes 6-9:1 Fours/Fives 8-10:1	Infants 3/4:1 Ones 3/4:1 Twos 4-6:1 Threes 6-9:1 Fours/Fives 8-10:1	Infants 4:1 Ones 4:1 Twos 5:1 Threes 8:1 Fours/Fives 13/15:1
Group Size[c]	Infants 8 Ones 12 Twos 12 Threes 18 Fours/Fives 20	Infants 6-8 Ones 6-8 Twos 6-12 Threes 12-18 Fours/Fives 16-20	Infants 6-8 Ones 6-8 Twos 6-12 Threes 12-18 Fours/Fives 16-20	Infants 6-8 Ones 6-8 Twos 6-12 Threes 12-18 Fours/Fives 16-20	Infants 8 Ones 12 Twos 10 Threes 16 Fours/Fives 20
Health and Safety	n/a	Staff certification in CPR and first aid	n/a	Illnesses and injuries tracked, first aid certification of staff	n/a
Curriculum	Full implementation of approved curriculum	Curriculum aligned with standards[b]	Curriculum aligned with state ELS	Curriculum aligned with state ELS	Curriculum aligned with state ELS
Environment	ERS Score 5.5	ERS Score 5.0	Classroom arrangement, activities, and schedule, literacy materials	ERS Score 5.25	ERS Score 5.0
Child Assessment	Screening and referral process in place	Assessment process in place[b]	Assessment process in place	Twice/year	n/a
Staff Qualifications[d] (% of Staff)					
Education Level/Credential	Associate's (50%)	Associate's or Some college (30%)	CDA or equivalent (50%)	Associate's	CDA to BA (50%)
ECE Credits	Yes, not specified	15 hours	0-60 hours	18 credits	--
Continuing Education/In-Service Training	10 hours	--	20 hours (50%)	--	15 hours
Years of Experience	--	--	--	--	0 to 4 years (50%)

Table II.19 *(continued)*

QRS Component Category	Quality Counts, Miami-Dade County[a]	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality[a]
Family Partnerships					
Family Resources	Family handbook, Modes of communication	Modes of communication[b]	Modes of communication	Family handbook, information on transitions	Modes of communication, Community resource list
Family Activities	Parent meetings, Conferences, School events	Orientation, Parent meetings, School events[b]	--	Orientation, Conferences, Parent meetings, School events	Parent meetings, Conferences
Family Participation In Planning	Family evaluation	Family evaluation, Parent Advisory Council[b]	Family evaluation	--	Family evaluation, Parent Advisory Council
Administration and Management					
Staff Management	Differentiated salary scale, performance evaluation	Differentiated salary scale, benefits, performance evaluation[b]	Benefits	PD plan, staff meetings	Benefits, differentiated salary scale
Fiscal Management	Annual budget	Financial records, budget plan[b]	--	Annual business plan	--
Administrative Management	--	Risk management plan, Marketing and business plan[b]	Risk management plan, strategic plan	Risk management plan, strategic plan	--
Provisions for Special Needs	Activities for children with potential delays	Individualized plan, screening and referral process[b]	Plans and accommodations available	Individualized plan, screening and referral process	n/a
Cultural and Linguistic Diversity	Resources in families' primary language	Resources in families' primary language[b]	n/a	n/a	n/a
Accreditation	No	Yes	Yes	Optional	No

Source: QRIS Profiles developed for the Compendium of Quality Rating Systems and Evaluations (Tout et al. 2010); site visits conducted as part of the QRS Assessment project.

n/a = not applicable (component not required by QRIS)

[a]Miami-Dade and Tennessee use combination systems so we provide the required indicators to receive the maximum score for each component.

[b]Indicators assessed using the Program Administration Scale.

[c]Illinois, Indiana, and Pennsylvania do not specify ratio and group size requirements for QRIS but require or accept accreditation at the highest rating level. For comparison purposes, we present NAEYC-recommended ratios and group size.

[d]For simplicity, we compare staff qualifications for a teacher.

THIS PAGE LEFT BLANK FOR DOUBLE- SIDED PRINTING

III. PROCESSES FOR QUALITY MEASUREMENT

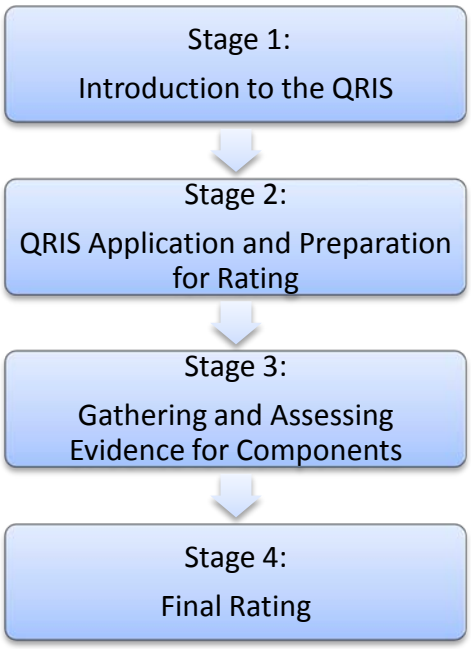
The contribution and ultimate effect of QRIS rests heavily on the measurement of quality, but little is known about the processes and formulas states use to construct and modify the ratings over time. For both practice and policy, it is essential to address the significant gap in the field's understanding of how rigorous these accountability systems are. This information is also critical to evaluation; it is necessary to know if the measurement is conducted systematically and with rigor in order to make better sense of analyses and findings.

In the previous chapter, we examined each site's definition of quality by delving into the composition of quality ratings. The next step in our investigation of quality measurement is an analysis of the processes that sites implement to collect information on each component and its indicators. We begin by describing how providers become involved with QRIS then go on to examine how the different quality components are measured and verified, and how components come together to constitute a quality rating. We also describe the different entities and staff members involved in the process and how staff members are trained. This provides important information for evaluating levels of reliability and validity of ratings at each site, as well as understanding comparability of ratings across sites.

A. Overview of the Quality Rating Process

Four common stages occur in the quality rating process across the five QRIS (Figure III.1). The order of the different procedures within each stage varies slightly by site. All QRIS introduce providers (stage 1) to the system and its processes through an overview and orientation. Tennessee's QRIS orientation is embedded in a required orientation session about the licensing process. In stage 2, providers prepare and submit a QRIS application and prepare for the rating. In four of the QRIS, the preparation for the rating occurs after the application; but in Illinois this preparation occurs

Figure III.1 Stages of the QRIS Rating Process



Source Site visits conducted as part of the QRS Assessment project.

prior to the point of QRIS application. In Miami-Dade not all providers that apply to the QRIS are enrolled. Rather the QRIS selects providers from among the applicants based on program size, geographic location, and percentage of subsidized children served. Stage 3 is the gathering and assessment of evidence for each of the QRIS components. The difference across the QRIS is the timing in the assignment of individual component ratings; in three QRIS this occurs prior to the Environment Rating Scale (ERS) assessment while in two QRIS this occurs after the ERS assessment. The final rating in stage 4 is the same across all five QRIS.

Differences in the rating process are driven, in part, by the infrastructure of each system—this includes the resources and staffing structure not just of the lead QRIS administrative agency, but also of various agencies that serve as system partners. Table III.1 lists the agencies involved in the administration and implementation of each QRIS. At least three entities are involved in the QRIS process: (1) the lead agency, which is the QRIS administrator, (2) the agencies that employ QRIS specialists who provide technical assistance and supports for the pre-rating process, and (3) one to three entities that spearhead the formal rating process including reviewing evidence for individual components ("raters"), conducting standardized assessments ("assessors"), and training raters and assessors. The number of unique agencies involved is larger if we consider the fact that some are subdivided further. For example, in four of the QRIS, specialists are spread out across multiple service delivery areas (SDAs). The number of SDAs ranges from six Regional Keys in Pennsylvania to as many as sixteen local child care resource and referral agencies (CCR&Rs) in Illinois. Even Miami's county-level system has four different agencies involved—two to work with centers in separate geographic areas, one to work with family child care providers, and one to provide support on professional development. And, in Indiana, the pre-rating support is provided by staff in two agencies based on levels of the QRIS rather than the type of care setting.

The breadth of operations for each QRIS makes it important to understand whether and how procedures are implemented to maintain consistency and reliability across entities within the system and throughout the rating process. The sections that follow will examine the degree to which QRIS standards and rating criteria are explained and applied consistently from the first time providers hear about ratings in an orientation session until a final rating is assigned.

B. Pre- rating Process

Examining how providers enter into the QRIS and the processes that take place before a quality rating is assigned is crucial for understanding how the population of QRIS participants varies by site. Differences in the pre-rating process and subsequent population of participants impacts the validity of ratings—that is, what a quality rating denotes in each site depends at least in part on what provider is being rated. A complete investigation of quality measurement processes in a QRIS requires an understanding of these nuances.

1. Introduction to the QRIS

The introduction period consists of sessions during which providers receive an overview of the QRIS and learn about requirements and procedures before formally enrolling or applying to participate in the system. The QRIS use these overview sessions to deliver information about the goals of QRIS, the system standards, expectations of participants, and resources and supports available to help providers at each stage of the process. Attendance of an orientation session is required for participation in all five sites and is documented in the system database (Table III.2).

Table III.1. Agencies Involved in the QRIS Rating Process

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
QRIS Administrator: Oversees operations	Early Learning Coalition (ELC) of Miami-Dade/Monroe	Illinois Department of Human Services (IL DHS)	Indiana Bureau of Child Care (BCC)	Pennsylvania Office of Child Development and Early Learning (OCDEL)	Tennessee Department of Human Services (TN DHS)
QRIS Specialists: Provides supports in preparation for rating	Independent Contractors (3 agencies[a]) Quality Counts Career Center (QCCC)	Local CCR&Rs (16 SDAs)	Local CCR&Rs (11 SDAs) Indiana Association for the Education of Young Children (IAEYC)	Six Regional Keys	Local CCR&Rs (10 SDAs)
Raters: Gather evidence for individual components	Devereux Validation Team QCCC	Illinois Network of Child Care Resource and Referral Agencies (INCCRRA)	TCC Software Solutions	Regional Keys (Designators)	TN DHS Program Evaluators (Licensing Unit)
Assessors: Conduct standardized assessments	Devereux Validation Team	National Louis University (NLU)	n/a	Pennsylvania Key	TN DHS Program Evaluators (Assessment Unit) University of Tennessee Social Work Office of Research and Public Service (UT-SWORPS)

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable

[a]Independent contractors include Florida International University, University of Miami, Family Central, Inc.

Table III.2. QRIS Orientation and Application Process

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Orientation Required	✓	✓	✓	✓	✓
Orientation Attendance Recorded	✓	✓	✓	✓	✓
Application Required	✓	✓	✓	✓	
Apply for Specific Level		✓			
Voluntary Participation	✓	✓	✓	✓	✓

Source: Site visits conducted as part of the QRS Assessment project.

Orientation sessions are provided by representatives from local child care resource and referral agencies in three sites (Illinois, Indiana, and Tennessee). Respondents perceived this to be a successful strategy because it builds on relationships these agencies tend to already have established with local providers. Both Pennsylvania and Miami-Dade contract with local agencies or universities to provide orientation and technical assistance rather than with CCR&R agencies.

2. Application

QRIS applications are fielded by either local CCR&R agencies or the lead oversight agency for the QRIS. All four QRIS that have a formal application process use paper application forms (Table III.2). Tennessee is the only QRIS that does not require a formal application because the Star Quality program is administered in conjunction with the licensing process and all Star participants must be licensed in order to participate. When assessments as part of the licensing evaluation process are complete, licensing staff members inform a provider of their eligibility to participate in the Star Quality program. Providers are automatically enrolled in Star Quality if eligible, but may choose not to participate. Illinois is the only QRIS in which the application stage comes after the preparation for ratings. This is necessary because they require providers to specify a level for which they are applying.

Participation is voluntary in all five QRIS (Table III.2). In Illinois, Indiana, and Pennsylvania, all eligible providers are able to participate in the system and submission of an application constitutes entry into the system. Miami has an additional selection process for all providers except for Head Start programs. All Head Start programs participate in Quality Counts. Other providers who submit applications are placed on a waiting list. From this pool of applicants, administrators select providers to participate in Quality Counts using an automated algorithm that is designed to enroll a representative group of providers based on program size, geographic location (including high poverty zip codes), and percentage of subsidized children served. The number of providers selected depends on the availability of resources to provide technical assistance and incentives. The Early Learning Coalition of Miami-Dade (ELC) negotiates a yearly contract with partner agencies that specifies target numbers for QRIS applicants and participants to be served.

Respondents from newer QRIS (Miami, Illinois, and Indiana) noted that certain issues seem to affect providers' decision to participate in the system. For example, in all three QRIS, participation rates for family child care providers are lower than that of center-based providers. Among the possible reasons cited for this difference is that QRIS requirements, particularly in terms of staff qualifications and standardized assessments, can be intimidating for family child care providers. Illinois administrators also expressed concern that providers (across types) are delaying participation until they are confident that they will be able to achieve a high rating. Indiana respondents observed that because their system requires all providers to begin at level one, there is less trepidation among potential QRIS participants regarding achieving a high rating right away and more value ascribed to the process of improving over time.

3. Preparation for Rating

Beyond a general orientation to QRIS, each site has a process in place to help providers learn about QRIS standards and ways to meet requirements. The objective of the preparation stage is to help providers gauge whether they are ready to undergo the rating process. The different supports available during this stage are summarized in Table III.3. Supports include self-study materials such as workbooks and worksheets to help understand the standards and conduct self-assessments,

additional training sessions, and individualized technical assistance provided through consultation with the QRIS specialist.

Table III.3. Supports Provided During Preparation for QRIS Rating

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Self-Study Materials	Yes	No	Yes	Yes	No
Self-Assessment Training	Required	Required	Optional	Optional	Optional
Self-Assessment	Required	Optional	Required	Required	Optional
Self-Assessment Verified by QRIS Specialist	Required	Optional	Required	Required	Optional
Consultation with QRIS Specialist	Optional	Optional	Optional	Optional	Optional

Source: Site visits conducted as part of the QRS Assessment project.

Self-study materials and training. All QRIS distribute materials that providers can use to familiarize themselves with the rating criteria and conduct a self-assessment (Table III.3)—an exercise to measure themselves against the QRIS standards and determine which level they are likely to receive. Miami, Indiana, and Pennsylvania have standardized materials that are used system-wide. Miami-Dade's Self-study Packet for centers includes definitions of key terms and formulas for calculating the percentage of staff meeting requirements for the staff qualifications component. It also lists various documents that assessors will accept as evidence. Indiana and Pennsylvania have materials tailored for each rating level. Indiana's Provider Workbook includes suggestions and tips for steps that providers can take to meet the requirements, and a bibliography of resources. Pennsylvania's Stars Worksheets are designed for use from the preparation stage until the actual rating process begins. For each level, worksheets list rating criteria and allot spaces to denote status at each stage. Self-study materials in Illinois and Tennessee are created by local CCR&Rs.

Along with the written materials distributed to providers, there are also self-assessment training sessions that providers can attend. These sessions discuss the QRIS standards in more depth than the overview sessions, particularly in terms of the use of standardized assessments and interpreting the results. Miami and Illinois require attendance in self-assessment training sessions. Miami's sessions provide guidance on all QRIS requirements, while in Illinois, providers are required to attend separate training sessions on the ERS and the Program Administration Scale (PAS) if they are applying for a level that requires it. Indiana and Pennsylvania only require attendance for providers who have chosen not to work one-on-one with a QRIS specialist, although all directors and family child care providers in Pennsylvania must attend training on the ERS. Although self-assessments are only required in three QRIS, respondents universally acknowledged the importance of this process for maintaining transparency of the rating process and enabling providers to use results to make improvements in quality. Respondents in Illinois noted that the self-assessment is particularly necessary because providers are asked to apply for a specific level.

Technical assistance on the QRIS process. All five QRIS offer the services of a specialist to help providers as they proceed through the preparation process. QRIS specialists offer guidance on the interpretation of standards and can connect providers to additional resources for more specific needs, such as specialized technical assistance to address the content of the standards (for example, on curriculum or setting up the physical environment). In two QRIS, the specialist's role is further

subdivided based on specific functions. In Miami, providers also receive support from career advisors who assist with staff qualifications and professional development requirements. In Indiana, specialists vary depending on the provider's current rating level. Providers at levels one and two have access to mentors from local CCR&Rs. When providers are at level three and preparing for accreditation to attain level four, they have access to quality advisors from the IAEYC.

The extent of support during the preparation stage tends to be driven by individual providers' needs. The role of the QRIS specialist may be informal as a point person that providers can contact as questions arise. All five QRIS have individuals designated for this purpose. In four QRIS (Indiana, Miami, Pennsylvania, Tennessee), specialists may also take on a formal role as a mentor if the provider prefers additional, more structured guidance. Indiana and Pennsylvania have specified a number of hours available to each provider to receive this guidance and support on the rating process. In Indiana, providers can receive up to 25 hours of this support per rating cycle; in Pennsylvania, providers can receive up to 40 hours over a six-month period. Whether or not a formal mentoring relationship is established, one of the key services specialists offer across the five QRIS is reviewing self-study materials and documents to determine whether requirements are complete and accurate. This step, which is essentially a dry-run of the rating process, is required in three QRIS sites and optional in two (Table III.3). Across the QRIS, respondents noted that the review conducted by QRIS specialists can minimize confusion when the rating process begins and ensure that required documents are available and organized for review. And, if missing documents are noted, providers have time to obtain them prior to the actual rating process.

The QRIS have invested effort and resources in training QRIS specialists so that the information given to providers during the preparation stage is consistent with the expectations of raters and assessors during the rating process. Pennsylvania has recently developed a "strategic conversations" protocol to help specialists provide clarification about the standards. This tool covers topics such as how to calculate director qualifications when there are two directors and which pieces of information might take longer to obtain and should be asked about early in the preparation process. In Illinois, the assessment team at National Louis University (NLU) recently began providing ERS training to CCR&R staff so that QRIS specialists can provide targeted supports on the ERS. In Indiana, QRIS specialists have monthly meetings with raters to address questions the specialists may have about the interpretation of standards. Respondents in each QRIS noted the importance of having QRIS specialists, raters, and assessors, and often licensing staff, on the same page in interpreting rating criteria and standards. While each QRIS stresses the importance of consistency in messaging, there is equal import placed on objectivity; each QRIS maintains a firewall between the different roles in order to maintain objectivity in assigning ratings.

C. Gathering Evidence for Individual Components

When providers, with the help of QRIS specialists, have examined rating criteria, compiled necessary documentation, and prepared facilities and staff, the rating process can begin. The first step in the formal rating process is an evidence review to determine whether a provider meets requirements for individual quality components.

While there is variation in which entity conducts this stage of the process across the QRIS, the commonality is that the raters, who gather and review evidence for individual components, are distinct staff from QRIS specialists who perform the pre-rating and supportive roles with providers (Table III.4). In Miami and Indiana, the evidence review is conducted by independent contractors. In Pennsylvania, raters are housed within Regional Keys. Raters may serve as QRIS specialists for some providers but never provide both technical assistance and evidence review for the same

provider. Tennessee raters are licensing staff who, along with reviewing evidence for QRIS, also check other licensing requirements including health and fire safety inspections. Finally, in Illinois, raters are staff members of the Illinois Network of Child Care Resource & Referral Agencies (INCCRRA).

Table III.4. Staff Responsible for Formal QRIS Rating Process

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Gather Evidence for Individual Components	Rater	Rater	Rater	Rater	Rater
Conduct Standardized Assessments	Assessor	Assessor	n/a	Assessor	Assessor
Assign Component Ratings	Automated in database	Rater	Automated in database	Rater	Rater
Calculate Final Rating	Automated in database	Rater	Automated in database	Rater	Automated in database

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable

1. Qualifications and Staffing Structure of Rating Team

It is important to understand the background of individuals who conduct evidence reviews and how they are trained and supervised because these factors contribute to the reliability of ratings—the consistency with which a single rater assigns ratings over time, and the comparability of ratings between raters and across providers.

Rating teams vary substantially in size due to the nature and amount of work that raters are required to do in each site (Table III.5) For example, there are 140 raters across 17 field units in Tennessee because they also perform other tasks associated with general licensing of facilities. In Illinois, there are two raters because they do not have to conduct site visits and have fewer documents to inspect due to the smaller number of components reviewed in this stage. In Illinois, the two raters divide caseloads based on provider type; one rater focuses on

Table III.5. Number, Caseload, and Qualifications of QRIS Raters

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Number of Raters	7 (Devereux); 2 (QCCC)	2	4	Varies by Regional Key	140
Approximate Caseload Per Rater	10 cases per month (Devereux)	n/a	55 – 65 cases per month	15 – 20 per month	25 – 30 providers
Rater Qualifications	BA in ECE, classroom experience, bilingual	BA	BA in ECE or equivalent	BA in ECE, Master's preferred	BA

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable

reviewing the QRIS applications of licensed centers and license-exempt family child care providers while the other focuses on licensed family child care providers (along with other tasks outside of QRIS). They report this strategy to be helpful because the requirements and applications vary slightly depending on type of provider. Miami raters from Devereux have the smallest caseload, but they conduct both the evidence review and observational assessments. In addition, two raters from the QCCC share the task of verifying staff qualifications for all QRIS participants. Indiana raters have the largest caseload, conducting about 55-65 ratings per month.

The educational background and experience of raters can influence their effectiveness in applying the QRIS rating criteria and critically reviewing supporting evidence. All five QRIS require raters to have a bachelor's degree and three require that this degree be in early childhood education or a related field (Table III.5). A background in early childhood was noted to be helpful for consistent interpretation and application of QRIS standards.

2. Rater Training

None of the five QRIS has a formal protocol for training new raters; however, three have developed materials to improve the consistency of the evidence review process (Table III.6). Indiana and Pennsylvania provide written references for procedures to follow on-site. Illinois raters can refer to sample provider files and annotated screenshots of database forms, which provide additional guidance on different criteria as linked to fields on the form.

Table III.6. Training of QRIS Raters

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Rater Materials	None specified	Sample provider files, screenshots of database with explanation of fields	Rating Procedures Manual	Designation protocol and tip sheet	Scoring guidelines in development
Rater Training Protocol	Shadowing current rater for at least one month (about 10 visits)	No standardized protocol	Shadowing on at least four visits	Shadow at least one visit of a current rater	Shadowing
Required Initial Reliability	No specific threshold	No specific threshold	No specific threshold but a check between existing and new rater occurs during shadowing	No specific threshold	No specific threshold
Required Ongoing Reliability	No specific threshold	No specific threshold	No specific threshold	No specific threshold but designator reliability protocol is used	No specific threshold
Frequency of Ongoing Reliability Checks	None specified	None specified	Quarterly	Every 15 rating visits	None specified

Source: Site visits conducted as part of the QRS Assessment project.

Four of the QRIS require new raters to shadow a current rater to familiarize themselves with procedures and gain field experience. The amount of time a new assessor spends shadowing ranges from at least one visit in Pennsylvania, to approximately 10 visits in Miami. Illinois does not require this level of training because raters do not go to the provider site and the staff is small.

None of the QRIS has instituted formal guidelines for the initial or ongoing reliability required of raters. However, Indiana and Pennsylvania do require raters to undergo reliability checks by conducting paired reviews with another rater. In Indiana, reliability checks occur during the shadowing process in which two raters independently assess the provider and then check agreement. Pennsylvania's protocol for monitoring reliability of raters was only recently developed when administrators learned of inconsistencies in the interpretation of criteria among raters from different Regional Keys. A workgroup was established to address the issue and the designator (Pennsylvania's term for rater) reliability protocol was one of the group's recommendations. The current protocol requires that raters conduct an initial reliability visit with another rater before doing independent ratings. All raters are also required to conduct a reliability visit between every 15 rating visits.

Raters in Tennessee reported challenges in consistently interpreting standards. The current training for raters in Tennessee varies by field unit. Licensing staff throughout the state use policy and procedures manuals, but respondents noted that they only cover licensing procedures broadly and do not offer specific guidance on assigning ratings for the Report Card and Star-Quality program. Administrators are currently working on improving the process by creating scoring guidelines to provide additional information on QRIS requirements and corresponding evidence. Currently, respondents noted that inconsistencies are reconciled informally and discussed during meetings and supervisor review. There is also constant informal communication between raters via email or instant messaging. Inconsistencies are also reportedly caught across the different layers of supervisor reviews. The scoring guidelines have been developed in response to the noted inconsistencies and the desire to document decisions made in connection with them. These similar processes (discussions among raters and supervisory reviews) also resulted in updates to the rater materials in Indiana and Pennsylvania.

3. Procedures for Reviewing Evidence

Across the five QRIS, raters review evidence for at least 2 and as many as 10 components for each provider (Table III.7). Individual component ratings are assigned through review of evidence gathered by providers during the preparation process. Evidence is usually obtained through direct observation, director/provider interview, document review, or a combination of the three (Table III.8). The required evidence for some components is fairly straightforward—for example, providers need only present current certificates to demonstrate licensing compliance and accreditation status. Other components can be much more cumbersome, according to raters we interviewed from each QRIS. For example, reviewing staff qualifications requires multiple steps because the paperwork provided often pertains to individual staff members. Raters check that the documentation provided is legitimate and that the course or training attended is acceptable. Then the rater must verify the position of each individual, and after reviewing each document they must calculate the number of staff meeting requirements and whether that number matches the set criteria. Evidence for family partnerships and administration and management also includes a wide range of documents including family handbooks, sign-in sheets for school events, administrative records including financial statements, staff evaluations, and risk management.

Table III.7. QRIS Components Considered During Evidence Review by Raters

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Licensing Compliance		✓	✓		✓
Ratio and Group Size	✓				✓
Health and Safety		✓		✓	
Staff Qualifications	✓	✓	✓	✓	✓
Administration and Management	✓		✓	✓	✓
Family Partnerships	✓		✓	✓	✓
Community Involvement				✓	
Environment			✓	✓	
Curriculum	✓		✓	✓	✓
Child Assessment	✓		✓	✓	
Cultural and Linguistic Diversity	✓		✓	✓	
Provisions for Special Needs	✓		✓	✓	
Accreditation	✓ (if applicable)	✓	✓	✓ (if applicable)	✓ (if applicable)

Source: Site visits conducted as part of the QRS Assessment project.

Table III.8. Sources of Evidence for Quality Rating Components

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Licensing Compliance	n/a	Document review, existing database	Document review, existing database	Existing database	Document review, existing database
Ratio and Group Size	Document review	n/a	n/a	n/a	Observation
Health and Safety	n/a	Document review	n/a	Document review	n/a
Staff Qualifications	Document review, existing database	Document review, existing database	Document review	Document review	Document review
Administration and Management	Document review	Standardized Assessment	Director interview, document review	Document review	Document review, Staff interview
Family Partnerships	Document review	Standardized Assessment	Observation, director Interview, document review	Observation, document review, staff interview	Document review
Community Involvement	n/a	Standardized Assessment	n/a	Document review	n/a
Environment	Standardized Assessment	Standardized Assessment	Observation	Standardized Assessment	Standardized Assessment
Curriculum	Document review	Standardized Assessment	Observation, director interview, document review	Observation, director interview	Observation, document review
Child Assessment	Document review	Standardized Assessment	Director interview	Observation, document review	n/a
Cultural and Linguistic Diversity	Document review	Standardized Assessment	Observation	n/a	n/a
Provisions for Special Needs	Document review	Standardized Assessment	Observation, director interview, document review	Observation, document review	n/a
Accreditation	Document review	Document review	Document review	Document review	Document review

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable

Given the large number of documents to be inspected as well as items that need observation on-site (aside from ERS assessments), four of the five QRIS conduct evidence reviews during a visit to the provider (Table III.9). Illinois is the only site that requires providers to enclose supporting documents when they mail in their Quality Counts application because the only components requiring supporting documents are licensing compliance, staff qualifications, and accreditation. All other components are assessed using the ERS and the PAS/BAS.

Table III.9. Procedures for Reviewing Evidence on Individual QRIS Components

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Location of Evidence Review	Provider facility	Rater office, documents submitted by mail	Provider facility	Provider facility	Provider facility
Mode of Evidence Review Data Collection	Electronic	Electronic	Electronic	Paper	Paper

Source: Site visits conducted as part of the QRS Assessment project.

The QRIS employ a number of strategies to increase the efficiency of the evidence review process, given the large number of documents involved. For evidence of staff qualifications, for example, Miami, Illinois, and Pennsylvania capitalize on existing professional development registries as resources. In Miami, raters do not have to review staff qualification documents because the Quality Counts Career Center takes care of verifying staff qualifications and entering data into the PD registry, which is linked to Miami's QRIS database. In Illinois, raters save time and resources by searching the Data Tracking Program (DTP) for staff members identified by providers as meeting Quality Counts requirements (by meeting appropriate Great START levels). The DTP database lists the Great START level that each staff member has met. In Pennsylvania, the PD registry is used as a source of evidence to support professional development records for each individual staff in a child care facility. There are other components for which sites have streamlined the review process. Pennsylvania, for example, requires that child observations be conducted for all children in the program. Rather than reviewing the records for every child, raters will typically review files for two children selected at random. Administration and management requirements also require extensive review. In Miami, providers are asked to compile all materials such as staff handbooks and accounting records and denote specific page numbers in those documents to facilitate the review.

Raters use a variety of tools, both in paper or electronic form, to record results of their review (Table III.9). In Miami, Illinois, and Indiana, raters enter data directly into QRIS databases during the evidence review. Raters in Miami and Indiana use tablet computers to enter information into the QRIS database as they are reviewing providers' self-study forms. Raters denote whether each criterion has been met and the different documents that were reviewed. In Illinois, raters enter data directly into the QRIS database as they review supporting documents sent in by mail.

Raters in Pennsylvania and Tennessee use paper forms to record results of the evidence review process while on site. In Pennsylvania, raters use a worksheet that includes notes from the provider and QRIS specialist completed during the preparation process and a column for raters to complete during evidence review to denote whether each criterion has been met. Component level results from the evidence review are not entered into the QRIS database. In Tennessee, there is no

standardized form used by raters statewide apart from the actual report card, which lists requirements but does not allot space for note-taking. Some of the raters we interviewed said that they have created tools for their own use while others simply take notes on the report card itself. After the evidence review visit, Tennessee raters submit reports to their supervisor for review and approval after which the rater or supervisor enters data into the QRIS database.

D. Conducting Assessments Using Standardized Measures

Once the evidence review is complete and all the necessary paperwork is found to be in order, an assessment is scheduled, if required. In three of the five QRIS, assessors (i.e., staff who conduct standardized assessments) are distinct from raters (i.e., staff who gather and review other sources of evidence) and the assessments occur on a different day from the evidence review. In Miami, the two procedures occur on the same day and are conducted by the same staff. Indiana does not employ a standardized assessment to measure quality although their readiness checklists include indicators that are assessed through observation (as discussed in Chapter II).

Four of the QRIS assess the quality of the environment with the use of a standardized measure, while Illinois also assesses a number of other components with a standardized tool (Table III.10). Indiana does not assess any components using standardized measures. The four QRIS use the ERS to measure the quality of the environment. In addition to the ERS, Illinois uses the PAS and the Business Administration Scale (BAS) to measure the quality of administrative practices.

Table III.10. QRIS Components Assessed Using Standardized Measures

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
QRS Component Category	Environment	Environment Curriculum Child assessment Family partnerships Administration and management Cultural and linguistic diversity Provisions for special needs Community involvement	None	Environment	Environment

Source: Site visits conducted as part of the QRS Assessment project.

The use of standardized measures entails additional training efforts in order to adhere to guidelines provided by instrument developers and ensure that assessments are conducted as intended. We provide background characteristics of assessors and describe sites' procedures for training assessors in the following section. We specifically focus on training procedures for the ERS because that measure is common to four QRIS.

1. Qualifications and Staffing Structure of Assessment Team

The role of assessors in QRIS is critical given that provider scores on standardized assessments must often meet a specified threshold or are weighted more heavily in the rating process (as discussed in Chapter II). For this reason, it is also important to understand the background and training process of assessors that are typically hired, trained, and supervised by a different entity from that of the raters.

Similar to the pattern seen among raters, the number of assessors per QRIS ranges widely, from a low of 7 assessors in Miami-Dade to a high of 60 assessors in Tennessee (Table III.11). The workload for assessors is similar across sites, with assessors conducting between 8 to 12 assessments per month. The total number of assessments per month can vary within each QRIS depending on the amount of travel required and the need to conduct the assessment over multiple days (for larger facilities). In each of the four QRIS, the assessment team includes lead assessors who supervise groups of assessors and train new assessors. Lead assessors monitor the overall quality of assessments but also conduct assessments themselves, albeit with a smaller caseload. Lead assessors also serve as the anchors for the different ERS scales in three of the four QRIS. Anchors are experienced assessors upon whom the ratings of other assessors are measured for consistency. In Miami, Illinois, and Tennessee, anchors serve the entire assessment team. In Pennsylvania, anchors serve regional teams of assessors. Tennessee is the only QRIS in which anchors and supervisors are distinct. In Tennessee, lead assessors handle supervision duties and also conduct assessments. However, the training of new assessors and monitoring of reliability is conducted by anchors from the University of Tennessee Social Work Office of Research and Public Service (UT-SWORPS), a longstanding partner and contractor that also performs data process, analysis, and evaluation tasks for the Department of Human Services.

Table III.11. Number, Caseload, and Qualifications of QRIS Assessors

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Number of Assessors	7	10	n/a	16	60
Approximate Workload of Assessors	10 assessment per month	8-12 assessments per month	n/a	12 assessments per month	10-12 assessments per month
Number of Lead Assessors	4	4	n/a	4	9
Lead Assessors Are Anchors	Yes	Yes	n/a	Yes	No
Educational Qualifications of Assessors	BA in ECE	BA minimum, Master's degree preferred	n/a	BA in ECE, experience in early education	BA, preferably in early childhood
Additional Qualifications	Classroom experience, bilingual, demonstrated writing skills	None	n/a	Demonstrated writing skills	None

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable

BA=Bachelor of Arts

ECE = Early childhood education

In the four QRIS, assessors are required to have a minimum of a bachelor's degree, and are required or preferred to have this degree in early childhood education. Illinois prefers that assessors hold a master's degree, a qualification that all of their current assessors meet. Tennessee is unique in their structure because assessors are state employees and technically hold the same position as licensing staff. As a result, they are unable to specify the qualifications they would prefer. However, they do prefer to hire assessors with a background in early childhood education, when possible. In addition to education level, two QRIS require experience in early childhood settings and Miami requires assessors to be bilingual due to the demographics of the providers and families in their locality. Miami and Pennsylvania also gauge the writing skills of assessor candidates. In Miami, assessor candidates are asked to compose a written response to a given scenario. In Pennsylvania, applicants must submit a writing sample.

2. Assessor Training

Few of the current assessors in the five QRIS have received direct training from the authors of the ERS (Table III.12). Respondents noted that, given the large number of assessors involved, it is not feasible in terms of time and resources to send each assessor to the North Carolina location for training with the ERS authors. Miami respondents noted that they capitalize on opportunities to send assessors to local training sessions conducted by ERS authors when the opportunities arise (typically once per year). Tennessee's assessment staff received direct training in previous years but this is no longer the standard practice as their system has expanded.

In lieu of sending all assessors for training with ERS authors, three QRIS have sent at least some of their lead assessors to receive training. Whether or not they receive training directly from the authors, each QRIS has built upon publisher-provided materials and guidelines to design protocols for conducting training sessions in-house. The basic parameters of the training process are similar across sites. In particular, the objective of training is to achieve 85 percent agreement with an established benchmark for ratings. Below, we describe the steps sites take to train assessors to this optimal level and ensure that the standard is maintained.

Pre-service training. Training for new assessors begins with in-depth study of training literature and practice vignettes. Respondents from the four QRIS mentioned using publisher-provided training materials at this stage, including training videos and workbooks, U. S. Department of Agriculture meal guidelines, and the Caring for Our Children: National Health and Safety Performance Standards. According to respondents in each QRIS, trainees also examine sample assessment reports and site-specific materials on ERS policies and procedures. In Tennessee, trainees are provided with a model classroom to familiarize themselves with classroom arrangements and materials.

Across the four QRIS, respondents stressed that an essential ingredient of training is shadowing experienced assessors. The bulk of the assessor training period is spent with trainees conducting practice observations in actual child care facilities, accompanied by anchors. Trainees begin by observing assessment procedures and eventually complete assessments of their own. Anchors and trainees review individual items together and discuss the scores to assign. As the number of practice observations increases, trainees progress to assigning scores independently and compare those scores with those of the anchor, once the observation has concluded.

In addition to training on the administration of the ERS, new assessors must also develop other skills related to conducting assessments. In Tennessee, assessors are evaluated on their skills in

Table III.12. Training of QRIS Assessors

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Assessors Trained by Authors	Some	Some	n/a	No	No
Anchors Trained by Authors	Yes	Yes	n/a	No	Some
Number of Anchors	4 anchors, each anchor on 1-2 scales	4 anchors, each anchor on 2 scales	n/a	4 anchors, each anchor on all scales	3 anchors, each anchor on all scales
Number of Practice Observations for Training	~7	~2	n/a	~2	~6
Required Initial Reliability	85% average across 3 consecutive ratings	85 % average across 3 consecutive ratings Each of the last 2 observations must be 85 percent or higher	n/a	85% on each of 5 consecutive observations with anchor	85% average across 3 consecutive ratings
Number of Scales Each Assessor is Trained and Reliable On	1-3 scales	2 ERS scales 1 PAS/BAS	n/a	All scales	1-4 scales

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable

introducing and concluding the observation visit, selecting classrooms, documenting adequately with informal notes, assessing the playground, conducting the teacher interview, writing assessor notes, and completing the ERS form. Other sites provide training on writing assessment reports as well. Like Tennessee, Miami and Illinois integrate report writing training during practice observations. In Pennsylvania, new assessors do not start report writing until they are reliable on the scales so that they can first focus on scoring.

Initial reliability. After a certain number of practice observations (ranging from about two in Pennsylvania to about seven in Miami), anchors begin tracking the agreement between their scores and those of the trainee. In each of the four QRIS, the standard for reliability of a new assessor is 85 percent agreement with an anchor. However, the number of assessments over which this threshold must be met ranges from three (in Miami, Illinois, and Tennessee) to five (in Pennsylvania). There are additional subtleties across the systems as to how this standard is enforced. For example, in Miami and Tennessee, an assessor must achieve an 85 percent agreement, on average, across three consecutive assessments. In Illinois, there is an additional requirement that the last two of the three observations be at 85 percent agreement or higher.

Assessors may be trained to administer several scales (Table III.12). The same training protocol is generally followed for each scale. In Illinois and Pennsylvania, all assessors begin by training on the ECERS. Assessors do not proceed with training on other scales until reliability is achieved on the first ERS.

3. Ensuring Ongoing Reliability

Across sites, assessment teams maintain constant communication with each other to discuss issues that come up while out in the field, particularly as it relates to the interpretation of scale items. This communication occurs both through informal means as well as regularly scheduled staff meetings. Lead assessors also consult with ERS authors on an interim basis to discuss issues related to training, reliability, and use of the ERS.

In addition to regular communication, sites employ several strategies to ensure that the quality of assessments is consistent over time. First, across sites, all assessors undergo periodic reliability checks to avoid rater drift and ensure that the 85 percent reliability standard is maintained (Table III.13). For new assessors, these ongoing checks occur every 6 to 10 observations in Miami, Illinois, and Tennessee, and quarterly in Pennsylvania. Two QRIS reduce the frequency of ongoing checks for more experienced assessors. For example, in Pennsylvania, ongoing checks are reduced to twice a year for more experienced assessors, and in Tennessee, an assessor who has worked on a scale for six months and maintained an average of 90 percent agreement for three consecutive reliability checks is designated as a veteran assessor. Tennessee also designates veteran extended assessors who have worked on a scale for 18 months and demonstrated an average of 90 percent agreement across six consecutive reliability checks. Veteran assessors undergo a reliability check every 12th to 15th observation for a given scale and veteran extended assessors undergo a reliability check every 18th to 21st observation visit.[10]

[10] In addition to the number of observations conducted, the need for a reliability check is also determined by the number of months since the previous reliability check. Both veteran and veteran extended must have an ongoing

Two of the larger assessment teams, from Tennessee and Pennsylvania, have developed additional documentation that assessors use as a resource in conducting ERS assessments. Tennessee has additional notes on the ERS, which incorporate updated notes from the ERS authors but also address ongoing local issues. The original impetus for creating the notes was a request from the Tennessee QRIS administrator to review some ERS items to ensure they were contemporary and in keeping with Tennessee licensing requirements. They have since evolved as a document that is revised or updated periodically to reflect clarifications about interpreting and scoring certain ERS items. There are separate notes for each scale, and all notes are shared with and approved by the ERS authors. Pennsylvania has a similar set of documents called "position statements" for each scale. Licensing staff and assessors in Pennsylvania work together to draft these position statements that provide guidance on the interpretation of ERS items based on licensing standards.

Table III.13. QRIS Strategies for Maintaining Ongoing Reliability of ERS Assessments

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Required Ongoing Reliability	Same as initial	Same as initial	n/a	85% on 2 consecutive observations with anchor	Same as initial
Frequency of Ongoing Reliability Checks	Every 6 -10 observations	Every 6 -10 observations	n/a	Quarterly to bi-annually	Every 6 - 9 observations
Site-Specific Notes on ERS Administration	No	No	n/a	Yes	Yes
Cross-Region Checks	No	No	n/a	Yes	No

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable

An additional concern in Pennsylvania is ensuring consistency of ERS assessments across regions because the reliability of assessors is checked against a region-specific anchor. On the advice of ERS authors, Pennsylvania conducts at least two cross-region reliability checks per year. Assessors from different regions within the state conduct paired observations and compare their ratings to investigate potential discrepancies in scoring.

4. Procedures for Conducting Classroom Observations

In addition to the qualifications and training of assessors, the procedures each site implements in conducting observations affects the subsequent reliability of assessment results. We now discuss

(continued)

reliability check at least every six months. Even if the threshold for the number of observations has not been met at the end of six months, a reliability check is still conducted.

how assessors prepare to conduct classroom observations, when observations are conducted, how the appropriate ERS scale is selected, and other factors that affect how ERS data are collected. Table III.14 presents a summary of procedures that sites implement in conducting classroom observations. Although Indiana does not use a standardized measure, we include Indiana in our comparison of procedures that sites implement in conducting classroom observations. Examining these procedures informs our understanding of the extent to which the results of these observations consistently capture the quality of classroom environments and whether differences in procedures may subsequently impact the comparability of ratings across sites.

Preparing for the on-site visit. When a provider requires an observational assessment, assessment teams are notified to begin preparing for an on-site visit. This notification is sent by fax, email, or phone in Illinois, Miami, and Tennessee. Assessors in Pennsylvania receive assessment requests through an electronic list of providers linked to their QRIS databases. The list is updated in real time, enabling assessors to claim cases from the queue as they become available.

As a first step in arranging an assessment visit, assessors obtain information on the ages of the children in each classroom at the facility. In Miami and Illinois, this information is gathered during a scheduling call. In Pennsylvania, the information is obtained by raters during their evidence review visit with providers; however, assessors find it helpful to make a phone call to providers in order to answer any questions, build rapport, and set the tone for the upcoming visit. Information on children's ages is used to determine which measures will be used for the observation. All four QRIS use the Infant/Toddler Environment Rating Scale-Revised (ITERS) in infant/toddler classrooms, ECERS-R in preschool classrooms, School-Age Care Environment Rating Scale (SACERS) for school-age classrooms, and Family Child Care Environment Rating Scale (FCCERS)[11] for family child care. In mixed age classrooms, sites typically use the measure appropriate for the age of the majority of children in the room. Pennsylvania and Tennessee have additional guidelines for mixed-age classrooms in which the age distribution is split equally. In Pennsylvania, assessors use the scale appropriate for the older age group in the classroom. In Tennessee, program directors are allowed to choose which measure will be administered.

Three sites do not inform providers of the exact date of the visit (Table III.14). Instead, Illinois, Miami, and Pennsylvania give providers a window of three to four weeks during which they can expect the assessment visit to take place, and allow providers to designate blackout dates during which they cannot be observed due to scheduling conflicts. According to the respondents, the reasoning behind this approach is that they are more likely to observe what a typical day is like if the visit is unannounced. Tennessee provides a specific date beforehand, although assessors did note some isolated incidents of programs attempting to falsify an assessment by buying new equipment, or asking some children to stay home. Depending on the gravity of the offense, providers in these situations receive a zero-star rating or may receive an unscheduled reassessment at a later date. Sometimes licensing staff in Tennessee will accompany the assessment staff on a visit to a provider because they have previously seen the state of the facility from conducting unannounced licensing inspections and may be able to detect irregularities, if any. Assessors also have access to photos taken by licensing staff that they can refer to as a previous benchmark if needed.

[11] Tennessee previously used the FDCRS but transitioned to the FCCERS-R in August 2010.

Table III.14. Procedures for Conducting QRIS Classroom Observations

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Observations Scheduled in Advance	No, 3-week window	No, 3-week window	Yes	No, 4-week window	Yes
Number of Classrooms Observed	1/3 of the classrooms per age group and at least one assessment for each age group	1/3 of the classrooms per age group and at least one assessment for each age group	1/3 of the classrooms per age group and at least one assessment for each age group	1/3 of the classrooms per age group and at least one assessment for each age group	1/3 of the classrooms per age group and at least one assessment for each age group
Selection of Classrooms	Random	Random	No set guidelines	Random	Random
Measure Used in Mixed Age Classrooms	Age of the majority of children in the classroom	No information	Specific items for infant/toddlers	No information	Age of the majority of children in the classroom; if equal distribution, provider's choice
How is Observation Data Recorded	Electronic (Branagh)	Electronic (Branagh)	Electronic (Wireless webforms)	Electronic (Branagh)	Paper (Scannable form)
Composition of Facility Environment Score	Average by age group	Average across classrooms	Each classroom observed must meet the standards. (If one classroom fails, the entire center fails.)	Average across classrooms	Average across classrooms OR Lowest classroom score if <3

Source: Site visits conducted as part of the QRS Assessment project.

During the visit. Assessors in each of the four QRIS observe one-third of the classrooms for each age group served and conduct at least one assessment for each age group. For example, a center serving children from ages 0 to 5 would receive at least one ECERS and one ITERS, and more than one ECERS would be administered if, for example, there are more than three preschool classrooms in the facility.

In the case of multiple classrooms, the classrooms observed are selected randomly. Classroom selection occurs on the morning of the assessments. Typically, the program director or a staff member randomly draws slips of paper denoting different classrooms. As classrooms are selected, assessors must determine whether a particular classroom cannot be observed due to atypical circumstances. For example, three of the QRIS require that at least half of enrolled children are present in a particular classroom (Miami, Illinois, and Tennessee), two QRIS exclude classrooms that are staffed by a substitute teacher (Miami and Pennsylvania), and two QRIS exclude classrooms if the teacher is new (Pennsylvania and Tennessee).

Miami, Illinois, and Pennsylvania record data in electronic form, using a tablet computer (Table III.14). These three QRIS use the ERS Data System (Branagh Information Group, 2011) to record observational ratings. Tennessee uses scannable paper forms designed for use with their own in-house database developed by UT-SWORPS.

After the visit. When observations are complete, assessors in each of the four QRIS write summary reports for every classroom observed. Assessors reported that the ideal is to complete the draft reports within one to two days of conducting an assessment. Assessments are typically scheduled every other day to accommodate report writing. Assessors in Miami, Illinois, and Pennsylvania use built-in capabilities of the ERS Data System to facilitate report writing. Assessors noted that electronic methods of scoring and note-taking is helpful for recording notes during the observation and importing notes directly into reports. It also makes review by supervisors more efficient because they have instant electronic access to the data, an advantage when assessors spend so much time at provider locations rather than at an office. Lead assessors review reports and scores and assessors complete revisions as necessary. When all reports for a facility have been finalized, they are compiled and transmitted back to raters (in Illinois and Pennsylvania) or to the QRIS administrator (in Miami).

Tennessee assessors submit draft assessment reports (including rating sheets and notes) to the lead assessor who inspects the files for accuracy and completeness. Each assessment report is also required to be reviewed by two other reliable assessors. Assessment teams usually conduct these reviews at weekly team meetings. After review, the lead assessor sends all reports from assessments conducted during a particular week to UT-SWORPS for processing and final review. The UT-SWORPS team processes ERS score sheets by putting them through a scanner. The scanner software reads the data, generates ERS scores, and produces a strengths sheet listing items on which a provider received a score of 4.0 or higher. UT-SWORPS sends a compiled report to the rater team so that rating scores for the environment component can be entered into the QRIS database.

Calculating facility-level ERS scores. Among the contents of the facility-level assessment reports that providers receive is a facility-level ERS score. Sites calculate this facility score in a number of ways. Illinois and Pennsylvania take the average score across all classrooms and scales administered (Table III.14). Tennessee also calculates an average across classrooms. However, if any individual classroom receives an ERS score below 3.0, the entire facility assumes that classroom's score. Miami produces separate averages for each scale administered (such as an ECERS average and an ITERS average).

Indiana classroom observation procedures. As previously mentioned, Indiana does not administer ERS assessments but includes some observational indicators in their quality rating tool (indicators are described in Chapter II). As Table III.14 shows, Indiana raters follow similar guidelines as assessors in other QRIS in determining how many classrooms to observe. However, there are no specific guidelines for selecting which classrooms to observe. In terms of determining a facility-level rating for environment, all observed classrooms must comply with standards in order for an entire facility to meet requirements for particular rating level.

5. Procedures for Administering Other Standardized Measures

Of the five QRIS, only Illinois administers another standardized measure, the PAS and BAS, in addition to the ERS. The PAS and BAS, as described in Chapter II, measure the quality of administration in child care programs. These measures are administered by the same assessors who conduct the ERS and are usually conducted on the same day as classroom observations. Administration of the PAS and BAS entails an interview with the program director or provider (in the case of family child care). Assessors also review many of the same program documents examined by raters in other QRIS—including staff and family handbooks, management plans, and financial records. Illinois assessors record PAS scores using the ERS Data System, which offers the same note-taking and automated scoring capabilities for the PAS and BAS as it does for the ERS.

E. Assigning Component Ratings

After the evidence review and all required assessments have been completed, ratings for individual quality components can be calculated. As we described early in this chapter, there is some variation between sites in when component ratings are assigned (Table III.1). In Illinois, Indiana, and Pennsylvania, component ratings are determined immediately after the evidence review is conducted but before assessments are carried out. Notably, these three QRIS are all building block systems that require providers to meet all criteria for a particular rating in order to receive it. Conducting assessments is costly in terms of time and resources. Therefore, before an assessment is administered, it is helpful in these sites to determine whether a provider meets requirements on all components (other than those requiring an assessment). If there is a missing or erroneous piece of evidence, providers can correct and complete documents before an assessment is scheduled. Component ratings may be determined automatically in QRIS databases or assigned manually by raters (Table III.5). In Indiana, component ratings are automatically assigned based on data entered by raters. In Illinois and Pennsylvania, raters enter data denoting whether a provider has met requirements for each component based on their reviews.

In Miami and Tennessee, component ratings are assigned after both evidence review and assessments have been completed. Both these sites implement combination systems where each component is assigned a number rating (as opposed to building block systems above where components are assigned dichotomous ratings to indicate whether or not requirements for that component have been met). Table III.15 displays point allocations for rating components in Miami and Tennessee. In Miami, data is populated automatically in the QRIS database as the evidence review and assessments are completed. However, for documentation purposes, the rater/assessment team also sends paper copies of assessment reports and signed copies of provider's self-study checklists to the QRIS administrative staff to denote that the rating process is complete. QRIS administrators then review materials and compare information against that entered into the database. Miami's QRIS database automatically assigns individual component ratings based on data entered by rater/assessment team for individual indicators. In Tennessee, raters manually calculate ratings for each component by tallying checks on the paper forms they complete during the evidence review.

The one exception is for the environment component—UT-SWORPS calculates the rating for this component based on facility ERS scores and includes that rating in facility assessment reports. Raters then enter the environment component score into the QRIS database based on the reports.

Table III.15. Component and Final Rating Point Allocations for Center- based Programs in Miami and Tennessee

QRS Component Category	Quality Counts, Miami-Dade County	Tennessee Star-Quality
Licensing Compliance	n/a	none
Ratio and Group Size	1 to 5 points	1 to 3 points
Health and Safety	n/a	n/a
Staff Qualifications	1 to 5 points	1 to 3 points
Family Partnerships	1 to 5 points	1 to 3 points
Community Involvement	n/a	n/a
Administration and Management	1 to 5 points	1 to 3 points
Environment	2 to 20 points[a]	1 to 3 points
Curriculum	1 to 5 points[b]	1 to 3 points
Child Assessment	Included in Curriculum score[b]	n/a
Cultural and Linguistic Diversity	Included in Family Partnerships score	n/a
Provisions for Special Needs	Included in Family Partnerships score	n/a
Accreditation	none	2 bonus points
Bonus Points	1 bonus point for each full-time teacher and/or curriculum specialist with a Bachelor's degree or higher and 18 ECE credits up to 6 bonus points allowed	Awarded for accreditation, see above
Formula for Final Rating	Total of component and bonus points[a]	Total of component and bonus points, divided by 7
Final Rating Cutoffs	Level 1: 4 – 12 points Level 2: 13 – 20 points Level 3: 21 – 28 points Level 4: 29 – 35 points Level 5: 36 and above[a]	Level 1: 1 – 1.49 points Level 2: 1.50 – 2.49 points Level 3: 2.50 and above

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable

ECE = Early childhood education

[a]Points are awarded separately for ECERS and ITERS. If one of the scales was not administered in a center, the point allotment for the administered scale is doubled to obtain the total points for the environment component.

[b]The curriculum component is currently being assessed and scored but points are not yet included in the final rating.

F. Assigning the Final Rating

Once individual component ratings have been determined, the final rating can be calculated. As with individual component ratings, overall quality ratings are either automatically calculated or manually determined by a rater (Table III.4). Miami, Indiana, and Tennessee have this process automated in their QRIS databases; that is, based on individual component ratings, the database automatically calculates the overall rating. This automation is particularly helpful in the combination sites where the formula for determining a provider's overall rating is a bit more complex.

In Miami, a provider can receive up to 40 total points across components, but can achieve the highest QRIS rating (level five) with a total score of 36 or higher. Up to six bonus points can be earned for each full-time teacher or curriculum specialist with at least a bachelor's degree and 18 credits in early childhood education. In Tennessee, scores across components are added (up to 21 total points), with bonus points awarded for accreditation. A provider's total score is then divided by the number of components to obtain the average score. The average component score is used to determine the overall QRIS rating—an average score of 2.5 or higher merits the highest QRIS rating (level three).

In Illinois and Pennsylvania, raters determine overall ratings manually by reviewing individual component ratings. Because both Illinois and Pennsylvania are building block systems as previously described, this process is fairly straightforward given the dichotomous nature of individual component ratings and the fact that failing to meet requirements for any individual component means that the overall rating is automatically downgraded. Whether or not the calculation of final ratings is automated in QRIS databases, each site requires that QRIS administrators review the final rating and formally approve it before providers are notified.

When a final rating has been approved, packets containing rating results are sent to providers. These packets contain a formal notification letter or certificate containing the final rating, classroom-level assessment reports, information about supports and resources for quality improvement (particularly pertaining to areas with low scores), and materials for notifying families about the QRIS rating level. All sites have a process in place that allows providers to appeal final ratings, although respondents universally noted that the various checks and quality assurance procedures they have integrated throughout the rating process have minimized the occurrence of appeals.

G. Renewals

Across the five QRIS, respondents noted that determining the frequency of renewals required a delicate balance between managing resources and maintaining accountability. Some QRIS have reconciled these competing needs by conducting observational assessments less frequently but asking providers to submit yearly documentation demonstrating that they have maintained the requirements for their current rating level on all other components. Table III.16 summarizes the validity period of ratings, evidence review results, and assessment scores in the five QRIS.

Table III.16. QRIS Renewal Timeline and Procedures

	Quality Counts, Miami-Dade County[a]	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Validity Period of Quality Rating	Initial: 2 years Subsequent: 13 months	1 year	1 year	2 years	1 year
Validity Period of Evidence Review	Initial: 2 years Subsequent: 13 months	1 year	1 year	2 years	1 year
Validity Period of Assessment Scores	Initial: 2 years Subsequent: 13 months	3 years	n/a	2 years	1 year

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable

[a]Miami is in the process of revising their renewal timelines. The new policy requires renewals every 18 months.

Miami originally designated providers' first rating as a baseline rating and allowed a two-year period during which providers could access supports and technical assistance but would not be required to undergo any assessments. However, administrators reported having concerns that providers were availing themselves of resources and supports without having to demonstrate or track quality improvements. At the time of our interviews, QRIS administrators were revising their procedures to end the deferment period but increase the time between assessments to 18 months (instead of requiring renewals every 13 months with a two-year deferment period). They were also considering requiring providers at the two highest levels to renew every three years.

In Illinois, providers must submit paperwork annually to document they have maintained requirements for their current rating level. If a provider wishes to apply for a higher rating level, both the evidence review and standardized assessments are conducted again. However, providers may only undergo a maximum of one observational assessment series (the ERS and PAS/BAS) per year. After three years of maintaining a rating level, all providers are required to undergo observational assessments again, regardless of whether they wish to maintain their level or apply for a higher rating. Indiana and Tennessee repeat their rating process yearly for all providers, although Indiana allows providers to apply for a higher rating before the year lapses if they so choose. In such cases, the evidence review would be conducted again. In Pennsylvania, the QRIS rating is valid for two years.

THIS PAGE LEFT BLANK FOR DOUBLE- SIDED PRINTING

IV. DATA COLLECTION, USE, AND ANALYSIS TO REFINE QUALITY MEASUREMENT IN QRIS

With the constraints on both financial and human resources that states face, planning for data collection that can inform system improvement and evaluation is not often at the forefront of policymakers' and program administrators' agendas. Yet, QRIS data collected for programmatic purposes (such as documenting components and constructing the ratings), can be useful for answering questions and resolving issues that were not initially planned for or considered during the development of data systems.

In this chapter, we describe the breadth and depth of the data collection that occurs in each of the sites and the ways in which QRIS data are currently used for monitoring, analysis, and evaluation. We explore how sites' data collection efforts have influenced the subsequent refinement of quality rating standards, policies and procedures, and whether future studies or data collection efforts are being planned to shed light on unresolved issues.

A. Overview of Data Systems

The five QRIS were selected, in part, because of their use of electronic data systems to record information on QRIS. All of the systems are web-based, providing access to multiple users across geographic areas. The data systems vary in scope, particularly in terms of covering state early childhood initiatives beyond QRIS. Table IV.1 summarizes information about the linkages between each QRIS database and other databases relevant to the care and early learning of young children.

Table IV.1. Linkages Between QRIS and Other Data Systems[a]

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Name of QRIS Database	Web-based Early Learning System (WELS)	Data Tracking Program (DTP)	Child Care Information System (CCIS)	Keys to Quality PELICAN (K2Q PELICAN)	Regulated Adult and Child Care System (RACCS)
Licensing	✓		✓	✓	✓
Provider-level Subsidy Receipt	✓	✓	✓	✓	✓
Child-level Subsidy Receipt			✓	✓	✓
Professional Development Registry	✓	✓		✓	
Child-level Outcomes				✓	

Source: Site visits conducted as part of the QRS Assessment project.

[a]We define a "linkage" as a built-in capacity to access and retrieve data from another existing database.

The QRIS data systems of Illinois and Tennessee were both developed in-house within the administrative agency that has oversight responsibility for the QRIS. Illinois' Data Tracking Program

(DTP) was developed by INCCRRA (under contract with the Illinois Department of Human Services) to monitor various child care initiatives in the state. In addition to QRIS data, DTP houses information on supports received by providers through state wage supplement and scholarship programs, as well attendance in various training opportunities. Tennessee's Regulated Adult and Child Care System (RACCS) database was introduced in 2008 to digitize licensing information that was previously collected through paper forms. It is used primarily by DHS licensing staff to track all licensing data for state adult and child care programs. These data include participation in the Report Card and Star Quality Programs. RACCS can be linked to provider- and child-level subsidy receipt data housed on a related database called the Tennessee Child Care Management System (TCCMS).

Miami's and Indiana's data systems were built under separate contracts with software development firms in partnership with individuals with previous experience in the child care field. Florida Miami Dade's Web-based Early Learning System (WELS) is used by a variety of staff and partners in the county, including QRIS administrators, specialists, raters, and assessors. WELS began as an Access database and was transformed into a web-based system in 2005. Overall, there are about 200 people with access to the database, and 35 to 40 individuals who enter data. WELS interfaces with Miami's professional development registry, which houses data on all training completed by child care staff. WELS also accesses information on licensing violations from the Florida Department of Children and Families licensing database and information on providers receiving child care subsidies from a separated database maintained by ELC. Indiana's Child Care Information System (CCIS) is a web-based system of databases covering a range of programs administered by the Bureau of Child Care. Different components of CCIS are accessed by various entities throughout the state that work with child care providers. Within CCIS, the Paths to Quality database tracks provider activity from the time of enrollment, TA activities utilized, indicators and insufficiencies noted at the time of formal assessment, and the final quality rating. Data on PTQ participation can be linked to information on licensing status and other provider-level information on the Regulated Child Care System database. The Automated Intake System within CCIS, which houses child-level subsidy data, can also be linked back to provider-level data.

Pennsylvania's Keys to Quality (K2Q) PELICAN is one component of the Office of Child Development and Early Learning's set of web-based PELICANs (Pennsylvania's Enterprise to Link Information for Children Across Networks) that catalog a broad set of information about programs and children. The PELICANs were created to enable examination of outcomes associated with various early childhood initiatives throughout the state. The K2Q PELICAN interfaces with other PELICANs containing information on child care licensing, subsidy administration, the state pre-kindergarten program, and professional development of individual staff in the early care and education field. In addition, K2Q PELICAN is the only QRIS database that can be linked to child-level enrollment and assessment data through Pennsylvania's Early Learning Network (ELN). The ELN database consists of information on approximately 60,000 children, including assessments collected by teachers three times a year, background information on risk factors, and child and family demographic characteristics.

All QRIS that include standardized quality assessments use separate databases to house item- and classroom-level data (Table IV.2). Miami, Illinois, and Pennsylvania all use the Branagh ERS Data System while Tennessee uses a database developed by UT-SWORPS. Each assessment database is maintained by the assessment teams in the respective sites. Only Miami's WELS database has a built-in linkage to their ERS Data System, with classroom- and item-level ERS data uploaded directly to WELS. QRIS databases in Illinois, Pennsylvania, and Tennessee are not currently designed to store item- and classroom-level assessment data. The databases for these three QRIS hold only facility-level assessment scores that are entered manually by raters based on assessment

reports from assessors. However, item- and classroom-level data can be merged with provider-level data from QRIS databases using provider identifiers.

Table IV.2. Databases Used to Store Standardized Assessment Data[a]

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Database Used to Collect Assessment Data	ERS Data System[b]	ERS Data System	n/a	ERS Data System	Site-specific database
Who Maintains Assessment Database?	Devereux	NLU	n/a	PA Key	UT-SWORPS
Assessment Database Linked to QRIS Database?	Yes	No	n/a	No	No

Source: Site visits conducted as part of the QRS Assessment project.

n/a not applicable

ERS = Environment Rating Scales

NLU = National Louis University

UT-SWORPS = University of Tennessee Social Work Office of Research and Public Service

[a]Assessment data includes the ERS in all QRIS and the PAS/BAS in Illinois.

[b]ERS Data System is a database and software package developed and distributed by the Branagh Information Group.

B. Availability of Data on Quality Measurement Ratings

At a minimum, each QRIS database stores information on current and historical quality ratings. That is, for each provider, there is a record of all quality ratings received since their initial participation. Beyond the ratings, all QRIS databases store component-level ratings for at least some quality rating components (Table IV.3). In Miami and Tennessee, these variables consist of the actual points a provider received on each component. Illinois' main QRIS database includes a dichotomous variable indicating whether or not requirements for the staff qualifications and environment components were met. There is also a variable denoting whether a provider met the cutoff score for the PAS or BAS. Indiana's database does not include component level ratings by the categories as we have defined them, but all indicator level data are available such that component ratings could be derived. Pennsylvania's database includes variables to denote compliance with health and safety requirements and accreditation status, but other component-level ratings are not available.

Three QRIS databases store indicator-level data. Miami's database contains the most detailed data on individual indicators. The record for each provider includes information on whether each indicator was or was not met, and the specific source of evidence the rater reviewed. This level of detail allows for automated calculation of points for each component. The QRIS database in Illinois includes indicator-level data only for staff qualifications and environment. For staff qualifications, data are available on the total administrative and teaching staff at a facility as well as the calculated percentage of staff meeting requirements. Indiana's database also tracks compliance with individual indicators but the sources of evidence are not recorded.

Table IV.3. Availability of Data Elements to Calculate Quality Ratings in QRIS Databases

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Final Rating	✓	✓	✓	✓	✓
Individual Component Ratings[a]	✓	Few[b]	✓	Few[c]	✓
Indicators Met/Not Met for Each Component	✓		✓		
Sources of Evidence for Each Indicator	✓				
Facility-level Assessment Scores	✓	✓	n/a	✓	✓
Classroom-level Assessment Scores	✓		n/a		
Item-level Assessment Data	✓		n/a		

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable (not required by QRIS)

[a] Components as defined for this report, not as defined by each QRIS.

[b] Component ratings available on licensing compliance, staff qualifications, and environment.

[c] Data available on licensing compliance, health and safety as it relates to core series training for practitioners, and accreditation status.

In addition to what is stored on QRIS databases, Miami, Illinois, Pennsylvania, and Tennessee have more detailed data on individual components stored in separate databases that can be linked to the QRIS database using provider identifiers (Table IV.4). Miami and Illinois have staff-level data on training and qualifications available through professional development registries. For Illinois providers, classroom- and item-level data from the ERS, PAS, and BAS are stored in the ERS Data System maintained by NLU. Component-level ratings for family partnerships, administration and management, curriculum, child assessment, cultural and linguistic diversity, provisions for special needs, and community involvement can be determined using indicator-level data from the PAS or BAS. Similarly, classroom- and item-level ERS data are available through the assessment databases in Pennsylvania and Tennessee.

C. Processes to Support Data Quality

As described above, each of the QRIS databases is accessed by a wide array of users. Individuals from different locations and agencies enter different pieces of information depending on their role within the QRIS. Having databases that are accessible to multiple users at once is important for immediate processing of data and close monitoring of operations. However, it also increases the potential for errors. Given the wide range of database users, administrators and data managers emphasized the importance of instituting procedures and implementing strategies to monitor and maintain data quality. We describe some common strategies reported by respondents from each of the QRIS below.

Table IV.4. Availability of Data on Individual Components in QRIS Databases, Assessment Databases, and Staff Databases

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Licensing Compliance	✓	✓	✓	✓	✓
Ratio and Group Size	✓	n/a	n/a	n/a	✓
Health and Safety	n/a	✓	n/a	✓	n/a
Staff Qualifications	✓[a]	✓	✓	✓	✓
Family Partnerships	✓	✓[b]	✓		✓
Community Involvement	n/a	✓[b]	n/a		n/a
Administration and Management	✓	✓[b]	✓		✓
Environment	✓	✓	✓	✓[c]	✓[d]
Curriculum	✓	✓[b]	✓		✓
Child Assessment	✓	✓[b]	✓	✓	n/a
Cultural and Linguistic Diversity	✓	✓[b]	✓	n/a	n/a
Provisions for Special Needs	✓	✓[b]	✓		n/a
Accreditation	✓	✓	✓	✓	✓

Source: Site visits conducted as part of the QRS Assessment project.

n/a = not applicable (not required by QRIS)

[a]Data housed in Miami-Dade Professional Development Registry.

[b]Data housed in ERS Data System maintained by NLU.

[c]Data housed in ERS Data System maintained by PA Key.

[d]Data housed in assessment database maintained by UT-SWORPS.

1. Database User Guides and Training

All QRIS have some documentation available describing how to use the database—typically in the form of user guides, which cover basic navigation instructions and provide guidance on what different fields mean. None of the sites has a formal training protocol in place other than the materials distributed, although ad hoc trainings are conducted on an as-needed basis. In Tennessee, database training is incorporated when new raters shadow more experienced raters during the training period. Respondents noted that in the process of a new rater shadowing a more experienced one, the use of the database naturally arises. They did note, however, that this method is less than systematic across the licensing units because individual raters often devise their own tools to make data collection easier.

2. Minimizing Duplicate Records

An important element for maintaining data integrity is minimizing the number of duplicate entries in the database. Having multiple records for a provider results in inaccurate and/or missing information and reconciling different pieces of data can be draining on resources. Duplicate records are sometimes created when data users search for a provider record and are unable to find a match due to a typographical error. This can result in multiple records for one provider if users create another record. To avoid such situations, Miami's, Indiana's, and Pennsylvania's systems only permit the creation of new records at the administrative level. In Indiana and Pennsylvania, provider records are typically pulled into the QRIS database from central databases containing information on all regulated child care providers. If a specialist or rater is unable to find a provider record in the database or based on their license number, they must contact the database manager, who then conducts the necessary verification before creating a new record. In Illinois, there have been occasions when a duplicate entry for an organization is generated (usually at the local CCR&R level) because the person entering data erroneously entered a search term and thought that a record did not already exist. INCCRRA staff are typically able to track down these duplicates and reconcile them early on.

Duplicate records can also arise when providers change license numbers due to a change of address or change of director or owner. Miami's WELS system has built-in exception reports that checks records for new license numbers and compares addresses on file of an existing provider with a different license number. If it is determined that an existing provider changed license numbers, WELS is updated manually.

3. Customized Access for Different Users

To minimize potential for error, all sites tailor database access to the roles of specific data users. Certain fields or entire records may be hidden from view if a particular user does not need access to that information to fulfill his or her responsibilities. Some users may only be able to view information but not edit it. For example, in Pennsylvania, only assessors can enter ERS scores. Indiana's CCIS and Pennsylvania's K2Q systems provide different levels of "read" and "write" access. Administrators have full access to view and edit records, but local staff are only able to access records of providers they are working with or the providers within their service area.

4. Built-in Data Entry Quality Control

Each database incorporates features to minimize erroneous data entries and provide guidance on possible values for each data field. One method for flagging potential errors is by linking data entry rules to certain fields. For example, a record may not be saved if mandatory fields, such as a license number, have not been filled. Similarly, a message may pop up on screen if a user enters text data into a field that is expected to be numeric, such as child-staff ratios. Drop-down menus or checklists are also used to provide users with instant access to the acceptable values for a particular component. Miami's WELS database, for example, provides a drop-down list of approved curricula from which assessors can select the appropriate entry. Using this list, there is less risk of raters mistakenly giving credit for curricula that are not approved. Further, the list can be updated from a central location so that it is always current—administrators do not have to be concerned with sending written updates to individual users because the change is immediately reflected in the database.

5. Pre-populated or Automated Data Fields

Respondents emphasized the need to minimize the amount of data entry required by using already existing information to populate certain fields. This reduces potential for error and saves time and resources. Sites take advantage of built-in linkages with existing databases to retrieve pertinent information rather than repeating data entry. For example, Illinois and Miami-Dade's systems populate the QRIS database automatically with staff qualifications data from their PD registries. Miami's and Tennessee's databases also reduce the need for raters to calculate quality ratings based on points per component by having the process automated in their systems. This is another way to ensure that point allocations and cutoff rules are assigned consistently.

6. Regular Data Checks

Finally, in addition to built-in database features designed to minimize data errors, sites regularly examine database entries for potential issues. Miami and Illinois have staff members designated to inspect data for possible errors. These individuals rely on automated reports to flag potential inconsistencies, particularly in terms of outlying data points. In Tennessee, inspection is conducted within each rater unit and each unit may have slightly different procedures for checking the data. Field supervisors and raters within each unit meet regularly to review provider files. During these meetings, they verify scoring accuracy, completeness of notes, and ensure that demographic information has been provided and that all required items have been completed.

For assessment data, the ERS Data System used by several sites includes built-in features designed to reduce data errors. Each of the sites using this system also have lead assessors check assessment records for potential errors before assessment reports are finalized. In Tennessee, prior to data entry, UT-SWORPS staff inspect score sheets submitted by program assessors to ensure there are no discrepancies, such as errors in identification numbers, typographical errors, or mismatches between scores and assessment notes.

D. Use of Data to Monitor and Evaluate QRIS

Among the advantages of large-scale web-based data systems is that data are continuously updated. As a result, administrators, in conjunction with external evaluation partners, can monitor various QRIS activities, forecast future needs, and address pertinent policy questions. In this section, we describe how administrators have typically used and analyzed data to inform and refine the QRIS.

1. Monitoring Participation and Quality

A primary use of QRIS data is to monitor provider participation and quality improvements over time. At the aggregate level, administrators use such data to examine how QRIS participants are progressing overall. All sites examine distributions of quality ratings at least annually and calculate the number of providers whose ratings have increased, decreased, or stayed the same. Administrators also examine data at the component or indicator levels to identify areas showing substantial progress and areas where large numbers of providers tend to underperform.

Respondents in Miami and Pennsylvania noted that their databases have improved the capacity to respond to specific questions about QRIS and child care in general from state legislators and other stakeholders. Administrators in Illinois and Indiana described how they include data about

participation in materials disseminated to other providers and parents to generate greater awareness about QRIS.

QRIS specialists use data on individual providers to track strengths and weaknesses. Miami's and Indiana's databases have built-in capabilities to generate such reports. Miami's WELS database generates a "strengths and needs" report based on the components on which a provider received high and low scores. Indiana's PTQ database provides a list of "insufficiencies," that is, all the requirements for the desired level a provider was not able to meet.

2. Informing the Allocation of Resources

Tracking participation and quality over time as described above allows sites to monitor the ongoing use and projected need for resources. For example, they can anticipate the number of providers who will be submitting renewal applications in a particular period and adjust resources as necessary. Respondents in Miami, Illinois, Indiana, and Pennsylvania mentioned using the data in this manner.

Several sites have also capitalized on their data systems to monitor the supports that providers access in preparation for the rating process. In Miami, Illinois, Indiana, and Pennsylvania, QRIS specialists log their interactions with providers, including the mode and length of each interaction. Illinois, Indiana, and Pennsylvania also collect data on the topics covered during interactions. Miami, Indiana, and Pennsylvania monitor the financial supports, incentives, and grants that providers receive (Table IV.5). Collecting this data enables administrators and supervisors to monitor caseloads of QRIS specialists. Respondents also discussed plans to eventually use these data to link the use of resources to quality improvements made over time and determine which components have required the most support from specialists. Examining component- and indicator-level data aggregated across providers also allows administrators to identify areas of critical need and shape decisions about the types of trainings, grants, and professional development opportunities to offer.

Table IV.5. Data Collected on QRIS Pre- rating Process

	Quality Counts, Miami-Dade County	Illinois Quality Counts	Indiana Paths to Quality	Pennsylvania Keystone Stars	Tennessee Star-Quality
Hours of Contact with QRIS Specialist	✓	✓	✓	✓	
Content of Contact with QRIS Specialist	✓	✓	✓	✓	
Incentives/Support Grants Received	✓		✓	✓	

Source: Site visits conducted as part of the QRS Assessment project.

3. Assessing Implementation

Several sites have conducted additional data collection to assess the experiences of administrators, field staff, providers and/or parents with QRIS. Miami and Illinois conduct ongoing provider and/or parent surveys. In Miami, assessors distribute and collect provider satisfaction surveys during the assessment visit. Illinois has an online QRIS participant survey that gathers feedback from providers and parents. Results of these surveys are analyzed regularly to identify issues that need to be addressed based on negative feedback.

Tennessee commissioned a qualitative implementation study of their QRIS in 2005 (Pope et al. 2006). This study revealed some concerns on the part of providers about inconsistencies in how expectations and results of assessments are communicated by various QRIS staff. In response to the findings of the implementation study, changes were made in the fifth and sixth years of QRIS implementation. These included increases in targeted technical assistance efforts (including introduction of Provider Self-Assessment and Monitoring Services) and ERS training for QRIS specialists. In Indiana, Purdue University is currently conducting a study to assess providers' experiences with Paths to Quality, including barriers to participation and their use of resources. The study is also investigating parent awareness of the Paths to Quality system and whether it has affected their child care decisions.

4. Examining Possible Changes to Quality Measurement

Pennsylvania and Tennessee have conducted studies to explore the utility of administering additional (or alternative) observational assessments. In Pennsylvania, there is a small pilot study under way to examine relationships between the Classroom Assessment Scoring System (CLASS) (Pianta et al. 2007) and the ERS measures currently used to assess the environment component of quality ratings. In Tennessee, UT-SWORPS recently conducted a study to compare ERS measures with the CLASS and the Early Childhood Environment Rating Scale Extension (ECERS-E; Sylva et al., 2003). In 2010, analyses of results from that study were under way.

5. Linking QRIS Participation to Changes in Quality and Child Outcomes

Of the sites we visited, only Indiana and Tennessee had explored (or plan to) relationships between quality ratings, observational assessments of quality, and child outcomes. In Tennessee, UT-SWORPS recently conducted a study of 114 center-based programs to compare QRIS component ratings with classroom quality assessments. A sample of children from each classroom was assessed on measures of language, literacy, numeracy and social skills. Although results were not yet available in 2010, the study will be examining associations between classroom quality and children's skills. In Indiana, evaluators from Purdue University are collecting data on cognitive, language, and social emotional development of a sample of 1,040 children over a three-year period. The study is also comparing scores on ERS and the Caregiver Interaction Scale (CIS) (Arnett 1989) to Paths to Quality ratings for 540 child care providers (Langill et al. 2009). Because Indiana does not use the ERS in constructing their ratings, they are able to use the ERS, along with the CIS, as an independent benchmark for quality to validate their quality ratings.

Respondents in Illinois and Pennsylvania also expressed an interest in examining relationships between quality ratings and child outcomes but noted that the cost of conducting child assessments was the primary barrier. Pennsylvania does, however, require the collection of child assessment data for providers in the upper levels of their QRIS in their ELN database. Respondents noted that they have just begun to explore the use of these outcomes for research purposes. Pennsylvania recently released a summary of child assessment outcomes for QRIS providers at levels three and four (the two highest), Head Start programs, and state-funded pre-kindergarten programs (OCDEL and The Pennsylvania Build Initiative, 2011).

E. Challenges to Using Data for Monitoring and Evaluation of QRIS

Despite the range of data being collected in each QRIS, respondents across the board noted challenges in using the available data for research and evaluation purposes. The issues raised center

around the scarcity of time and resources and the need for both better integration and more detailed information on quality and outcomes of interest.

To facilitate the timely and efficient analysis of data, each QRIS has made an effort to build automated reports into their data systems. To date, most of these reports are used for monitoring purposes. Across the five QRIS, there is limited in-house capacity for research and evaluation. Miami, Indiana, Pennsylvania, and Tennessee currently have established partnerships with external contractors to provide data-analytic and evaluation services. Miami and Pennsylvania also have in-house staff members who oversee research efforts. Respondents in Illinois noted that efforts to evaluate their QRIS have been limited to date, both due to their QRIS being relatively new as well as a lack of financial resources.

Building data systems that are suitable for both monitoring and research purposes requires ongoing dialogue with administrators and evaluators because, according to the respondents, it is difficult to anticipate all the potential uses of the data from the beginning. Because QRIS databases play a central role in implementation and administration, there is rarely time to pilot systems with users. This not only constrains the amount of time available to develop research aspects of the system, but also the time to investigate and fix all the bugs. Because data errors would jeopardize the utility of information—for research purposes or otherwise—efforts are first focused on making sure that key pieces of information are recorded accurately. Staff from Miami and Indiana particularly emphasized the latter point. Both Miami's and Indiana's databases were built with the guidance of individuals with experience in child care and early education implementation. Respondents noted that the substantive knowledge these individuals brought to the table was critical to building data systems that are responsive to program monitoring needs. Nevertheless, they also noted that rather than attempting to build a perfect system from the beginning, they built systems that can be modified and continuously refined to respond to new issues and questions.

As a more mature QRIS, respondents in Tennessee noted that the technology available for building databases has changed much since the inception of their system. They have therefore faced the constraint of working with a data system that is not as easily adaptable. Another barrier they face, according to respondents, is the lack of integration between individual databases. The QRIS database and assessment database used in Tennessee were developed separately at different points and thus operate independently. Having one integrated database to ease the linkage of data and the ability to update information between the two would be more suited for research purposes.

Pennsylvania's QRIS database is already part of an integrated system. However, respondents reported that they would like to add depth to the database by capturing electronic data at the indicator- and component- levels, for example. These data are already being collected in electronic form to a certain extent (in electronic spreadsheet form), but currently not stored in the database. Respondents noted that this would not only improve their ability to answer questions of interest, but would also reduce the data processing burden on staff.

V. EMERGING THEMES AND DIRECTIONS FOR FUTURE RESEARCH

This in-depth study examined five select QRIS to describe what is conceptualized as quality and how it is measured. The study provides detailed information about the inclusion and definitions of quality components within each QRIS, the processes for measuring them and determining the final rating level, and the staff that carry out the process. Such an examination was intended to provide information about the consistency, reliability, and validity in the quality measurement process within and across QRIS. The five QRIS were purposefully selected because each had a comprehensive data collection approach in terms of coverage of important quality dimensions and there are linkages with other data systems to support the quality measurement (rating) process. As described in Chapter I, documenting how rigorously and consistently quality measures and ratings are collected and assigned is central to furthering progress in QRIS development and management, as well as research.

In this chapter, we summarize what we have learned about the conceptualization and measurement of quality in the QRIS and the factors that contribute to their validity and reliability. We also summarize the approach the five systems have taken to ensure the quality of the data collected throughout the rating process. We conclude with a discussion of potential directions for future research.

A. Defining Quality: Factors that Affect the Validity of Quality Ratings

This study described and compared how quality was conceptualized by QRIS planners in five systems. The extent to which QRIS ratings capture multiple dimensions of child care quality has important implications for the utility of ratings as measures that parents can use to select high-quality settings and administrators can use to target resources to low-quality providers. Our analysis focused on factors that may affect the "content validity" of quality ratings—by describing which components and indicators each system includes and examining whether they are defined consistently across different types of child care providers.

Breadth, depth, and rigor of quality components. Unlike earlier iterations of QRIS, which tended to include a smaller number of components such as child-staff ratios and group size, staff qualifications, and environment (Zellman and Perlman 2008), the five QRIS profiled in this report incorporate more components in quality ratings. In particular, there is increased representation of family partnerships, administration and management, and individualization of services. A significant challenge for QRIS administrators is the lack of guidance from research that can help identify necessary components and appropriate thresholds at each rating level, especially in these emerging components. As a result, component indicators vary considerably across the five QRIS and none enters the rating scheme in the same way (such as at the same rating level or with similar specificity across the QRIS).

Five components are common to the five QRIS studied—staff qualifications, administration and management, family partnerships, environment, and curriculum. Within these common components, we found considerable variation in the definition of indicators:

- **Staff qualifications.** Each of the QRIS places an emphasis on teacher qualifications, but as with many of the components, requirements at each rating level are set based on what planners and administrators believe is reasonable for staff to achieve given existing professional development infrastructures and child care and early education workforce supports (such as wage supplements and financial assistance for education and training). Systems typically require a certain percentage of teaching staff to meet standards, with

actual requirements varying in terms of level of education (typically CDA to AA), or specialized training in early care and education. One consistent element is that all five QRIS require some credits in early care and education or a CDA starting at the lowest rating level. The emphasis has clearly evolved in this respect to require formal training in early care and education over experience.

- **Administration and management.** For the administration and management component, there is no indicator that is common to all five QRIS, although a few are included in four of them. The commonly used indicators tend to relate to supports for staff. They include the use of an annual professional development plan and of a salary scale (at the higher rating levels) that is differentiated based on staff experience and level of education. This reflects consistency with staff qualifications in the focus on building and retaining experienced staff.

- **Family partnerships.** There are few, if any, standardized measures for these components and systems typically rely on self-reported information from providers. Among the five QRIS, four require parent/teacher conferences and opportunities for family evaluation of the program but the specificity of modes and frequencies, as well as the levels at which they are required, varies substantially. For example, depending on the QRIS and level, providers may be required to demonstrate the provision of any family evaluation opportunity (yes/no), specify the frequency of the evaluation, and/or specify the modes for evaluation.

- **Environment.** An ERS assessment is the most time, labor, and resource intensive of the measures used, but the developers and administrators of the four QRIS that require it believe that it provides crucial information that would otherwise be difficult to obtain through other means. Study respondents from all five QRIS believe that an observational, objective assessment of quality (Indiana includes items similar to ERS items in their rating tool) lends credence to the process and ratings.

- **Curriculum.** The inclusion of a quality component to measure the adoption and use of a developmentally appropriate curriculum is just beginning to take shape within QRIS but the quality levels are not well differentiated and the specificity of the measures is still lacking. In the five QRIS studied, each includes a dichotomous (yes/no) indicator of the use of such a curriculum at the highest level; Tennessee just added this indicator in January 2010. The extent to which the curriculum is reviewed or approved against specific criteria, however, varies. Each QRIS now tends to use the Early Learning Guidelines of their states as a benchmark to assess curriculum content. Assessment of the degree of alignment to the guidelines is conducted by providers, a state entity, or curriculum developers, depending on the QRIS.

Although the measures for some components—particularly the newer ones focused on the individualization of services—are limited in terms of depth and detail, their inclusion signifies a growing recognition of these components' contribution to the quality of child care services. Our discussions with administrators and planners revealed that cost and resources tended to be as strong a driver in the selection of components as the research evidence backing a particular indicator. Nevertheless, there was a prevailing sense among respondents that including indicators or a component, however limited, serves as a meaningful signal to providers that quality care is a multi-dimensional construct and that all components are important. Although imposing requirements may not predict improved child outcomes or translate to measurable differences on standardized

measures of quality as currently constituted, bringing certain components to providers' attention may lead to more purposeful practices and, eventually, improved outcomes for families and children.

Equity of standards for different settings. The degree of equity in both defining and measuring QRIS standards across different types of child care settings (for example, child care centers, pre-kindergartens, Head Start programs, and family child care homes) also contributes to the validity of ratings. Parents and other consumers of the ratings must have a clear understanding of the extent to which QRIS ratings denote comparable levels of quality regardless of the type of provider to whom the rating is assigned. With few exceptions, these five QRIS have intentionally tried to develop standards that are equitable across different types of care One indication of the efforts to achieve and maintain equity is the time and resources that planners and administrators have invested in thinking about how licensing, accreditation, and Head Start program performance standards (HSPPS) should be incorporated in QRIS standards. Four of the five QRIS do not automatically qualify providers for a particular rating level given their status as a Head Start program or accredited center. Only in Illinois are accredited providers waived of QRIS requirements, such as an ERS assessment, at one rating level. In addition, there are no automatic level qualifications for Head Start programs in any of the five QRIS. Pennsylvania does accept compliance with HSPPS in lieu of licensing at the first rating level, but only did so after a crosswalk between the HSPPS and licensing requirements demonstrated strong alignment. The five QRIS have similarly stressed equivalency in the standards across care settings, and licensed and unlicensed care as much as possible. Again, Illinois is an exception in creating specific standards for unregulated home-based providers to reach this sizeable portion of providers that care for children receiving subsidized care.

B. Measuring Quality: Factors that Affect the Reliability of Quality Ratings

The quality measurement process in the five QRIS includes a standardized assessment of observed environment quality (except in Indiana), gathering and reviewing of evidence (through various modes) for other quality components, and calculating component and final ratings. Procedures employed at each of these stages can affect the accuracy and consistency with which ratings are assigned. There is generally greater consistency in the administration of the ERS across QRIS than in the procedures for gathering evidence on other quality components or calculating ratings.

Assessment of observed quality. The QRIS that use the ERS to measure observed quality employ procedures for training assessors and maintaining reliability based largely on the guidelines established by the instrument authors and other experts in the field (Hamre and Maxwell, 2011). The availability of external benchmarks for the administration of assessments has resulted in substantial overlap in procedures in the four QRIS that conduct standardized assessments. On an ongoing basis, there is a considerable amount of both formal and informal communication within assessment teams to come to agreement about the interpretation of particular assessment items and, in each QRIS, respondents communicate with the instrument developers to obtain answers as questions arise. In addition, the four QRIS that conduct standardized assessments were consistent in focusing the role and defining qualifications for assessors. Assessors in each QRIS are devoted full-time to conducting ERS assessments (and PAS/BAS in Illinois) and must hold at least a bachelor's degree, with a specialization in early care and education either required or preferred. Respondents noted that specialization in early care and education supports consistent interpretation and application of the assessment. The differences we noted in procedures were minimal. For example, initial reliability is based on three to five observations.

Nonetheless, there continue to be threats to the reliability of standardized assessments within specific QRIS, and more broadly for comparison across QRIS. Many assessors are not trained to reliability directly with the instrument developers, although as noted above, there is consistency in training through use of developer training materials. The cost of direct training with the developers is prohibitive, particularly for systems with large assessment teams. Whether these training approaches are sufficient for maintaining the validity and reliability of the ERS as it has moved from being primarily a research tool to being widely implemented for accountability purposes is not known. Also, many assessors are tasked with observing multiple classrooms (at times using multiple scales) during the same visit, which could compromise both reliability and validity.

The reliability of standardized assessments can also be affected by inconsistencies in how observations are conducted. The QRIS in this study use a mix of announced and unannounced visits and only observe a sample of classrooms in larger facilities. Variation in these procedures can make a big difference in the comparability of scores across QRIS. In addition, assessment teams and licensing staff in Tennessee and Pennsylvania have developed specific guidance for conducting ERS assessments in the context of licensing requirements. This could make ERS scores in these states less comparable to scores in states that do not employ the same guidelines.

Gathering evidence for other quality components. The measures of the quality components, beyond that of observed environment quality are generally not collected through standardized assessments and many present challenges to consistent, reliable data collection and interpretation. For example, staff qualification requirements are complicated and require raters to look at the records of multiple staff, decide the legitimacy of courses or in-service training sessions to meet requirements, and calculate percentages of all staff members who meet the requirements. The QRIS with PD registries that can either fully automate or at least aid this process may be better able to produce reliable results for this quality component. The administration and management component also tends to require review of a substantial number of records and documents. Similarly, the family partnerships component can be difficult to measure because it is not always well-specified in terms of the ranges in modes of activities (handbooks, sign-in sheets for events) or methods to capture frequency.

Multiple modes of data collection—such as observation, interview, and document review—could serve to confirm the presence of quality components (and increase reliability) but introduces tradeoffs in terms of cost. Indiana and Illinois used multiple modes to collect information on the quality components beyond observed quality. Indiana does not conduct ERS assessments, but raters seem to go beyond the role of raters in the other QRIS by combining modes of data collection. Illinois uses the standardized assessment tools of the PAS/BAS to capture many quality components—such as family partnerships, administration and management, and individualization of services—which may enhance the reliability in measurement. Assessors are trained by instrument developers and follow standard procedures similar to that of the ERS assessments.

In contrast to the process for conducting standardized assessments, procedures for gathering evidence for other quality components are just beginning to be standardized. None of the five QRIS has formal training processes for raters in the same way that ERS assessors are trained, but most require that new raters begin by shadowing experienced ones. Similar to assessors, however, raters are expected to hold, at minimum, a bachelor's degree, and early care and education specialties are preferred. Three QRIS have developed standard tools for data collection and two have created reliability procedures either for the initial rating period (for new raters) or on an on-going basis to ensure that components and their requirements are understood and measured in a similar way. Across the QRIS, building reliability in the rating process has been a learning process that is now

moving toward increased formality in procedures. Standardization efforts have resulted from administrators' efforts to monitor implementation and were often directly informed by feedback from raters regarding their experiences out on the field or from providers that noted the lack of consistency.

Transparency in the process. Transparency can support the reliability of the measurement process. If providers have a clear understanding of the information that raters are seeking, then they may be better able to present the necessary supporting evidence, and the formal rating can more accurately capture the content of the standards (the actual presence of the quality component), rather than reflect a lack of knowledge or organization on the part of the provider (due to missing information even if the quality component is implemented). The five QRIS have all made substantial efforts to educate providers about rating requirements and procedures. Orientation sessions introduce the full set of standards on which providers will be assessed. Providers also receive information about the manner in which components will be measured and an overall rating produced. Each site also has a pre-rating process in place to further support providers in gaining an understanding of quality components and ways to meet rating requirements. Specialists in each of the five QRIS are available to work individually with providers to build this understanding and help them prepare for the formal rating.

Accuracy in calculation of ratings. Reliability of the quality ratings can also be affected by the methods employed for calculating them. Only two of the five QRIS have fully automated the calculation of each component and the final rating (Miami and Indiana). In Tennessee, the other combination system aside from Miami, raters determine the score for each component, but the QRIS database calculates the final rating. The automation of the final rating is particularly important in combination systems in which errors in assigning and calculating scores can affect the final rating determination.

Consistency across the rating process. Reliability throughout the rating process is becoming more challenging as the systems grow and more people are involved in the process. Despite this growth, the QRIS have kept objectivity in the process by maintaining distinct roles between staff who assist providers during the pre-rating process and those who are raters and/or assessors. In all the QRIS, respondents described the many formal and informal mechanisms for communication between and among teams of specialists, raters, and assessors that contribute to a common understanding and interpretations of quality components, sources of evidence, and messages to providers. Nonetheless, increased efforts to develop standard training procedures and protocols for use in measurement are needed to support the reliability and confidence in the process.

C. Defining Levels

Among the five QRIS studied, there is generally greater consistency in the definitions of the quality components at the highest rating levels than at the baseline levels. For example, in each of four QRIS, the highest level (or the highest score in a combined rating structure) requires an ERS score of 5.0 or higher. There is greater variation in the threshold scores required at the lower levels, generally due to QRIS planners' perceptions or knowledge of the quality levels in their state at the time the QRIS was launched. This pattern is also evident for child to staff ratios, group size, and other components. At the highest level, the QRIS we studied either require accreditation, or set standards that overlap considerably with recommendations of accrediting organizations such as NAEYC.

The disparity at the lower levels and commonalities at the highest levels result from a number of contributing factors based on what respondents in each QRIS conveyed. First, QRIS planners and key stakeholders in each of the sites were interested in implementing homegrown systems, rather than an exact replica of another QRIS. They held this as important in order to be responsive to the context of their state or locality—the patterns in use of child care, the licensing standards, and the political and economic environment. This development in the state/local context was also important to QRIS planners in order to build the buy-in and ownership necessary to garner support of the QRIS for its launch and implementation. Second, QRIS developers tended to have a shared sense of what the goal for the highest level ought to be, generally based on what they know from accreditation. As a result, the planning process focused on building the QRIS as a lattice between licensing and accreditation (in at least four of the five) even if accreditation is not required at the highest level in all systems. There is not compelling evidence to support the cut-points for each level other than administrators and planners wanting to emphasize the importance of continuous quality improvement and making the steps accessible for providers to make.

D. Data Quality, Coverage, and Use for Research

The five QRIS were purposefully selected due to indications that their data collection coverage and practices may be further along than other QRIS. From these five QRIS, the research team selected three for inclusion in a secondary data analysis also focused on quality measurement, described below.

Data quality. Each QRIS has taken steps in various ways to ensure the quality of the data that are collected and the capacity of their data systems to support program monitoring and, to some extent, evaluation. These QRIS have maximized linkages across databases in order to (1) restrict new record creation and avoid the problems that duplicate records can bring, and (2) pull in common data elements from other databases (such as licensing, subsidy, and PD registries) to minimize the occasions for data entry error and inconsistencies. They also limit data entry across users to localize the entry to those staff most closely connected to any specific data collection process (such as specialists for the pre-rating process, raters, and assessors). They further support consistency in data entry through the use of standardized drop-down menus and preventing the exit from data entry screens if required fields are not entered or are not entered in the correct format.

Data coverage, accessibility, and use in research. Each of the five QRIS collect at least some data on each of the specific quality components included in the QRIS rating. Through the process of this study, we learned the details of the coverage and level of specificity to the data elements collected, how they are stored, and their accessibility for analysis. Based on information about the availability of component-level data (summarized in Table IV.4), we selected three QRIS for inclusion in a secondary data analysis. Table IV.4 shows that four of the five QRIS electronically store component-level data on the QRIS or other accessible databases. However, of these four, Indiana does not conduct ERS assessments. The planned secondary data analysis relies on ERS as the quality outcome because it is the only independent assessment for which data are available in administrative data systems. As a result, we proceeded with the analysis of data from Miami-Dade, Illinois, and Tennessee. The secondary data analysis uses a common metric to capture the shared concepts of quality components and define them across the three QRIS and examines whether the inclusion (or not) of a particular component has an effect on observed quality. The analysis describes the prevalence of quality components using the common metric for comparability across sites, examines the unique effect of each quality component in predicting observed quality, and presents profiles of providers based on the patterns seen within the quality components. This in-depth study report is a companion piece to the secondary data analysis report (Malone et al. 2011) in

that it serves as a foundation for understanding the differences in measurement across the QRIS and the methods of data collection.

E. Research Directions

The five QRIS have reflected an interest and willingness to seek out and respond to research to help guide their development and refinement efforts. Each system has included child care experts and researchers in various aspects of their system development, such as validating the alignment of the QRIS standards with the research base, creating data systems to support program monitoring and evaluation, and evaluating the implementation of the QRIS. To their credit, they have used initial guidance and subsequent findings to continue to inform practice and system refinement. For example, as the research community has focused more recently on the need for individualization of services—such as conducting child assessments and using the information to guide planning, adopting provisions for children with special needs, and incorporating practices to enhance cultural and linguistic diversity—the newer QRIS have included components that capture these dimensions in their standards, and existing systems are making refinements to their standards. These newer components can still benefit from greater specificity; they typically are measured by the presence of some activity, but are not often specified in terms of the mode, specific tool (in the case of child assessments or screening for special needs), or frequency of the activity.

States have also been responsive to the research on observed quality. The majority of QRIS—not just included in this study, but nationwide—have adopted the ERS as the measure of observed quality because this was the predominant measure used in research. As researchers continue to explore the associations between observed quality and child outcomes, a greater focus has been placed on the need for measurement tools that capture the dimensions of quality that may be more likely to affect child outcomes, such as teacher-child interactions. Two of the QRIS included in this study have piloted other measures, such as the CLASS. Future research to identify and refine the measures of observed quality will further benefit the measurement validity and rigor within QRIS.

There remain many unanswered questions about which quality components to include, and how, within the rating systems. Even if measures were developed or better defined within the many dimensions that the quality components cover, we do not yet know which of them—or which combination—are most likely to lead to positive outcomes for children. A first step is better understanding which components contribute to quality improvements. The companion work of the secondary data analysis will examine this issue, as is other ongoing research within and across systems. The challenge in cross-system analysis is the wide variation in the definitions and measurement of the quality components, as described in this report.

A focus for future research on helping to define thresholds, particularly for the highest level, may be the best guidance to support all QRIS given that there is greater similarity in the quality components at this level. The Q-DOT project (Child Care Quality Dosage, Thresholds and Features and Child Outcomes) is examining whether there are certain thresholds of quality above or below which there are stronger associations with child outcomes. In secondary analyses completed across a range of large-scale studies using a variety of quality measures, the project found several threshold effects. Analyses did not confirm that there was a "good enough" level of quality and it was not possible to identify optimal cut-points with certainty. Instead, the project team posits an "active range" of quality, suggesting that programs should focus on a two-pronged approach: supporting lower-quality programs to bring them into the active range where there is a relationship to child outcomes, and encouraging continuous improvement among providers within the higher-quality

range (Burchinal et al. 2011). Further research on what constitutes the active range of quality will help inform planning and ongoing refinement of QRIS.

REFERENCES

Arnett, Jeffrey. "Caregivers in Day-Care Centers: Does Training Matter?" *Journal of Applied Developmental Psychology*, vol. 10, 1989, pp. 541-552.

Barnett, W. Steven, Dale J. Epstein, Megan E. Carolan, Jen Fitzgerald, Debra J. Ackerman, Allison H. Friedman. "The State of Preschool 2010." New Brunswick, NJ: National Institute for Early Education Research. Accessed May 24, 2011 from [http://nieer.org/yearbook/].

Branagh Information Group website. "ERS Data System." Accessed June 27, 2011 from [http://www.ersdata.com/qr.html].

Burchinal, Peg, Kirsten Kainz, Karen Cai, Kathryn Tout, Martha Zaslow, Ivelisse Martinez-Beck, and Colleen Rathgeb. "Early Care and Education Quality and Child Outcomes." Washington, DC: Child Trends, 2009.

Burchinal, Margaret, Yange Xue, Hsiao-Chuan Tien, Anamarie Auger, and Andrew Mashburn. "Testing for threshold in associations between child care quality and child outcomes." Presentation at the Society for Research in Child Development Biennial Meeting, Montreal, March 31, 2011.

Early Learning Coalition of Miami-Dade/Monroe (ELC), Quality Counts website. "Trends from Miami-Dade's Quality Rating Improvement System (QRIS)." Accessed on June 27, 2011 from [http://www.elcmdm.org/QualityCounts/TrendsQRIS710.pdf].

Elicker, James, Carolyn C. Langill, Karen Ruprecht, and Kyong-Ah Kwon. "Paths to Quality: A Child Care Quality Rating System for Indiana. What is its Scientific Basis?" West Lafayette, IN: Purdue University. 2007.

Family and Social Services Administration (FSSA). "Paths to QUALITY Program Enrollment by Provider Type: Monthly Management Report May 2011." Accessed June 27, 2011 from [http://www.in.gov/fssa/files/PTQ_Grid_for_Sharepoint.pdf].

Fiene, Richard, Mark Greenberg, Martha Bergsten, Christopher Fegley, Barbara Carl, and Elizabeth Gibbons. "The Pennsylvania Early Childhood Quality Settings Study." Harrisburg, PA: Governor's Task Force on Early Care and Education. November 2002.

Florida Department of Children and Families. "List of Approved Accrediting Agencies." Accessed May 10, 2011 from [http://www.dcf.state.fl.us/programs/childcare/goldseal.shtml].

Hamre, Bridget K., and Kelly L. Maxwell. "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. June 2011.

Harms, Thelma, Richard M. Clifford, and Debby Cryer. "Early Childhood Environment Rating Scale, Revised Edition." New York, NY: Teachers College Press. 2005.

Harms, Thelma, Debby Cryer, and Richard M. Clifford. "Infant/Toddler Environment Rating Scale, Revised Edition." New York, NY: Teachers College Press. 2006.

Harms, Thelma, Debby Cryer, and Richard M. Clifford. "Family Child Care Environment Rating Scale, Revised Edition." New York, NY: Teachers College Press. 2007.

Harms, Thelma, Ellen V. Jacobs, and Donna Romano. "The School-Age Care Environment Rating Scale." New York, NY: Teachers College Press. 1995.

Illinois Department of Human Services. "IDHS Child Care Program Tiered Reimbursement System: Recommendations of the Child Care and Development Advisory Council Work Group." December 2004. Obtained during site visit in August 2010 for the QRS Assessment project.

Illinois Department of Human Services. "QRS Data." Accessed June 27, 2011 from [http://www.inccrra.org/component/docman/doc_download/174-qrs-data].

Langill, Carolyn, James Elicker, Karen Ruprecht, Kyong-Ah Kwon, and Joellen Guenin. "Paths to QUALITY – A Child Care Quality Rating & Improvement System for Indiana: Technical Report no. 2 Evaluation Methods and Measures." West Lafayette, IN: Purdue University. January 2009.

Lugo-Gil, J., Samina Sattar, Christine Ross, Kimberly Boller, and Gretchen Kirby. "The Quality Rating and Improvement System (QRIS) Evaluation Toolkit." Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. August 2011.

Malone, Lizabeth, Gretchen Kirby, Pia Caronongan, Kimberly Boller, and Kathryn Tout. "Measuring Quality Across Three Child Care Quality Rating and Improvement Systems: Findings from Secondary Analyses." Washington, DC: U.S. Department of Health and Human Services, Administration for Children and Families, Office of Planning, Research and Evaluation. August 2011.

Mitchell, Anne W. "Stair Steps to Quality: A Guide for States and Communities Developing Quality Rating Systems for Early Care and Education." Alexandria, VA: United Way of America, 2005.

National Association of Child Care Resource and Referral Agencies. "2011 Child Care in the State of Florida." Accessed May 24, 2011 from [http://www.naccrra.org/publications/naccrra-publications/publications/SFS-Florida.pdf].

National Association of Child Care Resource and Referral Agencies. "2011 Child Care in the State of Illinois." Accessed May 24, 2011 from [http://www.naccrra.org/publications/naccrra-publications/publications/SFS-Illinois.pdf].

National Association of Child Care Resource and Referral Agencies. "2011 Child Care in the State of Indiana." Accessed May 24, 2011 from [http://www.naccrra.org/publications/naccrra-publications/publications/SFS-Indiana.pdf].

National Association of Child Care Resource and Referral Agencies. "2011 Child Care in the State of Pennsylvania." Accessed May 24, 2011 from [http://www.naccrra.org/publications/naccrra-publications/publications/SFS-Pennsylvania.pdf].

National Association of Child Care Resource and Referral Agencies. "2011 Child Care in the State of Tennessee." Accessed May 24, 2011 from [http://www.naccrra.org/publications/naccrra-publications/publications/SFS-Tennessee.pdf].

National Association for the Education of Young Children. "NAEYC All Criteria Document." 2010. Accessed October 15, 2010 from [http://www.naeyc.org/academy/primary/viewstandards].

National Association for the Education of Young Children. "Teacher-Child Ratios Within Group Size." 2008. Accessed April 6, 2011 from [http://www.naeyc.org/files/academy/file/Teacher-Child_Ratio_Chart_9_16_08.pdf].

National Center for Children in Poverty. "Florida: Demographics of Young, Poor Children." Accessed May 24, 2011 from [http://nccp.org/profiles/].

National Child Care Information Center. "Good Start Grow Smart History." Accessed May 24, 2011 from [http://nccic.acf.hhs.gov/poptopics/gsgs_history.html].

National Child Care Information Center (NCCIC) and the National Association for Regulatory Administrators (NARA). "2008 Child Care Licensing Study." 2010. Accessed October 14, 2010 from: [http://www.naralicensing.org/].

National Center for Children in Poverty. "Illinois: Demographics of Young, Poor Children." Accessed May 24, 2011 from [http://nccp.org/profiles/].

National Center for Children in Poverty. "Indiana: Demographics of Young, Poor Children." Accessed May 24, 2011 from [http://nccp.org/profiles/].

National Center for Children in Poverty. "Pennsylvania: Demographics of Young, Poor Children." Accessed May 24, 2011 from [http://nccp.org/profiles/].

National Center for Children in Poverty. "Tennessee: Demographics of Young, Poor Children." Accessed May 24, 2011 from [http://nccp.org/profiles/].

Office of Child Care, Administration for Children and Families, U.S. Department of Health and Human Services. "Child Care and Development Fund Statistics – 2009 CCDF Data Tables." Accessed June 24, 2010 from [http://www.acf.hhs.gov/programs/ccb/data/index.htm].

Office of Child Development and Early Learning (OCDEL), Pennsylvania Departments of Education and Public Welfare. "Demonstrating Quality: Pennsylvania Keystone STARS 2010 Program Report." Harrisburg, PA: Pennsylvania Department of Public Welfare. November 2010.

Office of Child Development and Early Learning (OCDEL), Pennsylvania Departments of Education and Public Welfare. "Keystone STARS: Reaching higher for quality early learning. Program Report 2010." Harrisburg, PA: Pennsylvania Department of Public Welfare, 2010.

Office of Child Development and Early Learning (OCDEL), Pennsylvania Departments of Education and Public Welfare and The Pennsylvania Build Initiative. "Build Announcement: Child Outcomes Released for Three Pennsylvania Early Education Programs, June 13, 2011." Accessed June 27, 2011 from [http://paprom.convio.net/site/MessageViewer?em_id=10041.0]

Office of Head Start (OHS), Administration for Children and Families, U.S. Department of Health and Human Services. Head Start Program Fact Sheet, FY 2010. Accessed June 24, 2011 from: [http://eclkc.ohs.acf.hhs.gov/hslc/Head%20Start%20Program/Head%20Start%20Program%20Factsheets/fHeadStartProgr.htm].

Office of Planning, Research, and Evaluation (OPRE), Administration for Children and Families, U.S. Department of Health and Human Services. "Miami-Dade Quality Counts: QRS Profile." Washington, DC: Child Trends, April 2010.

Office of Planning, Research, and Evaluation (OPRE), Administration for Children and Families, U.S. Department of Health and Human Services. "Illinois Quality Counts: QRS Profile." Washington, DC: Child Trends, April 2010.

Office of Planning, Research, and Evaluation (OPRE), Administration for Children and Families, U.S. Department of Health and Human Services. "Indiana Paths to Quality: QRS Profile." Washington, DC: Child Trends, April 2010.

Office of Planning, Research, and Evaluation (OPRE), Administration for Children and Families, U.S. Department of Health and Human Services. "Pennsylvania Keystone STARS: QRS Profile." Washington, DC: Child Trends, April 2010.

Office of Planning, Research, and Evaluation (OPRE), Administration for Children and Families, U.S. Department of Health and Human Services. "Tennessee Star-Quality Child Care Program: QRS Profile." Washington, DC: Child Trends, April 2010.

Pianta, Robert C., Karen M. La Paro, and Bridget K. Hamre. "CLASS Classroom Assessment Scoring System Manual." Baltimore, MD: Brookes Publishing. 2007.

Pope, Bingham and Julianna Magda. "Tennessee Report Card & Star Quality Program Year 8 Annual Report." Knoxville, TN: The University of Tennessee College of Social Work Office of Research and Public Service. June 2010.

Pope, Bingham, Joanna H. Denny, Karen Homer, and Kay Ricci. "What is Working? What is not Working? Report on the Qualitative Study of the Tennessee Report Card and Star-Quality Program and Support System." Knoxville, TN: The University of Tennessee College of Social Work Office of Research and Public Service. November 2006.

Raudenbush, S. W., and S. Sadoff. "Statistical Inference when Classroom Quality is Measured with Error." *Journal of Research on Educational Effectiveness,* vol. 1, 2008, pp. 138–154.

Scientific Software Development. "Atlas.ti: Visual Qualitative Data Analysis, Management, and Model Building in Education Research and Business." Berlin, Germany: Scientific Software Development, 1997.

Sylva, K., I. Siraj-Blatchford, and B. Taggart. "Assessing Quality in the Early Years. Early Childhood Environment Rating Scale Extension (ECERS-E): Four Curricular Subscales." Stoke on Trent, UK: Trentham Books, 2004

Talan, Teri N. and Paula J. Bloom. "Program Administration Scale." New York, NY: Teachers College Press, 2004.

Talan, Teri N. and Paula J. Bloom. "Business Administration Scale." New York, NY: Teachers College Press, 2009.

Thornburg, K. "Evaluation Strategies Focusing on Implementation and Outputs." Presentation at the meeting on Evaluation of State Quality Rating Systems, Washington, DC, April 2008.

Tout, Kathryn, Rebecca Starr, Margaret Soli, Shannon Moodie, Gretchen Kirby, and Kimberly Boller. "Compendium of Quality Rating Systems and Evaluations." Report prepared for the Office of Planning, Research, and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. Washington, DC: Child Trends, April 2010.

U.S. Census Bureau. Florida QuickFacts. Accessed May 24, 2011 from [http://quickfacts.census.gov/qfd/states/12000.html].

U.S. Census Bureau. Illinois QuickFacts. Accessed May 24, 2011 from [http://quickfacts.census.gov/qfd/states/17000.html].

U.S. Census Bureau. Indiana QuickFacts. Accessed May 24, 2011 from [http://quickfacts.census.gov/qfd/states/18000.html].

U.S. Census Bureau. Pennsylvania QuickFacts. Accessed May 24, 2011 from [http://quickfacts.census.gov/qfd/states/42000.html].

U.S. Census Bureau. Tennessee QuickFacts. Accessed May 24, 2011 from [http://quickfacts.census.gov/qfd/states/47000.html].

Yin, Robert. Case Study Research: Design and Methods, Fourth Edition. Sage Publications: Thousand Oaks, CA, 2009.

Zaslow, Martha, Kathryn Tout, Tamara Halle, and Nicole Forry. "Multiple Purposes for Measuring Quality in Early Childhood Settings: Implications for Collecting and Communicating Information on Quality." Washington, DC: Child Trends, 2009.

Zellman, Gail L., and Michal Perlman. "Child-Care Quality Rating and Improvement Systems in Five Pioneer States: Implementation Issues and Lessons Learned." Santa Monica, CA: Rand, 2008.

Zellman, Gail L., Richard Brandon, Kimberly Boller, and J. Lee Kreader. "Systems for Early Care and Education and School-Age Care." Research-to-Policy, Research-to-Practice Brief OPRE 2011-11a. Washington, DC: Office of Planning, Research and Evaluation, Administration for Children and Families, U.S. Department of Health and Human Services. June 2011.

THIS PAGE LEFT BLANK FOR DOUBLE- SIDED PRINTING

APPENDIX A

LICENSING AND NAEYC ACCREDITATION REQUIREMENTS FOR CENTER- BASED PROGRAMS

THIS PAGE LEFT BLANK FOR DOUBLE- SIDED PRINTING

Table A.1. State Licensing Standards for Center- Based Programs

QRS Component Category	Florida[a]	Illinois	Indiana	Pennsylvania	Tennessee
Child-Staff Ratio					
Infants	4:1	4:1	4:1	4:1	4:1
Ones	6:1	5:1	5:1	5:1	6:1
Twos	11:1	8:1	5:1	6:1	7:1
Threes	15:1	10:1	10:1	10:1	9:1
Fours	20:1	10:1	12:1	10:1	13:1
Fives	25:1	20:1	15:1	10:1	16:1
Group Size					
Infants	--	12	8	8	8
Ones	--	15	10	10	12
Twos	--	16	10	12	14
Threes	--	20	20	20	18
Fours	--	20	24	20	20
Fives	--	20	30	20	20
Health and Safety[b]	✓	✓	✓	✓	✓
Director Qualifications					
Preservice Qualifications	State-specific credential	CDA or CCP credential; ECE credits	Bachelor's degree	Associate's degree with ECE credits	Previous experience; ECE credits
Ongoing Training	10 hours	15 hours	12 hours	6 hours	18 hours
Staff Qualifications					
Director	✓	✓	✓	✓	✓
Master Teacher	✓	None	✓	✓	None
Teacher	✓	✓	✓	✓	✓
Assistant Teachers	None	✓	None	✓	None
Aide	None	✓	None	None	None

Source: 2008 Child Care Licensing Study by the National Child Care Information Center (NCCIC) and the National Association for Regulatory Administrators (NARA), 2010

[a] 7 counties in Florida have their own licensing standards, but Miami-Dade is not one of those counties.

[b] Criminal background checks including review of criminal history records, fingerprints, child abuse and neglect and sex offender registries

Table A.2. Components Necessary to Meet NAEYC Accreditation for Center- based Programs[a]

QRS Component Category	NAEYC Accreditation
Licensing Compliance	Pre-requisite; License-exempt must demonstrate voluntary compliance
Child-Staff Ratio[b]	Infants 3/4:1 Ones 3/4:1 Twos 4-6:1 Threes 6-9:1 Fours/Fives 8-10:1
Group Size[b]	Infants 6-8 Ones 6-8 Twos 6-12 Threes 12-18 Fours/Fives 16-20
Health and Safety	First aid training for one staff member per group of children; Illnesses and injuries tracked; Updated info from local health authorities; Safe infant sleep practices; Sanitation and food safety; Procedures for child abuse reports
Environment	Indoor and outdoor equipment, materials, furnishings, activities, teacher-child interactions, peer interactions
Staff Qualifications	75% of staff with one of the ff:
Education Level/Credential	CDA or equivalent
ECE Credits	Enrolled in degree program
Continuing Education/In-Service Training	30 hours in past 3 years (plus AA or higher)
Years of Experience	3 years in NAEYC accredited program (plus AA or higher)
Family Partnerships	
Family Resources	Modes of communication and frequency; Family handbook; Information on transitions; Lists of community services
Family Activities	Conferences or home visits; Opportunities for families to gather and plan events
Family Participation in Planning	Methods to solicit information from families; Families' receive assessment results and are involved in planning for child
Administration and Management	
Staff Management	Staff work environment and orientation procedures; Differentiated salary scale and benefits; Professional development plan; Termination and grievance procedures
Fiscal Management	Annual budget, financial record, no evidence of deficit
Administrative Management	Written policies guide program operations, strategic plan, risk management, program evaluation
Curriculum	Clear written framework for curriculum; Activities foster multiple domains of development
Child Assessment	Assessment plan in place with multiple methods and multiple points in time; Tools assess multiple developmental domains are culturally-sensitive and psychometrically-sound; Screening and referral process in place; Individual child assessments inform planning of activities
Provisions for Special Needs	Steps to prevent and respond to behavior problems; individualized plans for children with special needs
Cultural and Linguistic Diversity	Resources available in families' home language; Methods to discuss cultural background with families; Activities and discussions teach the value of diversity
Community Outreach	Collaborations with community agencies and organizations that provide family support services and enrichment activities through cross-trainings, coalitions, joint events, etc.

Source: NAEYC All Criteria Document

[a]NAEYC has 3 sets of criteria for accreditation criteria: required, always assessed, randomly assessed, and emerging practice. In order to become accredited, a program must meet all required criteria, 80 percent of criteria for each standard, and 70 percent of criteria across all standards.

[b]NAEYC recommendations for child ratios depend on group size, and vice versa. The ranges presented show possible combinations of the two features.